

2024-05-06 -- Appendix I

Brief historical overview of the philosophical perspectives that bring us to the current moment. The current moment? understanding mindfulness using the tools of neuroscience, cognitive psychology, and artificial intelligence.

Dualism

- "Mind things" don't have color, size, physical weight - they seem to be different than the physical things we can interact with
- Based on introspection alone, we might be inclined to conclude that the mind, therefore, is made up of a different *stuff* than physical things.
- This is dualism

Problem

- this view is uninformative, tells us what the mind is *not*, as opposed to what the mind is.
- still, knowing that the mind is not equal to the physical things we can hold, throw, hug, see, hear, is important.

The Three Core Dualism Contenders:

1. Parallelism

Mind and body are distinct and causally isolated. Neither is capable of affecting the other.

To explain the illusion of a causal linkage, the theory relies on *synchronization*: Some entity, maybe God, set up two parallel worlds to happen side by side - a physical world and a mental one. "like two ideally accurate clocks set to the same initial time and left to run for eternity"

2. Epiphenomenalism

Physical world can cause changes in the mental world, but mental cannot impact physical world.

The mind is like the exhaust fumes from a car; caused by the engine of the car but unable to impact the car itself.

This feels counter intuitive, because it *feels* like desires cause us to do physical things, the the thought/experience of **thirst** driving us to grab a drink of water.

3. Interactionism

Treats the mental and physical as distinct but causally integrated items.

This perspective avoids the problems of the first two - the most famous form of this is *Cartesian Dualism*.

The question that still remains:

How do two distinct items (mental & physical) interact in a way that makes up a single causal network?

Examples like iron filings and magnetic fields show us that things can interact without physically touching; *Cartesian interactionism is conceptually possible*. So, Clark asks, why give up on Dualism?

Dualism Downfall

- drugs affect the brain, which change our mental states (physical causing physical brain changes, which in turn cause mental changes)
- evolution of intelligent creatures correlates with evolution in brain structure
- physical damage to the brain disrupts mental functioning

Materialism

There is only one "stuff". This argument seems to win on the grounds of simplicity.

Positive Arguments of Dualism (below) are unconvincing

- "how could *mere physical stuff* do that?" An argument that in light of materialism, seems to fall flat. How could physical interactions cause something like Gaudi's Cathedral in Barcelona?
- argument from introspection: our mind *feels different* than physical stuff.

re-connecting to materialism: until someone shows that a mental experience change is completely unconnected to a bodily state, we don't need to accept that introspection as decisive evidence of dualism. Note: brain imaging makes this even harder; changes in thoughts reflect changes in neural activity.

Behaviorism

The first major rejection of Dualism came from the "Linguistic Turn", a philosophical puzzle by Gilbert Ryle - interestingly, the behaviorism movement didn't come from the inadequacies of dualism itself.

From Mindware (253):

Philosophers, Ryle claimed, were failing to see the significance of mental talk, in much the same way as someone fails to see the significance of talk about a university who, on being shown the library and colleges and playing fields and accommodation, goes on to complain, "Yes. I see all that. But where is the university?"

The University is not something extra, it is *just* the organization of those items.

Ryle concludes: the mind is not something extra beyond the behavioral manifestations. "Mind-talk is just a way of talking about the organization of the behavior itself."

The problems of behaviorism (as presented by Clark):

1. Dispositional analysis looks either infinite, or circular

- it would be infinite if we have to list what every given belief will "dispose" an agent to do in every possible situation they could be in

I, Trent, don't see a problem with this. Who needs to list things? A university isn't a university just because someone comes along and lists out the piece components and what they mean. *They just are.*

- it would be circular if behavioral dispositions reference other mental states; e.g., Mary (who likes teaching), will continue to teach as long as she is happy and does not believe teaching is ruining her life.

2. Any truth to the "inner sanctum"?

- The dispositional account wants to rule out the internal mental space that is special (the *inner sanctum*), but don't we have internal experiences?

3. Behaviorism is explanatorily shallow.

- If we want to understand why someone loves teaching, we might want to know *how* the person came to love teaching beyond the surface behavior of the teacher.
- "method actors' fallacy" (Putnam 1980): We deny mental experience to anyone who suppresses behavior - this seems ridiculous. Take pain, as an example.

It seems to me, Trent, that the mistake of this account of behaviorism is to discount *neural behavior* as genuine behavior. If we were to take the behaviorist perspective alongside the idea that neural activity is as worthy of noting as whole-body-in-world activity, we might have a theory that is better. I also don't think the disposition listing is necessary. Importantly: Neural activity will "do its own thing" - often relating to activity of the body but operates at a faster time scale.

Identity Theory

Mental states *are* brain processes

- Schemas (like "minds are brain processes") could be used as a coherent scientific theory
- Identity theory is a schema developed by UT Place, JJC Smar, D Armstrong
 - The job of philosophers (in this case) is **not** to decide whether or not mental states are brain processes.
 - The job of philosophers are to determine *if the schema is possible*.

Identity Theory Doubts:

1. Leibniz' law problems

if $X=Y$, what can be said of X must be able to be said of Y

Problem list

- *Spatial location*. Mental state "I like donuts" is located in the center right of my brain. Does that mean, that the belief, "I like donuts", is located in that same spot?

I would have said, yes... at least partially. I would have said that the believe that comes from a particular neural organization is technically "located" there. Experienced "there", though we don't have sensory neurons to know where in our brains states are experienced.

- *Truth Value*. A belief may be true or false. How can a brain state be said to be true or false?

Again, I don't see this as a problem: It is true/false in the same way that a belief is, because they are one in the same.

- *Sensational content*. A pain may be sharp or tingly, but could a brain state be sharp or tingly?

This is a weird one: if we consider that the brain state includes the sensations of the perceptual system, then... yes, it is?

- *Authority.* I have authority over my mental states. But I do not seem to have authority over my brain states.

I would actually say that you don't complete authority over either, this simplifies what brain processes/mental states are.

Response to Leibniz' law problems:

One way to respond is to say that it might not *seem* like brain states might be spatially located/true-or-false/sensational/authoritative, but they *are*.

Leibniz' law is unreliable in contexts that involve peoples beliefs about object properties, rather than what they actually are.

Identity theory survives Leibniz' law!

Some sophisticated takes on identity theory are popular today.

2. Species Chauvinism

A strong reading of ID Theory implies that brain states involving certain fibers/organs are equal to the beliefs themselves.

Example: the firing of C-fibers is associated with pain. Therefore, no being without C-fibers can be in pain. This seems clearly untrue. (Animals, aliens, as examples of this)

A "way out" is to say that each individual brain state is equivalent to the individual sensation; one person's pain is theirs, another's, theirs. Not the same pain, but both can be thought of as pain none-the-less. The way to categorize what mental states are, then, could be by their functional role.

Intro: Putnam & *Functionalism*

Functionalism

- Dennett's problem: ID theorists had to "explain what all clocks have in common"
- Functionalism: it is their *role, purpose, or function*. (<- this is a scientific schema)

Example: computer programs

- A grocery list app may be written in different computer languages, and run on different pieces of hardware, but functionally it accomplishes the same thing in all instances. It might even look the same.
- The same program can run on different machines.
- The functionalist claim is that the mind is to the body/brain as the program is to the hardware

Machine Functionalism: The mind is a program, run (in humans) with the brain as its supporting hardware.

A lot of Mindware focuses on versions of machine functionalism, the appendix won't explore this any further.

Eliminativism

The question so far has been "What kind of scientific theory (schema) could possibly count as a theory of the mind?" -- Maybe this is the wrong question, wrong starting point because it assumes commonsense ideas about mental states which could be mistaken.

Example: people used to believe in vampires, ghosts, *ether*. Spending time investigating the scientific schema of "ether" would be a complete waste of time.

New question: **Is the commonsense view of the mind *scientifically sound*?**

Eliminative Materialists: Those who believe that the mind is like ether - it doesn't exist.

Most extreme EM's believe that nothing in our commonsense framework of the mind will be preserved; believes, desires, hopes, fears will all be abandoned in a future science of the mind.

Clark: It is extremely hard to make sense of this claim before that future science of the mind is here. Question: can we really drop all of these commonsense ideas? Does their legitimacy depend on finding a place in a scientific theory, or are they valuable by themselves?

2024-05-07 -- Appendix II

Consciousness is hot... but it's unclear what it even is.

Possible "Targets" For Consciousness:

- *Simple awokeness.* The state of being sensitive to our surroundings, able to process info and respond
- *Self awareness.* ID ourselves distinctly
- *Availability of verbal reports,* with capacity to access internal states

- *Availability for the control of intentional action*, indicating an "internal poise"
- *Qualia*. redness of an apple, taste of a peach, piercingness of sound

The only "special" thing on this list, that we cannot currently explain, is *Qualia*...

Consciousness-related Research in Cog Neuroscience:

Blindsight

- while subjects are "blind", and experience blindness, they can identify things in their blind field at rates above chance.
- e.g., detecting a light is flashed in a blind region, or detecting the orientation of a wrist or hand.
- Explanation: mid brain is processing the information, but the "experience", from higher levels of cognitive activity, is absent
- Alt Explanation: visual hotspots of that offer a cortical route of information processing
Regardless: Visual information appears to be processed **without the accompanying experience**

Dual-Streams Hypothesis (Milner & Goodale)

- visual awareness is separated from fine motor control
- this can explain why some humans and monkeys can achieve visually guided action without visual experience
 - the visual stream is impaired, while the dorsal stream is unaffected

Neural Oscillations

- Crick & Koch's work on 40HZ oscillations, & consciousness
- "neural mechanisms that achieve *binding*", where *binding* means that neural populations represent information - e.g., binding MT motion detectors to v4 hue detector representing a certain face, speaking.
- Bindings might be possible through "frequency locked neural oscillations in the various neural populations"
- Synchronized neural populations are then depicted as joining the various neurally represented features into a coherent whole, placed in working memory, which in turn renders the coherent percept - ready to be acted/reported upon...

Koch 2012 for a review & story of this account... :D

Global Workspace Theory

- Baars
- **this is connected to the oscillations & binding!**
 - GW Theory depicts consciousness as involving the "poising of information for widespread dissemination" -> control of action and response.
 - Dehaene and Changeux, 2004

Dynamical Core Hypothesis

- Tononi & Edelman
- uses tools and perspectives from complexity theory to argue that consciousness involves the use of re-entrant connectivity to form temporary clusters of neural activity; combining "integration" and "differentiation"
 - "differentiation" - uniqueness of a state
 - "integration" - state combination to form "complex wholes"
- the amount of neural activity satisfying this criterion, then allows us to **quantify consciousness**

I, Trent, think that it's the emergence of these complex wholes that would satisfy the "consciousness" target

What evidence theory or conjecture tells us about the phenomenon of consciousness?

The answer depends on the definition of "Consciousness"

Ned Block splits consciousness into two parts:

1. Access-Consciousness (A-Consciousness)
 - This is the "information is poised to act upon" thing
2. Phenomenal-Consciousness (P-Consciousness)
 - the qualitative experience of things (qualia)

Block would claim that the neuroscientific and experimental claims that Clark just reviewed can only touch a-consciousness.

Example: we could build a Dual-Stream Perception robot. This would have a-consciousness, but it would lack p-consciousness. It would be a "dual stream zombie, acting just like us but lacking all 'feeling'"

the David Chalmers distinction:

- "easy problem" -> how does the brain recognize objects (other things like this, a-con)
- "hard problem" -> does that recognition process involve any "what-it's-likeness" (p-consciousness)? Why is a particular "what-it's-likeness" the way that it is?
- It's possible that a creature behaves but has no conscious experience ("zombie") or experiences things totally differently ("inversions")
- Functional behaviors/profiles/systems cannot explain p-con
- Levine calls this an "explanatory gap"

Chalmers: even if we had perfect correlation between a-con and p-con, we still wouldn't understand *why* the two go together. "Correlation isn't explanation". Maybe science cannot move us further than that.

Representationalism, Narrationism

Representationalist: *the mental is exhausted by the representational*

- pain = an internal representation: "tissue damage at local X"
- An interesting and important take: Tye (researcher, 1997) claims that a sensory representation is sufficient.

| Is sensory activation a sufficient "sensory representation"?

- First order Representationalism, described above, and there is *higher-order thought theory*
 - H-O Thought Theory: A neural state is p-con when it is *itself* the object of a thought...
 - To feel a pain requires a thought about the representation of tissue damage
 - Rosenthal: "a neural state will be conscious if it is accompanied by a thought about the state"

Narrationism: *Consciousness is constructed by "mind-tools" and made available via our immersion in culture and language.*

Dennett's Moves:

1. Intentional Stance: systems have beliefs ready for behavior
2. Multiple Drafts: there are lots of processes working in parallel, quasi-independent (contrast with "central processors")
3. The Narrative Twist: feeling qualia, is a particular achievement of the human biological brain... "our kind of consciousness"...
 - an artifact of our immersion in culture and language
 - our immersion allows us to form a new kind of cognitive organization that allows us to make cognitive objects of our own thought processes and to weave a kind of ongoing narrative that artificially "fixes" the cognitive contents.

- The "fixing" of content isn't really fixed; there are parallel processes ongoing
- **it is the serial fixation on ideas that gives us a tendency to believe in qualia**, but "qualia" here really is just the series of judgements made at the top-level

There are other views of consciousness that will for now - be left until later or further reading.

If we explain all parts of a-consciousness, is there really anything left over?

Clark: If a-con perfectly correlates with p-con, why not say they are the same?

The argument against (Kim): the first person perspective cannot be "so easily ignored"

Clark: Kim—following Block—quotes Louis Armstrong on the appeal of jazz: “If you got to ask, you ain’t never gonna know” (Kim, 1996, p. 180 citing Block, 1980). Such a response should, however, give us pause. In no other scientific or philosophical debate would such a move be acceptable. Why here?

The Meta-Hard Problem

Clark: Just as “nothing in physics tells us why there is matter in the first place” (Chalmers, 1997a, p. 20), so nothing will tell us why there is consciousness in the physical world. But that does not stop us seeking correlations of the kind mentioned earlier in the chapter. More radically, Chalmers suggests that we might need to recognize a kind of fundamental “double aspect” to physical states that carry information, with the result that where there is information there is always some degree of phenomenal content (op. cit., pp. 26–28).

The "Ant" argument -> humans might have no chance at understanding consciousness, no more than we'd expect an Ant to understand their own existence/experience.

Clark finds this argument unsatisfying, I kind of do too. Language, culture, and technology extend our cognition.

Clark:

If we allow (as we surely should) that some representational states have no phenomenological dimension, then why suppose it is the representational content and not the missing “extra ingredient” that is making the other states phenomeno-logically conscious? even if representational content is part of the story, it does not look like the whole thing.

Second-order representationalism (a.k.a. “higher-order thought theory”) may look like a better bet here. For the idea here is to identify the “missing ingredient” as an extra layer of

thought.

Interesting idea; consciousness as a sense of internal mental activity... what if consciousness is sensation of neural activity?

Is there a “hard-problem” of consciousness, or isn’t there? Is there something special about phenomenal consciousness that places it outside the reach of current scientific approaches (as Chalmers and others believe), or is it just a matter of explaining a pattern of responsiveness and report (as Dennett and others suggest)?

The "Meta-hard problem" of consciousness is deciding between the perspectives of Chalmers and Dennett.

Psychologizing Consciousness:

Price: It might take time (apparently like in physics) for the current, seemingly unsatisfying explanations for consciousness to become accepted, satisfying explanations. *Our intuitive understanding is a poor guide to progress.*

Explanatory Gap: consciousness refers to the challenge of explaining how physical events give rise to conscious phenomenal experiences.

Co-pilot Summary of Chalmers v. Dennett:

Let’s briefly explore the contrasting views of David Chalmers and Daniel Dennett on consciousness:

1. David Chalmers:

- **Hard Problem of Consciousness:** Chalmers famously introduced the concept of the “hard problem” of consciousness. He argues that there is an intrinsic, subjective aspect to consciousness that cannot be fully explained by physical processes alone. This “hard problem” pertains to why and how certain physical processes give rise to conscious experience.
- **Qualia and Phenomenal Consciousness:** Chalmers emphasizes the existence of **qualia**, which are the raw, ineffable qualities of our subjective experiences (e.g., the redness of red, the taste of chocolate). He believes that these qualia cannot be reduced to physical explanations.
- **Dualism and Panpsychism:** While not a dualist himself, Chalmers entertains the possibility of panpsychism—a view that consciousness is a fundamental property of the universe, even at the level of particles.

2. Daniel Dennett:

- **Consciousness as an Illusion:** Dennett takes a more reductionist stance. He argues that consciousness is an “illusion” created by the brain’s information-processing mechanisms. According to him, there is no hard problem; instead, consciousness can be explained through cognitive processes.
- **Multiple Drafts Model:** Dennett proposes the “multiple drafts” model, where consciousness emerges from the brain’s ongoing processing of information. Our experiences are like drafts of a story, constantly revised and updated.
- **Eliminativism:** Dennett is critical of qualia and argues that they are unnecessary for explaining consciousness. He suggests that we should eliminate the concept of qualia altogether.

In summary, Chalmers emphasizes the mystery of consciousness and the hard problem, while Dennett seeks to demystify it by explaining it as an emergent product of cognitive processes. Their differing perspectives continue to fuel debates in the philosophy of mind and cognitive science