The „hard problem“ of consciousness

The hard problem of consciousness is to explain how qualia arise from physical processes in the brain.

(➔David Chalmers)
Some thought experiments underscoring the hard problem of consciousness
0. Aliens
Aliens & consciousness

- Imagine highly intelligent, but unconscious creatures visiting earth. They have unlimited time to study our planet and its people.
- Would they find out that human beings have consciousness?
Aliens & consciousness

• If you think they'd discover consciousness: why and how?
• If you think not: why not?

• Optional:
  • First assume they are not able to speak with humans.
  • Then assume that they could: would it help them to discover consciousness?
I. The Javelin
An ordinary explanation

- Explain the distance the javelin traveled from the point where I let go of it to the point where it hit the ground.

- Could the aliens find this explanation?
An ordinary explanation

- Two ingredients:
  - Initial conditions: force of my throw, acceleration at launch, height at launch, angle and spin of throw, shape of javelin, air resistance...
  - Natural laws: Newton‘s laws of motion and gravity
An ordinary explanation

• The application of (universal) laws to the (particular) initial conditions yield a mathematical prediction of the javelin's behavior
• Given the laws and the initial conditions, it seems necessary / inevitable that the javelin travels exactly as far as it does
An ordinary explanation

- It seems *inconceivable* that the javelin would travel farther or less far given that we have accounted for all relevant details that affect the outcome.
- In particular, it seems utterly inconceivable that the javelin would travel zero distance.
An ordinary explanation for qualia?

• With consciousness/qualia, this necessity / inevitability / inconceivability is absent: it does not seem necessary that a particular brain process $P$ gives rise to exactly the qualia $Q$ that it in fact does, instead of different qualia $R$ or no qualia at all.

• Phrased positively: we can conceive of a situation in which a certain brain process is present but accompanied by different qualia than usual or no qualia at all.

• That’s the core of the hard problem of consciousness.
II. NCCs
NCC = Neural correlate of consciousness

= the brain state or process that occurs whenever a given conscious experience occurs
We would know:

Whenever we experience green, then neurons in region A24 of visual cortex fire at 38Hz, which indicates NCC for "green".

Whenever we experience red, then neurons in region A16 of visual cortex fire at 45Hz, which indicates NCC for "red".
• Would this knowledge amount to a neuroscientific explanation of green and red qualia?

• After all, we would now know the exact conditions under which these qualia occur!

• And the fact that we’ve found a perfect correlation suggests that we’ve even found the neural causes of green and red qualia!
Questions:

• Are you satisfied with this knowledge as an explanation of consciousness (of red/green)?

• Would you not like to know why exactly it is that 45 Hz oscillations produce red experiences and 38 Hz oscillations produce green experiences and not the other way round?

• Would you not want to know what exactly it is about the neurons in these regions of the visual cortex that is responsible for producing conscious color experience?

• Why do not neurons in other regions produce these experiences?

• Would you not want to know why the firing of these neurons in this way produces some kind of experience at all?
Why knowing the NCC isn’t enough for an explanation

Because:

We wouldn’t know *what exactly it is* about 38Hz neural activity in region A24 of the visual cortex that leads to the conscious experience of green.

Likewise, we wouldn’t know *what exactly it is* about 45Hz neural activity in region A16 of the visual cortex that leads to the conscious experience of red.
One symptom of our lack of understanding is that we don’t see anything necessary or compelling in the fact that 38 Hz neural activity in region A24 leads to a green experience, and 45 Hz activity in A16 leads to a red experience.

We could just as well imagine 38 Hz neural activity in A24 to be associated with a red experience, and 45 Hz activity in A16 to be associated with a green experience.
We could even quite easily imagine neural activity in these two brain regions not to be associated with any conscious experience at all.
inverted qualia

absent qualia
This shows that even if we knew the NCC or even causes of conscious states, there would still be a gap in our understanding.

...because we still wouldn’t know what exactly it is about the neural activity that makes it necessary that it is associated with conscious experience, and that it is associated with exactly these and not other qualia.

This is in contrast to the explanation of (all?) other natural phenomena, where we know or at least are able to conceive of a necessary relation between the phenomenon and its implementation/instantiation.
III. Zombies
Absent Qualia: My zombie twin

• Imagine that you have a special kind of twin. Your twin is special in that his brain is a physical copy of yours, that is, its physical (and functional) organization is identical to yours in every respect.

• However, your twin completely lacks all conscious experience of the world. He has no qualia at all. His inner life takes place in complete „darkness“.

• Is such as scenario conceivable?
Me

My inverted twin

Inverted qualia

- Your twin is special in that his brain is a physical copy of yours, that is, its physical (and functional) organization is identical to yours in every respect.

- However, his color qualia are inverted with respect to yours. When looking at a ripe tomato, he has the color experience that you have when looking at grass. Of course, both of you will call the tomato 'red' and the grass 'green', but your color qualia will be inverted with respect to each other.
What, if accepted, do the absent/inverted qualia arguments prove?

- Less controversial: explanatory gap
- Very controversial: ontological gap
What, if accepted, do the absent/inverted qualia arguments prove?

• **Explanatory gap:**
  • Although consciousness may be grounded in physics (i.e. no *ontological* gap), at present we cannot understand or explain *how* it is grounded in physics- not even in principle. In other words, there is an explanatory gap between physical (or neurobiological) facts and facts about consciousness / qualia. This is a gap in our knowledge.

• The particular importance of the conceivability of zombies w. absent qualia for the explanatory gap is this:
  • The *conceivability* of absent qualia indicates...
  • ...the logical *possibility* of a situation which is...
  • ...*incompatible* with a physical *explanation* of qualia/consciousness
  • → Explanatory gap
The world with an explanatory gap

fundamental physical properties
- mass
- charge

physics

chemistry

biochemistry

biology

consciousness qualia

Unexplained, „brute“ fact

> explanatory gap <

organizational level

high

low
What, if accepted, do the absent/inverted qualia arguments prove?

- **Ontological gap**
- For some (e.g. Chalmers), the absent qualia/zombie argument is even taken to indicate that consciousness is not grounded in physics as we know it today. Maybe consciousness in some form is not derived from physics, but is an independent, fundamental property of reality (like fundamental physical properties such as spacetime, mass, charge...)
- Such a conclusion is widely regarded with suspicion
The world with new basic „qualia-properties“

Unexplained, „brute“ fact

consciousness
qualia

explanation

organizational level

high

low

biology

biochemistry

chemistry

physics

elementary
particles

fundamental properties

mass

charge

„qualia-like“ property
The move from an explanatory gap to an ontological gap may look less implausible once you consider how we ever conclude that we need to revise our views about the ontology (the „basic ingredients“) of the world.

It is when there are phenomena that cannot be reductively explained by our accepted physical theories that we consider to add them as basic properties of the world.

We cannot „see“ the ontology of the world directly, we can only get at it through our epistemic abilities, of which explanations are a crucial part.
Asking whether a phenomenon requires the addition of new fundamental properties is always legitimate since we have in fact several times in history seen it necessary to add new fundamental properties to advance our explanations of the world.
IV. Mary
Mary and the „Knowledge Argument“ (F. Jackson)

- Mary: Brillant Neuroscientist, who has never experienced color (has never had color qualia) in her life. But she has all the physical knowledge about how color experience is generated in the brain.

- Question: When Mary experiences color for the first time (e.g. the blue of the sky), will she learn a new fact about the world? Something like a „subjective fact“?

- If so, then physicalism is false, because the set of facts about the world is larger than the set of physical facts about the world.

- Epistemological or ontological consequences?
Summary & conclusion
• We know that physical/neurobiological processes in the brain appear to generate consciousness.
• Still, we don't have any idea why and how consciousness (qualia) is generated by these processes → that's the hard problem of consciousness.
• There's an explanatory gap between our knowledge of the physical states of the brain and our knowledge of its conscious states.
• One way to illustrate the explanatory gap is via thought experiments about zombies and absent or inverted qualia. These thought experiments want to show that we can conceive of situations that are incompatible with a physical explanation of consciousness/qualia, which leaves a fundamental gap in the explanatory edifice of the world.