Interventions, Mechanisms, and the Modularity of Mind

Matthew Haug
mchaug@wm.edu

[Draft; please do not cite; comments welcome]

Abstract

This paper takes as its starting point John Campbell's recent attempt to extend the interventionist approach to cover causation in psychology. I point out that Campbell's radical suggestion that causation between psychological variables may not be grounded in biochemical mechanisms conflicts with the completeness of physics. I then use a case study involving the effects of a nurturing environment on memory ability to argue that accepting the existence of underlying biochemical mechanisms does not commit one to the equally radical reductive view according to which mental causation is reduced to, or eliminated in favor of, biochemical causation. I show how the biochemical variables in any mechanism underlying nurturance's effect on memory are likely not as accurate or precise as psychological variables, nor are they as effective as a means of intervening on memory ability. I then explore the relationships between modular systems, robust variables, and reduction, arguing inter alia, that even non-robust cognitive variables in modular systems cannot be replaced by biochemical variables. These results support a straightforward general argument that psychological variables will play an ineliminable role in the etiology and treatment of many mental phenomena.

In recent years, psychologists, neuroscientists, and scientifically-based psychiatrists have made great strides in discovering the etiology of depression and anxiety disorders. One focus of this research is pharmacological – the discovery and testing of drugs like selective serotonin reuptake inhibitors (SSRIs) and benzodiazepines that affect the biochemical systems that are thought to underlie these ailments. Another line of research has focused on the social and environmental causes of these disorders. For instance, recent work has shown that different kinds of stressful life events are predictors of different kinds of mood disorders (Kendler et al. 2003, 2005). This latter line of research overlaps with recent work on the causal basis of the robust correlations between socio-economic status and cognitive development (Farah et al. 2006, 2008; Noble et al. 2007). Both research programs see social events (e.g., the death of a loved one, parental care behavior, respectively) as producing changes in emotional or cognitive aspects

of a subject (standing mood, memory ability) partly via changes in psychological states (e.g., feelings of loss, security, and well-being).

In a recent paper which extends the interventionist analysis of causation to apply to mental causation,[1] John Campbell suggests that one virtue of this approach to causation is that it is compatible with what he calls the "radical" view that some cases of mental causation are not grounded in biochemical mechanisms. For instance, humiliation and serotonin deficiency are plausibly both causes of depression, but, according to Campbell, "there is no empirical support for the idea that all causation that involves both psychological and biological variables bringing about a psychological outcome must be sustained by biological mechanisms" (2007, 65).

After some preliminaries about interventions, modularity, and mechanisms in Section 1, I point out that this radical view conflicts with the completeness of physics, as that claim is commonly understood, in Section 2. However, in Section 3, I show that the radical anti-reductive view contains an important grain of truth that is in danger of being overlooked in the flood of neurobiological research on the brain: sometimes psychological variables provide more accurate, precise, and easier to use means of intervening on psychological outcomes than biochemical variables do. Further, as I argue in Section 4, there are reasons to think that even non-robust cognitive variables found in modular systems cannot be replaced by biochemical variables.

The discussion in Sections 3 and 4 supports the crucial third premise in the following straightforward argument for the claim that psychological variables are indispensable in our attempts to understand the mind and its disorders.

[1] I follow Campbell in using the phrase "interventionist account" to refer to Jim Woodward and Chris Hitchcock's work on causation, which builds on previous work by Judea Pearl and Clark Glymour, among others. This body of work is detailed and sometimes technical; I am just concerned with the broad outline of the theory in this paper.

(1) In order to provide the best treatment for individuals with mental disorders and developmental cognitive deficiencies, we need a comprehensive, yet practically usable, account of mental processes.

(2) In order to develop a comprehensive, yet practically usable, account of mental processes, we need a set of exhaustive control variables for mental phenomena that are as accurate, precise, and informative as possible.

(3) Although some set of biochemical (or other physical) variables will be exhaustive (if physicalism is true), as argued below, these variables will sometimes be less precise, less accurate, and less significant as means of intervening on and explaining mental outcomes than psychological variables are.

(4) Hence, in order to develop a comprehensive, yet practically usable, account of mental processes, biochemical variables are not sufficient. Distinct psychological variables must be included.

1. Interventions, Modularity, and Mechanisms

In the interventionist approach to causation, causal relationships are represented by graphs in which arrows between variables represent direct causal relationships. Roughly, variable A is a (total) cause of variable B if, in some cases, intervening on A changes the value of B.[2] The intervention on A is required to be "surgical" – it should disrupt the normal causal pathways leading to A and leave intact only the influence of A on B. Further, the intervention should affect the value of B *only* via its effect on A. Finally, the causal graph should be

[2] Variables are properties or magnitudes that can take more than one value. Causation between properties can be thought of as holding between binary variables (i.e. those that take only two values, depending on whether or not the property is instantiated).

exhaustive; there should be no unrepresented common causes of pairs of variables that are represented in the graph.

The interventionist approach assumes that a causal graph that gives a correct, exhaustive representation of the causal structure of a system is *modular* – i.e., that it is in principle possible to intervene on one variable while leaving the causal relationships between other variables undisturbed. In his recent book, Woodward notes that the modularity assumption is connected to the idea of a mechanism (2003, 48). Intuitively, each equation in a causal model (the complete set of arrows directed into each variable in a graph) should correspond to a distinct causal mechanism. If we assume that two mechanisms are distinct only if it is possible (in principle) to interfere with the operation of one without interfering with the operation of the other (and vice versa), then we have support for the requirement that correct and complete representations be modular. A different kind of modularity that applies to systems or networks of variables will be discussed below.

In asking for a mechanism for a causal relation we are asking for more detail about how, or by what means, the cause produces the effect. This could be provided by simply specifying intervening variables between the given cause and effect. Or, it could be provided by moving to a different level of analysis in order to replace some or all of the given causal variables with mechanisms (networks of variables from the new level of analysis) that account for (the instantiation of values of) the variables that have been analyzed away. This is the procedure that would be required to provide an underlying biological mechanism for the way in which humiliation and serotonin deficiency jointly cause depression, which, according to the radical view suggested by Campbell, is not possible.

2. The Radical Anti-Reductive View vs. the Completeness of Physics

Campbell suggests that the relation between humiliation, serotonin deficiency, and depression is illustrative of a general pattern emerging from recent scientific work: psychological outcomes are produced by the joint action of biological and psychological variables. He suggests that further refinement will not eliminate the psychological variables: "Explanation by means of mechanisms must bottom out somewhere, and then we are left with the bare facts about what would happen under interventions" (2007, 65). According to Campbell, we currently have "no reason to suppose that a comprehensive set of control variables for depression [and other mental phenomena] will ever be found at the biological level" (ibid.).[3]

For mental phenomena like depression, if we have no empirical evidence that their production via psychological variables is sustained by biological mechanisms, then presumably we have no empirical evidence for underlying physical mechanisms of any kind.[4] But this blanket denial of empirical evidence for underlying biological mechanisms seems to be false. The explosion of work in cognitive and affective neuroscience is discovering portions of those very mechanisms.

The radical anti-reductionist may demur and claim that this research is only discovering the supervenience bases of psychological properties like humiliation and depression. After all, they continue, their view is perfectly compatible with the claim that the mental supervenes on the physical, but that the biological or physical supervenience base of say, humiliation, is not a control variable for depression—not an effective means of intervening on depression. So, the

[3] As spelled out by Campbell, a "control variable" for Y is one that has a dose-dependent effect on Y, and one that often has a large effect specifically on Y, and no other variable.
[4] In principle, one could bypass biological variables and move directly to a microphysical mechanism, but I doubt that is what Campbell has in mind.

anti-reductionist can claim that humiliation and depression have physical supervenience bases while also holding that it is a brute fact that humiliation causes depression.

However, the claim that there is no comprehensive set of *physical* control variables for depression conflicts with the completeness of physics—roughly, the claim that every physical event has a physical sufficient cause at any given time prior to its occurrence. It is generally accepted that the completeness of physics is an essential part of any physicalist view and is itself supported by much empirical evidence. Many who have discussed the completeness of physics claim that it is only at the fundamental physical level where we reach a domain that is causally complete (e.g. Kim 2003, 173). However, when we ask what kinds of entities must be included in this causally complete physical domain, we quickly see that more than individual fundamental physical entities and their properties are needed. We also need to include aggregates of those entities and their micro-based properties:

> Plainly the physical domain must also include aggregates of basic particles, aggregates of these aggregates, and so on, without end; atoms, molecules, cells, tables, planets, computers, biological organisms, and all the rest must be, without question, part of the physical domain. … It is important that these [objects and their] micro-based properties are counted as physical, for otherwise the physical domain won't be causally closed. Having a mass of one kilogram has causal powers that no smaller masses have, and water molecules, or the property of being water, have causal powers not had by individual hydrogen and oxygen atoms. (Kim 1998, 113-4)

This is also the strongest claim that is supported by the kind of empirical evidence we have for the completeness of physics. This evidence does not come solely, or even largely, from fundamental physics, but rather from physiology and biochemistry, which has provided a physically acceptable account of previously suspect biological substances and functions (like reproduction, respiration, protein synthesis, urea, etc.) and has shown that energy is conserved in the interactions between living things and their environment (cf. Papineau 2001).

Consider the suggestion that humiliation's causing depression is a brute psychological causal relation. This does not directly contradict the completeness of physics, since depression is arguably not a physical effect. However, there are at least two ways to bring out the conflict. First way: assuming that the anti-reductionist accepts that the mental supervenes on the physical, one can use a Kim-style argument to show that mental-to-mental causation presupposes mental-to-physical causation. The only way to cause a supervenient property to be instantiated is to cause one of its base properties to be instantiated (cf. Kim 1998, 42). Given that the anti-reductionist believes that it is brute fact that humiliation causes depression, she is also committed to it being a brute fact that humiliation causes the supervenience base of depression. (If she denied this, her view would collapse into the competing view which sees all mental causation as grounded in physical causation.) Hence, at some time simultaneous with or after the onset of the humiliation, there is no physical cause of the physical supervenience base of depression—contra the completeness of physics.

Second way: even though depression is not a physical event, depression will itself have physical effects (to take a lurid example, the bodily injury that depressed individuals may inflict on themselves). Suppose that the radical anti-reductionist claims that it also a brute fact that intervening on depression affects whether or not bodily injury occurs. So, there is some time simultaneous with or after the onset of depression at which there is no physical sufficient cause of the bodily injury. Hence, we have a physical effect with no physical sufficient cause at some time prior to that effect. On the other hand, suppose that the anti-reductionist claims that it is not a brute fact that intervening on depression affects whether or not bodily injury occurs. However, as noted above, she must still claim that there is no physical mechanism that grounds humiliation's causing depression. So, then a fortiori, there is no larger physical mechanism that

grounds the causal relations between humiliation, depression, and bodily injury. Again, at some time after the onset of humiliation but prior to the bodily injury there is no physical sufficient cause of that injury. In other words, if there is a time prior to the onset of depression at which there is no physical sufficient cause of depression, and depression in turn causes a physical effect, then there is a time prior to the occurrence of that physical effect at which it has no physical sufficient cause.

So, the radical anti-reductionist view is committed to cases of mental causation that are ungrounded in physical variables, not merely to cases of "autonomous" explanation in terms of mental variables.

3. The Virtues of Psychological Variables

Campbell himself may not be troubled by this result. However, those who have physicalist sympathies may also be interested in extending the interventionist framework to cover mental causation, and they may suspect that this will not involve simply discovering biochemical variables to replace the mental variables in the relevant causal graphs.

So, if we accept that some biochemical or other physical mechanisms underlie all causal relations between psychological variables, have we then committed ourselves to an equally radical, reductionist view according to which all causal explanation and intervention should be performed at the biochemical level (see, e.g., Bickle 2003)? Adopting this view apparently commits one to the irreality of mental variables that are distinct from biochemical or microphysical ones (see, e.g., Kim 1998, 119-120). Further, this view arguably also includes the claim that there is no unified, relatively species-independent science of psychology (Kim 1998, 109-110). On this view, psychological terms are at best merely designators for biochemical realizers which may differ from species to species or from individual to individual.

Even if physicalism commits one to the existence of physical or biochemical mechanisms that ground causal relations between psychological variables, I claim that the radical reductive view just sketched is false for at least four reasons: biochemical variables (and microphysical ones) will sometimes be (1) harder to use (2) less precise, (3) less accurate, and (4) less significant as means of intervening on psychological phenomena than psychological variables are. Together, these facts show that purely biochemical mechanisms will sometimes be less informative than mechanisms involving psychological variables.

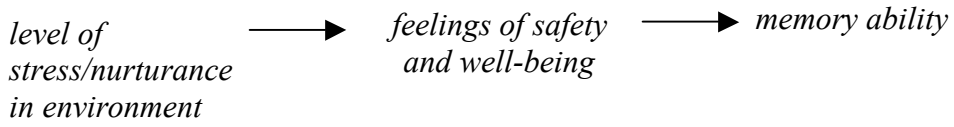
Many studies in non-human animals have shown that prolonged, chronic stress has detrimental effects on (declarative) memory ability (e.g, McEwan 2000 and the references therein) and that maternal care (e.g., licking and grooming in rats) has long-term positive effects on memory performance (e.g., Liu et al. 2000).[5] Further, recent correlational studies in humans suggests that a nurturing environment (or perhaps absence of a stressful environment) causes an increase in memory performance (Farah et al. 2008).[6] This research suggests that a causal relationship holds between the following variables:

level of stress/nurturance in environment ⟶ *memory ability*

[5] Note that in the literature the term "stress" is used to refer to both the external event (stressor) and the subjective response (stress response). A similar slippage is present in the work on stressful life events and depression mentioned above, humiliation and loss are characterized in terms of subjective feelings while entrapment and danger are characterized in terms of external circumstances (Kendler et al. 2003, 791).

[6] Nurturance was measured using a structured interview and observational checklist developed to ascertain warmth of parental interaction and availability of parental care. For 4-year-olds, the subscales included: Warmth and affection (e.g., 'parent holds child close 10–15 minutes per day') and acceptance (e.g., 'parent neither slaps nor spanks child during visit'). For 8-year-olds, the subscales used were: Emotional and verbal responsivity (e.g., 'Child has been praised at least twice during past week for doing something'), emotional climate ('parent has not lost temper with child more than once during previous week'), and paternal involvement (e.g., 'Child eats at least one meal per day, on most days, with mother and father [or mother and father figure]') (Farah et al. 2008, 4).

One way of providing a mechanism for this causal relationship is to spell out the way in which a change in nurturance brings about a change in memory ability. Although there is still much work to be done in this area, one hypothesis is that nurturance's influence on memory is mediated by a complex suite of emotional states, e.g., feelings of safety and well-being. This is merely one of many ways in which emotional states have been shown to affect cognitive performance. In particular, positive affect has been shown to facilitate creative problem solving and enhance memory ability (e.g., refs. in Ashby et al. 1999). These findings suggest that the following simple mechanism is at work:



Is there a biochemical mechanism that underlies this psychological mechanism? Contra Campbell, physicalism seems to demand it. It is not a brute or bare fact that increased nurturance causes certain feelings and emotions which in turn improve memory function. Yet specifying that mechanism is very difficult. Biological variables, like hippocampal size and levels of neurotransmitters and hormones (such as dopamine, acetylcholine, and glucocorticoids), which are believed to influence memory ability supplement, rather than replace, the emotional variable, in the same way that serotonin deficiency appears to be a distinct cause of depression alongside feelings of humiliation. There is no simple linear relationship between these biological variables and emotional states, and the former can often be manipulated independently of the latter (Gross and Hen 2004, 547). Hence, the biochemical variables are not plausible candidates for total realizers that could replace the emotional variables. Further, hippocampal size is likely a standing condition that mediates susceptibility to memory

degradation (and to affective disorders) rather than being the biological basis for the mediating emotional variable (cf. Gross and Hen 2004, 547).

Any mechanism that is arrived at by replacing the emotional variable with its biochemical realizer will be incredibly complicated, likely involving a large number of fine-grained variables of a highly distributed neural system. Thus, reproducing the effects of interventions on the emotional variable at the biochemical level would require subtle and complex coordination among interventions on countless biochemical variables. Accomplishing this feat of coordination would be immensely challenging and much more difficult than intervening directly (via behavioral therapy or social interaction) on a subject's emotional states.[7] In doing the latter, this immensely complicated suite of biochemical changes is produced in one fell swoop.

In addition, given the countless causal relationships between the biological variables, it will be much harder to determine if a given biochemical causal model is modular and thus an accurate guide for interventions. It will be difficult to determine if the represented mechanisms are genuinely distinct or rather if some distinct biochemical causal pathways are conflated in the representation. By contrast, it is relatively easy to determine that the psychological relations represented above are distinct.

Despite these practical difficulties, one might think that if we could identify and intervene directly on the biochemical variables, this would provide more precise, significant, and accurate manipulations of memory ability. However, there are several reasons to think that this is not

[7] The notion of "intervening directly" on mental variables presupposes merely that mental variables are sometimes control variables for certain mental phenomena. The current evidence suggests that the effects of manipulating these variables (e.g. via psychotherapy) cannot be exactly mimicked by any existing biochemical interventions (e.g. taking an SSRI). Of course, if physicalism is true, mental variables and manipulations are "nothing over and above" biochemical variables and manipulations.

true. First, as is well-known, pharmacological treatments have varying rates of success within a population. This variation in efficacy is likely due to the fact that the given emotional variable is realized by slightly different biochemical pathways in different individuals. Thus, within a population, a specific biochemical intervention that is successful in one subject will likely not be universally effective. By contrast, intervening directly on the emotional variable is likely to have similar effects throughout the population, since the intervention works only on the psychological aspects that are common to all of the various realizers.[8] Hence, manipulating biochemical variables will likely often have less precise effects on psychological outcomes than manipulating psychological variables will.

Second, although varying some of the myriad underlying biochemical variables will produce many changes in the memory system, many of these will not be significant at the psychological level; they will not show up in everyday life or in psychological tests. Further, pharmacological treatments of mood disorders do not appear to continue to be effective (e.g. they do not reduce the risk of relapse) after their use is terminated (Hollon et al. 2002). By contrast, treatments which intervene directly on psychological states (e.g. cognitive-behavior therapy or parental care during development) have long-term effects on emotional and cognitive states even after treatment has ended (ibid.). Thus, manipulating biochemical variables will often have less significant consequences for psychological outcomes than manipulation of psychological variables.

Third, barring the kind of sophisticated and massive coordination on thousands of biochemical variables mentioned above (which is currently beyond our technological

[8] This is easiest to see on the "subset" model of realization, according to which, roughly, the causal powers of the realized property are a proper subset of the powers of its various realizers (Shoemaker 2001). Psychological interventions affect only the causal powers in the subset that is common to all of the biochemical realizers.

capabilities), most interventions on individual biochemical variables will also produce effects that are irrelevant to memory performance—various physiological changes, including detrimental side effects. And, although changes in emotional states will also produce a wide variety of psychological effects, these will likely interact in complex feedback cycles with memory and other cognitive phenomena (cf. Fredickson's 2001 "broaden-and-build" theory of positive emotions). Hence, psychological mechanisms sometimes provide more accurate means of affecting psychological outcomes than biochemical mechanisms do.

Together, these results show that explanations of mental phenomena in terms of psychological variables are sometimes more informative than those solely in terms of biochemical variables. Emotional and cognitive variables are characterized by different kinds of aspects and vary along different dimensions than biochemical and microphysical variables. For instance, emotional states like anxiety are characterized by dimensions like levels of intensity, irritability, worry, and restlessness, which do not distinguish between biochemical variables. If physicalism is true, (particular values of) these high-level psychological variables are realized by (complicated states of affairs involving) biochemical and microphysical variables. However, the high-level aspects allow for generalizations and interventions that capture and exploit patterns that are not visible at the fine-grained, biochemical level. Pulling back from the welter of biochemical details facilitates explanations and treatments that are not provided by biochemical variables.

4. Robustness and Modularity

As Campbell points out, many personal-level, psychological states, like *feeling calm and secure*, *feeling humiliated*, and particular beliefs and desires, are robust. Robust properties or variables are ones that figure in a virtually unlimited number of different kinds of causal

processes. In addition, although Campbell does not mention this, robust properties are those that are detectable by a variety of different independent methods (see Wimsatt 1981). By contrast, many sub-personal psychological variables are parts of modular systems. For example, there is good reason to think that perceptual systems, like the early visual system, are modular. Since modular mechanisms are, in some sense, isolated from other systems (more on this below), it seems that the variables that figure in them will not be robust.

There are, in fact, a wide variety of different concepts of modularity. One notion is what Gabriel Segal calls "intentional modularity" (1996, 142-3). Roughly, an intentional module is a "component of the mind." According to Segal, a body of interrelated psychological states is an intentional module if it is either *encapsulated*—information from outside the system (e.g. about the subject's beliefs and desires) is not accessible from inside the module, or *inaccessible*—information available within the module is proprietary and not available to other modules or to non-modular systems. In contrast to this disjunctive characterization, Fodor and others take encapsulation to be the essential or fundamental feature of modular systems (1983, 37, 71), perhaps serving as the explanatory ground for other features of these systems (Currie and Sterelny 2000, 147-8).[9] Other authors, especially those who, unlike Fodor, believe that the mind is "massively modular," claim that a mental system can be modular even if it is not encapsulated in the sense described above (e.g., Carruthers 2006). Rather, modular systems need only have their internal *processing* encapsulated from other systems, or perhaps modular systems need only exhibit a non-standard "wide-scope" informational encapsulation, according to which those

[9] Fodor is cagey on this point. Sometimes he seems to claim that neither encapsulation, nor inaccessibility, nor any of the features listed below, are necessary aspects of modularity. Max Coltheart argues that domain specificity is the only necessary feature of modules (1999). Currie and Sterelny (2000) argue, however, that domains cannot be specified independently of the modules themselves.

systems cannot be affected by most of the information in the mind, "without there being a determinate sub-division between the information that can affect the system and the information that can't" (Carruthers 2006, 58). However, if one requires that modular systems merely be encapsulated in one of these weaker senses, this drains modularity of much of its significance. For now, there is nothing to prevent *any* mental system from being modular. Considerations of computational tractability suggest that any system must be "wide-scope" encapsulated (ibid., 56-9). And, even the practical reasoning system, which is clearly unencapsulated in the standard "narrow-scope" sense (i.e. there is no determinate, large body of information that the system cannot access), may be processing encapsulated (Nichols and Stich 2003, 122). Regardless of how these debates turn out, it seems clear that some kind of encapsulation is a necessary feature of modular systems.

Now, some intentional modules are realized by computational modules, systems in which the output produced by any given input is completely determined by the syntactic properties of the input (and perhaps some intermediate representations). Since the 1980s, Jerry Fodor has championed the idea that some computational modules are likely to possess a number of distinguishing properties. In addition to being informationally encapsulated and being inaccessibile, a Fodorian module is also often:

> fast,
>
> one whose operation is mandatory,
>
> one with "shallow," non-conceptual or "low level" conceptual, outputs,
>
> *domain specific*: it responds only to inputs concerning a specific subject matter, and the information and mechanisms at work within it are specialized to perform a particular task,

genetically determined with a characteristic ontogeny,

realized by a fixed, often localized, neural architecture, and

one that has a characteristic pattern of breakdown.

By definition, then, every Fodorian module is realized by a neural module – "a functional component of the brain, describable in purely neural terms" (Segal 1996, 145). However, intentional and computational modules need not be realized by localized, discrete neural modules; they could be realized by distributed, nearly global neural systems.

Intentional, computational, and Fodorian modularity form a nested sequence of increasingly strong concepts of modularity. Every Fodorian module is computational, and every computational modular is intentional, but the converse claims do not hold.

Campbell suggests that it is only within a Fodorian module that we are likely to move cleanly between psychological and neurological levels. For, according to Campbell, cognitive variables within in a Fodorian module are not robust; we need to descend to the biological level of realization in order to find robust variables. By contrast, if we have (a set of) robust psychological variables, then they cannot be involved in Fodorian modules. And, if Campbell is right, then there is no demand to search for biological variables underlying these psychological variables.

Why should we think that cognitive variables involved in Fodorian modules are not robust? Domain specificity by itself does not preclude robustness, since, as discussed below, a variable could be involved in a domain specific mechanism and yet also participate in operations that are unrelated to that domain (in which case the variable would be a subcomponent shared by multiple mechanisms). However, the inaccessibility and encapsulation of modular systems does seem to preclude the robustness of cognitive variables involved in those systems. The

intentional content of these variables is determined by their role in the processing of that particular modular system. Since at least the processing in this system is insulated from outside influence, and this processing is essentially involved in the individuation of the content of the variables, these cognitive variables do not have application outside of the system. In particular, cognitive variables involved in an inaccessible system are not detectable by a wide variety of psychological means, and cognitive variables involved in an encapsulated system do not play a role in many different kinds of psychological causal processes. Since cognitive variables in modular systems are not robust, cognitive explanations in, e.g., vision science, must be grounded in neural and biochemical mechanisms. Assuming that cogent explanations must involve robust variables, a cognitive characterization of modular systems is at best a sketch or placeholder for a biochemical explanation. And, other features of modular systems suggest that these biochemical mechanistic explanations will be in the offing. If a cognitive system is realized by a localized brain region, then each variable within it is likely to correspond to a unique (set of) neurological or biochemical variable(s).

I believe that exploring the relation between psychological and neurological variables and the connections between robustness and modularity is important and has the potential to reveal important implications for which methods are most effective for studying the mind. However, I think that the relations between robustness, modularity and reductive explanations are more complicated than is indicated by Campbell's brief, suggestive remarks. While the intentional, cognitive variables within a Fodorian module may be non-robust, there is no guarantee that they can be cleanly replaced with biochemical variables. Even within a Fodorian module, it may not be possible to move seamlessly between the cognitive and biochemical levels. And, even if the mind as a whole is "massively modular," this does not guarantee that these modular systems map

cleanly onto our best biochemical representations of the brain. In the remainder of this section, I argue for these claims and, use the considerations in this argument to *explain* why robust psychological variables in a non-modular system are not replaceable with biochemical or other neural variables.

Francois Jacob is famous for claiming that evolution by natural selection is a "tinkerer" that works with whatever materials are on hand. For this reason, novel cognitive systems are likely to consist of "new way[s] of putting together and modifying a broad array of previously existing subsystems. Different parts of the brain are probably specialized for different functions, but most of those functions are likely to be shared *subcomponents* for computation, not complete systems for single-handedly solving complex cognitive tasks" (Marcus 2004, 133). There are other evolutionary reasons to think that a kind of functional pleiotropy, what I have elsewhere called "multiple determinativity," will be common in the mind, especially in recently evolved, high-level cognitive systems. In most animals, it appears that modularity itself evolved via the parcellation of previously integrated and multi-functional parts (Wagner and Altenberg 1996, 972). Since this parcellation must take place on an evolutionary timescale and high-level cognitive systems are relatively new, it is likely that some aspects of the mind and brain remain "unparcellated," wherein an individual neural structure plays, or has been co-opted, for different cognitive functions.

This kind of pleiotropy or multiple determinativity can hold at (at least) two different scales: (1) within a given neural module, a particular biochemical variable may play multiple psychological functions, what we might call *fine-grained multiple determinativity*; (2) within an entire modular neurological theory, a particular neural module may play more than one psychological function or may be involved in different intentional modules, a kind of *course-*

grained multiple determinativity. There are of course type and token versions of each of these. When we have token multiple determinativity (of either individual variables or entire modules), we have a simple sharing of parts (which could be either synchronic or diachronic). This kind of multiple determinativity is a plausible explanation of the SNARC effect (when subjects are asked whether a number is odd or even, they tend to respond faster to large numbers with their right hand, and faster to small numbers with their left hand). There may be a high degree of overlap between the neural circuits responsible for spatial representation and those responsible for number representation. When we have type multiple determinativity, different tokens of the same type of variable or module are used for different purposes in the service of distinct functional systems.

Even if a psychological variable is part of an intentional module and that intentional module is (uniquely) realized by a neural module, this does not guarantee that the psychological variable is can be replaced by some biological variable. For, suppose that we have fine-grained multiple determinativity, where a single biochemical variable plays several different cognitive roles. In that case, if we tried to replace these cognitive variables with the multiply determinative biological variable, we would end up conflating some causal relationships that hold between these distinct psychological variables. The one-many relation entailed by multiple determinativity prevents informative reductions that preserve all of the psychological causal relations at the biological level.

This indicates that even *non-robust* psychological variables may not be replaceable by biological variables. Even if a psychological system possesses all of the properties of a Fodorian module, the variables involved in it may not be replaceable by biochemical variables. We will

not be able to move cleanly and easily between the intentional, computational, and realization levels.

Now, suppose that the mind is massively modular—that all of our psychology is underwritten by intentional modules. As noted above, this is compatible with any given intentional module not being realized by a single, localized neural module. There are several options here. First, a given intentional module may be realized by different parts of the brain at different times (e.g. the kind of neural plasticity that is sometimes evidenced after brain trauma). If the neural architecture underlying the psychological module is highly plastic, then we have a kind of multiple realization of the intentional module and its component variables. Thus, many neural variables correspond to a single psychological variable, and the only aspects that may group these neural variables into a kind are intentional or cognitive ones. Alternatively, we may have the kind of coarse-grained multiple determinativity introduced above. Suppose that even though a given intentional module is realized by a collection of fixed, localized neural modules, some of those neural modules play roles in other intentional modules. If this kind of mismatch between modularity at different levels occurs, then even though a given set of psychological variables might not be robust, there will be no set of biological variables that can be mapped uniquely to that set of psychological variables and no other. We have a structurally similar situation to fine-grained multiple determinativity, except here the one-many relation obtains between neural submodules and intentional submodules.

Finally, if we have robust psychological variables, Campbell suggests that have no need to search for robust biological variables, but he gives us no positive argument against the replacement of these robust psychological variables. He bases his contention solely on the supposed lack of empirical evidence for the existence of underlying biological mechanisms. The

varieties of multiple determinativity give us an argument against reducibility (in addition to the traditional multiple realizability argument). Further, an additional kind of very coarse-grained multiple determinativity may play a role in the realization of robust psychological variables. A non-modular, central cognitive system will likely be realized by a distributed neural system that involves a significant portion of the cortex. If a given psychological variable is realized by such a highly distributed neural system, it is likely that that same neural system will also realize many other robust, non-modular psychological variables. Of course, different aspects of these systems may be involved in realizing the different psychological variables, but identifying these aspects will likely itself rely on psychological factors.

5. Concluding Remarks

Biochemical and neurobiological research has exploded in recent years, and it is hard to deny that broadly reductive research strategies have taken the lead in scientific exploration of the mind and brain. The discussion in this paper suggests ways in which social and cognitive psychology put constraints on neuroscientific research. However, this does not support a traditional anti-reductive view according to which psychological research is "autonomous" and can be pursued completely independently of neurology and biochemistry. On the other hand, it is not compatible with a reductionist view, either, according to which psychological variables are uniformly replaced with biochemical ones. Rather, the relation between psychology and neurobiology suggested here is a reciprocal, "bootstrapping" one, where results in each field put constraints on the other.

For example, the discussion of nurturance and memory and of modularity suggests that "personal level" psychological research, such as cognitive and psychophysical testing, has a role to play even in the investigation of non-robust variables in Fodorian modular systems. Further,

individuation of some neural states should be sensitive to the environmental, social, and psychological context. Even if neural structures are individuated functionally (cf. Bechtel and Mundale 1999), the functional nature of these neural structures cannot be completely described in neural terms. Functional individuation will draw essentially on psychological theory, if, as is likely, different neural features realize the same psychological feature (in different individuals or in the same individual at different times) or a single neural feature realizes several different psychological features (as discussed above). Finally, in any neuroscientific theory that accurately models psychological phenomena, the set of biochemical variables will likely have many interaction terms with social and environmental variables (cf. Caspi and Moffitt 2006). Investigating these interaction terms will require work in social and cognitive psychology and will be crucial for identifying and mapping the causes of certain psychological phenomena. Future research into the biochemical variables underlying our cognitive and affective capacities should be guided by our best research methods. If this paper is on the right track, then this research will continue to support the conclusion that biochemical variables supplement, but do not replace, psychological variables.

The radical anti-reductionist view which is open to the non-existence of biochemical mechanisms for mental causation is likely false because it is incompatible with the completeness of physics. However, even physicalists can accept the weaker, more general claim lying behind it: psychological variables will continue to be an ineliminable part of our best accounts of the mind.

References

Ashby, F. Gregory, Alice M. Isen, and And U. Turken. (1999) "A Neuropsychological Theory of Positive Affect and Its Influence on Cognition." *Psychological Review*. 106: 529-550.

Bechtel, William and Jennifer Mundale. (1999) "Multiple Realizability Revisited: Linking Cognitive and Neural States." *Philosophy of Science*. 66: 175-207.

Bickle, John. (2003) *Philosophy and Neuroscience: A Ruthlessly Reductive Account*. Dordrecht: Kluwer.

Campbell, John. (2007) "An Interventionist Approach to Causation in Psychology." in *Causal Learning: Psychology, Philosophy and Computation*. Ed. Alison Gopnik and Laura Schulz, eds. Oxford: Oxford UP. 58-66.

Carruthers, Peter. (2006) *The Architecture of the Mind*. New York: Oxford UP.

Caspi, Avshalom and Terrie E. Moffitt. (2006) "Gene-environment Interactions in Psychiatry: Joining Forces with Neuroscience." *Nature Reviews. Neuroscience*. 7: 583-590.

Currie, Gregory and Kim Sterelny. (2000) "How to Think About the Modularity of Mind Reading." *Philosophical Quarterly*. 50: 145-160.

Dehaene, Stanislas et al. (1993) "The Mental Representation of Parity and Number Magnitude." *Journal of Experimental Psychology*. 122: 371-396.

Coltheart, Max. (1999) "Modularity and Cognition." *Trends in Cognitive Sciences*. 3: 115-120.

Farah, Martha J. et al. (2006) "Childhood Poverty: Specific Associations with Neurocognitive Development." *Brain Research*. 1110: 166-174.

Farah, Martha J. et al. (2008) "Environmental Stimulation, Parental Nurturance and Cognitive Development in Humans." *Developmental Science*. Published online, 17 Apr. 2008: 1-9.

Fodor, Jerry. (1983) *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fredrickson, Barbara L. (2001) "The Role of Positive Emotions in Positive Psychology: The Broaden-and-Build Theory of Positive Emotions." *American Psychologist*. 56: 218-226.

Gross, Cornelius and Rene Hen. (2004) "The Developmental Origins of Anxiety." *Nature Reviews Neuroscience*. 5: 545:552.

Hollon, S.D. et al. (2002) "Treatment and Prevention of Depression." *Psychological Science in the Public Interest*. 3: 39-77.

Kendler, Kenneth S. et al. (2003) "Life Event Dimensions of Loss, Humiliation, Entrapment, and Danger in the Prediction of Onsets of Major Depression and Generalized Anxiety." *Archives of General Psychiatry*. 60: 789-796.

Kendler, Kenneth S. et al. (2005) "The Interaction of Stressful Life Events and a Serotonin Transporter Polymorphism in the Prediction of Episodes of Major Depression. A Replication." *Archives of General Psychiatry*. 62: 529-535.

Kim, Jaegwon. (1998) *Mind in a Physical World*. Cambridge, MA: MIT Press.

Liu, Dong et al. (2000) "Maternal Care, Hippocampal Synaptogenesis and Cognitive Development in Rats." *Nature Neuroscience*. 3: 799-806.

Marcus, Gary (2004) *The Birth of the Mind*. New York: Basic Books.

McEwan, Bruce S. (2000) "The Neurobiology of Stress: From Serendipity to Clinical Relevance." *Brain Research*. 886: 172-189.

Nichols, Shaun and Steven Stich. (2003) *Mindreading*. New York: Oxford UP.

Noble, Kimberly G. et al. (2007) "Socioeconomic Gradients Predict Individual Differences in Neurocognitive Abilities." *Developmental Science*. 10: 464-480.

Papineau, David. (2001) "The Rise of Physicalism." in *Physicalism and Its Discontents*. Ed. Carl Gillett and Barry Loewer. Cambridge: Cambridge UP, 3-36.

Segal, Gabriel. (1996) "The Modularity of Theory of Mind." in *Theories of Theories of Mind*. Ed. Peter Carruthers and Peter K. Smith. Cambridge: Cambridge UP.

Shoemaker, Sydney. (2001) "Realization and Mental Causation." in *Physicalism and Its Discontents*. Ed. Carl Gillett and Barry Loewer. Cambridge: Cambridge UP. 74-98.

Wagner, Gunter P. and Lee Altenberg. (1996) "Complex Adaptations and the Evolution of Evolvability." *Evolution*. 50: 967-976.

Wimsatt, William. (1981) "Robustness, Reliability and Overdetermination." in *Scientific Inquiry and the Social Sciences*. Ed. M.B. Brewer and B.E. Collins. San Francisco: Jossey-Bass.

Woodward, James. (2003) *Making Things Happen*. Oxford: Oxford UP.