Against the “system” module

John Zerilli

To cite this article: John Zerilli (2017) Against the “system” module, *Philosophical Psychology*, 30:3, 235-250, DOI: [10.1080/09515089.2017.1280145](https://doi.org/10.1080/09515089.2017.1280145)

To link to this article: <https://doi.org/10.1080/09515089.2017.1280145>

Published online: 25 Jan 2017.

Submit your article to this journal [↗](#)

Article views: 149

View related articles [↗](#)

View Crossmark data [↗](#)

Against the “system” module

John Zerilli

College of Arts and Sciences, Australian National University, Canberra, Australia

ABSTRACT

Modularity is a fundamental doctrine in the cognitive sciences. It holds a preeminent position in cognitive psychology and generative linguistics, as well as a long history in neurophysiology, with roots going all the way back to the early nineteenth century. But a mature field of neuroscience is a comparatively recent phenomenon and has challenged orthodox conceptions of the modular mind. One way of accommodating modularity within the new framework suggested by these developments is to go for increasingly soft versions of modularity. One such version, which I call the “system” view, is so soft that it promises to meet practically any challenge neuroscience can throw at it. In this paper, I reconsider afresh what we ought to regard as the *sine qua non* of modularity and offer a few arguments against the view that an insipid “system” module could be the legitimate successor of the traditional notion.

ARTICLE HISTORY

Received 20 July 2016

Accepted 29 December 2016

KEYWORDS

Cognitive dissociations;
modularity; neural reuse;
neuroplasticity

1. Introduction

A familiar trope of cognitive science, linguistics, and the philosophy of psychology over the past forty or so years has been the idea of the mind as a modular system. In the context of contemporary psychology, a modular system is, typically, one consisting of functionally specialized subsystems responsible for processing different classes of input (e.g., vision, hearing, human faces, etc.), or at any rate for handling specific cognitive tasks. The general motivation for this hypothesis is the belief that the mind is both physiologically and anatomically versatile. Though no doubt a plausible, methodologically fruitful, and highly influential idea in its own right – thanks in no small part to Fodor (1983), whose pioneering effort gave it contemporary theoretical expression and substance – modularity entered the scene in a big way at just about the time that saw the arrival of a new and potentially subversive force in the behavioral and brain sciences, the mature field of neuroscience, which despite earlier beginnings only really came of age in the late twentieth century. One of the outstanding achievements of neuroscience in the past thirty years has been the discovery of the brain’s lifelong powers of renewal

CONTACT John Zerilli john.zerilli@anu.edu.au

This article was originally published with errors. This version has been corrected. Please see Erratum [10.1080/09515089.2017.1366517](https://doi.org/10.1080/09515089.2017.1366517).

© 2017 Informa UK Limited, trading as Taylor & Francis Group

and reorganization. Neuroplasticity has, for better or worse, challenged many of the orthodox conceptions of the mind that originally led cognitive scientists to postulate mental faculties. Similarly, rapidly accumulating neuroscientific evidence of the reuse or redeployment of neural circuits, revealing the domain-general and interactive structure of brain regions, has upset basic assumptions about the relation of function to structure upon which modularity – not to say neuroscience itself – originally depended (Anderson, 2010, 2014; Dehaene, 2005).

There is a conspicuous lack of consensus surrounding the status of modules as neuroanatomical entities, in part because modularity has proven itself to be a highly versatile concept sustaining different research agendas across the biological and mind sciences. Are they functionally dedicated, innately designated (species-constant) regions of wetware whose operations may be described by algorithms (Quartz & Sejnowski, 1994, p. 726) or are they of the nature of software systems, having no phylogenetically necessary relation to specific cortical sites, be they dedicated or otherwise? Is there indeed room for both types, or for hybrids combining features of both types (Horst, 2011, pp. 224–225, 261–262)? Fodor, perhaps sensing that the real interest of modules lies partly in their functional/neural dedication and ontogenetic robustness, considered that the first description could serve as a paradigm of modularity – a view which has the merit of being in broad agreement with the neurosciences (Anderson & Finlay, 2014, p. 5; Bechtel & Mundale, 1999; but cf. Doidge, 2007, pp. 291–297; Gold & Roskies, 2008, p. 354). Nevertheless, in recent decades, enthusiasts of modularity have been more willing to throw their lot in with alternative proposals or otherwise endorse increasingly anodyne suggestions about what a module really amounts to. Apart from the general explosion of discoveries in the neurosciences, new and dramatic evidence of the precise extent of neuroplasticity and neural reuse has necessitated a shift of emphasis away from implementation. The innateness hypothesis alone looks to be disastrously discredited if the potential for neuroplasticity is indeed as advanced as it appears, since it underscores the crucial role that learning must play in the acquisition of competencies otherwise presumed fixed or defined by characteristic ontogenetic pace and sequencing. The evidence of neural reuse, for its part, indicates that high-level cognitive tasks such as language processing are enabled by highly distributed neural networks comprised of very many smaller brain regions or nodes that are themselves multifunctional and domain-general: the selfsame circuits are redeployed over and again across different tasks and task categories. This discovery potentially undermines the claim that such high-level cognitive feats reflect domain-specific competencies. Quandaries like these have understandably motivated the attempt to rescue the theory through a renewed emphasis on computational design (Anderson, 2010; Anderson & Finlay, 2014, p. 5; Jungé & Dennett, 2010). My concern in this paper is to question this move and to advocate for the view that functional specialization qua dissociability ought to be regarded the *sine qua non* of modularity.

2. The “system” module and dissociability

A common objection to accounts of cognitive architecture which downplay or question the modular hypothesis is that modularity has not been given due credit as the uniquely versatile concept that it is, and that the dissenters have simply fettered themselves with an impossibly narrow and needlessly structuralist conception of cognitive architecture that is unwarranted in all the circumstances (the circumstances being the Cognitive Revolution, the fact that no one seriously denies that the mind has a rich internal structure, the unquestionable boon of functional decomposition as an effective research strategy in the cognitive sciences, etc.). Jungé and Dennett appear sympathetic to this point of view: “A software theory of massive modularity – programs evolved to serve particular adaptive functions within brains – without commitments about implementation (unlike anatomical modularity) could survive [the evidence of neural reuse] largely untouched” (2010, p. 278). At issue here is whether a nondissociable “system” of some variety could be regarded as modular – whether, say, a language or norm acquisition device comprised of very many smaller, domain-general neural regions could in some sense be a module. Against this suggestion is the claim that functional dissociability ought to fix a definite threshold beneath which a system cannot be regarded as modular.

Recent developments in neuroscience have no doubt added to the luster of the “system” module, as I shall call it, and even encouraged the view that such modules represent what was always the most important contribution of modular theories to our understanding of the mind (see section 3). But actually, the system module has been around for a long time. Its fortunes can nowhere be more illuminatingly charted than in the annals of generative grammar. Generative grammarians are notorious for prevaricating on the issue of linguistic modularity – one can easily locate passages which would suggest the modularity in question is anatomical or at the very least functional (Chomsky, 1975, pp. 40–41; Chomsky, 1980, pp. 39, 44; Chomsky, 1988, p. 159; Chomsky, 2002, pp. 84–86; Collins, 2008, p. 155; Fitch, Hauser, & Chomsky, 2005, p. 182; Pinker & Jackendoff, 2005, p. 207) and others where what they seem to have in mind is little more than a “domain of inquiry” – “the view that internal cognitive systems can fruitfully (for purposes of successful theory-construction) be studied independently of other such systems” (McGilvray, 2014, p. 235; see also Chomsky, 2005, p. 5). Notice that the module-as-domain-of-inquiry very effectively neutralizes the sting of neuroscientific evidence, since there is really no evidence that neuroscientists can adduce against the existence of such a module (a point to which I return in section 3). Indeed, the system module is frequently endorsed by playing down the significance of implementation and emphasizing its “methodological value as a research heuristic” (Badcock, Ploeger, & Allen, 2016, p. 11; see also Scholl, 1997). But let us return to the other theme of this section, the notion of dissociability.

In a straightforward sense, a system is dissociable if it is functionally specialized – if it can (in principle) be modified without directly impeding the operation of a comparable system.¹ More precisely, if a neural system n consisting of neural primitives $\{p_1, p_2, p_3 \dots p_n\}$ contributes some specific and functionally discrete operation f such that all (or most) elements of the set $\{p_1, p_2, p_3 \dots p_n\}$ are dedicated to f , then n will be dissociable. On this understanding, a speech production center will be dissociable if its impairment has no direct effect on any system “considered with the same grain of analysis” (Carruthers, 2008, p. 295) (e.g., numeracy, rhythm, speech comprehension, episodic memory, IQ, etc.) even though it might ramify to compromise a higher level functional system that draws upon the speech production center for processing (e.g., singing, signing, etc.).

As I argue in detail elsewhere (Zerilli, 2017), the evidence of neural reuse strongly suggests that the only such dissociable unit we are likely to encounter in the brain will be one that resembles the neuroscientific notion of a module (and even *that* much is not guaranteed).² The neuroscientific module is sometimes called a “brain module” or “cortical module” (Gold & Roskies, 2008; Mountcastle, 1978, 1997; Pascual-Leone & Hamilton, 2001; Rowland & Moser, 2014; Zador, 2015), other times a “cortical column” or “columnar module” (Amaral & Strick, 2013; Buxhoeveden & Casanova, 2002; Mountcastle, 1978, 1997; Zador, 2015), still at other times an “elementary processing unit” (Kandel & Hudspeth, 2013), or simply an “operator” (Pascual-Leone & Hamilton, 2001; Pascual-Leone, Amedi, Fregni, & Merabet, 2005). It corresponds roughly with the node of a neural coactivation graph,³ and is known to perform only exiguous subfunctions such as aspects of edge detection or depth discrimination – certainly nothing as high-order as language acquisition or norm acquisition per se. In the context of neural reuse, then, we may presume that the impairment of a cortical module will ramify to all higher level functional composites in which it plays an active role, and yet so long as no other cortical module is directly put out by such an intervention, it remains sufficiently discrete to be regarded as dissociable in the relevant sense.

Notice that when spelled out in this way – and all I have done is follow through with the logic of dissociability as it is commonly understood (see, e.g., Carruthers, 2008, p. 258; Coltheart, 2011, pp. 227–228) – the requirement potentially loses some of its explanatory power. For what it entails is that the smaller and more functionally promiscuous a neural system gets – remembering that neural reuse itself implies that the typical brain region will be both extremely small and highly versatile – the more difficult to quarantine the effects of regional impairment, since those effects will presumably ramify to all affected distributed systems. Nothing theoretically significant (from the point of view of evolutionary psychology) can follow from the fact that a tiny brain region is dissociable if its impairment will disturb the operation of (potentially) many higher level cognitive systems. It is only when modules directly implement high-level cognitive functions (e.g.,

sentence parsing, cheater detection, face recognition, and the like) that the effects of modification can be contained in a way that makes dissociability an important constraint on cognitive theory. For then evolution itself can have a clear role to play in shaping cognitive systems by selectively modifying brain regions in a way that does not reverberate detrimentally across distributed systems. This indeed was thought to be a major argument in favor of modularity – the neat solution it offered to the so-called evolutionary “debugging” problem (Cosmides & Tooby, 1994). By contrast, any account of modularity in which modules come out as small and promiscuous is an account of modularity which no longer promises to solve the debugging problem. And any criterion of modularity which casts modules in such a diminutive role cannot be considered especially salient (at least from the point of view of evolutionary psychology).

Now I am defending dissociability as a criterion of modularity. My position must therefore seem a little curious, for am I not by defending dissociability actually defending the wrong sorts of modules – given the sorts of modules that this criterion delivers if neural reuse is pervasive? I can certainly see how an evolutionary psychologist would be puzzled by my position. But, as has been argued extensively elsewhere, the evolutionary psychologist’s own position on the question cannot be easily defended. Frankly, the sorts of modules she is after are very unlikely to be found anywhere beyond the most primitive domains, and the search for them at all reflects a misunderstanding of the brain and its evolution (see, e.g., Anderson & Finlay, 2014). Like it or not, therefore, it looks as if we are going to have to rest content with a diminutive role for modules – which may not be such a bad thing anyway, for while dissociability may not ultimately meet the desiderata for a theory of evolutionary psychology, it ought to safeguard a respectable threshold for modularity nonetheless. It furnishes a kind of cognitive movable type for the mind, and mechanisms that can support robust laws, generalizations, and predictions (e.g., “forward” inferences from cognitive tasks to brain areas) (Burnston, 2016). If, for a given neural area A, there is some univocal description D such that D explains the functional role of A’s activity whenever A functions, it should be possible to formulate a theory tokening A providing “functional descriptions that apply over a range of instances of functioning,” and “functional explanations in particular contexts that are relevant to contexts not yet explored” (Burnston, 2016, pp. 529, 531). This would be a “*very* powerful theory in terms of generalizability and projectability” (Burnston, 2016, p. 531). Brain regions that are domain-general in the way envisaged by theories of neural reuse may of course ultimately prove *not* to sustain completely generalizable and projectable accounts of local function in view of the effects of both neural and ecological context (see, e.g., Klein, 2012). Even so, this would not affect my argument: I am not arguing that the mind/brain is modular; I am arguing rather that an optimal interpretation can be given to this claim.

3. Against the system module

So what, then, of Jungé and Dennett's suggestion? The problem, as I see it, is that it confuses modularity with faculty psychology more generally, and so reduces it to a platitude. Being neither controversial, falsifiable, or particularly interesting, the system view fails to live up to the theory's venerable reputation. On such an expansive definition, who would *not* emerge as a defender of modularity? Certainly few theorists in the cognitive neurosciences would deny the utility of functional decomposition as an effective research strategy (Anderson, 2010, p. 249; Anderson, 2014, p. 106; Prinz, 2006, cf. chapter 4; Boone & Piccinini, 2016; Piccinini & Craver, 2011).⁴ And of course a high-level cognitive system composed of shared neural elements might well exhibit natural kind properties, such as a systematic set of procedures for dealing with typical inputs (Chomsky, 1980, 2006; Pinker, 1997). But it is difficult to see how such a definition could have any substantively worthwhile theoretical upshots, certainly of a kind that could possibly justify the enormous effort spent in advancing modularity as some sort of solution to a deep and longstanding set of issues. On this weak view, what would the modularity of cognition explain about cognition beyond the simple fact that the mind, too, may be investigated using the techniques of natural science (i.e., “divide and conquer” works here too)? If the answer is “not much,” this cannot be a good account of modularity – assuming that by “modularity” we mean a substantive doctrine.

On the weak construal, modules turn out to be little more than fruitful perspectives on the mind, the mind considered from this or that particular point of view, say, the point of view of its linguistic capabilities, its pitch discrimination capabilities, its problem-solving capabilities, and so on (in principle without limit). Such perspectives unquestionably give us useful entry points into what would otherwise be intractably complex, and allow us to figure out what it is that the mind actually does. But it is hardly surprising that a targeted coming-to-grips with a complex object should yield significant insights. The same strategy is familiar in one form or another in virtually all domains of rational inquiry, be they physical, chemical, biological, psychological, or otherwise. That “science works here too” I take not to be an interesting claim, if it comes to that, because it does provide a theory of cognition at all: if anything it says more about science than it does about cognition. Furthermore, it is not entirely clear that the behaviorists would have spurned the sort of modules in view here. What they denied was the existence of *sui generis* principles, or the computational/ architectural/ modality independence of certain capacities (e.g., language), which I regard as evidence of true modules. They would not as a rule have denied that a partition of their subject matter could lead to interesting results. Recall the title of Skinner's *Verbal Behavior* – I think it is fair to say that Skinner sought quite literally to explain *the language faculty*, albeit in terms of general associationist learning mechanisms, and is this not nearly comparable to the system sense of a module now under discussion? The

behaviorists may have offered a shallow theory of human capacities, but even it did not appear to preclude modules in this sense (see, e.g., Chomsky, 1979, pp. 49–51). Nor, for that matter, is there any logical reason why a connectionist or holist would have to rule them out either. Contemporary PDP models of cognitive architecture in fact *do* have a sort of generic componential motivation behind them (Jilk, Lebiere, O'Reilly, & Anderson, 2008; O'Reilly, 1998).

It is worth being clear about exactly why system modularity fails the test of “interestingness”. It is easy to be misled here by the genuinely “interesting” results which have been achieved as a consequence of adopting the system view, that is, by what has been learned about distinct domains of psychology as a result of iterative functional analysis (task analysis, decomposition, “boxology,” etc.). The system module’s notable successes, as well as its historical association with the computational theory of mind and the view of the mind as richly and intricately structured, are apt to lead to an exaggerated estimate of its true significance. Any theory of the mind pitched at the level of faculties (or analyzable parts, components, units, etc.) – as modularity most assuredly is – must tell us what it does about the mind *through what it tells us about the faculties* (or whatever the relevant units of analysis happen to be). If it does not speak “through the faculties,” as it were, it cannot so much as count as a faculty psychology, since the properties of the mind to which a faculty psychology brings our attention are, in the first instance, properties of the mind’s faculties.

This point is at once obvious and yet so readily overlooked that it needs to be emphasized. That the mind is richly structured, that the mind is a computer, that the mind obeys laws exhibiting a clear mathematical structure, and so forth – these statements, if they are true, are true of the mind *generally*, meaning that it ought to be unsurprising if the *divisions* of the mind are correspondingly rich, intricate, computational, systematic, and so on. None of these properties attach to faculties *per se*. Moreover, learning that the *language* faculty has such and such features, or that *vision* operates in this or that fashion, need not tell us much about faculties *qua* faculties either, as against telling us about this or that particular faculty. Thus, neither *general* claims about the mind associated with the Cognitive Revolution, nor *specific* claims about *specific* faculties, hard won though these insights may have been, automatically get reckoned as among the distinctive insights marking out a truly *general theory* of faculties, which is after all what a faculty psychology aims to be. Contrast such claims with those of a well-developed faculty psychology (e.g., Fodor, 1983). The roster of properties associated with Fodorian modularity (domain specificity, encapsulation, shallow, fast, and mandatory processing, being hardwired, etc.) do not amount to a list of properties pertaining to the mind generally, nor to specific faculties considered independently, but to *all* faculties *qua* faculties. This is what made his theory interesting. So as easy as it is to roll the system module in the glitter of the Cognitive Revolution, a frank assessment of this module demands that we isolate clearly what it is the theory which posits such modules says about the mind *at the level of faculties* – and when we do this, I

maintain, we will be hard put to find anything that would not heartily be conceded by anyone who believes in the power of science (be they classical modularists, connectionists, holists, or, as I suggested, probably even behaviorists, *mutatis mutandis*).

Furthermore, lest it be thought that the very idea of functional decomposition can underwrite the system view – for one must admit that decomposition proceeds in a curious fashion in the context of computational systems, namely the execution of subroutines by homunculi, surely a nontrivial design feature of such systems – it need only be pointed out that *homuncularity* is not the same thing as *modularity*. Careful psychologists have always understood the difference, and that modularity is really a special type of homuncularity (Mahon & Cantlon, 2011, pp. 149–151), just as homuncularity is a special type of decomposition (van Gelder, 1995, p. 351). It is interesting to observe in this connection that Marr, one of the chief architects of the computational theory of mind, did not see computationalism (and therefore, we may surmise, homuncular functionalism) as providing a free pass to his “principle of modular design.” Modularity seems for Marr to be an added feature that *some* computational systems, for largely heuristic reasons, might be thought to possess:

Any large computation should be split up and implemented as a collection of small sub-parts that are as nearly independent of one another as the overall task allows. If a process is not designed in this way, a small change in one place will have consequences in many other places. This means that the process as a whole becomes extremely difficult to debug or to improve, whether by a human designer or in the course of natural evolution, because a small change to improve one part has to be accompanied by many simultaneous compensating changes elsewhere. (Marr, 1976, p. 485)

So although homuncularity is not as generic as “mere decomposition,” it is nowhere near as important a principle as modularity either. Accordingly (and for additional reasons I canvass below), we should withhold the more serious designation from generic system subcomponents and procedures that are non-dissociable.

So far I have said nothing about two important features of classically modular systems – domain specificity and informational encapsulation. Can they get the system module over the line? Actually, the question itself is incoherent. Consider that once a module is allowed to consist of shared parts, it can no longer be domain-specific, except perhaps in an abstract sense. This is because the “module” will be sensitive to potentially as many domains as there are parts comprising it, and quite likely more, since the parts themselves are presumably modular and domain-general. Put another way, domain specificity requires a functionally integrated unit that can respond to specified inputs. While the component modules of a composite consisting of shared parts would be functionally integrated, it is not obvious that the composite itself would be, although it might be said to have a sort of ad hoc integrity when in use. Notice also that a composite is unlikely to be informationally encapsulated “precisely because in sharing parts [it] will have access to the information stored and manipulated by [other high-level systems]”

(Anderson, 2010, p. 300). Anatomically distributed and overlapping brain networks simply must share information on some level (Pessoa, 2016, p. 23). Lacking both of these properties, then, one or the other of which has been considered definitive (Coltheart, 1999; Fodor, 1983, p. 71), its postulation does not quite serve the purposes many would assume. One might have supposed that the system module could be more strongly motivated if at least it had the property of either domain specificity or encapsulation (in a concrete and unambiguous sense). And yet just *because* it is a composite it can be neither truly domain-specific nor (in all likelihood) informationally encapsulated.

In something of a *reductio*, then, the system view of modularity leads only to the claim that the mind can do different things at different times. Certainly a more ambitious and theoretically interesting claim than this is available, namely that the mind can do different things at the *same* time; but as far as we know this requires functional specialization, that is, separate moving parts (*real* modules), since the prospect of neural time-sharing appears low. The pervasiveness of cognitive interference effects and processing bottlenecks in stimulus-rich environments that impose overwhelming attentional demands are enough to make this clear (Anderson, 2010, p. 250). In short, either these debates are trifling, or the claims at stake are more adventurous than the system view permits.

Bear in mind also that in the context of cognitive neuropsychology modules have been defined largely by reference to what the dissociation evidence has revealed, that is, “on the basis of the specific behavioral effects of brain lesions” (Bergeron, 2007, p. 177). Bergeron calls this inferential strategy the “functional modularity inference.” Basically, “the presence of highly selective cognitive impairments (dissociations) like prosopagnosia and various linguistic processing deficits suggest the functional independence of at least some cognitive processes,” and this in turn licenses the postulation of functionally independent modules subserving those processes (Bergeron, 2007, pp. 176, 177; see also Gazzaniga, 1989, p. 147). The fact that brain lesions are often also relatively localized suggests that such modules reside in a “specific and relatively small portion of the brain” (the “anatomical modularity inference”) (Bergeron, 2007, p. 176; see also Gazzaniga, 1989, pp. 947, 950). Make no mistake, the legitimacy of these inferences is hotly contested, since noisy dissociations are compatible with a system’s being dissociable, and clean dissociations compatible with a system’s being substantially non-dissociable. In the first instance, “there are a variety of reasons, well explored in the neuropsychology literature, for which lesions to brain systems can produce noisy rather than clean patterns of breakdown even when the systems required to complete the task are modular” (Barrett & Kurzban, 2006, p. 642). A good example would be a focal lesion at the border of two adjacent modules – the breakdown would not be clean, and yet the two systems would be modular. In the second instance, even perfect (or “double”) dissociations cannot conclusively establish that the affected systems are modular, for a lesion might only compromise a small, isolable component of an otherwise highly interpenetrative circuit. Damage to

this component might result in the system depending on that component being independently impaired, but it does not follow from this that the system would be functionally dedicated (although admittedly it would be dissociable at the margins). In light of this, it may come as something of a surprise to be told that these arguments ...

... have failed to deter theorists from employing either of the inferential strategies. Indeed, the functional modularity inference continues to be one of the most common approaches among cognitive neuropsychologists for inquiring about the structure of cognition. Similarly, the recent cognitive neuroscience literature abounds more than ever with cases involving the use of the anatomical modularity inference. (Bergeron, 2007, p. 177)

But what the persistence of these inferences bears witness to is the fundamental role that dissociation evidence plays in the search for modules, and that functional specificity itself continues to be the lodestar for deciding upon whether and to what extent the mind is modular within the major disciplinary fields concerned with this question.

As we have seen, the same assumption underwrites evolutionary psychology and massive modularity, the central claim of which is that the mind is predominantly composed of parts selectively shaped by evolutionary pressures (Carruthers, 2006). As two prominent evolutionary psychologists state their position, “We intend an explicitly evolutionary reading of the concepts of function and specialization: modules evolved through a process of descent with modification, due to the effects they had on an organism’s fitness” (Barrett & Kurzban, 2006, p. 630). This view affirms the existence of systems which, though perhaps spatially extended and neurally interspersed, are dissociable in principle:

Psychologists generally agree—as do we—that because cognitive architecture is instantiated in brain architecture, the two will be isomorphic at some level ... However, at a larger, macroscopic level, there is no reason to assume that there must be spatial units or chunks of brain tissue that neatly correspond to information-processing units. An analogy might be to the wiring in a stereo, a computer, or other electronic system: Individual wires have specific functions, but at the level of the entire machine, wires with different functions might cross and overlap. For this reason, removing, say, a three-inch square chunk from the machine would not necessarily remove just one of the machine’s functions and leave the rest intact. In brain terms, it could be, and probably is, that macroscopic regions of brain tissue include neurons from multiple information-processing systems with multiple functions. (Barrett & Kurzban, 2006, p. 641)

And of course other contemporary models of cognitive architecture, such as the successful ACT-R model, also posit the existence of independently modifiable subsystems of the brain.

When it comes to clarifying just what makes modularity interesting, one final set of considerations may be suggestive. While terminological nuances can hardly be decisive in an area like this, I think it is no coincidence that massive modularity bottoms out in claims about the separate modifiability of functional components. This is because the very word “module” evokes images of movable parts that can

be assembled and reassembled in a variety of distinct combinations, and which may be affected independently of one another. If all modularity amounts to is the claim that the mind can do different things at different times (rather than the stronger claim that it can do different things at the *same* time), and this suffices to call it modular, it ought to be permissible to say that a knife which cuts both meat and bread is modular. And yet no one thinks of knives as modular (unless they are Swiss army knives which actually come with different blades). It is instructive that other anatomically nondissociable systems with shared parts, such as nervous systems, reproductive systems, endocrine systems and the like – all of which may be singled out for their natural kind properties – *are* termed “systems.” One never hears of digestive modules or reproductive modules. The “modules” of developmental biology and neuroscience which do have shared and reusable elements are an anomaly of network science (see section 4). Most biologists, including developmental biologists, continue to think of modules as “anatomically distinct parts that can evolve independently” (Wolpert, 2011, p. 115). Limbs and vertebrae would be modular on this view (being organs), but not the larger anatomical systems they comprise.

It is well worth stressing here that my argument should not be read as an instance of mere carping or terminological pedantry. There are certainly occasions when scruples over the use of words reveal carping tendencies, and nothing much beyond that, but this is not one of them. Philosophers and cognitive scientists who allege a “module” for this capacity and a “module” for that capacity must be taken to be saying something substantial – something more than merely the fact that we *have* the capacities in question. To dignify these capacities with the honorific title “module” is, I would suggest, an attempt to invest the capacities with special-purpose, special-structure status. If philosophers and cognitive scientists persist in referring to modules for X and Y in the face of contrary evidence (i.e., evidence suggesting that the X and Y “modules” are not special-purpose, special-structure devices), they betray a willingness to exploit the connotations of a powerful term for rhetorical purposes. For if by alleging that there is a module for X or Y the speaker intends only to say that we can give systematic accounts of X and Y – where X and Y represent particular foci of the scientific gaze upon the mind – the speaker is only avowing a belief in the efficacy of the scientific method in the realm of cognition, which I take no naturalist to deny. In such circumstances, it would be better to drop the term “module” altogether and settle for a less loaded (and therefore more honest) term like “capacity,” “faculty,” or “system.”

So while it is true that I am insisting on correct usage, this insistence is not without justification and not without consequences should laxity prevail. In some ways, the issues here are analogous to those which have arisen in the philosophy of biology over the proper use of the word “innate” (Griffiths, 2002; Collins, 2005; Griffiths & Machery, 2008; Mameli & Bateson, 2006, 2011; Marcus, 2004; Mameli & Papineau, 2006). Neither those who urge elimination – because the word engenders confusion and fallacies of ambiguity amid a plethora of conflicting

folk-biological intuitions – nor those arguing that a technical definition can be given should be seen as engaging in a merely feeble semantic dispute.

4. Whither modularity?

How the foregoing considerations constrain the future of modular theorizing is an important question. On one level, I have suggested that if a certain condition for modularity cannot be satisfied (dissociability in principle), we have little reason to be sanguine about its prospects. This may seem stiflingly austere, so let me round out the analysis by offering a somewhat more optimistic picture that is nonetheless consistent with the overall thrust of my remarks. For what it is worth, though, I still think it is salutary to remember what it was that made modularity an intriguing hypothesis in cognitive science. If it turns out that that hypothesis cannot be sustained, I would suggest that there is nothing inherently wrong with adopting an eliminativist stance. After all, no one is *obliged* to take modularity seriously, come what may. Its prospects should be assessed in accordance with the same standards as would be applied to any other theory in the cognitive sciences whose fortunes are at issue. All the same, here are some thoughts striking what I hope will be a less gloomy note.

First, let me repeat and expand slightly on what I said earlier. Neural reuse at a minimum implies a shift of focus away from understanding modules as unimodal high-level cognitive mechanisms toward a conception of them as metamodal nodes subserving exiguous, low-level subfunctions. Entities like this have been studied extensively for many decades, forming part of the neocortical columnar hypothesis in neuroscience (Mountcastle, 1978, 1997), and it is likely that at least *some* brain regions of this sort will satisfy the conditions of functional specialization (and hence of modularity), even if many of them are domain-general/metamodal (Pascual-Leone & Hamilton, 2001). Of course, precisely to what extent they populate the cortex is very much an open question (Zerilli, 2017; Burnston, 2016).

Second, it is testament to the immense versatility of modularity that it has descriptive utility well beyond the confines of cognitive science. Modules play an important role in fields as diverse as developmental and systems biology, ecology, mathematics, computer science, robotics, and industrial design, and none of these varied uses of the term come with anything like the philosophical baggage of modules in cognitive science. Might the induction of one of these other senses of the term into the lexicon of cognitive science be a legitimate way of salvaging a hallowed vocabulary? It seems so. One interesting application of the term occurs in the study of networks, and neural networks in particular. In network graph theory, a module is defined as a community of “densely interconnected nodes” where “the existence of several [such] communities is characteristic of [a] modular [network]” (Bullmore & Sporns, 2012, p. 342). In network neuroscience specifically, network models take the form of neural coactivation graphs, where modules are identified as communities of nodes that are functionally coactive. In the context of neural

networks, then, “modularity refers to the existence of multiple communities of neurons or brain regions as defined by patterns of [functional] connectivity” (Bullmore & Sporns, 2012, p. 342). The upshot is that because the abstract topological features of these coactivation graphs frequently (if only roughly) track the functional taxonomies of cognitive psychology, it seems that network neuroscientists are *already*, by all accounts *harmlessly*, referring to something like the system module *as* a module. The usage here is clearly *sui generis*, originating outside the tradition which gave rise to modular theorizing in cognitive science. It is difficult to see how confusion could arise from this highly context-specific usage.

5. Conclusion

In recent decades, neuroscience has challenged the orthodox account of the modular mind. As I have shown, one way of meeting this challenge has been to go for increasingly “soft” versions of modularity, and one version in particular, which I dub the “system” view, is so soft that it promises to meet practically any challenge neuroscience can throw at it. In this paper, I have reconsidered afresh what we ought to regard as the sine qua non of modularity, and have offered a few arguments against the view that an insipid “system” module could be the legitimate successor of the traditional notion. In part, my paper can be read as a plea for the precise use of language. But there is more than pettifoggery behind this plea.

Notes

1. Its modification may of course *indirectly* impede a comparable “downstream” system, that is, one at the receiving end of its efferent projections.
2. The reason why it is not guaranteed is because the effects of the neural network context may compromise the region’s ability to maintain a set of stable input-output functions. This is an important issue which I do not pursue here. I consider it in Zerilli (2017).
3. Slight variations in the meanings of these terms are not important here. See section 4, on neural coactivation graphs.
4. Decomposability and modularity do come apart. Anderson, who has “become pessimistic that it will *generally* be possible to identify the specific neural mechanisms ... to which we can assign cognitive operations,” and is no defender of modularity, nevertheless states, “I have no particular argument with the practice of mechanistic analysis and decomposition, especially when it is sensitive to the importance of high-level dynamic interactions ... When it works it can provide elegant models that offer genuine insight” (2014, p. 106). Similarly, Boone and Piccinini outline “a mechanistic version of homuncular functionalism, whereby higher-level cognitive capacities are iteratively explained by lower-level capacities until we reach a level in which the lower-level capacities are no longer cognitive in the relevant sense” (2016, p. 1524). While this might entail modularity for some lower-level elements (they do not say as much), it does not entail modularity for higher-level elements composed predominantly of shared parts (indeed the word “modularity” or “module” appears nowhere in their paper): see McGeer’s (2007) helpful discussion of the cognitive neuropsychologist’s

understanding of modularity. Prinz (2006) is actually explicit that so long as the units of decomposition do not exhibit the properties associated with Fodorian modularity, we should proceed with decomposition but abandon the label of modularity. See my remarks, below, for further clarification of this point.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

This research was supported by an Australian Government Research Training Program (RTP) Scholarship.

References

- Amaral, D. G., & Strick, P. L. (2013). The organization of the central nervous system. In E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, & A. J. Hudspeth (Eds.), *Principles of neural science* (pp. 337–355). New York, NY: McGraw-Hill.
- Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences*, 33, 245–266.
- Anderson, M. L. (2014). *After phrenology: Neural reuse and the interactive brain*. Cambridge, MA: MIT Press.
- Anderson, M. L., & Finlay, B. L. (2014). Allocating structure to function: The strong links between neuroplasticity and natural selection. *Frontiers in Human Neuroscience*, 7, 1–16.
- Badcock, P. B., Ploeger, A., & Allen, N. B. (2016). After phrenology: Time for a paradigm shift in cognitive science. *Behavioral and Brain Sciences*, 39, 10–11.
- Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review*, 113, 628–647.
- Bechtel, W., & Mundale, J. (1999). Multiple realizability revisited: Linking cognitive and neural states. *Philosophy of Science*, 66, 175–207.
- Bergeron, V. (2007). Anatomical and functional modularity in cognitive science: Shifting the focus. *Philosophical Psychology*, 20, 175–195.
- Boone, W., & Piccinini, G. (2016). The cognitive neuroscience revolution. *Synthese*, 193, 1509–1534.
- Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, 13, 336–349.
- Burnston, D. C. (2016). A contextualist approach to functional localization in the brain. *Biology and Philosophy*, 31, 527–550.
- Buxhoeveden, D. P., & Casanova, M. F. (2002). The minicolumn hypothesis in neuroscience. *Brain*, 125, 935–951.
- Carruthers, P. (2006). *The architecture of the mind: Massive modularity and the flexibility of thought*. Oxford: Oxford University Press.
- Carruthers, P. (2008). Précis of the architecture of the mind: Massive modularity and the flexibility of thought. *Mind and Language*, 23, 257–262.
- Chomsky, N. (1975). *Reflections on language*. New York, NY: Pantheon.
- Chomsky, N. (1979). *Language and responsibility*. New York, NY: Pantheon.
- Chomsky, N. (1980). *Rules and representations*. New York, NY: Columbia University Press.

- Chomsky, N. (1988). *Language and problems of knowledge: The Managua lectures*. Cambridge, MA: MIT Press.
- Chomsky, N. (2002). *On nature and language*. New York, NY: Cambridge University Press.
- Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, 36(1), 1–22.
- Chomsky, N. (2006). *Language and mind* (3rd ed.). New York, NY: Cambridge University Press.
- Collins, J. (2005). Nativism: In defense of a biological understanding. *Philosophical Psychology*, 18, 157–177.
- Collins, J. (2008). *Chomsky: A guide for the perplexed*. London: Continuum.
- Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Sciences*, 3, 115–120.
- Coltheart, M. (2011). Methods for modular modelling: Additive factors and cognitive neuropsychology. *Cognitive Neuropsychology*, 28, 224–240.
- Cosmides, L., & Tooby, J. (1994). Origins of domain specificity: The evolution of functional organization. In L. Hirschfield & S. Gelman (Eds.), *Mapping the world: Domain specificity in cognition and culture* (pp. 85–116). New York, NY: Cambridge University Press.
- Dehaene, S. (2005). Evolution of human cortical circuits for reading and arithmetic: The “neuronal recycling” hypothesis. In S. Dehaene, J. R. Duhamel, M. D. Hauser, & G. Rizzolatti (Eds.), *From monkey brain to human brain* (pp. 133–157). Cambridge, MA: MIT Press.
- Doidge, N. (2007). *The brain that changes itself*. New York, NY: Penguin.
- Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, 97, 179–210.
- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Gazzaniga, M. S. (1989). Organization of the human brain. *Science*, 245, 947–952.
- Griffiths, P. E. (2002). What is innateness? *Monist*, 85, 70–85.
- Griffiths, P. E., & Machery, E. (2008). Innateness, canalization, and ‘biologizing the mind’. *Philosophical Psychology*, 21, 397–414.
- Gold, I., & Roskies, A. L. (2008). Philosophy of neuroscience. In M. Ruse (Ed.), *The Oxford handbook of philosophy of biology* (pp. 349–380). New York, NY: Oxford University Press.
- Horst, S. W. (2011). *Laws, mind, and free will*. Cambridge, MA: MIT Press.
- Jilk, D. J., Lebiere, C., O’Reilly, R. C., & Anderson, J. R. (2008). SAL: An explicitly pluralistic cognitive architecture. *Journal of Experimental and Theoretical Artificial Intelligence*, 20, 197–218.
- Jungé, J. A., & Dennett, D. C. (2010). Multi-use and constraints from original use. *Behavioral and Brain Sciences*, 33, 277–278.
- Kandel, E. R., & Hudspeth, A. J. (2013). The brain and behavior. In E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, & A. J. Hudspeth (Eds.), *Principles of neural science* (pp. 5–20). New York, NY: McGraw-Hill.
- Klein, C. (2012). Cognitive ontology and region- versus network-oriented analyses. *Philosophy of Science*, 79, 952–960.
- Mahon, B. Z., & Cantlon, J. F. (2011). The specialization of function: Cognitive and neural perspectives. *Cognitive Neuropsychology*, 28, 147–155.
- Mameli, M., & Bateson, P. (2006). Innateness and the sciences. *Biology and Philosophy*, 21, 155–188.
- Mameli, M., & Bateson, P. (2011). An evaluation of the concept of innateness. *Philosophical Transactions of the Royal Society B*, 366, 436–443.
- Mameli, M., & Papineau, D. (2006). The new nativism: A commentary on Gary Marcus’s *The birth of the mind*. *Biology and Philosophy*, 21, 559–573.
- Marcus, G. (2004). *The birth of the mind: How a tiny number of genes creates the complexities of human thought*. New York, NY: Basic Books.

- Marr, D. (1976). Early processing of visual information. *Philosophical Transactions of the Royal Society B*, 275, 483–519.
- McGeer, V. (2007). Why neuroscience matters to cognitive neuropsychology. *Synthese*, 159, 347–371.
- McGilvray, J. (2014). *Chomsky: Language, mind, politics* (2nd ed.). Cambridge: Polity Press.
- Mountcastle, V. B. (1978). An organizing principle for cerebral function: The unit module and the distributed system. In G. Edelman & V. B. Mountcastle (Eds.), *The mindful brain* (pp. 7–50). Cambridge, MA: MIT Press.
- Mountcastle, V. B. (1997). The columnar organization of the neocortex. *Brain*, 120, 701–722.
- O'Reilly, R. C. (1998). Six principles for biologically-based computational models of cortical cognition. *Trends in Cognitive Sciences*, 2, 455–462.
- Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L. B. (2005). The plastic human brain cortex. *Annual Review of Neuroscience*, 28, 377–401.
- Pascual-Leone, A., & Hamilton, R. (2001). The metamodal organization of the brain. *Progress in Brain Research*, 134, 427–445.
- Pessoa, L. (2016). Beyond disjoint brain networks: Overlapping networks for cognition and emotion. *Behavioral and Brain Sciences*, 39, 22–24.
- Piccinini, G., & Craver, C. (2011). Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese*, 183, 283–311.
- Pinker, S. (1997). *How the mind works*. New York, NY: Norton.
- Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition*, 95, 201–236.
- Prinz, J. J. (2006). Is the mind really modular? In R. Stainton (Ed.), *Contemporary debates in cognitive science* (pp. 22–36). Oxford: Blackwell.
- Quartz, S. R., & Sejnowski, T. J. (1994). Neural evidence for constructivist principles in development. *Behavioral and Brain Sciences*, 17, 725–726.
- Rowland, D. C., & Moser, M. B. (2014). From cortical modules to memories. *Current Opinion in Neurobiology*, 24, 22–27.
- Scholl, B. J. (1997). Neural constraints on cognitive modularity? *Behavioral and Brain Sciences*, 20, 575–576.
- van Gelder, T. (1995). What might cognition be, if not computation? *Journal of Philosophy*, 92, 345–381.
- Wolpert, L. (2011). *Developmental biology: A very short introduction*. New York, NY: Oxford University Press.
- Zador, A. (2015). The connectome as a DNA sequencing problem. In G. Marcus & J. Freeman (Eds.), *The future of the brain* (pp. 40–49). Princeton: Princeton University Press.
- Zerilli, J. (2017). *Neuroplasticity, neural reuse, and the language module* Unpublished doctoral dissertation). Australian National University, Canberra.