

Don't Mind the Gap

David Rosenthal
CUNY Graduate Center
Philosophy and Cognitive Science
<http://davidrosenthal1.googlepages.com/>

CUNY Cognitive Science
December 12, 2008

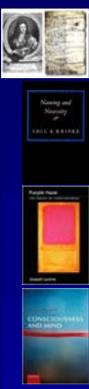


OVERVIEW

- I. Leibniz's Mill Argument
- II. Kripke's Modal Argument
- III. Levine's Explanatory Gap
- IV. A Theory of Mental Qualities

Acknowledgment: This talk was first presented, in a slightly earlier version, at a conference at McGill University in honor of Fabrizio Mondadori.

Don't Mind the Gap December 12, 2008 2





I. Leibniz's Mill Argument

- In Monadology, Leibniz advances a well-known argument for the view "that perception and that which depends upon it are inexplicable on mechanical grounds, that is to say, by means of figures and motions."
- He invites us to imagine "a machine, so constructed as to think, feel, and have perception, ... [but] increased in size, while keeping the same proportions, so that one might go into it as into a mill" (Monadology §17).

Don't Mind the Gap December 12, 2008 3

- Leibniz argues that "we should, on examining [the] interior [of this machine], find only parts which work one upon another, and never anything by which to explain a perception."
- He concludes that this shows that perceiving can occur only in "a simple substance, and not in a compound or in a machine."
- Descartes had argued (M6) from *the unity of consciousness* that a thinking substance must be indivisible; Leibniz uses his Mill Argument in part to reach a similar conclusion for perceiving.

Don't Mind the Gap December 12, 2008 4

- But I won't discuss mental unity today, but rather two other claims that are both suggested by the Mill Argument.
- One is that a physicalist theory of perceiving is simply false.
In Leibniz's terms: Perceiving cannot occur in compound, mechanistic objects—i.e., in objects that operate "by means of figures and motions."
- But the Mill Argument supports this view at best obliquely, arguing directly only for a more modest view: not that physicalism isn't true, but that we cannot understand how it could be.

Don't Mind the Gap

December 12, 2008

5

- Leibniz imagines entering a perceiving machine, but "never finding anything by which to explain a perception." Nothing in its physical makeup or operation would enable us to understand perceiving.
- Joseph Levine has called this more modest claim the explanatory gap: Even if perceiving is due to a physical process, we cannot understand how that could be (cf. David Chalmers's Hard Problem, 1995a, b, 1996).
- Why would any physical process result in a qualitative experience of red—rather than an experience of green? And why would it result in any experience at all?

Don't Mind the Gap

December 12, 2008

6

- Leibniz countenances perceiving that isn't conscious—as he explains it, perceiving in the absence of *apperception* (§14).
- And because the Mill Argument does not mention apperception, it presumably applies to nonconscious as well as to conscious perceiving (thanks to Massimo Mugnai and Fabrizio Mondadori for pressing this).
- Levine's explanatory gap, by contrast, applies only to *conscious perceiving*, since perceiving involves mental qualities, which he holds are *always conscious*.
- My main concern today is Levine's gap, but I'll return briefly to Leibniz in closing.

Don't Mind the Gap

December 12, 2008

7

- I'll argue that there is no explanatory gap, and that the temptation to think there is *rests on certain assumptions that are not only unwarranted but untrue*.
- But I'll also raise, and argue against, some considerations due to Saul Kripke that are intended to support the *stronger* conclusion— not just that we *can't understand* how mental qualities could be some type of physical process, but that mental qualities *cannot be* any type of physical occurrence.

Don't Mind the Gap

December 12, 2008

8

- In §II I'll consider Kripke's argument—in the process laying some ground for my subsequent discussion.
- I'll then argue in §III that we must reject the idea of an explanatory gap—which casts some doubt on Leibniz's Mill Argument as well.
- In my closing §IV, then, I'll sketch a theory of mental qualities, which I'll argue helps dispel whatever intuitive resistance we have to the idea that states with qualitative character are physical states—states whose properties are exclusively physical.

Don't Mind the Gap

December 12, 2008

9



II. Kripke's Modal Argument

- Perceiving involves the registration of information about one's environment or about one's own bodily condition.
- But it's not that aspect of perceiving that seems problematic—and that the Mill Argument seems to show cannot be explained in mechanistic terms.
- Rather, it's that perceiving takes place by way of *mental qualities*—and it seems unclear how to explain mental qualities in physical terms.

Don't Mind the Gap

December 12, 2008

10

- Just as nothing that one sees inside Leibniz's enlarged machine seems to explain mental qualities, so it seems to many that physicalist "theories must 'leave out' the qualitative, conscious side of mental life" (Levine 1993/1997, 543).
- Kripke's argument is a *metaphysical* version of this worry, whereas Levine's is in effect an *epistemological* version.
- But both appeal to considerations about *conceivability*—and both in that way echo Leibniz's Mill thought experiment.

Don't Mind the Gap

December 12, 2008

11

- Here is Kripke's argument. Rigid designators refer to the same things in every possible world in which those things exist.
- Kripke holds, plausibly, that the terms in (1) and (2) are all rigid:
 - (1) Water = H₂O.
 - (2) Pain = the firing of C-fibers.
- So both are necessary if true at all.
- But both *appear* to be contingent. Still, we can *explain why* (1) appears to be contingent: Something can *appear* to be water without *being* H₂O (Levine; Kripke uses heat).

Don't Mind the Gap

December 12, 2008

12

- But we can't, Kripke urges, explain why (2) Pain = the firing of C-fibers. seems contingent, since whatever appears to be pain is pain.
- "For a sensation to be felt as pain is for it to be pain" (IN 163, n. 18; Kripke's emphasis).
- And also: "[F]or [something] to exist without being felt as pain is for it to exist without there being any pain" (NN 151).
- So, "although we can say that we pick out [physical] heat [or water] contingently by the contingent property that it affects us in such and such a way, we cannot similarly say that we pick out pain contingently by the fact that it affects us in such and such a way" (IN 161).

Don't Mind the Gap

December 12, 2008

13

- And if we can't explain away the appearance of contingency in (2) Pain = the firing of C-fibers. then (2) simply is contingent.
- If so, (2) isn't necessary—and if both designators are rigid, it's not true at all.
- A standard response to this argument is to urge that 'pain' is not rigid after all—that 'pain' refers to different states in different counterfactual situations.
- I'll return shortly to that issue. But I want first to question the alleged disanalogy between pain, on the one hand, and water or heat, on the other.

Don't Mind the Gap

December 12, 2008

14

- Something can *appear* to be water and yet not *be* water—and *conversely*.
- But Kripke claims that neither thing can happen with pain. Appearing to be pain is being *felt as* pain, and “the way we identify pain is by feeling it” (IN 163, n. 18).
- And Kripke also claims, conversely, that “for [something] to exist without being *felt as pain* is for it to exist without there *being any pain*” (NN 151; Kripke’s emphasis).
- If so, pain has no *epistemic intermediary*. And imagining C-fibers’ firing without any feeling of pain would then show that C-fibers’ firing is distinct from pain itself.

Don't Mind the Gap

December 12, 2008

15

- If *being* pain and *appearing* to be pain coincide in this way, then how a pain feels is “an essential (indeed *necessary and sufficient*) property” of pain (IN 163, n. 18; my emphasis).
- And if Kripke is right about this, our *third-person* grasp of pain is due solely to *contingent* correlations with pain—where pain is identified *just in a first-person way*.
- And the contingency of those correlations would mean that our third-person grasp of pain *varies from one counterfactual situation to another*.

Don't Mind the Gap

December 12, 2008

16

- But that pretty clearly isn't so.
- It's central to something's being a pain that people have an aversive reaction to it, that it distracts them, and that it results from characteristic stimuli and leads to characteristic behavior, etc.
- Not all pains have all these features. In pain asymbolia (a result of cingulotomy or morphine), pains lack aversive effect. So we might identify such a pain just by its subjective appearance—its feel.
- But that feel is arguably *a symptom of a state that in the typical case has the third-person properties just mentioned.*

Don't Mind the Gap

December 12, 2008

17

- Another reason to deny that pain is essentially as it appears is that some pains occur *without being conscious*—and so without there being anything that *it's like for one* to have the pain.
- Consider a headache that, as we say, lasts all day—though often during that period one isn't *at all* conscious of it. Or the slight painful pinch of a shoe of which one is seldom *in any way* aware.
- Here there is no conscious feel at all; it seems subjectively that *there is no pain*—no state that *seems* to be a pain. Appearance and reality thus diverge.

Don't Mind the Gap

December 12, 2008

18

- Not every pain appears to be pain. But how about the converse?
- In so-called *dental fear*, patients report pain in the drilling of a tooth whose nerve is fully anaesthetized or even absent. The standard explanation is that the patient experiences vibration and fear, and *mistakes those for pain*. Some confirmation: When given this explanation and drilling resumes, patients no longer report feeling pain, but only fear and vibration.
- So the converse also fails: Apparent pain does occur without any actual pain.

Don't Mind the Gap

December 12, 2008

19

- Indeed, our first-person access to pain is itself sometimes erroneous.
- Pain of a fixed intensity *introspectively seems more intense* when one thinks it results from an intention to hurt one than when one doesn't think that (Gray and Wegner, *Psychological Science* 2008).
- Psychological factors *other than* the pain itself lead introspection to *misrepresent* the intensity of the pain.
- So third-person factors can trump our first-person access to pain, driving yet another wedge between pain and how the pain feels.

Don't Mind the Gap

December 12, 2008

20

- Some theorists approach the mind only in a first-person way (Nagel, Searle), others in just a third-person way (Dennett?).
- But pain occurs without its typical third-person concomitants, as in pain asymbolia, and it also occurs without its typical first-person properties.
- The concept of pain is, as Hilary Putnam has urged ("Brains and Behaviour," 1965, 5), a cluster concept:
It does not rely on one or even several dispositive criteria, but on a weighted collection of many indicators, both first- and third-person in nature.

Don't Mind the Gap

December 12, 2008

21

- The same arguably holds for all mental states, qualitative as well as intentional states.
- Neither first- nor third-person properties are, alone, essential to pain—nor to mental states generally.
- So we must guard against accounts that posit either first- and third-person properties as being decisive indicators of pain or other qualitative states at the expense of the other—
and we must also *reject any account on which first- and third-person properties don't fit comfortably together.*

Don't Mind the Gap

December 12, 2008

22

- These considerations all suggest we must distinguish pain—as well as other qualitative states—from *our awareness of those states*.
- And that suggests a way to explain why (2) Pain = the firing of C-fibers. *appears* to be contingent even if it is in fact necessarily true (assuming that 'pain' is a rigid designator).
- *Seeming* to be pain ≠ pain. Just as the appearance of water is tied only contingently to H₂O, so *the appearance of pain* is tied only contingently to the firing of C-fibers.

Don't Mind the Gap

December 12, 2008

23

- But there's a problem. This *appearance of pain* should itself be associated with *yet another neural state*. And we can then imagine this appearance of pain as occurring without *that* neural state.
- So the pain itself aside, the appearance of pain seems tied only *contingently to its associated neural state*. If so, they're not necessarily identical—so perhaps not identical at all (Dan Shargel, unpublished, to whom thanks for useful conversation).
- Still, we have access to this appearance of pain, in turn, by our *introspective awareness* of how pain appears to us.

Don't Mind the Gap

December 12, 2008

24

- And that suggests a reply. Recall that when it *seems* we imagine pain as independent of C-fibers' firing, we're *really* imagining *the way pain appears to us* as independent of C-fibers' firing.
- So when it seems that we imagine the appearance of pain as independent of its associated neural state, *we're really imagining only our introspecting of that appearance without the neural state that's associated with that appearance.*
- At each stage, we imagine a mental state without the neural state *associated with the mental state one level down.*

Don't Mind the Gap

December 12, 2008

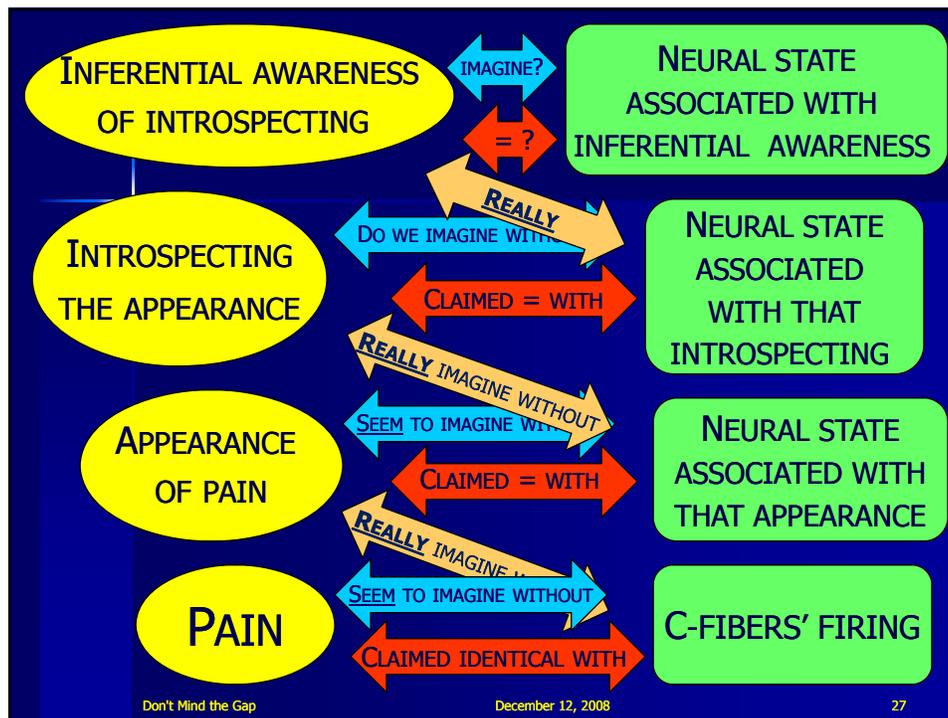
25

- This goes no higher. There is no *hyper-introspection* by which we're aware of our introspective awareness itself.
- This is evident because introspecting does not sometimes seem subjectively to be conscious and sometimes not.
- Rather, we're aware of introspecting indirectly, by inference: We *infer* we're introspecting because we come to be aware not just of our conscious states, *but also of our being conscious of them.* (Cf. Dretske's [1994/5, 1999] account of introspection as *displaced perception*—though he explains *all* first-person access that way, not just access to introspecting itself.)

Don't Mind the Gap

December 12, 2008

26



- If we're inferentially aware of introspecting the subjective appearance of pain, we can imagine it without its associated neural state. And perhaps if we infer that the inferential awareness occurs, we can go even one step higher.
- But we can also just bypass this entire sequence. At each step we can imagine a mental state as occurring without *some particular neural state*.
- So perhaps we can *also* imagine each mental state as occurring, not just without some particular neural state, but *independently of any neural state at all*.

- If so, a single imagining, at whatever level, would show that some mental state is tied contingently to all neural states—and so isn't identical with any.
- So let's return to whether 'pain' is rigid. Why think it is? Who knows what 'pain' refers to in counterfactual situations?
- 'Pain' will seem rigid if we see our access to pains as limited to the way we're conscious of them.
- If we have no other access to pains, there's nothing in virtue of which 'pain' might refer to the different things across different counterfactual situations.

Don't Mind the Gap

December 12, 2008

29

- But since pains do occur independently of their being conscious, there's no reason to think that 'pain' is rigid.
- And then Kripke's argument doesn't go through even if we can conceive, in a relevantly trustworthy way, that pain or a suitable awareness of pain can occur without any neural states at all.
- N.B.: Kripke's argument fails at bottom because pain and awareness of pain are independent: Either that awareness is an epistemic intermediary between pain and C-fibers—or that independence undermines the claim that 'pain' is rigid.

Don't Mind the Gap

December 12, 2008

30

- Kripke's argument rests on our inferring from its being conceivable that pains occur independently of neural states to its being possible that they do.
- So we can also contest Kripke's argument by disputing that inference.
- The inference from what's conceivable to what's possible is least convincing when appearance and reality diverge, since then it's credible that conceivability outstrips possibility:
Perhaps, e.g., we can conceive that water exists without H₂O's existing even if it isn't possible that water ≠ H₂O.

Don't Mind the Gap

December 12, 2008

31

- Perhaps of course conceivability implies possibility even then—perhaps one can conceive only of something that seems to be water as existing without H₂O.
- But whatever the case about that, the inference from what's conceivable to what's possible is at its best when appearance and reality do coincide—as Kripke claims is true of pain.
- So rejecting the inference from what's conceivable about pain to what's possible about it will likely again mean disputing the claim that the reality and appearance of pain coincide.

Don't Mind the Gap

December 12, 2008

32

- For continuity with Levine, I've used his example of water and H₂O, instead of Kripke's example of the identity of heat with mean molecular kinetic energy.
- Heat does make for a more striking disanalogy with pain, since the appearance of heat *just is the feeling of heat*, and that occurs without heat itself—in contrast with what Kripke urges about pain and the feeling of pain.
- Levine appeals to the *superficial perceptible properties* of water. But we can equivalently appeal, as with heat, to our *awareness* of the relevant properties.

Don't Mind the Gap

December 12, 2008

33

- Independent of the foregoing, Putnam has suggested another way to avoid Kripke's argument (personal communication).
- We can seek an *epistemic intermediary* not for 'pain', but for 'firing of C-fibers'.
- Just as water or heat can appear to be present when it isn't, so the firing of C-fibers can *appear* to occur even when it doesn't really occur.
- So when it seems that one is imagining pain as independent of C-fibers' firing, it can be that one is instead imagining only that pain is independent of the *appearance* of C-fibers' firing.

Don't Mind the Gap

December 12, 2008

34



III. Levine's Explanatory Gap

- Kripke's modal argument purports to show that the qualitative character of mental states cannot be physical.
- But even if such qualitative character is physical, it might still be that we *cannot understand* how it *could* be physical, or even result from anything physical.
- It's this more cautious conclusion that Levine labels the explanatory gap and that Leibniz's Mill Argument is in the first instance intended to establish.

Don't Mind the Gap

December 12, 2008

35

- Levine again formulates things in terms of identities. Consider our familiar pair:
 - (1) Water = H₂O.
 - (2) Pain = the firing of C-fibers.
- Even if both are true, Levine argues, they're on different footings. We have *no difficulty understanding how (1) can be true*, as we do with (2).
- And like Kripke, Levine puts the contrast in terms of *conceivability*. On the one hand, he urges, "it is ... conceivable that there should exist a pain without the firing of C-fibers, and the firing of C-fibers without pain" (1993/1997, 548).

Don't Mind the Gap

December 12, 2008

36

- By contrast, he insists:
“While it is conceivable that something other than H₂O should manifest the superficial macro properties of water, ... it is not conceivable ... that H₂O should fail to manifest those properties (assuming, of course, that we keep the rest of chemistry constant)” (548).
- But note the parenthetical qualification:
Perhaps it's *also not* conceivable that the firing of C-fibers should fail to manifest the property of being pain—assuming that we feed in a suitable future neuropsychology.

Don't Mind the Gap

December 12, 2008

37

- Descartes held that we're warranted in positing unextended substances other than ourselves because we can't explain linguistic behavior physiologically (to Newcastle, AT IV 573-576, to More, V 275-279; Discourse, VI 58-9).
- Similarly, our current vast ignorance of neuropsychology may well induce the sense that—no matter what we came to know—the firing of C-fibers could fail to manifest the property of being pain.
- But Descartes was rash: We can now foresee a physiological explanation of linguistic behavior. Caution about a full account of pain is similarly appropriate.

Don't Mind the Gap

December 12, 2008

38

- Levine epitomizes the situation as follows:
“Once all the standard superficial properties are explained by reference to the structure of H₂O molecules and general chemical laws, there seems to be no substantive cognitive significance to the question how water could be H₂O.”
- “On the other hand, even after all the causal role properties of experience are explained by reference to its neurophysiological or functional structure, still there seems to be genuine cognitive significance to the question how reddish-ness [e.g.] could be a neurophysiological or functional property” (2001, 83).

Don't Mind the Gap

December 12, 2008

39

- But even if it does *seem* that way, why trust that intuitive appearance? Why think it would still seem that way *given the full neuropsychological story?*
- According to Levine:
“[I]t seems to make sense that a creature could instantiate the same physiological states as me and yet have different qualia, or none at all” (2001, 79).
- Making sense is again *conceivability*; it's conceivable, he thinks, that this should occur. And if it is conceivable, there's presumably *no rational tie* between pain and neural functioning.

Don't Mind the Gap

December 12, 2008

40

- But if pain is C-fibers' firing, the rational tie pain has to neural functioning is no less than that between water and H₂O.
- In *both* cases, the rational tie would stem *only* from some theory—chemistry for being water and H₂O and neuropsychology for mental qualities and neural functioning. *Such rational ties are always due to theory—folk or scientific.*
- Cf. Eddington's two tables (Gifford Lectures, 1927/8): Before atomic theory became second nature, the *2-table gap* seemed unbridgeable; once the theory is taken in, the gap no longer seems serious.

Don't Mind the Gap

December 12, 2008

41

- Here's what Eddington wrote:
Two tables! Yes; there are duplicates of every object about me—two tables, two chairs, two pens. ...
One of them ... is comparatively permanent; it is coloured; above all it is substantial. By substantial I do not merely mean that it does not collapse when I lean upon it; I mean that it is constituted of "substance" and by that word I am trying to convey to you some conception of its intrinsic nature. It is a thing; not like space, which is a mere negation. ...
There is nothing substantial about my second table. It is nearly all empty space—space pervaded, it is true, by fields of force, but these are assigned to the category of "influences," not of "things" (Introduction).

Don't Mind the Gap

December 12, 2008

42

- Today we may find this worry somewhat quaint. Eddington's first table simply is, at bottom, his second table.
- But we can resuscitate his concern: How can something permanent, solid, and colored be mainly colorless empty space? (E.g., Sellars, *Science, Perception and Reality*, ch. 1.)
- We dispel the worry not as Levine might like, by tracing a *theory-independent* rational path from the scientific to the commonsense table, but simply *by coming to accept that the ties physics forges between scientific posits and commonsense tables is just the way it is.*

Don't Mind the Gap

December 12, 2008

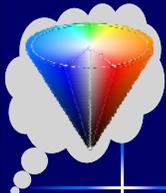
43

- Levine notes that the explanatory gap is due to the conceivability of physical duplicates of us that have no (*conscious*) qualitative states (79; cf. Chalmers's "zombies").
- Similarly, Leibniz's Mill Argument purports to show that whatever machinery may operate within us, we can conceive of it as not resulting in perceiving—whether conscious or not.
- In my closing section, I'll argue that this *apparent conceivability* is due to a mistake in how theorists often think about qualitative character—as being essentially or necessarily conscious.

Don't Mind the Gap

December 12, 2008

44



IV. A Theory of Mental Qualities

- Descartes saw all mental phenomena as cases of conscious thinking, and all physical, bodily phenomena as inhering in a substance whose only essential property is extension.
- These disparate characterizations make it seem at best mysterious how mental phenomena could be physical.
- Few if any accept those ideas today. But something similar is at work in much current thinking about mental qualities.

Don't Mind the Gap December 12, 2008 45

- Just as the distance between conscious thinking and pure extension may seem too great to bridge rationally, so too for the intuitive distance between neural functioning and conscious qualitative character.
- But we needn't go straight from neural functioning to conscious qualitative character.
- Pains—and qualitative states generally—need not be conscious. So we can go from neural functioning to nonconscious qualitative character, and from there to conscious mental qualities.

Don't Mind the Gap December 12, 2008 46

- For this to work, we need *two things*:
 - ➔ 1. An account of mental qualities that's *independent* of those qualities' being conscious.

And we also need:

- ➔ 2. An account of what *additional factor* results in mental qualities' sometimes being conscious.
- The two accounts *must fit together*.
And together they must allow for a credible *theoretical bridge* from neural functioning to the relevant mental phenomenon—thereby enabling us to cross the intuitive explanatory gap.

Don't Mind the Gap

December 12, 2008

47

- Perceiving requires discriminating among perceptible properties—properties an individual can perceive. And to do that, the individual must be able to be in *states that resemble and differ among themselves in ways that are homomorphic to the perceptible similarities and differences among the relevant objective properties*.
- The relevant similarities and differences among perceptible properties are *those perceptible by the individual*—not as described, e.g., by psychophysics or by physical theory.

Don't Mind the Gap

December 12, 2008

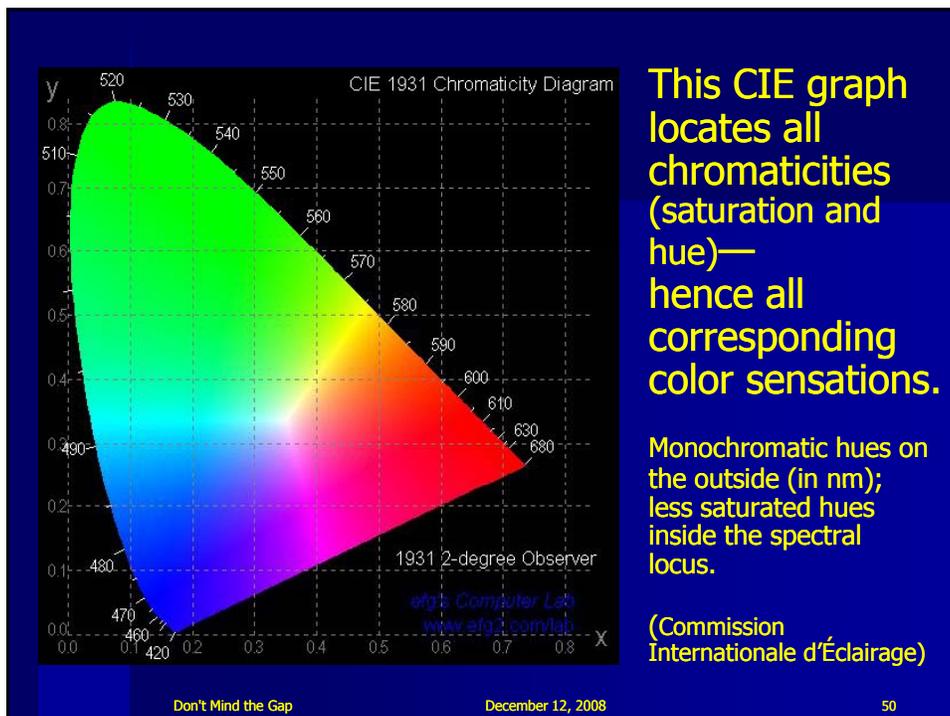
48

- We can then identify mental qualities as the differential properties of states that enable discrimination among properties accessible by each perceptual modality.
- Since discrimination among perceptible properties can occur subliminally—and hence without being conscious—the foregoing account is independent of qualitative states' being conscious.
- So we can individuate mental qualities independently of their being conscious—and hence independently of any first-person appearances that figure in an intuitive explanatory gap.

Don't Mind the Gap

December 12, 2008

49

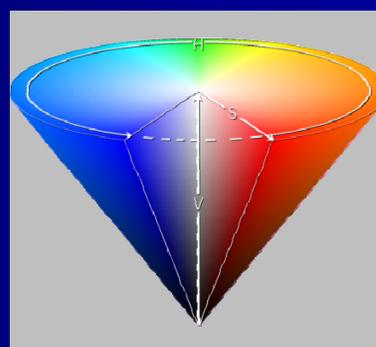
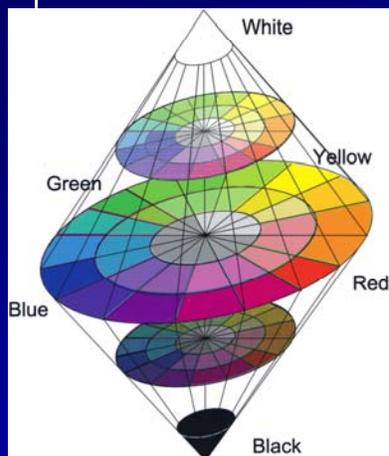


Don't Mind the Gap

December 12, 2008

50

For locating color sensations in respect of brightness as well as hue and saturation, we need a 3D color solid; this HSV color cone permits more intuitive 3D locating.



Don't Mind the Gap

December 12, 2008

51

- Identifying mental qualities by locations in a quality space homomorphic to the locations of sensed perceptible properties in their quality space is not new (see, e.g., Wilfrid Sellars, Science, Perception and Reality, chs. 2, 4, 5, and Sydney Shoemaker, "Functionalism and Qualia").
- What's new is the use of this technique to give an account of mental qualities that's independent of their being conscious (Rosenthal, Consciousness and Mind, chs. 5-7, esp. 7).
- The technique also explains the spatial aspects of mental qualities (Douglas B. Meehan, 2001, 2002, unpublished), as well as qualities that figure in bodily sensations.

Don't Mind the Gap

December 12, 2008

52

- So there's no obstacle to understanding *nonconscious* qualitative states as identical with neural states. But what about when the qualitative states *are* conscious?
- No state is conscious only if one isn't aware of it in some suitable way. So a qualitative state's being conscious consists in one's being aware of that state *in some suitable way*:
- A qualitative state is conscious if one is (suitably) aware of oneself as being in the relevant type of qualitative state.
- What, then, is that "suitable way"?

Don't Mind the Gap

December 12, 2008

53

- One must be aware of the state in a way that's subjectively unmediated and subjectively "from the inside"—i.e., independent of any inference or observation *of which one is aware*.
- And one must also be aware of the qualitative state *as such*—i.e., in respect of the relevant mental quality.
- That means being aware of oneself as being in a state whose *location in its quality space* corresponds to the location of the target perceptible property in the quality space of properties accessible by the relevant sensory modality.

Don't Mind the Gap

December 12, 2008

54

- This higher-order awareness (HOA) results in the target qualitative state's being conscious because a state's being conscious *consists* in its being a state one is, in the specified way, aware of oneself as being in.
- It's not that the HOA *transfers* its own consciousness to the target.
- So the HOA need not consist in a state that is itself conscious. And a mental state need not be conscious to make one conscious of its target; subliminally perceiving a thing makes one aware of it—though not consciously aware.

Don't Mind the Gap

December 12, 2008

55

- Since the HOA need not itself be conscious, there's no intuitive obstacle to understanding that HOA as being some neural or functional process. So there's no problem at that level.
- I've argued elsewhere that the higher-order awareness is a matter of one's having a thought that describes oneself as being in such a state—what I've called a higher-order thought.
- But any mental implementation on which the HOA is independent of the target state will serve to avoid the intuitive explanatory gap.

Don't Mind the Gap

December 12, 2008

56

- An explanatory gap obtains if we cannot understand why a particular neural process should yield a mental quality of, say, red—as against a mental quality of green, or none at all.
- But a perceiving organism has access to a range of perceptible properties by way of particular neural processes. So identifying those neural processes with *nonconscious* qualitative states yields a robust theoretical tie between neural processes and mental qualities.
- So our theory of mental qualities crosses *the first step* of the explanatory gap.

Don't Mind the Gap

December 12, 2008

57

- There's *no further mystery* when those mental qualities are conscious: They're conscious if one is aware of oneself as being in states with those mental qualities—by way of a suitable HOA.
- That's because the HOA represents the individual as being in a state that has a particular relation to what the individual can perceive—in respect of parallel locations in the relevant quality spaces.
- Combining the *homomorphism theory of mental qualities* with an acceptable *higher-order theory of consciousness* bridges the apparent explanatory gap.

Don't Mind the Gap

December 12, 2008

58

- This *two-stage theory* does not fit comfortably with the idea that qualitative states are intrinsically conscious (e.g., Levine, Kriegel, Chalmers).
- On the two-stage theory, the properties of *having mental quality* and *being conscious* are independent of one another.
- That does not *entail* that consciousness is not an intrinsic property of qualitative states.
 But *if a state's having mental qualities is independent of its being conscious, we should expect that being conscious is not an intrinsic property of that state.*

Don't Mind the Gap

December 12, 2008

59

- We taxonomize qualitative states by their mental qualities. So *mental qualities are intrinsic to the states.*
- And since mental qualities are independent of consciousness, it's *unmotivated* then to hold that being conscious is also intrinsic to qualitative states—and even unclear what that could amount to.
- And there is *evidence* that qualitative states are not intrinsically conscious. We're far more likely to be *conscious of two slightly different mental qualities as distinct* when they occur together than one after the other (see, e.g., Raffman, 1995).

Don't Mind the Gap

December 12, 2008

60

- But if consciousness were *intrinsic* to the qualitative states, it wouldn't matter whether they occur together or in succession.
- The best explanation of this striking disparity is the two-stage theory—on which *we're conscious of mental qualities in respect of their location in the relevant quality space*.
- We're conscious of finer differences among simultaneous qualities because we can locate them *comparatively* within their quality space, which is a lot harder when the qualities don't occur together.

Don't Mind the Gap

December 12, 2008

61

- If consciousness were *intrinsically tied to mental qualities*, 1st-person access to them would trump 3rd-person access.
- And since my mental qualities aren't conscious for you, mine might, as far as you can know, be inverted relative to yours or even absent altogether—which *the two-stage theory blocks*.
- If quality inversion and absence were conceivable, *mental qualities would seem independent of nonmental reality*, which would not only suggest an explanatory gap, but also preclude an *informative* account of mental quality.

Don't Mind the Gap

December 12, 2008

62

- The two-stage theory enables us to bridge the explanatory gap. But as noted earlier, Leibniz takes his Mill Argument to show that we can't understand mechanistic explanation even for qualitative states that aren't conscious.
- Why did Leibniz think that it applies even to nonconscious mental qualities? And should his reasons worry us?
- As noted in §I, Leibniz countenances perceiving that isn't conscious, using apperception to explain why perceiving is sometimes conscious (Monadology, §14)—a higher-order theory (inner sense).

Don't Mind the Gap

December 12, 2008

63

- I've argued that mental qualities do not by themselves yield an explanatory gap, since we can give an informative theory of those qualities that's independent of their being conscious.
- But Leibniz had no such theory. So he had no way to think about mental qualities—whether conscious or not—except by reference to how we access those qualities in a distinctively first-person way when they are conscious.
- And thinking of them solely in that first-person way makes it hard to see how there can be any rational explanation.

Don't Mind the Gap

December 12, 2008

64

- Hence Leibniz applied the Mill Argument even to nonconscious perceiving.
- But a mechanistic explanation can forge rational ties if we have an informative account of mental qualities—such as homomorphism theory—which is independent of consciousness.
- Because seeing mental qualities as accessible only by way of consciousness precludes a mechanistic explanation, we need a *two-stage theory*, which combines a suitable higher-order theory of consciousness with *an independent theory of mental qualities*.

Don't Mind the Gap

December 12, 2008

65



Summary

- Kripke's modal argument fails because mental qualities are not intrinsically conscious.
- And we can forge an explanatory link between neural and qualitative states by way of a *two-stage theory*—first explaining mental qualities by appeal to states that enable one to discriminate among perceptible properties and then explaining qualitative states' sometimes being conscious by appeal to a suitable higher-order awareness.

Don't Mind the Gap

December 12, 2008

66