

Manchester Lit & Phil

Philosophy Forum



Conscious Realism

A new kid on the metaphysics block

Focus paper for the Manchester Lit & Phil Philosophy Forum

Christopher Burke

August 2023

If you are having a tea break while reading this, you assume the cup is a separate solid thing happily sitting there waiting for you to pick it up. Very sensible. But that view of reality as persistent objects in spacetime, **Physicalism**, despite being the successful basis of our understanding the world since forever, has become increasingly contested ... even in physics. Neuroscience can explain brains but has completely failed to explain consciousness. Why? One neuroscientist and philosopher thinks he knows ... **consciousness is the reality and believing in the physical is a delusion!** But is it he who is delusional?



This paper deals with a contemporary philosophy which challenges our profound intuition that external reality consists of mind-independent separate objects in spacetime. **Conscious Realism** proposes that there is no objective reality, merely consciousness. The main proponent of this particular metaphysical theory is **Donald Hoffman**, an eminent Professor of Cognitive Studies at the University of California.



His philosophy consists of **three interconnected theories**:

Fitness Beats Truth (FBT) Theory
Interface Theory of Perception (ITP)
Conscious Realism (CR)



Below we take a critical look at Hoffman's final theory, Conscious Realism. The previous two were explored in an earlier paper: *Reality – virtual or veridical?* But this paper can be treated independently.

The concepts are challenging, but technical terms are explained in side panels close to their first usage, where they are highlighted in a **bold orange** font. We will also clarify the ideas at the Philosophy Forum.

There are **discussion questions in section 6**.

Manchester
Lit&Phil

1 The landscape of metaphysics

1.1 **Defining metaphysics** has been difficult since the term's first coinage as a posthumous collective title for the fourteen books written by Aristotle (384-322 BCE). Perhaps the simplest modern definition is given by the Cambridge Dictionary:

“The part of philosophy that is about understanding existence and knowledge”

Metaphysics can be partitioned into two main areas: **ontology**, which theorises about the *nature* of reality, and **epistemology**, which theorises about how we *know* about reality. Obviously there is an intimate connection between the two.

1.2 There is a **spectrum of possible metaphysical stances**, which can be broadly summarised under four positions:

Empiricism or Metaphysical Realism

There is a mind-independent reality and you know about it as it is in itself.

Representationalism

There is a mind-independent reality but you know about it only via your mental **representations** of it, not as it is in itself.

Weak Idealism

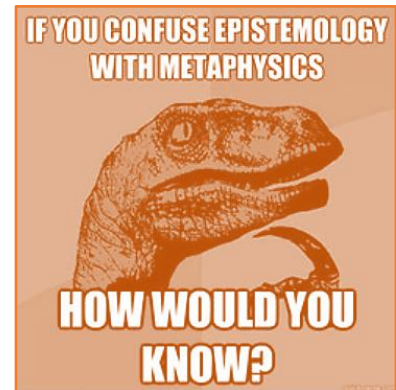
There is a mind-independent reality but you can't know about it in any way.

Strong Idealism or Solipsism

There is no mind-independent reality, merely your consciousness.

1.3 These are **not four discrete categories**: they blend into one another. It will come as no surprise to anyone familiar with a bit of philosophy that there are a myriad subtly intervening positions on this spectrum, but this simplification will suffice for our purposes. A much more magisterial overview can be viewed at [Stanford Encyclopedia of Philosophy - Metaphysics](#).

1.4 So how do **Hoffman's Conscious Realism (CR)**, with its two supporting theories – Fitness Beats Truth (FBT) and the Interface Theory of Perception (ITP) – fit into this landscape?



Representation:

It has a very broad meaning in philosophy: a bit of reality which contains information about another bit of reality. A representation refers to something else: its representatum.

Psychical types include sensations, percepts, concepts, theories. Even hopes and feelings can be regarded as representational.

Physical types include all signs and symbols, both alphanumeric and iconic, and the systems of which they are a part. Thus all our brain functions relating to the above psychical types are representations. As is writing, computer models, photography, most visual art, drama and even some music and dance.

2 What is Conscious Realism?

2.1 **Conscious Realism** (CR) suggests that reality consists not of separate objects in spacetime, but solely of **conscious entities interacting with one another**. The physical world is an illusion created by our perception. It has been developed over the last couple of decades by Donald Hoffman, a distinguished Professor of Cognitive Studies at the University of California, Irvine, and a team of his colleagues.

2.2 Leslie Allan provides a useful and concise summary of CR, which I quote at length:

2.2.1 “In a wealth of academic papers, videos and interviews spanning more than a decade, respected neuroscience researcher, Donald Hoffman, has proposed a novel solution to the mind-body problem. In the philosophy of mind, the mind-body problem specifically is the challenge of finding out what exists fundamentally; whether it is mind or physical bodies or some combination of both. Whatever we take to be ontologically fundamental, philosophers of mind are also tasked with explaining the nature of the relationship between mind and body. After more than two millennia of philosophers and scientists sweating over this problem, it appears no less intractable.

2.2.2 Hoffman’s solution to the problem consists of the following three interconnected theories.

1. **Fitness Beats Truth (FBT) Theorem**
2. **Interface Theory of Perception (ITP)**
3. **Conscious Realism**

2.2.3 Working with other researchers, Hoffman’s **Fitness Beats Truth** (FBT) Theorem posits that during the course of the evolution of species, organisms whose perceptual apparatus are tuned for fitness for reproduction always win out against organisms that are tuned to perceive reality accurately. The theorem purportedly results from mathematical modelling of the selective pressures operating during the evolutionary process. This leads Hoffman to propose a pictorial representation theory of sensory perception, named the **Interface Theory of Perception** (ITP). According to ITP, every organism sports a species-specific perceptual interface modelled on the metaphor of icons on a computer desktop. Just as icons on our computer desktop do not accurately mimic the underlying complex objects they represent, so do our perceptual representations of external physical objects hide their enormous complexity.

2.2.4 The combination of these two theories is consistent with a realist view of the external world; i.e. the view that physical objects and processes exist independently of minds that perceive them. It is with the third theory in Hoffman’s tripartite synthesis that he recommends a radical departure from both common sense and the dominant scientific realist view of what actually exists. The first two posits provide the theoretical underpinning for Hoffman’s **Conscious Realism**; the view that the real world consists solely of conscious agents.”

Leslie Allan, 2022, [*Hoffman's Conscious Realism: A Critical Review*](#)

2.3 So where does **Conscious Realism** fit on our spectrum of metaphysical beliefs? Hoffman – quite correctly in my opinion – **dismisses Metaphysical Realism** on two grounds:

- It is not evolutionarily plausible for our brains/minds to record the world veridically. It would be far too inefficient and informationally overwhelming. Hence his **Fitness Before Truth (FBT)** theory.
- It is not logically coherent. How could we know what **extramental** reality is like *in itself* when we can only represent it **intramentally**; ie through thoughts and neural configurations? Our percepts and concepts are part of us – a different bit of reality from the objects or events they refer to. Hence his **Interface Theory of Perception (ITP)** where percepts and concepts are like computer desktop ‘icons’.

Extramental:

Existing independently of the mind/brain.

Intramental:

Referring to thought and feelings ‘within’ the mind/brain.

2.4 This moves Conscious Realism decisively along the spectrum into **Representationalism**. Thus far, I wouldn’t disagree with the theory, and nor would many philosophers and most neuroscientists and psychologists. But Hoffman pushes CR *much* further into more problematic and controversial territory via his proposition that **only conscious entities exist**. What are his reasons for such a bold move?

2.5 Before going into details, let’s see how Hoffman’s theory fits into the metaphysical landscape of paragraph 1.2. Here is a **summary of his case**:

- Cognition is about fitness benefits not ‘truth’ [FBT]
- Our perceptions are ‘icons’ on our interface [ITP’s first proposition]
- Those ‘icons’ don’t represent anything physical [ITP’s second proposition]
- Nothing is physical: reality consists solely of conscious entities [CR]

The **first two steps** could be a form of Representationalism (there is a mind-independent reality but you know about it only via your mental representations of it). The **third step** could be a form of weak Idealism (there is a mind-independent reality but you can’t know about it in any way). But the **last step**, Conscious Realism, is where Hoffman moves into a strong Idealism (there is no mind-independent reality, merely consciousness).

3 Why Conscious Realism?

“Reality is merely an illusion, albeit a very persistent one.”

Albert Einstein, 1934, *Mein Weltbild* (English translation 1935, *The World as I See It*)

- 3.1 What philosophical problems is Hoffman trying to solve by claiming that *only* conscious entities exist? Basically he is trying to **put Physicalism out of its misery**. Physicalism is the idea that *all* existence can be explained by scientific theories. Notwithstanding the immense achievements of scientific understanding over recent centuries, its success with explaining phenomenal consciousness – *what it’s like* to think and feel, to *experience* life – has been zero. All we have are neural correlates of consciousness – parts of the brain that are active during specific types of experience.

“In each attempt so far, at just the moment when consciousness pops out of unconscious ingredients, a miracle occurs, and the metaphorical rabbit pops out of a hat. The failure, I think, is principled: you simply cannot cook up consciousness from unconscious ingredients.”

Donald Hoffman, 2019, *The Case Against Reality*

- 3.2 Despite ... or perhaps because of ... his long career as a neuroscientist, he ridicules the idea that physical events can cause phenomenal consciousness, dubbing it the ‘Astonishing Hypothesis’ and quoting the renowned cognitive psychologist Steven Pinker approvingly:

“The last dollop in the theory [of Physicalism] – that it subjectively feels like something to be such [neural] circuitry – may have to be stipulated as a fact about reality where explanation stops.”

Steven Pinker, 2018, *Enlightenment Now: the Case for Reason, Science, Humanism and Progress*

- 3.3 The huge recent advances in our understanding of brain function merely serve to make our **lack of a physical explanation of consciousness** even more glaring. Observing neural and endocrinal behaviour in brains is *never* going to tell anyone *what it’s like to be* the accompanying experience. This omission is usually obscured by much talk of minds magically emerging from, being facilitated by, or supervening upon, brains. These proposed relations may be descriptively useful but have **no causal power whatsoever**.

- 3.4 An outline of a **logical argument against Physicalism** might look like this:

Physical explanations are based on the observation of behaviour.

#1 *premise*

Consciousness is not observable behaviour.

#2 *premise*

There can be no physical explanation of consciousness.

#3 \Leftarrow (#1 \wedge #2)

- 3.5 In response to this impasse, Hoffman could have proceeded down the **Panpsychism** route to address Physicalism’s deficiencies. Panpsychism claims that consciousness is not just restricted to complex organisms but is a fundamental and universal aspect of the cosmos ... it is part of the whole fabric of reality like matter and energy. Everything is physical but also simultaneously psychical in one

form or another. To me, this leads into a **quagmire of incoherence** reminiscent of Descartes' original schism of existence into mind and matter. Saying that stuff is both physical and psychical doesn't solve the problem of how those two aspects are supposed to interact in bio-agents. It also makes the little local problem of consciousness a universal one. Hoffman seemingly concurs, because he eschews Panpsychism.

3.6 However, his attempt to regain coherence is by **moving to pure Psychism** and dispensing with physical reality altogether! Hence his hypothesis that there is *no* existence other than consciousness.

“Physical objects in spacetime are simply icons in our desktop.”
Donald Hoffman, 2019, *The Case Against Reality*

To spell this out very clearly, Hoffman is not just claiming that **each of us lives in a virtual world of our interface icons**, but also proposing that our normal intuition that these icons represent **mind-independent objects separate from us, is illusory**. Wow! We are definitely now deep into the Idealism bands of our metaphysical spectrum.

3.7



I refute it thus
(sculpture of Dr Johnson)
William Fawke (1950-2018)

An historical antecedent of Conscious Realism is George Berkeley's 1710, *A Treatise Concerning the Principles of Human Knowledge*, which declared that “**esse est percipi**” – **to be is to be perceived**. Berkeley (1685–1753) proposed that **physical objects only exist when observed**, an immaterialist philosophy subsequently called Subjective Idealism. Many of his contemporaries were predictably outraged at this, with essayist and lexicographer Dr Johnson (1709–1784) asserting “**I refute it thus!**” by kicking a nearby stone.

3.8 Dramatic ... but perhaps not philosophically convincing. So how did Berkeley explain the stability and persistence of these supposed apparitions – tables, chairs, other people? He had a big get-out clause: God! This **all-seeing being continuously observed everything**, thus ensuring the world stayed permanently in place and Samuel Johnson felt a pain in his toe. Well ... Berkeley was a bishop.

3.9 But Hoffman hasn't got that ecclesiastical backup. He cannot call on a god to rescue the physical, so his reality remains strictly virtual: Johnson's stone is just an interface icon and presumably so is his pain. The aphorism "**extraordinary claims require extraordinary evidence**" certainly flags up here. Does Hoffman's Conscious Realism spell the **sad death of objective reality**? Let's explore how coherent and plausible it is. Does CR stack up?

4 Conscious Realism – pros and cons

4.1 One major positive feature of CR is that it avoids the perils of **ontological dualism**; the idea that there are two different types of stuff, psychical and physical. Hence no insuperable problems with how two utterly different types of existence – mind and matter – are supposed to interrelate. CR is admirably **monist: consciousness is the sole ‘stuff’ of existence**. It is a form of pure Psychism (everything is psychical) – the diametrical opposite of Physicalism (everything is physical).



4.2 Another bonus is that this idea of existence being universal consciousness is quite **existentially comforting**. The cosmos is not the cold, soulless domain of the physical; it's much more like us. But then why do some aspects of our experience (eg fellow humans) seem more conscious than others (eg stones)? According to CR, this is not because of a deficiency of consciousness in reality but due to a **decline in the sensitivity of our interface**.

“Conscious realism pins the decline where it belongs—on our interface, not on an unconscious objective reality. Although each successive icon, in the sequence from human through ant to quark, offers a dimmer view of the conscious world that lies behind, this does not entail that consciousness itself is on a dimmer switch. The face I see in a mirror, being an icon, is not itself conscious. But behind that icon flourishes, I know firsthand, a living world of conscious experiences. Likewise, the stone I see in a riverbed, being an icon, is not conscious nor inhabited by consciousness. It is a pointer to a living world of conscious experiences no less vibrant than my own—just far more obscured by the limitations of my icon.”

Donald Hoffman, 2019, *The Case Against Reality*

4.3 But a question that begs to be answered is: **how does he know that reality is just consciousness** when all he has are his interface icons not representing anything? And just because he knows that consciousness flourishes behind his own image in a mirror, does that indicate *anything* about consciousness behind all the other ‘icons’ in his experience? This opens the door on many other philosophical problems for CR.

4.4 One might logically conclude that an *ontological* claim – like consciousness being the fundamental stuff of reality – is beyond our understanding according to CR. How can we magically **peer ‘behind’ our virtual interface** to directly experience what reality ‘really’ is? Aren’t we supposed to be stuck in the virtual world of our unrepresentative icons? But counterintuitively, CR apparently says we can; so Hoffman boldly goes where even idealist angels fear to tread, declaring reality as **“properly described as a network of conscious agents”**. The credence of such a hypothesis is more defensible if restricted to the icons which look somewhat like our mirror image (ie humans and animals), but how can anyone possibly know that a stone is “a pointer to a living world of conscious experiences no less vibrant than

my own”? This **insight beyond the interface** is a trick he doesn’t explain, so the charge of inconsistency remains unanswered ... and almost certainly unanswerable.

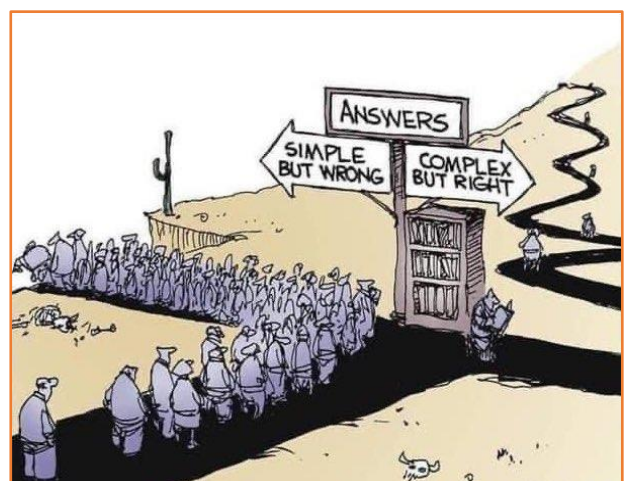
4.5 That isn’t the only problem for CR. Hoffman’s precursor Interface Theory of Perception (ITP) maintained that icons provided no information about intrinsic reality. But if that is the case, **how does the interface reveal fitness pay-offs** in line with his other precursor theory, Fitness Beats Truth (FBT)? Those pay-offs must now themselves be conscious agents since that is all there is, according to CR. So if I see my friend pick up a stone to use as a hammer (a fitness pay-off), this is all merely a virtual experience masking an underlying reality of interactions between consciousnesses. But my interface only reveals different degrees of the underlying consciousness, depending on the icon. The icon of my friend’s smiling face is supposed to be a portal granting me a glimpse into her consciousness, whereas my icon of the stone is opaque to any of the putative consciousness behind it. So **icons are now informative**, after all.

4.6 CR begs many more questions than it answers. Here are just a few:

- Why should the interface **filter out some conscious agents and not others**? Hoffman would argue that the closer the icon is to my own image, the more consciousness is revealed. Hence a human face icon reveals more than a stone icon. How early humans managed before mirrors to see their own reflection is an interesting question. Were there that many waters still enough for reflection?
- How does the **variation in the consciousness-detecting efficiency of icons deliver fitness pay-offs**, which are what icons must presumably be about?
- Assuming my own consciousness somehow detects the **different pay-off values of the other consciousnesses** I observe, how and why do different consciousnesses have those various values?
- Does the **concept of consciousness** retain any link to its original meaning – ie *what it’s like to be human*? Should *what it’s like to be a stone* – if anything – be called ‘consciousness’?

5 Conclusion: the root of the problems

5.1 Conscious Realism is a form of **Psychism**, which posits that reality fundamentally consists solely of conscious experience – thinking and feeling. But that only seems useful to apply to certain percepts that show complex behaviours, not those that are inanimate. Applying the term ‘consciousness’ to the latter would imply stretching the concept so far that it loses any meaning. **Physicalism**, on the other hand, posits that reality fundamentally consists solely of the interactions



between inanimate stuff, from which consciousness inexplicably emerges. But physical concepts don't apply to mentation: you can't put a thought or feeling on weighing scales or measure its length. But if neither Psychism nor Physicalism are adequate for the job of modelling existence, are we condemned forever to be in **metaphysical limbo**?

5.2 My diagnosis of the root of the problem is human hubris: the **persistence of anthropocentric thinking**. Adherents of both these stances believe that existence is fundamentally just how they think about it. This type of fallacy is a modern echo of historic anthropomorphic concepts of gods and spirits. But what gives humans the right to claim anything about what reality is fundamentally like, either psychical or physical or anything else? All we can do is mentally represent both our self and the rest of reality that is not our self, based on our sensations. Sensations are our baseline representations from which all others are derived. But our sensations – when interacting with a bit of reality – do *not* reveal the fundamental *nature* of that bit of reality, merely **how it affects us**.

5.3 This means that we need to realise that our understanding of existence is better cast as **epistemological rather than ontological**. Psychology and the arts are just a set of symbols which helps us navigate our own thoughts and feelings plus those we assume are experienced by other specially complex bits of reality similar to ourselves. The physical sciences are just a set of symbols which helps us navigate both ourselves to a certain extent (ie bodily) and the rest of reality. The psychical and the physical are not two different realities, merely **two different modes of representing one reality** using their different symbols. They are closely connected to the distinction between **1st and 3rd person perspectives** and that between **expression and explanation**. So our knowledge of the world can be conceived as either cultural artefacts (all our books, databases, research papers, etc) or neural configurations in brains or memorised cognitive/affective structures. But whatever approach we take, **it is inevitably representational**.

5.4 To know about something, from cursory experience to deep understanding, is to represent it. It is therefore meaningless to claim to know what reality fundamentally is, because that would imply knowledge without representation, which is contradictory. So the reason why both Psychism, including Conscious Realism, and Physicalism come adrift is because they make arrogant ontological claims they can never be in a position to justify. Indeed, if one accepts that knowledge is essentially representational, **both stances are logically incoherent**.

5.5 Representationalism, the middle of our metaphysical spectrum between the wild extremes of Empiricism and Idealism, is the stance I have been using to critique Psychism and Physicalism. But is it not hypocritical of me to denounce CR for making ontological claims, then propose one of my own: that representation is the fundamental reality of our understanding the world? It would be if Representationalism was claiming that, but it isn't. It recognises that it is itself a high level representation; merely a coherent and propitious *model* of bio-agents interacting with their world. It stays humbly epistemological and thereby it is the philosophy I have found to be the least incoherent. I think it acknowledges the human condition: we are *bits of whatever reality is* interacting with other *bits of whatever reality is*. To me, this offers a path to intellectual coherence in a way Conscious Realism doesn't.

“It is sometimes an appropriate response to reality to go insane.”
Philip K Dick, 1981, *Valis*

Thank you for reading this bit of reality ... whatever that is!



6 Summary and questions

Conscious Realism does ‘what it says on the can’: reality consists solely of consciousness. Our percepts and ideas of the world as mind-independent physical things are all illusory. Instead, there are just interactions between consciousnesses.

Q1

What aspects of Conscious Realism do you like? Why?

Q2

What aspects do you dislike? Why?

Q3

Which philosophical stance comes closest to your own:

- Psychism (consciousness is the only reality eg Conscious Realism)
- Physicalism (reality consists of mind-independent physical objects and events very like our percept and concepts)
- Representationalism (reality consists of mind-independent physical objects and events but we can only ever represent them, not know them as they are independent of us)

Why?



Whether or not you agree with him, Hoffman’s latest book is an exciting read:
Donald Hoffman, 2019, *The Case Against Reality*, Allen lane.

This can be downloaded via [Oceanofpdf](#).

There is also a lengthy 2022 discussion: *Lex Fridman Podcast #293*:

[Reality is an Illusion - How Evolution Hid the Truth](#).

Those without the time to tackle these can look at his 22 minute **TED Talk from 2015: [Do we see reality as it is?](#)**



Your notes