

In this paper I discuss functionalism, a relatively popular theory of the mind which asserts that mental states should be defined based on their cognitive role and causal relations to each other. Though the ensuing discussion is applicable to all variants of functionalism, I will concentrate on Jerry Fodor's "representational theory of the mind" to explain and analyze the merits of functionalism.

In this paper I argue that functional roles are not the defining feature of the mind, and that this characterization cannot explain qualia<sup>1</sup>. I will first explain what functionalist concepts add to the mind-body discussion and how they differ from their behaviorist antecedents. I will then discuss the fundamental flaws within functionalist paradigms that prevent them from giving adequate accounts of conscious experience. Finally, I will argue that functionalism's conceptual inability to analyze qualia renders it an implausible theory.

### Functionalism

Functionalism defines the mind solely based on its causal role: that is, how inputs in the form of stimuli interact with different mental states to ultimately cause behavioral outputs. An example of external stimuli triggering a behavioral motor response in a subject<sup>2</sup> can be found in the natural disposition of prey to avoid their natural predator; the screech of a hawk may cause a mouse to hide.

A *behaviorist* theory of the mind will characterize the mental state of the mouse, call it "fear," as simply the disposition to seek safety; certain publicly-observable stimuli (the auditory screech) cause this disposition which is effectuated by the associated "fleeing" motor response. However, as a *functionalist* will argue, an appeal to mental terminology is necessary to fully capture the complexities of observable behavior. Even in this relatively simple predator-prey example, there appears to be a sequence of mental states, a "mental process," that ultimately leads to the observable output. For example, the screech of a hawk may activate a certain memory within our rodent subject that corresponds to a belief that this presents a threat, which finally coincides with a desire to live; all of these mental states ultimately result in a disposition to flee. The advantage of a robust mental vocabulary over behaviorism's singular, dispositional mental descriptions becomes even more apparent when we consider more complex behavioral scenarios such as the Trolley Problem where subjects must grapple with conflicting moral beliefs and utilitarian desires amongst other considerations.

This necessary mental terminology is perhaps best defined in Jerry Fodor's theory of "common sense psychology." Essential to this theory is the concept of intentionality, or the *aboutness* of mental states. For example, the aboutness of my belief that the upcoming winter will be cold is simply just the upcoming winter. This aboutness, or intentionality, can further be reduced into an attitude, the relation we bear to the object of our mental state, and content, essentially the proposition or overall direction of your mental state. Returning to the above example, the attitude in this case corresponds to my "belief" and the content of my mental state is "the upcoming winter will be cold."

---

<sup>1</sup> Qualia refers to the experiential character of mental states. This term is synonymous with "conscious experience" and the "phenomenal character of the mind."

<sup>2</sup> A subject is a being with a mind.

In forming a coherent theory of the mind from these seemingly unrelated psychological observations Fodor applies a linguistic analysis to the structure of our mental states. A sentence is characterized by individual words, all with their own individual meanings or “symbols” that ultimately represent the attitude and proposition of a sentence. In a similar fashion, Fodor explains, our thoughts are more or less the language of the mind; different symbols within our minds, perhaps as a result of neuronal activity, ultimately result in our mental states, replete with attitudes and contents based on our external environment. This theory is known as the representational theory of the mind (RTM). It is functionalist in the sense that it defines mental activity by the causal roles of, and relations between various mental states, and materialist in the sense that it reduces our mental representations back down to symbols generated by observable brain functions. Though a further discussion into the psychological merits of RTM is beyond the scope of this paper, it is worth noting that this theory has reasonable scientific backing and is less theoretical than one may be initially led to believe.

### Against Functionalism

Like many other physicalist theories of the mind, the primary struggle functionalism stems from the inexplicable nature of conscious experience. More specifically, functionalism fails to give an account for why conscious experience arises and may attribute mental properties to states that lack an experiential character.

Beginning with my former grievance, it is not surprising that qualia would be left out of functionalist accounts of the mind; after all, this relational theory places a theoretical focus on the intentionality of mental states instead of their experiential character. As aforementioned, RTM takes the logical leap from mental terminology to physical terminology by describing the mind as the quasi-language of our cognition; just as mental states have a clearly defined, semantically evaluable intentionality, so do sentences. Accordingly, we can translate mental representations to their physical, symbol constituents. Right? Well, not without the loss of qualia. While there is something that it is like to be in a mental state, there is nothing that it is like to be a sentence. Put differently, qualia is not a functional role in and of itself, but a phenomena that coincides with the performance of functional roles. Accordingly, all attempts to translate mental phenomena into functionalist terminology will inevitably lose qualia in the process. Returning to our predator-prey example, one can imagine changing the qualitative experience of the mouse so that it perceives the hawk “screech” as a “roar” of equivalent volume. Assuming that the mouse has always had this slightly adjusted qualitative experience, it is not only metaphysically conceivable, but likely that there would be no change in the auditory, functional role of perceiving this sound as a threat.<sup>3</sup>

A common functionalist response to qualms regarding qualitative concepts is to simply acknowledge the inability of functionalism to explain qualia but still hold that it is an accurate theory of the mind. For one thing, given that conscious experience is a fundamental part of the mind, it is unclear how a theory of the mind can give an accurate account of the nature of mental states and their relation to the natural world without an appeal to qualia. Perhaps more importantly, however, functionalism’s theoretical neglect of qualia makes it guilty of liberalism. That is, functionalism may prescribe mental properties to things that do not have them. In his well-known “China brain” thought-experiment, Philosopher Ned Block demonstrates how a

---

<sup>3</sup> This is an illustration of what is commonly known as the “inverted qualia” argument.

functional system obviously void of qualia could meet the structural requirements of functionalism. In this demonstration, Block imagines replacing each neuron in the human brain with a citizen from the numerically similar population of China. With each person or “neuron” communicating via two-way radio, and the overall representation or “mental state” being projected by satellite, it can be reasonably argued that such a scenario meets the functional requirements of the mind: sensory inputs via an artificial body, causally connected mental states, and behavioral outputs.

The metaphysical conceivability of this “absent qualia” functional system stems from what is known as the “multiple realizability” property of functionalism. Since different physical structures can instantiate the same functional properties, all that is needed to discount functionalism is to imagine a physical system with the same functional roles as the mind, but void of qualia. If a persistent functionalist were to continue to argue that the absent qualia system still has a “mind,” then this characterization would likely be a definitional issue. However, as Descartes revealed in his meditations a priori, and as intuition strongly suggests, conscious experience is far more fundamental to our concept of the mind than any functional role is.

In conclusion, the fault within functionalism lies within its defining mental states primarily based on their functional roles. Intuitively, the functional roles of our cognition appears to be more of a characteristic of the mental than the defining quality. As I have argued throughout this paper, when a mental paradigm is founded on the performance of functions, or functional roles, it is bound to lack proper accounts of qualia and be guilty of liberalism.