Neil Melkot

Professor King

Minds, Machines, and Persons

4/7/2025

The Ghost in the Network: On the Edge of Function and Feeling

In the mid-twentieth century, philosophers sought new ways to understand the mind as a natural phenomenon, moving beyond Cartesian dualism. One approach was the identity theory, which held that mental states just are physical brain-states, for example, that the sensation of pain is identical to some specific neural firing pattern. However, Hilary Putnam soon proposed a compelling alternative known as machine functionalism, which treats mental states as functional states of a system rather than as particular material states of the brain. In his 1967 paper "The Nature of Mental States," Putnam argues that minds can be understood by analogy with Turing machines or computational automata. According to this view, what makes something a mental state, like pain or belief, is not what it is made of, but what role or function it plays in the overall system. This essay will explain Putnam's version of machine functionalism and how it conceives mental states as analogous to machine states, then contrast his view with the identity theory by showing why functionalism is superior to a brain-state identity view. Next, I will present Ned Block's famous China Brain thought experiment as an objection to functionalism and analyze how this scenario challenges Putnam's account. Finally, I will evaluate which position is stronger, considering both philosophical reasoning and the current consensus in the philosophy of mind and cognitive science, arguing that Putnam's functionalist framework, with some qualifications, remains more influential and broadly accepted today, even as Block's objection exposes important questions about consciousness that continue to be debated.

Hilary Putnam's machine functionalism views mental states as functional states of an organism, analogous to the states of a Turing machine. A Turing machine is an abstract model of computation that processes input and output according to a set of rules or a machine table of instructions. Putnam suggests that any creature with a mind can be seen as instantiating a sort of computational machine program. In his words, "the state of being in pain is a functional state of a whole organism." Essentially, a mental state like pain is defined by its causal relations to the organism's inputs, outputs, and other internal states rather than by the particular material that realizes it. For example, being in pain typically results from certain inputs, such as bodily injury, interacts with other mental states, like causing the desire to be rid of the pain, and produces characteristic outputs, such as wincing or saying "ouch." This network of causes and effects is what pain consists in, according to functionalism. The actual physical stuff implementing those relations, whether carbon-based neurons or silicon circuits, is not part of the definition of the mental state. Just as a software program can run on different hardware but implement the same computational state, a mental state is the program state the organism is in, defined by its functional organization.

This view marks a shift to a higher-level understanding of the mind. Instead of equating mental events with particular gray matter processes, Putnam's functionalism treats mental states as roles that can be multiply realized in various substrates. A helpful analogy is to consider a clock, a clock can be made of gears, electronics, or sand, as in an hourglass, yet what makes it a clock is the functional role of telling time. Similarly, what makes a state of an organism a pain or a memory is the functional role it plays in the organism's life, how it responds to stimuli and influences behavior, not the specific substance or structure implementing it. By introducing machine functionalism, Putnam provided a systematic way to think about minds in

computational terms. This approach went on to deeply influence cognitive science by suggesting that understanding the mind is akin to understanding the software running on the brain's hardware. Functionalism rose to prominence in the late 20th century as a leading theory of mind in both philosophy and psychology.

Putnam's functionalism stands in clear contrast to the earlier type, identity theory of mind. The identity theory, articulated by philosophers such as U.T. Place and J.J.C. Smart, claimed that each type of mental state is identical to some specific type of brain-state or physical, chemical process in the brain. For example, an identity theorist might suggest that "pain equals C, fiber firing" in humans, implying a one, to one correspondence between mental kinds and neurophysiological kinds. Putnam found this view too restrictive and species chauvinistic, because it seems to imply that only organisms with a human, like brain chemistry could have certain mental states. His most famous argument against the identity theory is the argument from multiple realizability. In short, the same mental state can be realized by different physical systems, not just the human brain's biology. Putnam asks us to consider nonhuman creatures or even hypothetical aliens. If an octopus, which has a very different nervous system from ours, feels pain, it is unlikely that its neurons are identical to human C, fibers, yet we do not want to deny that the octopus genuinely experiences pain. Similarly, one could imagine a Martian with a silicon, based neural system who exhibits all the behaviors indicative of pain; we would want to say the Martian can be in pain even though its brain-state might be silicon circuitry rather than C, fiber activation. The identity theory struggles here, it would have to claim either that these creatures do not really feel pain or try to redefine "pain" as disjunctive across species, an awkward and ad hoc move. To save the identity theory, one would have to identify a single physical state common to any organism that feels pain, an exceedingly ambitious and

implausible task. In contrast, functionalism easily handles these cases, pain is whatever state plays the pain role in a given organism, and both the octopus and the human can share the property of being in pain despite physical differences because each has an internal state that fulfills the same functional criteria, such as being caused by injury and leading to withdrawal. This multiple realizability of mental states is a core selling point of functionalism over the type, identity theory.

Despite its advantages, functionalism faces serious objections. Ned Block formulated one of the most famous challenges to Putnam's account with his China Brain thought experiment. This imaginative scenario tests the claim that duplicating the functional organization of a mind is enough to duplicate a mind itself. Block asks us to imagine that the entire population of China is enlisted to simulate the activity of a single human brain. Suppose there are roughly a billion Chinese individuals, each assigned to act like one neuron in the brain. They communicate with one another by radio or telephone in such a way that the pattern of signals exchanged mirrors the pattern of neural firing in an actual brain. They might even be connected to an artificial body, with sensors feeding in sensory inputs to the citizens and effectors producing motor outputs, like moving a robotic arm or answering questions, based on the messages passed around. In effect, the billion people plus their communication network collectively implement the same machine table that a real brain follows. The question then is, would this giant system have a mind, would there be a single, unified consciousness, a brain consisting of the whole system that has thoughts and feelings, such as experiencing pain or seeing the color red?

According to machine functionalism, if the functional organization is the same, then the mental states should be the same. So if the China Brain duplicates every relevant functional state that a human brain goes through when it feels pain, functionalism seems committed to saying

that the China Brain is in pain and has a conscious experience of pain when those citizens enact the corresponding state. Block, however, finds this conclusion deeply counterintuitive and uses it as a reduction argument for functionalism. He argues that while the Chinese Nation brain could certainly behave as if it had a mind, since by hypothesis it mirrors the functional responses of a person, it is implausible to believe that there would literally be a conscious mind emergent at the level of the whole system. It seems more reasonable to say that an elaborate simulation of a mind has been constructed, but not an actual thinking, feeling mind. Nowhere in the system is there a single person or brain receiving all the inputs and experiencing the outputs, it is a distributed network. Block contends that this functional duplicate of a mind would not actually have a mind in the way a person does.

Block's China Brain case illustrates a more general objection known as the absent qualia problem, wherein a system might replicate all of the relevant functional states of a conscious being yet seemingly lack genuine subjective experience. If such an entity is possible, functionalism alone cannot suffice to explain consciousness, for it would neglect the reality of what it feels like to have experiences. Block's thought experiment imagines that the functional organization of a conscious brain is reproduced by a vast network of human "agents", each corresponding to a single neuron, coordinating their signals to mirror the inputs, outputs, and internal transitions of a mind. Even if this enormous collective precisely instantiates the same functional roles that occur in a standard human brain, it is implausible that the giant homunculus system would actually experience what it is like to be in pain or to see colors. Thus, Block argues that mentality cannot consist solely in functional organization, implying there might be a crucial component tied to biology or integrated structure that functionalism overlooks. While Putnam's version of machine functionalism underscores how mental kinds operate at a higher functional

level, the China Brain scenario reveals an intuitive gap between replicating functional roles and capturing the raw feel of experience.

Having introduced Block's counter-challenge, we can now consider the broader philosophical reception of functionalism versus the identity theory. Putnam's functionalism, bolstered by the multiple realizability argument, has prevailed as the dominant framework in philosophy of mind and cognitive science, whereas strict brain-state identity theories have lost traction. Putnam famously demonstrated that tying each mental state to one specific physical type was restrictive and anthropocentric, especially in light of hypothetical aliens or AI systems that could exhibit mental states without sharing our exact neurochemistry. As research programs in cognitive science, artificial intelligence, and neuroscience advanced, functionalism's emphasis on causal organization and information processing became increasingly influential. Although Block's critique raises a serious question about whether functional characterization alone can guarantee genuine consciousness, it has not resuscitated the identity theory. Instead, it prompted refinements of the functionalist approach: many scholars propose that additional constraints, like embodiment, integration, or specific architectures, may be needed to account for qualitative experience. Today, most philosophers of mind remain physicalists, accepting that mental phenomena depend on physical processes, yet rejecting the claim that each mental kind must map directly onto a single brain-state type. Functionalism, taken in a broad sense, continues to guide the study of mind, while identity theories survive only in softened or heavily modified forms.

All in all, Hilary Putnam's machine functionalism redefined the philosophy of mind by treating mental states as functional roles rather than strictly biological processes. This perspective overcame the limitations of the identity theory by acknowledging multiple

realizability and shedding a narrow, species-based conception of mind. Ned Block's China Brain scenario, however, highlights a persistent worry: purely functional duplicates of mental states might not necessarily entail subjective awareness. Although Block's objection underscores the complexity of consciousness and the challenges of capturing phenomenal experience, it has not toppled functionalism's central position. Instead, it serves as a reminder that any comprehensive theory of mind must ultimately reconcile the functional roles of mental states with the intrinsic qualities of subjective experience. Putnam's functionalist framework, broadly construed, remains the most influential starting point for research in cognitive science and artificial intelligence, and ongoing debates continue to refine how best to integrate consciousness into a thoroughly functionalist picture of the mind.