Review of Carruthers’ Massive Modularity Thesis

Abstract: According to Carruthers’ (2006) massive modularity (MM) thesis, the central systems of the mind are widely encapsulated and operate via heuristics and approximation techniques similar to those found in computer science. It follows from this, he claims, that widely encapsulated central systems are feasibly tractable. I argue that insofar as Carruthers uses this weakened definition of encapsulation, his thesis faces a dilemma: either is a misnomer (Prinz, 2006) and therefore unrecognisable as a version of MM, or it isn’t, and must put forward a convincing version of MM (Samuels, 2006). I claim that Carruthers’ commitment to this claim about central systems meets this challenge by adopting an understanding of central systems whose information-frugal and processing-frugal operations allow for feasible tractability. I conclude that the CWT provides a plausible and distinctive account of MM.

Keywords: Massive Modularity Thesis (MM), Heuristics, Tractability, Carruthers

1. Introduction

In an elegant defence of Massive Modularity (MM), Carruthers (2006) sketches a revised account of the human mind. His thesis proceeds from two claims. The first claim is positive; it establishes that the input systems (e.g. those responsible for perception) and output systems (e.g. those responsible for motor movement) of the brain are modular in a strong sense.1 The second claim is less positive; it claims that the central systems (e.g. those dealing with reasoning and decision making) are modular in a very weakened, “wide-scope” sense of the word. I’ll call this the Central Wide-scope Thesis (CWT).

CWT: The processes of central cognition are weakly wide-scope encapsulated (WEM).

Carruthers’ focus is on the second of these two claims (i.e. CWT). This claim is often objected to on the grounds that it necessitates a denial of the connection between encapsulation and modularity. On this point Prinz (2006) and Samuels (2006) see Carruthers’ notion of module as a misnomer. I develop this objection into a novel challenge to defenders of CWT. I claim that those who hold CWT face a dilemma. The dilemma is:

The First Horn: WEM cannot, plausibly, be described as encapsulation, and is therefore a misnomer.

The Second Horn: WEM can, plausibly, be described as encapsulation, but must put forward a convincing account of modularity.

1 In what follows I shall use the word ‘system’ interchangeably with the terms module and mechanism.

*Corresponding author: Max Skipper Griffiths, University of Bristol

© 2016 Max Skipper Griffiths, published by De Gruyter Open.
This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 License.
So, either CWT is not close enough to the accepted definition of encapsulation and is therefore espousing a thesis based on a misnomer, or it is close enough to the accepted definition but must put forward a convincing modularity thesis.

So, insofar as Carruthers and others hold CWT, they face an explanatory burden: they must explain either why CWT is not a misnomer, or, if the misnomer challenge is met, why it is not an unconvincing account of MM. I claim Carruthers’ commitment to this negative claim about central systems meets this challenge by adopting an understanding of central systems whose information-frugal and processing-frugal operations overcome tractability problems posed by the Computational Tractability Argument (CTA). I conclude that the CWT provides a plausible and distinctive account of MM.

My argument has five stages. In the first stage, (section 2) I state some preliminary claims. In the second stage (section 3) I explain and clarify Carruthers’ thesis. In the third stage (section 4), I examine the two horns of the dilemma. Then, in the fourth stage (section 5) I look at objections and replies. I conclude in the fifth stage (section 6).

2. Preliminaries

As a preliminary, it is necessary to say something about how I shall view the position held by an advocate of massive modularity (MM). The plausibility of MM, as I will understand it, rests on three main claims. According to the first claim the human mind is composed (largely or completely) of mental modules. I will call this the Module Thesis. According to the second claim the human mind has many of these modules. I will call this the Multiplicity Thesis. According to the third claim modularity is found in central regions of the human brain responsible for cognitive capacities such as reasoning and problem solving. I will call this the Central Thesis. In what follows, I take these claims to represent the basic theoretical commitments that proponents of MM affirm. These claims may or may not be correct. I shall begin by assuming, however, that they are acceptable enough for the MM proponent to endorse.

I will also - following Carruthers (2006) - not take a position on the holistic arguments against MM. In particular, I will not defend against the challenge that central cognition is global: specifically, that central cognitive processes – paradigmatically, reasoning – are sensitive to the properties of the agent’s entire set of background beliefs. Instead, I shall assume for the sake of argument that these arguments pose no threat to the various theses of mind as conceived by proponents of MM.

Another point worth mentioning is how I shall view Carruthers’ position with regard to competing accounts. On this point I will remain agnostic. The Challenge, as I see it, is for Carruthers to put forward his thesis in a convincing way. So, insofar as Carruthers holds CWT, he offers a plausible alternative understanding of central system modules. The challenge, as I shall argue, is for him to present that alternative in a plausible way.

I put forward four conditions he must satisfy to meet this challenge. I propose four basic requirements. Firstly, a convincing account must meet an informative condition. That is to say, it should provide a clear explanation of the underpinning features of its account of modularity.

Secondly, the thesis must meet a ‘significance condition’. It should therefore say something nontrivial about modularity – a point expressed by both those who hold a non-modular view about central cognition (e.g. Samuels (2006)), and those who hold that central cognition is modular (e.g. Sperber (1994, 2002)). To see this more clearly, consider how an uninteresting modularity hypothesis might, for example, espouse the view that the mind is made up of dissociable modules. This may be true, but it is a trivial point and hence uninteresting and insignificant (cf. Samuels, 2006).

---

2 Adherents of this approach, whose original proponents are noted for their work in Evolutionary Psychology (Cosmides and Tooby, 1994; Sperber, 1994; Pinker, 1997; Barrett, 2005; Barrett & Kurzban, 2006), claim that the mind is like a Swiss army knife: assembled of specialised tools (or devices) designed for specific tasks.

3 This is contrasted with those modularity theorists who claim – qua Fodor (1987) – that modularity is only found in peripheral regions of brain architecture.
Thirdly, the thesis must be close enough to the ordinary understanding of modularity to satisfy a ‘consistency condition’. As a result, it cannot simply ignore standardly accepted concepts or be at odds with background theory. It must not draw invalid conclusions from unsound premises.

Fourthly, an adequate account must meet the ‘distinguishable condition’. This condition must be met in order to prevent, for example, a non-modular account masquerading as an MM account. For example, whereas Fodor (2000) thinks that the central regions of the mind are non-modular, Sperber (2006) explicitly denies - he claims that the central regions are modular. Carruthers must situate himself in this debate in a way that doesn't assume both modular and non-modular architecture.

Consequently, I take it that Carruthers must, when clarifying his thesis, satisfy all of these conditions in order to meet the explanatory burden. However, as I will show, whilst he may meet some of these requirements it is by no means clear that he meets all of them.

3. Carruthers on Modularity

In this section I explain Carruthers’ MM thesis. In particular, I focus on the key features of his revised Fodorian criteria. Carruthers believes that brain architecture consists entirely of mental modules. Each module is responsible for a specific feature of cognition, including high-level tasks like problem solving and planning. There are, he claims, five distinctive features that characterise a module. Specifically, a module must satisfy five criteria:
1. Dissociability
2. Domain specificity
3. Mandatory operations
4. Localizability
5. Central inaccessibility

Note firstly that Carruthers’ leaves encapsulation out of the above criteria. So understood, we might reasonably describe this as a unique feature of his criteria – a judgment of distinctiveness that is motivated by comparison with Fodor (1983). According to Fodor (1983) central systems are non-modular while peripheral systems are modular and, he claims, encapsulation is the most distinctive criterion of a module (Fodor, 1983, p. 72-73). I discuss why this is important in due course. But for now it will be useful to begin clarifying these concepts.5

Firstly, dissociability. This notion refers to idea that a cognitive system can become impaired through neurological damage with little or no effect on the operation of other systems.

Secondly, domain specificity. This notion refers to the determinate class of representations that a cognitive system can receive as input (Carruthers, 2006, p. 5). It is, therefore, this class of representations that activate a system to initiate it into operation. Along with the other four features of a cognitive mechanism, this notion places architectural constraints on the representations a cognitive system can calculate. Furthermore, in the course of its processing, the only information a module can access is what it receives as inputs together with a determinate body of information stored within its system, which affects its internal operations – paradigmatically, its ‘subsidiary database’. In addition, and perhaps most importantly, these two features – input to a system and the processing database of a system – should be distinguished from one another (Sperber, 2002; Carruthers, 2003). I discuss the importance of this distinction in due course.

Thirdly, mandatory operations. To say that a cognitive system is mandatory in its operations is to claim that it is automatic – working below the level of conscious control (cf. Bargh and Chartrand, 1999).

Fourthly, neural localizability. According to this modular feature, a cognitive system is associated with

---

4 Fodor’s (1983) criteria consists of nine conditions: mandatory operation, shallow outputs, fast operation, dissociability, fixed neural architecture, characteristic ontogenesis, informational encapsulation, inaccessibility, domain specificity.

5 As Carruthers writes, “if a thesis of massive modularity is to be remotely plausible, then by module we cannot mean ‘Fodor-module’” (Carruthers, 2006, p. 12).
neural localizability to the extent that it is instantiated in neural circuitry – circuitry that may perform more than one cognitive function.

Finally, central inaccessibility. According to the standard definition, an inaccessible system is one whose internal operations and proprietary database are inaccessible to other systems held in the mind (Carruthers, 2006, p. 5). 6

Overall these criteria seem to offer a promising foundation for the proponent of MM. Some of the details may of course be disputed, but for present purposes they serve simply to illustrate the general idea at the heart of Carruthers’ thesis. Furthermore, the criteria do not (at this stage, at least) appear to violate my requirements for an adequate account of MM. That is, the criteria are well explained and descriptively adequate so it meets the instructive condition; it’s consistent with background theory, so does not fail the consistency condition; and it is distinct from alternative accounts (e.g. Fodorian modularity) so it doesn’t violate the distinguishable condition.

Carruthers’ account centres on to two independently credible claims. First, he argues, the input and output systems of the mind are “narrowly” modular or ‘encapsulated’ with respect to the class of representations those systems can receive. Specifically, this class of representations determines the informational properties a module can access and internally process and, hence, the range of inputs it can compute. This information must, according to the standard definition, be “less than all of the information at the disposal of the organism whose cognitive faculty it is” (Fodor, 1983, p. 25). In contrast to (3) domain specificity, encapsulation concerns the class of representations that a system can use once it has been activated through a relevant class of domain specific representations.7

Low-level perception and phonology, neither of which draws on representational states like beliefs or goals, are paradigmatic examples in the literature of narrowly encapsulated modularity (NEM) – a claim accepted by nearly all modularity theorists (e.g. Fodor, 2000; Samuels, 2006; Sperber, 2002; see Prinz, 2006, for arguments against). 8 More precisely:

NEM: a cognitive mechanism M is narrowly encapsulated if, in the course of its processing it cannot be affected by most of the information held in the mind.

The Müller-Lyer illusion, construed as a paradigm of peripheral modularity and an example of NEM, is a case in point. 9 Taking this example we can – plausibly – infer that the system responsible for visual length perception is unable to draw on the belief that, visual appearances notwithstanding, the two lines are of equal length. This system would be unable to utilize other information even if it were relevant for length perception. This leaves us with an archetypal input-module10, which can access and process “sensorily-transduced” information, but can’t access any of the stored information held in the mind (cf. Carruthers, 2006, 59).

For Carruthers, however, the modular status of the central systems require refinement. The refinement concerns the ‘modular status’ of the central systems of the brain. (His focus, therefore, is to build upon the Central Thesis outlined in section 2). As Carruthers sees it, these central systems are WEM; that is, they have a limited access to an unrestricted body of information. As a result, they can receive most of the information held in mind, but cannot process it all at once. I will refer to this widely encapsulated sense of modularity as WEM (Carruthers, 2006, p. 58).

---

6 This argument is problematic, however. One may object, as Carruthers does, that all mechanisms are completely accessible to other mechanisms. But in doing so he fails to acknowledge the possibility that some mechanisms may be able to access some of the information contained within other modules.

7 Four points about encapsulation. First, encapsulation is an enduring and permanent characteristic. Second, it is not a product of performance factors, such as e.g. motivation, knowledge, or attention. Third, it is cognitively impenetrable; that is, it cannot be changed through beliefs or intentions. Finally, the access relations are modal in the sense that they relate to the set of information a system can make use of in the course of its processing (Cf. Samuels, 2006).

8 Another example of NEM is how facial recognition system(s) only has access to information about previously encountered faces.

9 Paradigmatic examples such as length perception or phonological processing are given in the literature.

10 As argued for by Fodor (1983).
**WEM:** a cognitive mechanism M is widely encapsulated if, in the course of its processing, it can access most of the information held in the mind, just not all at once.\(^{11}\)

To summarise, then: whereas \textit{NEM} cannot be \textit{affected} by most of the information held in the mind, \textit{WEM} can \textit{access} most of that information, just not all at once. Specifically, \textit{NEM} can, for instance, access and process “sensorily-transduced” information, but cannot be affected by most of the stored information held in the mind (cf. Carruthers, 2006, 59). And with respect to the abovementioned distinction between the input to a system and its processing data-base Carruthers would claim that within the processing data-base of a \textit{NEM} there would be some determinate body of information that can affect the operations of its system. The implication being that most of the remaining information held in the mind cannot affect the operations of that system (except, of course, that information which is taken as input).

According to Carruthers, \textit{NEM} cannot be a wide-ranging model for how we ought to understand modularity. This is because if we are aiming to characterise central system modularity, then we can no longer think of encapsulation as a matter of isolating the system from stored information. The reason is simple: central systems will often need to operate on stored information as input.

With this in mind, if central systems only require \textit{WEM}, then Fodor’s (1983) argument against central modularity, and in particular his claim in support of the conceptual connection between encapsulation and modularity is, following \textit{WEM}, presented with a challenge (see Fodor, 1983, pp. 107-110). The challenge is to explain why, given \textit{WEM}, we should not deny the following radical claim (RC):

\[\text{(RC) We can deny that NEM is required for modularity.}\]

In light of this, we can – contrary to Fodor – deny the claim that the more encapsulated a central system is (with respect to the amount of information it can process), the more modular the system required to accomplish a task will be (i.e. the nearer to \textit{NEM} it will be). According to Carruthers, either a system can access \textit{no} stored information in executing its algorithms, or:

\[\ldots\text{it can only access a limited data-base of information that is relevant to the execution of those algorithms (in which case the system is encapsulated to a degree inversely proportional to the size of the data-base)}\] (Carruthers, 2006, p. 19).

This may of course be rejected. But it appears to be a good strategy for proponents of MM. Non-modular accounts (e.g. Fodor, 2000; Samuels, 2006), by contrast, cannot take this approach. They explain the information processing of central systems in terms of being strongly global. For instance, Fodor (2000) agrees with Carruthers’ positive claim (outlined in the introduction) that the input and output systems of the mind are modular (the Central Thesis), but rejects CWT in favour of a non-modular approach (see Fodor, 1983, pp. 107–108). I will call this the Global Central Systems Argument (GCA). We can represent this as follows:

1. Central systems subserve global cognitive processes.
2. Global cognitive processes require less encapsulation; that is, they are inversely proportionate.
3. There exists a strong connection between encapsulation and modularity.
4. Hence, (1,2,3) there exists an inverse relationship between global cognitive processes and modularity (i.e. the more global the process, the less modular the system).
5. Hence, (4) Central systems process information globally.

‘Global’ here means, roughly speaking, those cognitive processes that have access to effectively all of our belief system i.e. beliefs, goals and intentions. This quick sketch of Fodor’s argument may of course be disputed. I will assume for present purposes, however, that it will suffice to warrant acceptance since my

\[^{11}\text{Carruthers’ own formulation is as follows: “the system is such that it can’t be affected by most of the information held in the mind in the course of its processing” (Carruthers, 2006, p. 19).}\]
aim is not to assess the plausibility of this argument. Rather, I use it instead firstly to illustrate the difference between Carruthers (2006) and Fodor’s (1983, 2000) approaches to modularity, and secondly, highlight the challenge it poses to a modular approach towards central systems.

One point to mention about this argument is that Carruthers’ RC directly challenges its third premise. Another point, related to the first, concerns how if Carruthers’ CWT is true (i.e. that the processes of central cognition are WEM) and his RC thereby sound, then it follows that Fodor’s GCA is false. And so, according to Carruthers, we can plausibly claim that central systems are modular in the sense of WEM.

However, given the importance of NEM to Fodor’s modularity thesis, and in the absence of strong supporting arguments for WEM, all this challenge represents, as I understand it, is that central cognition might be modular in a non-Fodorian sense – but, crucially, not that its systems are in fact WEM.

Let us take stock. As Carruthers’ sees it, WEM entails that central systems can access most of the information held in mind, but cannot process it all at once. Hence, whilst this does not immediately entail the falsity of GCA it does threaten to challenge it.

But what is the explanatory ground for this class of WEM central systems? We can summarise the answer as Carruthers explaining the existence of central systems in terms of “search heuristics and stopping rules”. Specifically, processing systems whose operations are “both information-frugal and processing-frugal” (and hence, encapsulated in the WEM sense). Which is to say, those central systems must only access a small sub-set of the total obtainable information while executing their tasks (Carruthers, 2006, pp. 16, 22). Carruthers’ stresses this point:

“[…] a module can be a system that must only consider a small sub-set of the information available. Whether it does this via encapsulation as traditionally understood (the narrow-scope variety), or via frugal search heuristics and stopping rules (wide-scope encapsulation), is inessential. The important thing is to be both information-frugal and processing-frugal” (Carruthers, 2006, p. 20).

He arrives at this conclusion via a Computational Tractability Argument (CTA). This argument is essential, I think, for forming his overall picture of brain architecture and function, and more specifically within that architecture, the state, structure and processes of a constituent module. As a consequence, the force of Carruthers’ thesis turns on the soundness of this argument. We can represent it – CTA – as follows:

1. The mind is computationally realised.
2. All computational mental processes are suitably tractable.
3. Only processes that are at least widely encapsulated are tractable.
4. Hence, if all mental processes are to be feasibly tractable and only (mental) processes that are at least WEM allow for such tractability, then all of those processes must be at least WEM.
5. Hence, (4) if all of those mental processes are at least WEM, then the mind is massively modular.
6. Hence, (5) the mind is massively modular.12

Before clarifying CTA, it is necessary to point out that this argument assumes the classical computational theory of mind, according to which cognitive processes can be realised as algorithmically specifiable mental representations. This premise may be rejected. But I shall assume – following orthodoxy – that it is warranted. According to this assumption, cognitive processes can be realised in computational terms. It follows from this that cognitive capacities – such as, for example, visual perception, practical judgments and analogical reasoning – can be modelled in terms of input-output mapping with algorithms on which the systems of those cognitive capacities run. And as a consequence, if all else is equal, we should expect a system with a greater computational load, required to check many items of information, would demand a more complex algorithm. However, simply stipulating algorithms that demand computational resources (memory, processing power, etc.) beyond the level attributed to humans, would be inconsistent with real-time human performance. So, cognitive systems must be computationally tractable in the sense that mental processes are successfully completed in a finite amount of time, and do not require more time and resources,

in the form of memory, information, and computational power, than human beings are reasonably understood to possess. This is called in-practice tractability. As a category of tractability—contrasted from in-principle tractability—this refers to the capacity of a system to undertake state changes at the same time as changes take place in the external environment.

It follows as a consequence of this that the computational theory of mind requires cognitive processes to be practically feasible. This means that the algorithms behind those cognitive processes must be tractable in two ways. First, the algorithms must be an appropriate level of complexity. Second, the computations of those algorithms should only draw on a limited set of information relevant for their computations (cf. Carruthers, 2006, p. 18). So, if a cognitive capacity, characterised by a specific function, cannot be computed within a realistic psychological and computational time limit (i.e. practically unfeasible), then it is said to be intractable.13

Proponents of MM claim that domain specific modules with the property of encapsulation and domain specificity are able to solve this problem. This, they claim, is because such modules only operate on the inputs of a restricted domain, thereby reducing the scope of information they must consider, select and revise in light of new information. And insofar as this is true, those modules process information faster than their non-modular, domain-general counterpart. By contrast, the non-modular view fails to overcome the challenge set because, with unconstrained access to the remainder of cognition a domain-general mechanism ostensibly fails to determine which operations, and which sets of information, are relevant for computing.

According to CTA, only at least WEM are feasibly tractable. Consequently, the mind must consist entirely of at least WEM. Hence, the mind must be MM. According to Carruthers, this argument presents a promising alternative to non-modular approaches (e.g. Fodor’s GCA), which claim that central processes are global. This alternative approach views central systems as systems whose operations are both information-frugal and processing-frugal.

Let us take stock. So far we have seen that a system’s algorithms are such that only a limited amount of information is ever consulted before a given task is finished. Assuming then that CTA is true, the following statement would be false: if cognitive processes are to be feasibly tractable, then the mind must be composed out of processing systems encapsulated in the sense of NEM (Carruthers, 2006, p. 13)). It follows from this, Carruthers claims, that the traditional strong argument for MM, which states the mind is composed entirely of NEM-type modules, would collapse too. But we still have the argument that if cognitive processes are to be tractably realized, then the mind must be WEM.

Hence, for Carruthers central systems do not engage in an exhaustive search concerning all of the elements of information held in the mind— if they did, then that would render those systems intractable. Instead in order to ensure feasible tractability central systems must, together with the sort of approximation procedures found in computer science and Artificial Intelligence (AI), use “search heuristics and stopping rules” (Carruthers, 2006, p. 16)). This approach to human decision-making is not optimal. Rather, it employs a method that is satisficing—producing solutions that reach some threshold level for satisfactory decision-making.

With heuristics Carruthers’ account can avoid time-consuming database searches by identifying a relevant subset of information (beliefs) relevant to a cognitive task. This procedure might not hit upon relevant solutions all of the time, but that is not a problem because human reasoning processes are mostly suboptimal. So, it follows that encapsulation is not the only way to ensure tractability. But what evidence is there for this?

That humans use heuristics in performing cognitive tasks is empirically supported. To see this, consider the Take the Best heuristic. This heuristic requires that a given system draw on additional information concerning the items of information related to the task in question. Crucially, though, it does not look at all the information related to those items. Instead, it “searches for the item of information concerning the two target items that has most often been found in the past to discriminate between items of that type along the

13 The roots of this intractability can be traced back to the well known ‘frame problem’, which Fodor describes as the problem “of putting a frame round the set of beliefs that may need to be revised in the light of specified newly available information” (Fodor, 1983, p. 112).
required dimension” (Carruthers, 2006, p. 14). Gigerenzer et al. (1999) have shown that this heuristic can perform at the same level as far more sophisticated processing algorithms, and all whilst being more frugal in the information that it uses and the requests that it places on the system’s resources (Gigerenzer et al., 1999).

Real-world computational systems can search through an enormous database without incurring any significant computational cost. Internet search engines are a case in point, for they highlight the sense in which frugal, algorithmic processes enable central systems of the mind to calculate information fast enough to be practically feasible; then, Prinz claims that with heuristics “we can avoid exhaustive database searches even when a complete database is at our disposal” (Prinz, 2006, p.33).

This brings into focus the sense in which frugal, algorithmic processes enable central systems of the mind to calculate information fast enough to be practically feasible. And this feasibility is necessary because, all things equal, humans respond in real time to stimuli, and therefore, require central systems with the capacity to execute cognitive processes virtually immediately (i.e. within milliseconds).

Before turning to the discussion of CTA, it is worth developing the account a little further in the context of encapsulation. Carruthers explains that it is frugality, not encapsulation, that is necessary to ensure feasible tractability in central systems, which appears to explain why he omits encapsulation from the five criteria a module must satisfy. On his view, search heuristics and approximation techniques which access no more than a small sub-set of total information available solve the problem of intractability. Therefore, given that this modular property – encapsulation – is left out of Carruthers’ list, it represents a prima facie case for saying that there are strands to CWT that are distinctive. In this sense CWT is distinguishable from non-modular accounts. So, at this stage, it appears to satisfy the distinguishable condition, previously outlined in section 2.

So if CTA is sound, then it strongly supports the core claim behind CWT: that central systems are WEM. As a result, if we are to accept the soundness of CTA we possess a sufficient, direct argument that would render non-modular accounts of central systems (e.g. Fodor’s GCA) computationally intractable. And so, at the very least advocates of those non-modular central systems, such as Fodor (2000), face a prima facie challenge to explain the tenability of those non-modular systems. Carruthers supports this position with a bottleneck argument. It runs as follows. Unlike a multiplicity of modules, general-purpose systems face a tractability problem because by having access to effectively all of our belief system (i.e. beliefs, goals and intentions) those general-purpose systems entail a processing ‘bottleneck’ that would prohibit the formation of beliefs in real time. One may object to this argument (see Samuels, 2006). But this clearly falls outside the scope of the present examination.

Carruthers’ thesis – and especially my brief outline of it – clearly requires further philosophical scrutiny. But, assuming that it withstands this scrutiny it can be thought to provide a promising response to Fodor’s GCA. This is because it gives a central place to the practical utility of believing in search heuristics in explaining how central systems can be responsive to the huge range of information within the remainder of cognition. But there is a complication. One might think that with respect to information processing, this notion of WEM is counter-intuitively unencapsulated – after all encapsulation entail sharp constraints on the amount of information a module can receive does it not?

4. The Dilemma

In this section I outline, clarify and assess the dilemma facing the defender of CWT. My initial focus will be on the first horn: that CWT is a misnomer. I claim that whilst this argument does not immediately entail the falsity of CWT, it presents a strong prima facie challenge to it. The challenge is to explain why WEM is modular. Following this I then claim that, assuming CWT is able to withstand this challenge (i.e. can be plausibly described as modular) it falls on the second horn of the dilemma according to which an explanation must be given as to why it is not an unconvincing account of MM.
4.1 Misnomer

Given the foregoing discussion, defenders of Carruthers might be tempted to think that CWT provides a plausible explanation of central cognitive systems. On this view, not only is CWT a sound metaphysical thesis in that it offers a plausible account of the sense in which properties of the mind exist, but also, in telling us how we should engage in normative enquiry about those properties, it presents us with an epistemological claim; and furthermore with a semantic one, since it tells us about the relationship between a specific set of theoretical words and their meanings.

Despite its promise however, Carruthers’ understanding of encapsulation (i.e. NEM and the non-standard WEM) is problematic. One obvious problem concerns the inaccuracy of the use of the term encapsulation. The idea here is that, insofar as WEM can “draw on most of the remainder of cognition, though not all at once,” one may object that it cannot be reasonably understood as an instance of encapsulation. As such CWT threatens to miss the consistency condition to make it an adequate account of MM. Therefore, Carruthers’ commitment to the view that central systems are, via heuristics, WEM belies an understanding of central systems at odds with encapsulation, as it is standardly understood. This is the first horn of the dilemma. I will call it the Misnomer argument (MA); it establishes that if Carruthers’ thesis works, and hence is true, then it follows that WEM is not, by definition, encapsulated. This argument – MA – can be represented as follows:

1. If Carruthers’ thesis is true, then central modules are WEM.
2. If central modules are WEM, then most of the information in the human mind is accessible, but not all of it can be accessed at once.
3. But encapsulation implies restrictions on information into a module.
4. Hence, WEM does not have the distinctive property of encapsulation required to make it modular.
5. Hence, either Carruthers’ thesis is false, or WEM is modular even though does not have the distinctive property of encapsulation required to make it modular.

In order to make this stand out, consider that to overcome the first horn of the dilemma Carruthers must somehow show that WEM (i.e. a limited access to an unrestricted body of information) can be legitimately described as an instance of encapsulation (and hence, modularity). This argument – appealed to by Samuels14 - therefore reflects a warranted scepticism about the plausibility of CWT, questioning whether Carruthers is entitled to call WEM modular.

I do not think, however, that MA is a knockout argument against CWT. It fails to be a knockout argument because premise 4) is not straightforward. That is to say, it is not obviously clear that WEM is unrecognisable as an instance of encapsulation. If it is not, then the resulting argument fails to establish premise 5). Indeed, defenders of Carruthers may object to MA on the basis that insofar as central systems, as they conceive of them, cannot at any one time process all of the information held in the mind, then those systems meet the core condition required for encapsulation. This core condition, for Carruthers, is that those central systems “can’t access more than a small sub-set of the total information available before completing their tasks” (Carruthers, 2006, p. 16). Specifically, if computational tractability requires WEM, then for each system and sub-system it must be the case that:

1. We are able to identify and differentiate its input from its processing database (assuming it has one).
2. The system’s processing database must contain only a small sub-set of the total amount of information available.

If heuristics propose frugality without either (1) or (2), then it follows that CTA would threaten to undermine CWT. But given that the operations of WEM are (via heuristics) both information-frugal and processing-frugal, we can reasonably claim that those systems only access a small sub-set of the total accessible information (from the total set they can access) during their processing. And therefore WEM arguably meets

---

14 Though Samuels (2006) only mentions this as a brief remark, in contrast to the lengthy formulation and discussion that I have given it here.
the abovementioned requirements (1) and (2). The result is that WEM appears to be both encapsulated and tractable. And therefore there is good reason to believe that MA is a straw man and a false argument.

Against this, however, one might claim that you cannot incorporate heuristics into a modular architecture. Hence, on this view, by acknowledging the merits of heuristics, Carruthers allows that the argument for encapsulation vis-à-vis central cognition is unsuccessful since encapsulation is replaced with heuristics.

Two responses. The first is to mention that, like an encapsulated module a heuristic module can ease the cognitive load through adopting the satisficing heuristic of stopping a search when it has found an item of information that is acceptable for use in its current task. Furthermore, simple heuristics offer a way to regulate interactions among modules, thereby engendering limits on the information a module can access during the course of its processing. For example, the mind-reading system, which is required to query a large range of other systems for data relating to social norms, values and behaviours, could employ heuristics to acquire that data (cf. Carruthers, 2006, p. 11).

The second response is to say that simple heuristics can fit within modular architecture, as shown by the ecological rationality heuristic. According to the ecological rationality heuristic, there exists a range of different environments in which heuristics will function in an accurate and reliable way. We can think of these heuristics as having been selected (under evolutionary pressures) to solve problems within those environments. This requires that they hold within them, or ‘encapsulate’, a level of information that is sufficient for its immediate objectives – in such a way as to be described modular.

The Recognition heuristic, according to which when an agent only recognises one of two items they should assign the recognized item a higher value, is another case in point. This is because the only information concerned with the recognition calculation needs to be accessed. And so, the recognition heuristic merits the standardly understood label, ‘module’.

However, one might object on behalf of MA that the problem here has to do with Carruthers explanation of CWT. The reason is this. If Carruthers does rely on this problematic understanding of WEM, then it seems strange that it is only briefly discussed near the end of the 2006 compared to his lengthy discussion of the supporting arguments from biology, task specificity and computational tractability. Furthermore, this is confounded by the fact that this is not a thesis either explained by or argued for in the writings of any other theorists of modularity. Carruthers is putting forward the non-standard view. It is he who has applied problematic vocabulary in the context of central systems. It follows, then, that it is he who must bear the explanatory burden that that necessitates.

One response to the first challenge is to concede that the CWT is vague and ill-defined but suggest that, properly understood, this concession does not undermine the account. And this is because lack of specificity is an objection that can (plausibly) be raised against other modularity architecture, since there are several ways in which we might describe the concept ‘module’ (see: Prinz, 2006).

Both the scope and precision of this defence are problematic, however. Simply pointing to the vagueness of other modular architectures is a weak form of counter argument. If ambiguity is a quality that can be proven to apply to those other accounts, then so be it; the proponent of this view can easily show this. However, this does not support the stronger claim that these rules should be substituted by the CWT.

One other means by which a defender of the CWT might answer this problem is to point out that Carruthers’ thesis does not appear to violate my criteria for an adequate account of MM. The reasons are as follows. The CWT is well explained and descriptively adequate, so can be said to meet the informative condition; it is consistent with background theory on MM, heuristics, and computational theory, so does not fail the consistency condition; and finally, it is distinct from alternative accounts (e.g. Fodorian (1983)), so does not violate the distinguishable condition. It seems to follow that WEM can overcome the first horn of the dilemma (that it can, plausibly, be described as encapsulation). As for the second horn, I shall now assess whether it meets this challenge, that is to say, whether it is a convincing account of MM.
4.2 Convincing Account

In order to assess whether CWT is a convincing account, I shall now look at whether CWT meets the remarkability and consistency conditions required to satisfy the explanatory burden. I focus only on these two for I take it that the distinguishable and informative conditions have been met.

**Remarkability**

One objection to CWT is that it is an uninteresting account of MM. This objection runs as follows. WEM is simply what you get by denying exhaustive search; but given that no one thinks exhaustive search is characteristic of human cognition, then WEM is simply stating the obvious point that the mind makes use of heuristics; therefore, CWT is neither distinctive nor interesting (cf. Samuels, 2006, p. 45).

Note firstly that “interesting” here may either be referring to CWT as a general hypothesis or its form of modularity (number and degree of modules) and its function (i.e. to avoid intractability).

Let us begin with its form and function. We can distil the essence of Carruthers account in terms of modules processing the inputs of a small sub-set of the information held within cognition. This is not a unique picture of the mind. But the use of heuristics is. MM formulated with this notion is interesting not only because it overcomes tractability problems via a programme of heuristics consistent with encapsulation, but also because it claims there are a large number of such mechanisms (Samuels, 2006, p. 39). This is interesting not only because CWT argues that central cognition is modular in the teeth of evidence that central cognitive processes are global, but also – as shown already – is distinguishable from its modular and non-modular counterparts. Insofar as this is true, CWT provides an interesting account of cognitive architecture.

I also take it that as a general hypothesis Carruthers’ MM thesis is distinctive and interesting. This is because it is sustained by, and consistent with, the understanding of the simple heuristics research program. And not only is employing heuristics to solve tractability problems interesting but the *way in which* it is argued for is, too. As such, we have good reason to describe the account as a convincing account, which answers the challenge of the second horn. I will now assess this claim, alongside questions concerning the consistency condition, by looking at two arguments given in support of WEM.

**Consistency**

At the outset of *Architecture of the Mind* Carruthers introduces three arguments for massive modularity – the argument from biology, the argument from task specificity, and the argument from computational tractability. Together with an analogy from AI, these arguments are essential for forming his overall picture of brain architecture and function, and more specifically within that architecture, the state, structure and processes of a constituent module. This, one can argue, shows just how convincing Carruthers account is.

There is reason for scepticism, however. This scepticism concerns how these arguments fail to marshal evidence in favour of the particulars of CWT. And given that the force of his thesis lies in the strength and range of this and the two other theoretical arguments for which he proposes a MM of brain architecture and function, a failure in any one of these arguments threatens to undermine the force of CWT. I shall consider one of these arguments, as well as the analogy from AI.

We can represent the argument from biology (AB) as follows (Cf. Carruthers, 2006, p. 25):
1. Biological systems are designed systems, constructed incrementally
2. Such systems, when complex, need a massively modular organisation
3. The brain is (i) a biological system, and (ii) complex
4. So the brain is massively modular in its organisation
In defence of premise (1) and (2) of his argument, Carruthers asserts that there exists a wide range of evidence from across biology that “complex functional systems are built up out of assemblies of sub-components”. Following the proposition behind this statement – that functional complexity demands structural complexity – and serving as the seat of brain architecture and function, Carruthers concludes that the brain is massively modular in its organisation.

This is an intuitively plausible argument. But, as I shall now suggest, it still does not follow from AB – as Carruthers intends – that the mind is massively modular. The reason is the following. All that AB establishes is a strong case for the multiplicity thesis (discussed earlier in the preliminaries). All it shows, therefore, is that the mind is hierarchically organised into dissociable subsystem. It does not, however, follow from this that the central systems of the mind are massively modular (i.e. dissociable, domain specificity, and so forth). Furthermore, the claim that central cognition contains decomposable, dissociable subsystems is not inconsistent with non-modular accounts, which also espouse such views.

In light of this, we lack reason to believe that Carruthers can legitimately infer massive modularity of mind (in function) from massively modular organisation and structure – and AB misses its mark.

One means by which Carruthers might respond is to turn to the argument from AI. The argument from AI proceeds by identifying that during the last decade, researchers within AI have converged on modular architectures. This, so the argument goes, has been something they have converged on through trial and error, with the implication that successful design systems are structurally organised in a modular way. Indeed, many theorists believe that to ensure tractability, computational processes need to be divided into modular systems and subsystems (McDermott, 2001). Gigerenzer suggests “smart robots need to be ... equipped with special-purpose abilities without a centralized representation and computational control system” (Gigerenzer, 2001, p. 43). In light of this, we have reason to believe that human cognitive architecture has a modular organization. Whilst this argument may not be sufficient to claim MM is true, it provides further justification for believing that conclusion.

Non-modular systems have in practice turned out to be intractable while modular systems have not had any such problems. As a result, AI and robotics provide a strong case for MM, and in particular the need for modules. These modules – which Brooks (1999) calls “reactive behaviours” – are strongly encapsulated devices whose internal processes engender specific “predefined responses to environmental conditions” (Brooks, 1999, p. 90). However, the problem that emerges is whether the notion of “module” in this argument from AI bears any resemblance to, or in any relevant way describes, the modular architecture found in CWT. At first blush, it appears to describe a NEM-type module, not WEM. So the burden of proof rests with Carruthers to show why this is not the case.

Despite these shortcomings I believe that this still does not constitute a significant enough case against CWT to warrant the claim that it is unconvincing. Instead, I take it to simply weaken the force of CWT’s argument to the extent that AB and the argument from AI are not sufficient to establish the truth of Carruthers’ thesis. In the context of justification, this means that CWT is not completely truth-tropic (i.e. that it will always, in every relevant instance, yield true beliefs or conclusions).

5. Objections and Replies

I’ll now consider two objections to my discussion of CWT:

Objection 1.
The objection concerns my failure to acknowledge the strength of objections to the MM thesis. Specifically, one strategy for undermining the MM thesis is via a developmental neurobiological argument. Consider, for example, Buller and Hardcastle’s (2000) argument that there exists a discrepancy between recent neuroscientific evidence of neurobiological plasticity and the view of innateness held by proponents of MM, where innateness is understood in the Fodorian sense of developing according to “determined patterns under the impact of environmental releasers” (Fodor, 1983, p. 100). It is a consequence of this, they claim, that the massively modular position is undermined and suffices to warrant rejection.
One response to this challenge is to reject the strength of plasticity premise of the above argument. Specifically, there exists strong evidence that the cerebral cortex develops under the control of a particular set of genes whose specific, endogenous operations produce the neurons responsible for cognitive processes such as thought, perception and memory. In light of this evidence, the developmental plasticity premise of Buller and Hardcastle’s argument is insufficient to preclude innate specification of brain architecture.

In this case, Buller and Hardcastle’s argument is too quick. I claim that the soundness of my counter argument in favour of innateness renders a more balanced understanding of brain architecture, namely that the above discussion illustrates how the judgment of likelihood associated with the influence of innateness and developmental plasticity on brain architecture admits of degrees. Hence, for the most part the view of innate specificity and developmental plasticity of the brain will depend upon empirical considerations from cognitive science/neuroscience, and the strength of their net contribution, measured in degrees. Given these considerations, a module is more likely the product of the joint net effect of both innate and developmental processes.

Suppose for the sake of charity to Buller and Hardcastle that the above response fails to defuse the developmental neurobiological argument. It would still not follow that MM is false. This is because the proponent of MM need not take a strong stance on innateness at all, as with for e.g. Carruthers (2006) and Tooby and Cosmides (1992). Hence, given that Buller and Hardcastle falsely presuppose innate specification as a necessary condition of MM, when no such condition exists, and given that nothing in my argument turns on this, we can put this matter to one side.

Objection 2.
The second objection concerns what these theoretical models bring to bear on scientific research of cognitive architecture. The argument is as follows. To the reader not familiar with the literature, discussions within the modularity debate may appear to be philosophers trying to answer empirical psychological questions through a priori reasoning.

One means by which a defender of modularity theory might answer this problem is to claim that on a theoretical level modularity specifies theoretical, evolutionary explanations of mental phenomena and on an experimental level modularity allows for explanations that are plausible in relation to data collected from comparative psychology. Furthermore, the modularity framework from which its theories develop can be fallible (i.e., believing something that is false, e.g., every event having a cause) and defeasible (i.e., shown to be wrong by further evidence).

In addition, modularity theory not only provides explanations of empirical data such as, for example, the speed of cognitive processes and behavioural dissociations at the behavioural level; it also obtains evidence at the psychological and neurological level as with, for example, functional dissociations corresponding to neuropsychological case studies of language (Prinz, 2006). Modularity theory must therefore comport with what we have good reason to believe about human reasoning from detailed empirical inquiry, theoretical grounds or both.

It is possible to argue against this and claim that a priori arguments should not be used since questions about cognitive architecture are to a very large degree empirical. In this case however, the sceptical challenge is too quick. The concept of modularity is not undermined. This is because in response the defender of modularity theory can argue that insofar as philosophy embraces naturalism their inquiry is continuous with scientific inquiry. As such we should not view philosophy and the natural sciences as non-overlapping forms of inquiry but rather as complimentary disciplines. And insofar as empirical neurophysiological evidence is limited the former arguably supplements the latter.

Not only, then, have a priori arguments been used but they have been correctly used, beginning with empirically based premises (derived from a posteriori observable facts) and ending with logically warranted conclusions. It is, therefore, subject to empirical falsification. Furthermore, it is fairly normal in science to use what are in a sense “a priori” arguments to argue for empirical conclusions. These involve considerations of general plausibility, descriptive adequacy and so forth and allow us to determine how well an explanation coheres with our current system of beliefs rather than directly appealing to empirical
evidence. With this in mind, I think that the appeal to computational tractability in the modularity of mind debate is not solely philosophical rather than empirical. Rather, it seems to be a combination of both forms of inquiry.

6. Conclusion

I have presented a challenge to Carruthers’ CWT. The challenge is that insofar as Carruthers holds CWT he faces an explanatory burden. The explanatory burden is to explain either why CWT is not a misnomer or, if the misnomer challenge is met, why it is not an unconvincing account of MM. I have shown that Carruthers successfully meets this challenge by adopting an understanding of central systems whose information-frugal and processing-frugal operations overcome tractability problems posed by CTA. I conclude that the CWT provides a plausible, distinctive and convincing account of MM.

References