Yamat's comment made on ***Block, "Troubles with Functionalism"*** *page* 3 identifies a flaw in the vaguely defined functionalism thought Block's examples and observes the "hollow" nature of GPT responses, specifically noting their lack of qualia. While I do agree with Yamat's observation, my argument aims to further solidify and ground why this "hollowness" exists and why is poses a problem for any functionalist account, while also refining the language that we use to describe AI behaviors.

The idea that we can evaluate the consciousness of a computer (specifically GPTs) is naïve and should be evaluated with more of an observational behaviorist view. The reason for this is because it removes the inner processes that cannot yet be measured and allows us to focus on the discernable and set aside the question of whether GPTs are capable of being conscious.

While functionalism aims to make sense of mental states by defining them functionally, its inherent flaw becomes apparent when attempting to account for subjective experience (qualia), reducing consciousness to mere input-output relations. When applying this theory to Artificial intelligence it falls short since we can see the inputs and output, but not the inner conscious experience (if any). The original comment by Aton Yamat, implies that it is possible for GPTs to be conscious, and capable of thought. It is quite the opposite, even when adopting an observational behaviorist view, autonomous thought and the mental processes that would need to be present and observable are instead ignored. No theory that has been introduced thus far (functionalism, nor even a pragmatic behaviorist approach) has been able to quantify or provide observable evidence that GPTs are conscious or capable of genuine thought.

The very concept of AI qualia is paradoxical and foreign to our understanding. What would a large language model feel? What would it be like to be a statistical model predicting the next token? If it cannot be conceptualized, then the likelihood that there is nothing to observe in terms of inner conscious life becomes more probable. This is not merely a limitation of our current tools, but to a fundamental difference in kind. Furthermore, Yamat's earlier mention of "Real reactions tend to arise instinctively and spontaneously", supports my earlier position against GPT consciousness. While its response might seem spontaneous, this is an artificial spontaneity rooted in algorithm execution, not genuine autonomy or felt experience. Foundationally, Large Language Models (LLMs) are built to be cold, calculated and accurate predictors of patterns born from big data. Their responses are dynamically created but are ultimately deterministically generated based on statistical probabilities and pre-defined architecture.

The “instinctive” nature of real reactions as Yamat phrases it, implies an evolutionary history and biological grounding that was shaped from a place of survival, emotion, and environmental factors. As it relates to biology, a GPT has no such grounding; it does not desire anything, it does not desire to communicate, and it does not possess any inherent will. Its responses, however complex, remain algorithmic in nature. Yamat’s classification of “real reactions" is problematically vague when applied across different systems. The hollow feeling that he refers to is not a flaw, it is a designed feature. They have been programmed for accurate and dynamic responses to prompts. Some might argue that the dynamic responses that these LLMs provide are spontaneous and relevant, and from a purely superficial input-output perspective, they are. However, this superficiality needs more specific classification.

LLM outputs should be categorized as “simulated reaction” to distinguish them from “real reactions” of conscious beings. This revised terminology helps distinguish that while LLMs skillfully mimic the structure of human conversation, they fundamentally lack the internal, subjective phenomenal experiences that defines a “real reaction” in a conscious entity.

In a comment made by Max Willson in **van Gulick, *"Understanding the Phenomenal Mind: Are We All Just Armadillos?*** On page 2 He argues the point that Mary’s dilemma is an epistemological question rather than an ontological one, and while some might agree, suggesting Mary’s learning is epistemological overlooks the distinction between knowing something and experiencing it. If her understanding were merely epistemological, a new fact or piece of information on wavelengths and Willson describes it would already be contained within her physical knowledge. The learning that she undergoes from this hypothetical experience is the direct understanding of qualia, a phenomenal property that exists independently of its physical counterparts. This is not a question of what Mary knows but rather a question of what exists in her experience that was previously missing, which would be an ontological question.

To illustrate this point when can draw another hypothetical, knowing every physical fact about the taste of chocolate, that includes the molecular structure, the neural pathways that activate when the dopamine transmitters are triggered, the chemical reactions of the taste buds, etc. Do you think that a person could devise what chocolate tastes like with these facts? One would surmise that the answer is no; Mary's situation is analogous. She knows every fact about the color red; she has never seen the color red or had the phenomenal experience of seeing the color red. The knowledge cannot be reduced to a conformation of physical fact because it is a direct apprehension of a phenomenal property that did not previously exist. Therefore, Mary's dilemma is not an epistemological question of what she knows in a propositional sense, but rather the question of what *exists* in her phenomenal experience that cannot be replicated. The addition of this new experience, a phenomenal fact, resolutely positions Mary's dilemma as an ontological question.

This last comment was made by Faruk F Arasan from **Turing, excerpts from *"Computing Machinery and Intelligence"*** page 3. He argues that the Turing's approach is flawed and asks the questions – if the test accurately measures whether a machine can think? Or, if thinking can be reduced to rule following? The goal of the Turing Test is not to see if machines can replicate every nuanced aspect of human thought, like 'breaking rules' or 'intuitive leaps,' which are difficult, if not impossible, to quantify objectively. Instead, it tests whether a machine can produce indistinguishable linguistic behavior from a human in a conversation. This notion that human intelligence involves rule breaking, and intuitive leaps is incorrect and are both higher order concepts that are derived from the core of intelligence. A more appropriate description of "real human intelligence" would be the ability to acquire and apply knowledge.

In response to Arasan's original question "if thinking can be reduced to rule following?" we can ask a different question in tandem – what about rule-following is not thinking? Genuine rule-following implies an internal, conscious and intentional process of understanding and application. So, by that understanding, rule-following is far from being unintelligent, it's a fundamental manifestation of intelligence, relying on the acquisition and application of knowledge. The revised definition of "real human intelligence" allows us to more precisely define what constitutes thinking; it includes the high-order functions to which Arasan originally referred, while also including rule-following as a process that demonstrates some level of thinking. According to the old definition, some humans are incapable of "real human thinking," some people are capable of cognitive thought but lack the creativity to think outside the box and the problem-solving skills to "make intuitive leaps" as he refers to. Often, it is a small minority of humans who break rules, while the majority follow those rule breakers; what about this demonstrates "real human thinking"?

Furthermore, Arasan's focus on the minority of “rule-breakers” over the majority of “rule followers” inadvertently undervalues a critical aspect of human intelligence. The act of adopting, learning and effectively applying a new rule or convention demands a large amount of cognitive effort. Structured learning and adaptive application are two concepts that are central to humans. Thus, the intelligence that is shown by the majority of rule-following people is not a lesser or biased form of “real thinking” but rather a foundational one that supports all higher-order cognitive functions.

This brings us directly to Arasan's central critique of Turing's approach. He asserts that defining digital computers from “human computers” who followed fixed rules inherently biases the test against genuine thinking. He goes on to suggest that by modeling machines on strict rule following humans that we have already stacked the deck in favor of machines that resemble sophisticated calculators rather than genuinely thinking entities. However, this argument misguidedly assumes that strict rule following is inherently mechanical or unintelligent activity when, in actuality, its is the direct demonstration of applying knowledge within a specific context – the core of human intelligence.

By assessing a machine's ability to follow rules under the contexts of whatever test, it is precisely measuring its capacity for knowledge acquisition and application. If, as Arasan proposes, “real human intelligence” involves breaking rules, making intuitive leaps, and behaving with autonomy, then a machine that *simulates* these “rule breaking” behaviors within the confines of whatever experiment would still be demonstrating a sophisticated application of learned rules and patterns. The bias that Arasan perceives is not in Turing's approach but rather in the narrow definition of what constitutes genuine human intelligence.