

Multiple Realizability

Set question: what, if anything, does multiple realizability reveal about the nature of mental states?

Written for UCL's Philosophy of Mind module. Putnam's multiple realizability objection targets Smart's identity theory: if pain just is C-fibres firing, what do you say about an octopus in pain, which has no C-fibres to fire? I trace the identity theorist's escape routes, the token/type distinction and then a retreat to functionalism, and argue functionalism is where multiple realizability actually lands: mental states end up defined by the functional role they play, not by which physical process happens to realise them.

ARGUMENT IN BRIEF

1. Smart's identity theory: a mental state simply is a physical brain process. Pain is C-fibres firing.
2. Putnam's objection: an octopus has no C-fibres. If it can be in pain, pain cannot be identical to any one physical process.
3. First escape route, the type/token distinction: give up type identity, keep token identity. This survives the objection but tells us almost nothing about what unites a mental state across creatures.
4. That is the real damage. The identity theorist can no longer say what makes two very different physical states both instances of pain.
5. Functionalism answers that question: a mental state is the functional role it plays between inputs and outputs.
6. So multiple realizability does not just refute Smart. It shows mental states are functional kinds rather than physical ones.

Multiple realizability is an objection to the physicalist identity theory in the philosophy of mind. It aims to prove that mental sensations cannot be defined simply in terms of brain processes. In this essay, I shall explain identity theory and the multiple realizability objection. I shall outline how a physicalist may attempt to salvage identity theory, but ultimately conclude that concerns about multiple realizability lead to a functionalist outlook.

As a result, multiple realisability reveals that mental states are best understood as the functional roles they play within a system, and shows that any theory of mind must show how the same mental kind can be realised by different physical processes.

To understand multiple realisability, one must first understand how it is used in the philosophy of mind. Hilary Putnam uses the theory of multiple realisability as a response to the identity theory of JJC Smart. Smart is a physical monist who argues for the equivalence of mind and body. The human being and the brain are arrangements of physical particles (Smart 143). Any notion of consciousness is an arrangement of physical particles, explained by a brain-state. Thus, any mental state can be explained as, and is, a physical brain-state (Smart 144). An example of this is seeing a yellow afterimage. The experience of this is in the brain, and one's reporting of it is a report of a brain process. Thus, sensations are nothing "over and above brain processes" (Smart 142). This claim pertains crucially to numerical identity: sensations are identical to brain processes. Smart argues for his physicalist mind-identity theory using Occam's razor:

1. There is a correlation between mental states and physical brain processes. When an individual experiences mental sensations, one can locate certain brain processes (various parts of the brain are associated with said experience)
2. This correlation is explained through claiming (i) the mental state is equivalent to the brain process, or (ii) the mental state and brain process are correlated (a further explanation is needed to explain this correlation)
3. (i) is a simpler claim than (ii), as it requires fewer assumptions
4. Occam's razor argues for the simpler claim, as it requires fewer assumptions
5. Thus, one must accept (i) – mental states are brain processes (Smart 145)

Having laid out Smart's identity theory, one can see how the theory of multiple realisability objects to it.

Putnam argues that Smart's identity theory is deficient as it fails to acknowledge that mental states are multiply realised by different physical processes (Putnam 436). Certain mental sensations, such as pain, can manifest differently in humans and in other animals. Putnam argues that the identity theorist must be able to show that (i) every psychological mental state is a brain-state, and that (ii) there is no instance of said mental state manifesting itself as X in one being and Y in another. Putnam argues that the identity theorist is likely unable to demonstrate (ii), since mental sensations, such as pain, appear one way in humans and another way in other beings.

In humans, it is commonly regarded that pain has the equivalent brain process of C-fibres firing. However, it is also commonly regarded that octopuses can feel pain; if an octopus were injected with acid, it would retract its tentacle, suggesting that it feels pain in response

to the acidic stimulus. However, octopuses lack brains with C-fibres. Thus, their pain cannot be equivalent to the firing of C-fibres. It may be viable to claim that the octopus's pain is equivalent to D or E fibres firing. This can be formalised in the following form:

1. Beings without C-fibres (such as octopuses) experience pain
2. The pain in these beings cannot be equivalent to C-fibres firing, since there are no C-fibres to fire
3. Thus, pain $\neq$ C-fibres firing

Thus, Smart's identity theory faces a serious problem. Pain can manifest itself in ways other than C-fibres firing in the brain – it is multiply realisable (it is both C-fibres and D-fibres firing). Multiple realisability suggests that Smart's identity theory is insufficient to reveal the nature of mental states; they cannot be equivalent to specific physical brain processes. Importantly, this does not show that mental states are non-physical. However, the identity-theorist may respond with a few points. The first is to distinguish between token-identity and type-identity theories. The second is to move to a different physicalist theory – functionalism.

The first move available to the identity theorist is distinguishing between token-type identity theory. Type identity theory argues that types of mental states are equivalent to types of physical brain processes; every type of mental state is equivalent to a kind of brain-state (the mental type 'pain' is the brain-state type of C-fibres firing) (Schneider). This is the identity theory Smart defends. By contrast, token identity theory holds that every instance of a mental state is equivalent to a specific physical process. A specific mental process is simply a specific physical process occurring in the brain. This differs from type identity theory, as the token identity theorist argues that every instance of a mental sensation is equivalent to a specific brain process, but not all specific instances can be categorised under the same physical process (Schneider). Not all specific instances of pain, therefore, are equivalent to the same specific brain process – pain in the human may be equivalent to C-fibres firing, and pain in an octopus can be equivalent to D-fibres firing. Thus, token identity theory remains a physicalist theory as it still argues that mental states can be explained in, and are simply equivalent to, physical processes.

Having distinguished between type and token identity, the physicalist can see how multiple realisability explains mental states. Type identity theory still faces the same criticisms of multiple realisability – the claim that the same mental sensations can manifest differently across different beings is incompatible with type identity theory. In response, the type identity theorist may dispute the first premise of the multiple realisability argument. More specifically, with the notion of 'pain'. The type identity theorist has two options: (i) outright deny that creatures such as octopuses can experience any pain, or (ii) argue that the pain felt by octopuses is not the same as we know it (in humans). Of these, (i) is weaker. To deny that octopuses can experience pain seems absurd. It is a basic empirical fact that when

exposed to a painful stimulus (like acid), octopuses react as if they had felt pain. The octopus withdraws its tentacle and tends to the targeted site of pain (Crook). Science suggests that octopuses experience pain. Thus, (i) seems untenable.

However, this does not eliminate (ii). It may be correct that octopuses behave in such a way that they experience 'pain'. However, this 'pain' does not have to be the same type of 'pain' that we, as humans, experience. Granted, the behavioural output displayed by both humans and octopuses suggests 'pain' of a sort. However, how this pain, as a mental sensation, exists within the minds and brains of beings need not be identical – the mental 'type' might not be the same. If true, then type identity theorists may have defended themselves against multiple realisability. It may be false that the 'type' of pain felt by humans and octopuses is the same, and so it need not manifest in the same way (C-fibres firing). Instead, pain in humans differs from pain in octopuses, as the way it actually exists as a mental sensation may differ, despite exhibiting the same behavioural disposition typically associated with pain. Thus, the human type of pain is not multiply realisable – it will always be C-fibres firing. Likewise, because octopus pain is a different type to human pain, it, too, is not multiply realisable (it will always be D-fibres firing).

Although type identity theory can defend itself against multiple realisability challenges, it must adapt. It must find workarounds, meaning that type identity theory can only discuss mental sensations within certain species – human pain is not the same as octopus pain, which is not the same as, say, dog pain. Yet, it is hard to imagine how pain can be different across animals, despite generating the same or similar behavioural output(s). Thus, one may consider token identity theory.

Token identity theory is compatible with the multiple realisability objection. Token identity theory claims that a specific instance of a mental state is equivalent to a specific instance of a physical process. Token identity theory does not claim that types of mental sensations are equivalent to types of physical processes. Thus, it allows pain to be realised in one way in humans and in another in octopuses. All that matters is that each instance (token) of a mental sensation is explained as a brain process (C-fibres, D-fibres, etc.). Thus, multiple realisability and token identity theory align. If the only way we can define pain is as something physical occurring in the brain, what do we learn about the nature of pain? Token identity theory does not tell us how each instance of pain (which may involve a different physical process) is related. It does not reveal how each of these 'pains' is unified in the notion of pain more generally.

Thus, to answer this, the physicalist may turn to functionalism, which works in tandem with multiple realisability to better reveal the nature of mental states. Still broadly physicalist, functionalism holds that the physical constitutes everything. However, it also claims that mental states are simply functional states (Putnam 433). This is best thought of as a combination of sensory inputs and corresponding implied motor outputs – internal mental

states are simply how these two are connected (Putnam 434). Thus, pain is defined as the sensory input of, say, a deep cut, the relation of believing that one is in pain, and finally the motor output of perhaps exclaiming "ouch". The same goes for octopuses, too. For them, pain may be the sensory input of being injected with acid, the relation of wanting to avoid further pain (damage), and the motor output of withdrawing a tentacle or tending to the wound.

In this regard, functionalism does not oppose the multiple realisability objection. Under functionalism, pain can be realised differently across species and people, but in each instance, the same functional processes are at work. In every scenario, there is still the underlying process of sensory input, relation, and motor output. This is how each of these instances of 'pain' is united – they each share a common underlying functional process. In this way, functionalism succeeds where token identity theory may fail. Additionally, the functionalist avoids the difficulties that the type-identity theorist may encounter. The functionalist need not deny that octopuses feel pain, nor need they claim that the pain octopuses feel is different from the pain humans feel (it is still part of a functional process). Therefore, it can be argued that the objection of multiple realisability helps motivate the functionalist position. The functionalist's concern is with the stages behind mental sensations. So long as a mental sensation has a (physical) functionalist framework, it satisfies the theory, even if pain is multiply realised. This is how multiple realisability works in tandem with the functionalist theory of the mind.

Having now seen how physicalists may respond to the initial objection laid out by multiple realisability, one can see the broader philosophical significance. The objection highlights the difficulty in defining mental states purely in physical terms. Thus, the physicalist is led to a functionalist position: mental states are best understood through their functional roles. This is done through the identity theory lens, as I have shown, by raising issues with the (type-identity) claim that all mental states are nothing other than physical brain processes. While this captures part of what the physicalist seeks to prove, it is incomplete; there are different ways in which the same mental sensation can manifest. Pain is both C-fibres firing (in humans) and yet also something else, perhaps D-fibres firing (in an octopus). This is incompatible with the initial identity view and thus reveals that mental sensations cannot be equivalent to a single type of physical process.

Although identity theory presents important physicalist views, multiple realisability shows that it cannot account for the unity of mental states across different organisms. This is why physicalists distinguish between type and token identity theories. However, all that type identity theory can reveal (in response to multiple realisability) is that either non-humans do not experience pain (or mental sensations more generally), or that mental sensations in humans are different from those in non-humans. However, both of these answers seem

inadequate. Denying pain in non-humans appears to be at odds with empirical evidence, and it seems counterintuitive to conclude that human pain differs from that of non-humans, given that they produce the same behavioural outputs and have the same functional roles.

This dissatisfaction with the type-identity theory response suggests that mental states are best defined as something one would expect from a functionalist. However, before this is done, it's worth examining whether multiple realisability and token identity theories can shed light on the nature of mental states. In response to the multiple realisability objection, token identity theory reveals that each instance of a mental state is equivalent to an instance of a physical process. This tells us something, but, similarly to type identity theory, it doesn't reveal much. The question of what a mental sensation actually is, still remains. Suppose one can isolate a mental sensation in each specific instance, such as pain. How can one define said mental sensation, giving it a definition to apply in every instance? Thus, while token identity theory appears compatible with the multiple realizability objection, an explanatory gap remains. This is where multiple realisability and functionalism respond.

Ultimately, multiple realisability objections lead to a functionalist response. Identity theory holds that mental sensations are each a physical process, but it does not explain how their nature is such that we can identify them in every instance. Functionalism holds that mental sensations are best understood as the functional roles they play. Pain, therefore, is a combination of the inputs and outputs involved, and this is how one can identify it. They are thus outlined by their functional role. Mental sensations are thus best thought of as functional kinds.

Multiple realisability outlines the weaknesses in Smart's initial identity theory. It fails to account for the fact that the same mental state can manifest as different physical processes in beings with differing physical brains. The token-type distinction attempts to salvage identity theory, but still doesn't reveal much about mental states. One is still unable to answer what unites the same mental sensation across different organisms and how we can even define a mental state. Thus, multiple realisability reveals that mental states cannot be defined solely by their physical makeup. Rather, mental states need to be defined as something else – the functions they fulfil within a system. Multiple realisability, therefore, reveals that mental states are not defined physically, but rather by their functional roles within systems and the mind.

WORKS CITED

Crook, Robyn J. "Behavioral and neurophysiological evidence suggests affective pain experience in octopus." *iScience* (2021).

Putnam, Hilary. "The Nature of Mental States." In *Philosophical Papers*. Cambridge University Press, 1975. 429–440.

Schneider, Steven. *Internet Encyclopedia of Philosophy – Identity Theory*.

Smart, J.J.C. "Sensations and Brain Processes." *The Philosophical Review* (1959): 141–156.

RELATED

[Is the Human Ability to Acquire Languages Innate?](#)

[The Gettier Problem](#)

Alexandros Kalogrianitis · UCL Philosophy, Philosophy of Mind module, 2025, all rights reserved