

The China Brain

by Ned Block

from *Troubles With Functionalism* (1978)

Functionalism, Behaviorism, and Physicalism

The functionalist view of the nature of the mind is now widely accepted. Like behaviorism and physicalism, functionalism seeks to answer the question “What are mental states?” I shall be concerned with identity thesis formulations of functionalism. They say, for example, that pain is a functional state, just as identity thesis formulations of physicalism say that pain is a physical state. ...

One major thesis of this article is that, because of this feature, functionalism fails to avoid the sort of problem for which it rightly condemns behaviorism. Functionalism, too, is guilty of liberalism, for much the same reasons as behaviorism.

The failing I speak of is the one that functionalism shows *physicalism* to be guilty of. By ‘physicalism’, I mean the doctrine that pain, for example, is identical to a physical (or physiological) state [e.g., C-fibers firing]. As many philosophers have argued, if functionalism is true, physicalism is probably false. The point is at its clearest with regard to Turing-machine versions of functionalism. Any given abstract Turing machine can be realized by a wide variety of physical devices. ... Therefore, if pain is a functional state, it cannot, for example, be a brain state, because creatures without brains can realize the same Turing machine as creatures with brains.

I must emphasize that the functionalist argument against physicalism does not appeal merely to the fact that one abstract Turing machine can be realized by systems of different *material composition* (wood, metal, glass, etc.). ... Rather, the functionalist argument against physicalism is that it is difficult to see how there *could be* a nontrivial first-order physical property in common to all and only the possible physical realizations of a given Turing-machine state. Try to think of a remotely plausible candidate! At the very least, the onus is on those who think such physical properties are conceivable to show us how to conceive of one.

One way of expressing this point is that, according to functionalism, physicalism is a chauvinist theory: it withholds mental properties from systems that in fact have them. In saying mental states are brain states, for example, physicalists unfairly exclude those poor brainless creatures who nonetheless have minds. ...

Homunculi-Headed Robots

In this section I shall describe a class of devices that are *prima facie* embarrassments for all versions of functionalism in that they indicate functionalism is guilty of liberalism—classifying systems that lack mentality as having mentality. ...

Imagine a body externally like a human body, say yours, but internally quite different. The neurons from sensory organs are connected to a bank of lights in a hollow cavity in the head. A set of buttons connects to the motor-output neurons. Inside the cavity resides a group of little men. Each has a very simple task: to implement a “square” of an adequate machine table that describes you. On one wall is a bulletin board on which is posted a state card, i.e., a card that bears a symbol designating one of the states specified in the machine table. Here is what the little men do: Suppose the posted card has a ‘G’ on it. This alerts the little men who implement G squares—‘G-men’ they call themselves. Suppose the light representing input I_{17} goes on. One of the G-men has the following as his sole task: when the card reads ‘G’ and the I_{17} light goes on, he presses output button O_{191} and changes the state card to ‘M’. This G-man is called upon to exercise his task only rarely. In spite of the low level of intelligence required of each little man, the system as a whole manages to simulate you because the functional organization they have been trained to realize is yours. A Turing machine can be represented as a finite set of quadruples (or quintuples, if the output is divided into two parts): current state, current input; next state, next output. Each little man has the task corresponding to a single quadruple. Through the efforts of the little men, the system realizes the same (reasonably adequate) machine table as you do and is thus functionally equivalent to you.

I shall describe a version of the homunculi-headed simulation, which has more chance of being nomologically possible. How many homunculi are required? Perhaps a billion are enough:

The China Brain: Suppose we convert the government of China to functionalism, and we convince its officials to realize a human mind for an hour. We provide each of the billion people in China (I chose China because it has a billion inhabitants) with a specially designed two-way radio that connects them in the appropriate way to other persons and to the artificial body mentioned in the previous example. We replace each of the little men with a citizen of China plus his radio. Instead of a bulletin board we arrange to have letters displayed on a series of satellites placed so that they can be seen from anywhere in China.

The system of a billion people communicating with one another plus satellites plays the role of an external “brain” connected to the artificial body by radio. There is nothing absurd about a person being connected to his brain by radio.

Perhaps the day will come when our brains will be periodically removed for cleaning and repairs. Imagine that this is done initially by treating neurons attaching the brain to the body with a chemical that allows them to stretch like rubber bands, thereby assuring that no brain-body connections are disrupted. Soon clever businessmen discover that they can attract more customers by replacing the stretched neurons with radio links so that brains can be cleaned without inconveniencing the customer by immobilizing his body.

It is not at all obvious that the China-body system is physically impossible. It could be functionally equivalent to you for a short time, say an hour. ... In describing the Chinese system as a Turing machine, I have drawn the line in such a way that it satisfies a certain type of functional description—one that you also satisfy, and one that, according to functionalism, justifies attributions of mentality. ...

Objection: The Chinese system would work too slowly. The kind of events and processes with which we normally have contact would pass by far too quickly for the system to detect them. Thus, we would be unable to converse with it, play bridge with it, etc.

Reply: It is hard to see why the system's time scale should matter. Is it really contradictory or nonsensical to suppose we could meet a race of intelligent beings with whom we could communicate only by devices such as time-lapse photography? When we observe these creatures, they seem almost inanimate. But when we view the time-lapse movies, we see them conversing with one another. Indeed, we find they are saying that the only way they can make any sense of us is by viewing movies greatly slowed down. To take time scale as at all important seems crudely behavioristic.

What makes the homunculi-headed system (count the two systems as variants of a single system) just described a prima facie counterexample to (machine) functionalism is that there is prima facie doubt whether it has any mental states at all—especially whether it has what philosophers have variously called “qualitative states,” “raw feels,” or “immediate phenomenological qualities.” (You ask: What is it that philosophers have called qualitative states? I answer, only half in jest: As Louis Armstrong said when asked what jazz is, “If you got to ask, you ain't never gonna get to know.”) In Nagel's terms (1974), there is a prima facie doubt whether there is anything which it is like to be the homunculi-headed system. ... Call this argument the Absent Qualia Argument. ...

Is the Prima Facie Doubt Merely Prima Facie?

The Absent Qualia Argument rested on an appeal to the intuition that the homunculi-headed simulations lacked mentality, or at least qualia. I said that this intuition gave rise to prima facie doubt that functionalism is true. ...

Appeal to intuitions when judging possession of mentality, however, is *especially* suspicious. No physical mechanism seems very intuitively plausible as a seat of qualia, least of all a brain. Is a hunk of quivering gray stuff more intuitively appropriate as a seat of qualia than a covey of little men? If not, perhaps there is a *prima facie* doubt about the qualia of brain-headed systems too?

However, there is a very important difference between brain-headed and homunculi-headed systems. Since we know that *we are brain-headed systems*, and that we have qualia, we know that brain-headed systems can have qualia. So even though we have no theory of qualia which explains how this *is possible*, we have overwhelming reason to disregard whatever *prima facie* doubt there is about the qualia of brain-headed systems. Of course, this makes my argument partly *empirical*—it depends on knowledge of what makes us tick. But since this is knowledge we in fact possess, dependence on this knowledge should not be regarded as a defect.

There is another difference between us meatheads and the homunculi-heads: they are systems designed to mimic us, but we are not designed to mimic anything (here I rely on another empirical fact). This fact forestalls any attempt to argue on the basis of an inference to the best explanation for the qualia of homunculi-heads. The best explanation of the homunculi-heads' screams and winces is not their pains, but that they were designed to mimic our screams and winces.

Some people seem to feel that the complex and subtle behavior of the homunculi-heads (behavior just as complex and subtle—even as “sensitive” to features of the environment, human and nonhuman, as your behavior) is itself sufficient reason to disregard the *prima facie* doubt that homunculi-heads have qualia. But this is just crude behaviorism.

My case against Functionalism depends on the following principle: if a doctrine has an absurd conclusion which there is no independent reason to believe, and if there is no way of explaining away the absurdity or showing it to be misleading or irrelevant, and if there is no good reason to believe the doctrine that leads to the absurdity in the first place, then don't accept the doctrine. I claim that there is no independent reason to believe in the mentality of the homunculi-head, and I know of no way of explaining away the absurdity of the conclusion that it has mentality (though of course, my argument is vulnerable to the introduction of such an explanation). ...

... Here is a summary of the argument so far: Functionalism has the bizarre consequence that a homunculi-headed simulation of you has qualia. This puts the burden of proof on the Functionalist to give us some reason for believing his doctrine. ...