Res Cogitans

Volume 8 | Issue 1 Article 2

2017

Two Objections to the Eliminativist Research Program

Lorenzo Nericcio
California Polytechnic State University, lorenzo_nericcio@yahoo.com

Recommended Citation

Nericcio, Lorenzo () "Two Objections to the Eliminativist Research Program," *Res Cogitans*: Vol. 8: Iss. 1, Article 2. https://doi.org/10.7710/2155-4838.1161

Two Objections to the Eliminativist Research Program

Lorenzo Nericcio
Portland State University

Published online: September 7 2017

Abstract

Eliminative materialist philosophers, like Paul and Patricia Churchland, argue that the common use of mental state language is confused. They hold that neurological descriptions of mental states, more accurate and scientifically rigorous than "folk psychology", should replace mental state language in a serious research program. In this paper, I argue that eliminative materialism instead poses an awkward and unwieldy research program. I take a computational functionalist position in order to demonstrate the way that mathematical descriptions of natural phenomena are useful in a scientific research program, and that mental states are in principle amenable to mathematical descriptions and modeling. I then argue that the eliminativist cannot avail herself of the same resources.

Introduction

The eliminative materialist maintains that the use of "mental talk," such as refers to mental states as identifiable entities or processes, is fundamentally confused. She holds that, as any good physicalist will reduce a mental state to its correspondent brain state anyway, a stronger science will not refer to loosely defined and essentially private reports of mental states at all—let alone attempt to reduce them. It is therefore the task of the neuroscientist to identify relevant brain phenomena corresponding to certain observed behaviors in order to provide a reproducible, public science of the brain; proceeding as such, she eliminates mind as an unnecessary posit. I will advance a computational, functionalist thesis that runs contrary to the eliminative account. I maintain that in eliminating a given mental state, one neglects to observe its computational state, which, though instantiated in a neurological system, is not justifiably eliminated (note: I do still maintain a physicalist assumption). Plainly, locating brain processes responsible for certain observed behavior is insufficient evidence to justify eliminating a computationally describable mental state. I will argue that mental states equal the causal roles and computational states they occupy; the same mental state instantiated in two different media is the same so long as the functional role and computational state is the same. From this position, I raise two objections: (A) the eliminativist denies the ontological status of functional states on a "higher" level of supervenience, and (B) due to this elimination, the eliminativist research program is unwieldy and unclear. A computational theory of mind does not struggle with either problem.

The Eliminativist's Thesis[1]

If the entities posited by a scientific theory continue to allow accurate predictions to be made, it becomes more and more likely that the entities quantified over by that theory actually exist.[2] Consider two theories about the causes of illness: witch theory and germ theory. Witch theory quantifies over witches. It posits that for any given illness there must have been a witch who caused it. Using this theory, one might make certain predictions about the world, but it is unsurprising that his predictions will turn out wildly inaccurate. The germ theorist applies her theory and finds that it allows for tremendous predictive success, and so it is very likely that the entities, germs, quantified over by the theory actually exist – and that witches do not. So it is, says the eliminativist, with mental states and neurological states. When one uses *folk psychology* (a term which refers to the ordinary language discourse about propositional attitudes) to predict the nature and outcomes of certain mental states, he is likely to be wildly inaccurate[3] (or so maintains the eliminativist). Accounts given by neuroscience and cognitive science are far more predictively accurate than those given by folk psychology, and so it is much more likely that the entities those sciences quantify over (neurological processes) exist than that the entities—that is, mental states (desires, beliefs, etc.)—quantified over by the folk psychologist exist. Moreover, the reason one is an *eliminativist* is that she does not attempt to identify, say, germs as *functionally witches;* she simply eliminates the entity *witch* from her ontology. Similarly, the eliminativist does not identify brain states as mental states; she simply eliminates the mental states from her ontology. This outlines the position that I aim to refute.

The Computational Model

The computational model is a species of functionalism that defines any given mental state as the computational state of an automaton.[4] My argument is informed by Hilary Putnam's[5] and Jerry Fodor's[6] work on computational functionalism. The thesis states:

---

[1] My treatment of the issue is informed by the claims in Paul Churchland's in *Eliminative Materialism and the Propositional Attitudes* (1981).

[2] This view is due to W.V.O Quine: *Posits and Reality* (1955)

[3] The predictive success of folk psychology is very obviously up for dispute. I will not, however, take it up here directly.

[4] These computational states account for the causal roles played by the mental state, as they may consist of a finite set of instructions—an algorithm.

[5] Putnam, Hilary. "Psychological Predicates." *Art, Mind, and Religion.* (1967).

[6] Fodor, Jerry A. "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology. *The Behavioral and Brain Sciences III.* (1980).

the human mind is a very complex automaton, and so mental states can be understood as the computational states of this automaton. An automaton is a theoretical machine that contains a finite number of states that are sequentially ordered. Inputs change the system from one state to the next, and produce an output based on a set of instructions specific to each state as it functions given an input at that state: e.g. an input of “1” at S_1 may instantiate the instruction, “change to S_2.” Simple finite automata are found in such common objects as vending machines and elevators. If one presses a button (the input), there is, say, a change in the state of the machine from S_1 to S_2, which initiates a specific output resulting from that change: delivering your bag of chips. This concept can be developed to the full-fledged Turing Machine. The Turing Machine, a theoretical representation of a very complex automaton, can account for a far greater level of complexity in calculation. It relies on a theoretically infinitely long tape divided into cells. Each of these cells contains binary (say, a 1 or 0) information that the Turing Machine “reads,” and then produces a certain output based on a set of instructions contained within the machine. (For instance, if the input cell contains “1” the instruction might be “move left, change S_2 to S_3, print a 0” which will produce some output O_1.)

These basic concepts in place, I will apply them to questions about mental states. Putnam takes the daunting task of applying computational functionalism to pain; I demonstrate it using a simple belief, which is rather less daunting. If Smith believes that tomorrow is Friday, and Smith believes that tomorrow is payday, then Smith understands that tomorrow is payday. This systematic structure is clearly a logical transition from one state to the next (not unlike the structure of a Turing Machine). The connection from one belief to its logical implication certainly can be mapped onto the celled tape of a Turing Machine. Because so much of thought occurs in an inferential way[7], there is good reason to believe that using mathematical, computational theories of this kind, one might map the entirety of every functional state of the human mind onto some sequence of logical inferences.[8]

Functionalism thus solves problems arising from the multiply realizable nature of mental states. Squids and humans can both react to light, but it is unlikely that squid-brains and human-brains have the same physical-chemical structure. Instead, observe that given a certain input, light, there is some change of the computational state in the mind of the squid or the human. To simplify for demonstration, given an input I_1, there might be a change from mental state M_1 to M_2, producing output O_1. This structure, just as with any automaton, need not rely on some chemically specific medium to occur so long as the defined inputs, outputs, and state transitions remain intact.

[7] Whether or not these inferences are accurate or true of the world is irrelevant to the matter that they still structure thought.
[8] This, while controversial, is a corollary of the Church-Turing thesis.

Another important feature of the computational theory is that it employs syntax to structure thought. Computers manipulate strings of syntactically ordered symbols to produce ever more complex logically valid systems. Minds appear to do much the same through language. An important feature of a syntactically valid statement is that it respects the semantic values of its symbols in any operation. No matter what semantic values are assigned the atomic sentences in, say, *modus ponens*, the valid structure will preserve them. Thus, though I will not speak directly to the production of semantic content in the mind here, suffice it to say that a benefit of this basic computational model is that it leaves its structure open to be filled by a well-developed theory of semantic content, where Eliminativism excludes the possibility of semantic content. I take this as a benefit of computational functionalism moving on to the following considerations.

All this in place, I may move on to explain the supervenience of mental states on physical states. It is upon the supervenience of mind on brain that my main argument against the eliminativists will rely.

Supervenience and Computational States[9]

It is trivially true of non-eliminative[10] physicalist theses that mental states supervene on physical states. If some organism O is in mental state M_1, then there must be some underlying brain state B_1 that has determined it. Any change from M_1 to M_2 would imply some change of B_1 to B_2. For the functionalist, one may assume that the change in M_1 and M_2 must mean some change in the *functional orientation* from $B_1 = F_1$ to $B_2 = F_2$ (where F_i denotes a functional state) in order to effect the change from M_1 to M_2. So, I will take it that fixing a mental state always fixes its corresponding functional state, and as functional states are instantiated by the brain (for humans, anyway), mental systems supervene on brain systems. Any computational state of the mind that is in principle describable must supervene on the brain state that instantiates it. I will use this point to object to the eliminative theory by first using a thought experiment.

Problems for the Eliminativist

Take as an example a different macroscopic system that supervenes on a microscopic system: a cat. Cats supervene on certain physiological and microbiological systems that make up their bodies. Thus, some change in the position, behavior, size, etc. of a cat indicates a change on the microbiological level; I take it that cat (C) states supervene on

[9] I rely on Kim, Jaegwon. "The Mind-Body Problem: Taking Stock after Forty Years." *Philosophical Perspectives 11: Mind, Causation and the World*. (1997) for my general conception of mind-brain supervenience.

[10] Eliminativists do not attempt to explain the way the mind supervenes on the brain, as they just do away with the mind.

microbiological (K) states. A team of cat scientists may take it upon themselves to develop a description of every possible physical orientation of a cat's body into its C states: C_1 through C_n. Each of these C states, according to the supervenience thesis, is fixed by some orientation of its microbiological components: K_1 through K_n states. If this is true, then these scientists are proposing two sets of theoretical entities: cat entities and microbiological entities. Now, Imagine that another team of *eliminative* cat scientists catches wind of the efforts of the first group and makes an audacious claim: If the orientation of a cat into some state C_2 (say, sleeping) may be described by some microbiological state K_2, then we may deny the existence of the purported C state altogether. Further, the eliminativists press, people are very often mistaken about whether cats are asleep, or whether they are sitting or laying down, or whether or not they are hungry. At the level of the K state, however, these facts can be discovered with more certainty. There are, we can assume, clear microbiological indicators that a cat is hungry, or whether its legs are bent such that it is sitting, or whether it is actually asleep. Thus, the entities quantified over by the eliminative cat theorist – microbiological entities – are much more likely to exist than the entities quantified over by the original cat theorist (if you like, the *folk* cat theorist).

Having recently received a grant, however, the original team of cat scientists expand their research and discover that dogs can also realize many C states: they can sleep, they can sit, they can be hungry, etc. After reading the subsequently published journal article, the eliminative cat theorists seek to account for this apparent correlation. Sticking to their guns, they employ their microscopes and discover that each apparent C state can be, in dogs, decomposed into some microbiological D state. Once again, they feel justified in rejecting the C states and maintaining that all that *really* exist are D states.

The problem for the eliminativist is that the C state still successfully describes the *functional* role of the cat or the dog. Sitting, sleeping, or eating can be described in terms of some function that the animal performs at a given time, the description of which is not limited to the properties upon which it supervenes: sitting is not a property of molecules. One could devise a geometric model that describes what sitting might mean for any four-legged, and thus isomorphic enough to be able to instantiate a similar set of functional states, animal—or even a well constructed robot. If this is the case, and provided the geometric model is correct, it is not justified to say that this functional state *does not exist*. The fact that a system's functional properties supervene on some other micro-level properties is not sufficient to propose that said system does not exist, especially if an accurate model of the macro-level states describes things that are not so clearly isomorphic at the micro-level (e.g. cat-sitting given in K terms vs. dog-sitting given in D terms).

The case is analogous for mind.[11] Mental states can clearly instantiate certain functional roles: performing mathematical operations, using language correctly, identifying flowers, making logical inferences, etc. These functional roles are likely to be states of the mind that are relations from the given input (say, a poppy flower), to an internal state (some internal function which identifies the poppy flower-perception with the word “poppy”), and giving some output (I say “Oh, a poppy!”). These inferential connections and causal roles can be described and instantiated in multiple individuals (or even in a robot, or some such non-human intelligence). A science that quantifies over certain mathematically describable states will likely enjoy a great deal of predictive success; consider the predictive success mathematically described quantum states and the electrons posited in those states. This could in principle meet the eliminative materialist on her own terms for the ontological status of a scientifically posited entity. Such a science would require that inferential networks are reproducible across many cases. I see no issue in this; most people employ the basic structure of an inferential network leading from background information to a probable conclusion regularly. I now bring my two aforementioned objections to bear: (A) and (B) above.

The thought experiment above demonstrates (A) as a problem for an eliminative program. Mathematical models can be used to describe many phenomena in nature, and provide a coherent structure for understanding what features of certain systems are equivalent across many examples. The orbit of planets, and the orbit of electrons, the radioactive decay of isotopes useful in radiocarbon dating, etc. all rely on a consistent mathematical schema that can be applied to many instances. Science consists of such mathematical schema.

Even simple, non-mathematically described biologically functional roles are consistently applied in scientific theory building. Consider the concept of convergent evolution in biology. Wings, for instance, appear in both birds and insects. However, birds and insects do not share any relevant winged common ancestor; they independently evolved wings in reaction to whatever environment happened to bring this about. Both function to cause the animal to fly, and so both function as a wing. The problem for Eliminativists that I demonstrate is that, even if all of these above examples (C-states, wings) are decomposable into theoretically more predictively accurate parts, the eliminative thesis has denied precisely the attribute which gave cause to investigate them in the first place: the functional features. Wings are able to produce flight because of their interaction with air such that lift is achieved. This lift is mathematically describable: ($p + 1/2pV^2 + pgh =$ constant). Thus, the, say, *pressure-to-volume state* of a wing can be determined in terms of inputs to the given formula, the output of which is the state of some wing’s lift. Analogously, mental states may be given in terms of their computational states because they process information; some formula describes the inputs and the subsequently produced output. Denying the existence of these

[11] Mind-states supervene on brain-states similarly to the way that my C-states supervene on K-states.

computational states, in my view, commits one to denying the existence of the mathematically describable relationship between wings and the air around them when considering lift.

Understanding the relationship between a mathematical model and the phenomenon for which it accounts brings about my objection (B). If one takes the eliminativist at her word, the resulting research program would be startlingly unwieldy. The various computational states present in a human mind at a given time, as established in section III, will realize some functional role with computable inputs, outputs, and state transitions. If this is the case, then these states are likely to be common among many individuals. Three people all in a state of *believing that Tower Bridge is in London* are likely to have analogous functional states. The functional role that this belief occupies will have logically analogous corollaries: e.g., "Tower Bridge is in London. I wish to see Tower Bridge. Therefore, I will travel to London." Understanding human motivation and belief at this level is predictively quite successful (given a normative set of cases[12]); psychologists engage with belief and motivation this way as a method of practice. To highlight to potential absurdity, imagine a case where one was required to understand behavior causation without these inferential steps taken into account at all. The psychologist would have to probe the neuroanatomical features of the subject until the underlying biochemical causes were revealed—and that is already assuming that such biochemical events could actually etiologically linked to a given behavior.

This seems highly implausible: as another example, consider an analogous case wherein one's web browser failed to function properly (perhaps pop-ups have clogged up the interface) and in response she calls an electrician. Of course, the computational states realized in a desktop computer while running a web browser program supervene on electrical structures built in to the computer's hardware, but this does not mean that predictive models of behavior at the software level are best understood in terms of hardware. The program running a web browser can be described independently (because it can be realized in many computers) of whatever particular desktop computer has instantiated it.

Concluding Remarks

I have illustrated the relationship between a functional role, which is described by a computational state given inputs and outputs, and a formal description of mental states. I argue that because a functional role can in principle be described in the terms of a mathematical theory—namely, by the set of computational states realized by the mind—to deny the mind's existence is analogous to denying the existence of any other mathematically describable phenomenon that is realized across multiple instances. This

[12] It is probable that in abnormal minds many of the computational states realized are vastly different from those in the normative subject; that does not mean they do not exist, however.

denial, I contend, presents an awkward and implausible research program for both philosophers and scientists. I have not considered the specific way by which, at the level of neurons, thought processes supervene on brain processes, but this is precisely because I hold that the causal and inferential descriptions of the thought processes (i.e. mental states) themselves do not rely on the chemical structure of neurons so much as the functional orientation that they realize. (Consider that the Bernoulli equation given above applies to wings made from feathers, wood, or aluminum. The pressure-to-volume function relation nevertheless obtains.) And it is this functional orientation that I propose is, in principle, amenable to a mathematical model.