

AYDEDE Talk 2 – Presented at Carleton College (May 17, 2005)

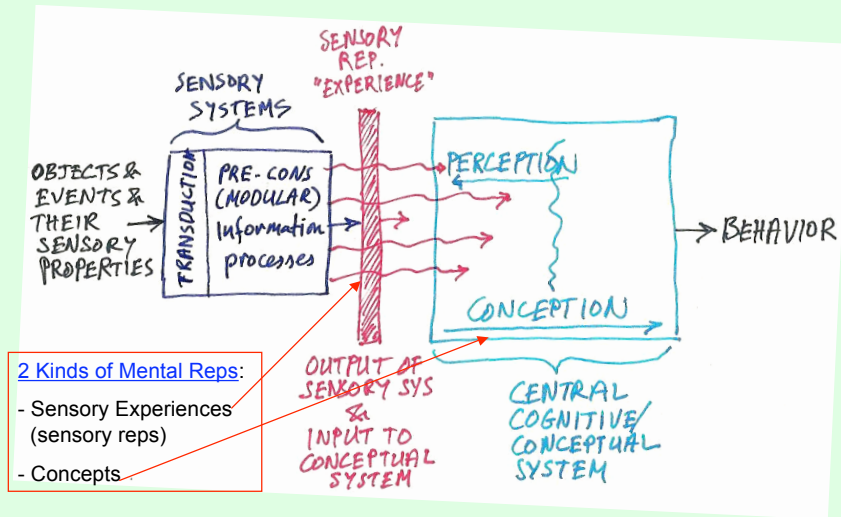
Have We Solved the Puzzle of Consciousness?

Have We Solved the Puzzle of Consciousness?

- T.H. Huxley (1866):
“How it is that anything so remarkable as a state of consciousness comes about as a result of irritated nervous tissue, is just as unaccountable as the appearance of the Djin, when Aladdin rubbed his lamp.”

Not yet...

REMINDER: Architectural distinction between sensory and central conceptual systems



A few REMINDERS:

- Both experiences and concepts carry information in both digital and analog form.
- The informational content is distinct from semantic content.
- The semantic content of mental reps is identified with their digital content

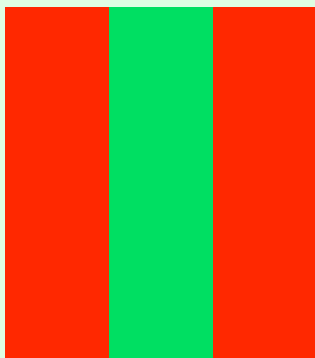
REMINDER: Intuitive Idea behind the Analog/Digital Distinction



- Specify at each point, according to an ordering scheme, the hue, saturation, and brightness (color, for short)
- Express this information in a conjunctive statement
- This is the **digital** content of this picture
- ⇒ What about the **analog** content?
- Compare: “here is a cube”

5

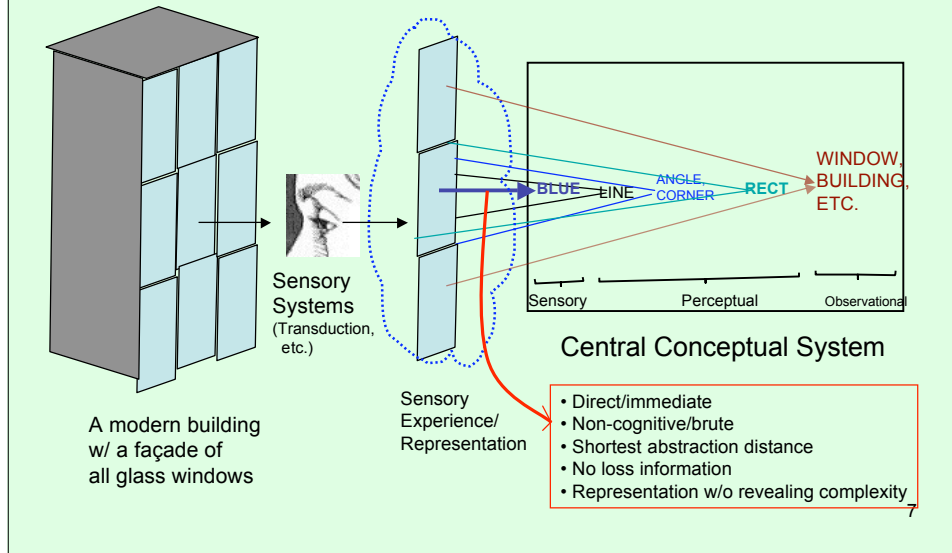
REMINDER: More about digital sensory information



- On the basis of your visual experience, you think that [there are three vertical colored stripes]
- Although this is a judgment thus conceptual, it also expresses a piece of analog information that your visual experience carries because it is nested in its digital information content.
- Here is a bit more specific info: [there are three vertical stripes colored red, green, red in that order]
- This information is nested in the color distribution at each point of the square area you are seeing.
- What information is available at each of these points, then?
- Color info, of course. But what can you know about colors just by experiencing it?

6

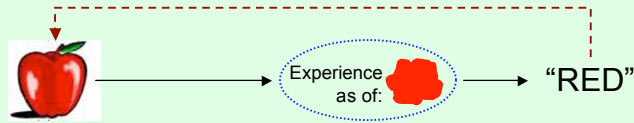
REMINDER: Cognitive Architecture of Consciousness



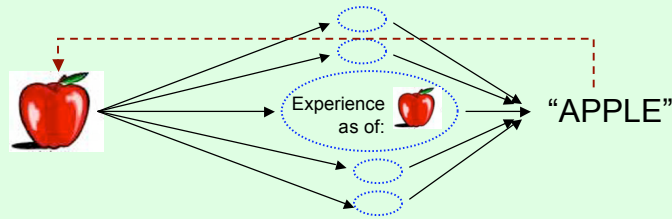
REMINDER: Sensory Concepts Summarized

- The acquisition of sensory concepts from their sensory bases is not mediated by any consciously available more specific information: the sustaining mechanisms for these concepts are non-cognitive. They can also be vertically applied as such, without cognitive mediation.
- I.e., these sensory concepts digitalize the most specific information carried by sensory experiences about the relevant values of a secondary quality they apply to. And this is to say that they have an abstraction/ digitalization distance that is minimal.

REMINDER: Manners of Tracking in Vertical Deployment of Concepts



There aren't very many ways in which *red* can visually strike us: the sensory concept "RED" tracks *red* by tracking a feature of red experiences – so, it has *dual information* content (IC)=<*redness*, e-*red*> (e-*red* = reddish experience)



There are indefinitely many ways in which an apple can perceptually strike us: SO the observational concept "APPLE" tracks apples without tracking apple experiences. 9

ENCORE: What is intuitively necessary and sufficient

for having phenomenally conscious experiences is the availability of information in the sensory array to a conceptual system sophisticated enough to be able to digitalize this information in the service of conceptually representing the world and thus interacting with it more selectively depending on the behavioral needs of the organism.

10

REMINDER:
Perception vs. Introspection

- So far, all the concepts we have examined are concepts that apply to extramental reality: we have been talking about perception, or better, exteroception.
- How do we detect, conceive and think of our experiences and their qualities?
- **How do we introspect?**
What concepts do we deploy in introspecting?

11

REMINDER: **Phenomenal Concepts**

- Phenomenal concepts are the concepts that give us the ability to detect and categorize the phenomenological qualities of our sensory experiences.
- It is through them we come to notice and thus come to know what it is like to see colors, hear sounds, smell odors, feel warmth, etc.
- The ability to thus reflect on our experiences is a sophisticated cognitive achievement.
- So, where do phenomenal concepts come from and what are they, exactly?

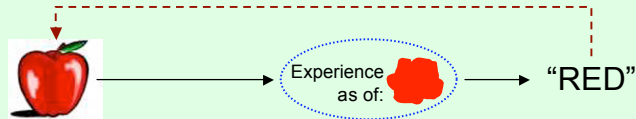
12

REMINDER: Intransitive Bodily Sensations

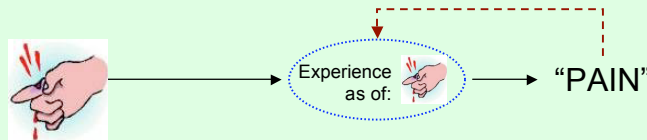
- We have a few phenomenal concepts (PC's) whose acquisition and deployment follow a different pattern than the rest, and do not require a lot cognitive sophistication.
- These are the concepts of intransitive bodily sensations such as pains, itches, tickles, tingles, orgasms, and a few others.
- Let's focus on pain briefly.

13

REMINDER: Identical Information Flow but Different Semantics



Dual IC("RED")=<redness, e-red>, but SC("RED")=redness
So "RED" is **sensory concept!**

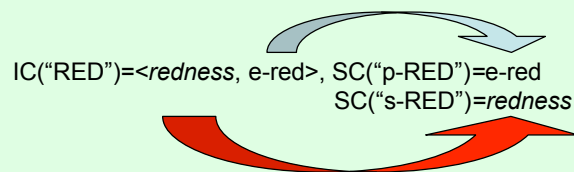


Dual IC("PAIN")=<damage, e-damage>, but SC("PAIN")=e-damage
So "PAIN" is a **phenomenal concept!**

14

REMINDER: Introspection: Sensory Concepts Becoming Phenomenal Concepts

- **Introspection** is precisely that mechanism which takes the second element in the informational content (IC) of **sensory concepts** and makes it their semantic content (SC), turning them into **phenomenal concepts** deployed in detecting and categorizing experiential qualities
- Let's denote phenomenal concepts by prefixing a 'p-' to the name of sensory concepts
 - E.g., The sensory concept, call it s-RED, becomes the phenomenal concept p-RED.
- **RED carries info about red by virtue of carrying info about e-red**



15

Phenomenal Concepts are closest, in terms of abstraction distance, to what they apply

- When PAIN or p-RED is vertically applied to the relevant aspects of their sensory bases, they carry the most specific information about those aspects (which are then made the concepts' semantic contents)
- These aspects are brain processes realizing the experiences.
- But the way PCs carry this piece of digital information about these brain processes makes the analog information nested in them non-extractable.
- Thus, the information about the internal complexity of these brain processes is not available for further digitalization: The abstraction distance between them is minimal.
- Thus, phenomenal concepts will present the relevant neurophysiological processes realizing experiences as structureless atoms -- primitive elements of our experiences

16

Further Immediate Consequences: Sensory and Perceptual Concepts

- It follows from this that our phenomenal concepts (just like sensory concepts) can pick out the experiential qualities they denote directly and immediately, and that they are independent of any other concepts in this sense.
- In particular, they are independent of any physical or functional concepts, and therefore not only cannot be defined in terms of them but also no such concepts are involved even in fixing their reference: i.e., none of them is involved in the sustaining mechanisms that determine their semantics.
- This means that phenomenal and sensory concepts cannot be derived from any other concepts or theories couched in terms of them.

17

No other option for autonomous intentional organisms!

- No other concepts work this way — all others involve *cognitive* sustaining mechanisms (abstraction distance is not minimal -- hence there is always loss of info)
- The reason for this is a nomologically necessary fact about autonomous intentional creatures like us:
- Any such creature that interacts with its environment has to have information entry mechanisms. These cannot deliver every piece of information in an analog form nested in more specific information. There will have to be a point where some information has to be delivered in digital form that encodes the information nested in it in a non-extractable format.
- **This is one of the most fundamental laws about such systems. This peculiar working of phenomenal concepts is not an optional feature.**

18

Some puzzling implications – 1

- Suppose this Naturalistic Story (NS) is true and we deploy phenomenal concepts when we think about consciousness.
- Since the semantics of phenomenal concepts, unlike all other concepts, is fixed non-cognitively, there is not only no semantic or conceptual analysis of phenomenal concepts but also no reference-fixing evidence base related to these concepts apart from the phenomenal properties themselves.
- Then, we cannot derive, *a priori* or otherwise, the existence of phenomenal consciousness from a complete physicalistic description of our world! [WHY?]
- So there is no formal contradiction in entertaining the conjunction of this physical description and statements denying the existence of consciousness expressed by deploying phenomenal concepts!

19

Meet Mary the Omniscient Scientist

Frank Jackson (1986):

"Mary is confined to a B/W room, is educated through B/W books and through lectures relayed on B/W television. In this way she learns everything there is to know about the physical nature of the world. She knows all the physical facts about us and our environment, in a wide sense of 'physical' which includes everything in completed physics, chemistry, and neurophysiology, and all there is to know about the causal and relational facts consequent upon this, including of course functional roles. If physicalism is true, she knows all there is to know."



Knowledge Argument (1986):

- (1) Mary (before her release) knows everything *physical* there is to know [about other people].
- (2) Mary (before her release) does not know everything there is to know [about other people] (because she learns something about them on her release).
- (3) Therefore, there are truths about other people (and herself) which escape the physical story.

20

What is Physicalism?

Jackson (1986):

“Physicalism is not the noncontroversial thesis that the actual world is largely physical, but the challenging thesis that it is entirely physical. This is why physicalists must hold that complete physical knowledge is complete knowledge simpliciter.”

Jackson (1994):

“Any world that is a minimal physical duplicate of our world is a duplicate simpliciter of our world.”

MPD = A world built with all & only the physical material of the actual world, put together in exactly the same way.

21

Information-Theoretic Response

- Before her release (BHR), Mary doesn't have any color sensations; so her visual experiences don't carry information about colors
- Thus, there is no information for her conceptual system to digitalize from her experiences
- Thus she can't have sensory concepts of colors; this implies that she won't have cognitive structures with double informational content
- This in turn implies that she can't acquire phenomenal concepts to represent color phenomenology, BHR.
- AHR, upon experiencing colors, she acquires their sensory concepts, and the corresponding phenomenal concepts, so she can represent the relevant old facts she already knew (under their scientific description) in a new way.
- So what is the problem?

22

Jackson's Version

- BHR, Mary doesn't have any color sensations; so her visual experiences don't carry information about colors
- Thus, there is no information for her conceptual system to digitalize from her experiences
- Thus she can't have sensory concepts of colors; this implies that she won't have cognitive structures with double informational content
- This in turn implies that she can't acquire phenomenal concepts to represent color phenomenology, BHR.
- AHR, upon experiencing colors, she acquires their sensory concepts, and the corresponding phenomenal concepts, so **she can now represent color phenomenology and thus come to know facts about it that she didn't know before.**
- This is the problem is for physicalism!
- **Is this credible given the story I have just told? NO!**

23

Jackson's (1994) Defense –1

- 1) Let P stand for the complete microphysical description of the world, and K for the complete description of all truths/facts entailed by P.
- 2) 'P&K' describes all facts about our world according to the physicalist, and that is what Mary knows.
- 3) Further: $\Box(P \rightarrow K)$, that is, the entailment is metaphysical necessitation.
- 4) So for any particular truth, K_i , $\Box(P \rightarrow K_i)$.
- 5) However, for Jackson, this can't be a brute fact, and needs explanation.
- 6) The only plausible explanation is that given our ordinary competence with the semantics of terms/concepts involved in the expression/judgment of K_i , and sufficient acumen, we can a priori figure out how K_i follows from P, that is ' $P \rightarrow K_i$ ' is a conceptual truth.
- 7) For example, we are able to figure out a priori how water facts are made true (explained) by H₂O facts. In the absence of a conceptual connection, all we would have are metaphysically brute necessary correlations. There would be no or very little intelligible connection between the two sets of truths.

24

Jackson's (1994) Defense – 2

- 8) If physicalism is not to be a mysterious doctrine, not only all Ks must follow this pattern, but as a matter of fact, they do follow this pattern, except when the the truths in question are phenomenological truths (call these, C) like 'people sometimes feel pain'
- 9) While in the B/W room, Mary, who knows P, can derive all Ks except the truths about color phenomenology that she comes to know only after her release.
- 10) But even after acquiring this body of phenomenological knowledge after her release, it remains conceptually isolated: Mary still cannot derive it a priori, cannot explain it by appeal to P; that is, the connection still is unintelligible, except that now she at least possesses this knowledge and enjoys it very much, delights in it, etc.
- 11) So there are some truths, C, not entailed by P; thus physicalism is false.
- 12) The reason why phenomenological truths cannot be derived from P is because phenomenal concepts have no a priori conceptual analyses; or put less strongly, they have no cognitive reference fixers.

25

Surprise?

- On the Natrualistic Story (NS) I have told, this result is to be expected!
- In fact, given the way we are cognitively structured and organized, the fact that phenomenal concepts lack conceptual analyses or cognitive reference fixers is nomologically necessary -- **it is a law of nature!**
- But given that NS is purely naturalistic, this result gives no support to the rejection of physicalism.
- On the contrary, physicalism of the sort I defend predicts and explains how phenomenal concepts will have these perplexing consequences.

26

Options for physicalists: A Priori Physicalism

' $P \rightarrow C$ ' is an a priori/conceptual truth, thus its necessity
(so C is a proper subset of K).

Problem: This solution requires

- either that given her physical omniscience Mary will be able to figure out a priori what it is like to see colors prior to her release (or at least, after her release she will be able to construct such an a priori derivation) -- this in turn requires that our concepts of phenomenal qualities have conceptual analyses in terms of physically respectable conditions like behavioral dispositions or functional/causal roles, etc.
- or that Mary doesn't in fact acquire any propositional knowledge after her release (the Ability Response to KA)

Neither of these claims are credible.

27

Options for physicalists: A Posteriori Physicalism

' $\Box(P \rightarrow C)$ ' is true but ' $P \rightarrow C$ ' is an a posteriori necessity,
just as ' $P \rightarrow K$ ' is! (So C is a proper subset of K).

Problem: Even if we grant this claim, there is certainly some epistemic relation between P and K that doesn't hold for the relation between P and C. This epistemic relation may not amount to a priori or conceptual entailment, but surely there is some intelligibility relation between, say, H₂O facts and water facts, that doesn't hold for, say, the brain facts and phenomenological facts. The mystery is removed in the former case, but the mystery deepens in the latter. This fact can easily be turned into an argument against physicalism in a way paralleling Jackson's argument (Levine).

28

Options for physicalists: Exceptionalist Physicalism

' $\Box(P \rightarrow C)$ ' is true but not a priori/conceptual, so ' $P \rightarrow C$ ' is an a posteriori necessity, which constitutes an exception to the fact that ' $P \rightarrow K$ ' a conceptual truth (so C is not a subset of K).

Problem: Claiming exceptions doesn't necessarily make theories false, but it certainly makes them arbitrary and ad hoc. What explains the exception? The anti-physicalist has a ready answer. The exceptionalist physicalist has a hard time explaining why it should be the case that in every possible world in which P is true, C is also true, especially when ' $P \& \sim C$ ' seems now perfectly conceivable. Hence the argument from the conceivability of phenomenal zombies against physicalism (Hart, Chalmers). To rule out their possibility now requires a principled, independently motivated explanation why conceivability in this particular case doesn't imply possibility (when it does in every other case).

29

Revised Thought-Experiment involving Mary

- BHR, suppose that Mary knows all there is physical to know not only about color vision but also all about introspection and concept formation. Then, supposing that NS is true, the complete details of this account are what she would know. But then she would automatically be in a position to know about the curious asymmetry involved in the epistemic access to phenomenal/physical facts.
- This body of knowledge she has BHR would not of course remove her curiosity (the surprise element) about coming to know in a first-person way facts she already knew under their scientific description. On the contrary, she would be even more curious and intrigued to *instantiate* those phenomenal/physical states herself, which are necessary for acquiring the peculiar perspectival concepts, and thus first-person knowledge.

30

Parallels with Indexicality

- In this revised thought-experiment, where we assume the truth of NS, Mary can derive from the physicalistic premise the existence of phenomenal concepts and their curious semantics, but since she herself doesn't yet have these concepts about color experiences, she cannot *use* them to attribute *color experiences* to others (she doesn't yet have color experiences herself). But she can *mention* these concepts as given by NS and attribute *them* to others.
- There is of course an obvious fact about Mary before her release: namely, since she doesn't have the relevant phenomenal concepts yet, she cannot derive claims couched in them from the physicalistic premise.
- But of course it isn't this fact that makes derivability impossible. Even after her release, when she comes to possess all the relevant phenomenal concepts, she still cannot derive them from the physicalist base. Any phenomenal claim couched in phenomenal terms/concepts will make *use* of phenomenal concepts, and she won't be able to derive these claims from the physicalistic base even when this base includes the details of NS, unless she uses a premise in which she identifies the referent of the theory-given phenomenal concepts with the referent of her first-person phenomenal concepts by *using* the latter.
- But this premise itself is not derivable from the physicalist premise with NS. The situation here is parallel to the impossibility of deriving an indexical claim from indexical-free premises (cf. Perry 1979).

31

Final Thoughts

- As in the case of Mary, suppose that a completed physicalistic cognitive science vindicates our information-theoretic account of concept acquisition and sensory/phenomenal concepts and we have a complete physical knowledge of our world. Then the physicalist premise base true of our world will entail that we have sensory and phenomenal concepts that we use to denote conscious qualities of our experiences and make claims about them that we cannot derive from this physicalist base.
- Nevertheless, we will know that our sensory/phenomenal concepts as we have acquired them from our experience pick out the same physical/functional properties that certain physical/functional description in our theory picks out.
- If I am right about NS, physicalism is not threatened by these conceivability arguments. On the contrary it is vindicated!

32

Textual Resource

A comprehensive and detailed examination of the basic issues covered in these two lectures can be found in an article forthcoming in Noûs, 39(2): 197-255, (co-authored by Güven Güzeldere).