Phil 296D: Phil Mind Seminar W 21 Prof. Aaron Zimmerman
Office Hours: By Appointment Office: https://ucsb.zoom.us/j/5972101028

Handout #3: Nagel and Knowing "What It's Like" to Experience Something

Descartes' dualism is immoral: The Cambridge Declaration on Consciousness

1. The First Sentence of "What is it like to be a bat?"

"Consciousness is what makes the mind-body problem really intractable" (1979, 165).

Questions: What is the mind-body problem? Which, if any, of the problems we might describe as "the mind-body problem" is intractable? What is consciousness? What does Nagel mean by "consciousness" when he uses it in this first sentence? If Nagel means what Block, Chalmers and others have called "phenomenal consciousness" in contrast with "access consciousness," "self-consciousness," "creature consciousness" and so on, does Nagel's claim in this first sentence distinguish him from Descartes and those who think that discursive consciousness or thought is the main source of philosophical trouble for materialist or physicalist theory of the mind?

Nagel's 1st Thesis: The mind-body problem is intractable because we cannot provide a materialist theory or reductive physicalist theory of consciousness.

After admitting that other animals have consciousness, asserting that we cannot be sure about its presence in the "simpler" creatures," and hypothesizing that there are radically different forms of consciousness on planets in distant solar systems, Nagel defines "consciousness" for the reader.

Nagel's definition of consciousness: "The fact that an organism has conscious experience *at all* means, basically, that there is something it is like to *be* that organism…we may call this the subjective character of experience" (1979, 166).

He also writes of consciousness in the sense of their being something it is like "for" an organism.

Note some features of this definition: consciousness is a property of the experience of an organism. The idea seems to be that some organisms have experience and that some or all of these experiences have a subjective character. The subjective character of an experience is identical to what it is like for the organism enjoying that experience (or suffering that experience) while she is enjoying (or suffering) it.

The idea is compatible with greater and lesser amounts of consciousness if we can make sense of there being more or less involved in what it is like to be a being at a particular time.

"Clinicians assessing patients with disorders of consciousness assign a level of consciousness, with coma at one end of the scale, conscious wakefulness at the other, and various intermediate grades (such as deep sleep and light sleep) in between" (Birch et al, 2020, 790; citing S. Laureys (2005) "The neural correlate of (un)awareness: lessons from the vegetative state," *Trends Cogn. Sci.* 9, 556–559.)

Questions: Are there experiences that lack any subjective character? If not, why doesn't Nagel just define consciousness as experience? And is it wrong or at least misleading to write of some specific property of experience in virtue of which the organism has that experience or in virtue of which the experience of an organism has a subjective character? How does Nagel's 1st thesis fare if we interpret "consciousness" in the "multi-dimensional" manner suggested by Birch, Schnell and Clayton, which "highlights" five significant dimensions of variation: (i) perceptual richness; (ii) evaluative richness; (iii) integration at a time (or unity); (iv) unity across time (temporality) and (v) self-consciousness? Which, if any of these properties resists identification with or reduction to the mechanisms described by cognitive neuroscientists?

The problem of how to define consciousness in order to classify disorders of consciousness:

Bayne et al Bernat

2. Nagel's Second Thesis

Nagel's 2nd Thesis: Identifications of mental phenomena with physical phenomena are unlike other "inter-level" identifications such as those identifying macroscopic phenomena with microscopic phenomena.

"The mind body problem [is] unique and unlike the water-H2O problem or the Turing machine-IBM machine problem or the lightning-electrical discharge problem or the gene-DNA problem or the oak tree-hydrocarbon problem…It is most unlikely that any of these unrelated examples of successful reduction will shed light on the relation of mind to brain" (1979, 165-6).

3. Nagel's 3rd Thesis: "We have at present no conception of what an explanation of the physical nature of a mental phenomenon would be" (1979, 166). "Physicalism is a position we cannot understand because we do not at present have any conception of how it might be true" (1979, 176).

This is stronger than Nagel's first thesis in some respects and weaker in others. It's stronger because he doesn't just say the problem is intractable; he says we have "no conception" of how it even might be solved. It's weaker insofar as he is focused on the situation at the time in question.

"Perhaps a new theoretical form can be devised for the purpose, but such a solution, if it exists, lies in the distant intellectual future" (1979, 166).

4. **Nagel's 4th Thesis**: None of the biochemical or functional analyses of "the mental" are correct because the absence of consciousness is "logically compatible" with the presence of the phenomena to which it has been supposedly reduced.

Comments: This is a bad argument because the reductions are supposed to be a posteriori. We should only expect logical impossibility of the sort Nagel demands here when the reduction is definitional or analytic. Nagel's confusion here comes out in footnote 5 where he admits it might

be impossible for a robot to realize a full functional description of a mammal's mind without there being something it is like to be that robot. He says "That, if true, is a fact which cannot be discovered merely by analyzing the concept experience." Why should the success of the reduction be judged by the a priority or analyticity of the relationship between the functional characteristics highlighted by, e.g. Birch et al in their account, and their being something it is like to be the organism who manifests those characteristics? Why isn't the plausibility of this hypothesis or its superiority to competing theses along standard dimensions of epistemic evaluation, sufficient?

5. The Rorty-Kripke Observation

Nagel credits Rorty (1965) with the observation that mind or consciousness cannot be reduced to a physical substrate in the same way scientists reduce heat or color. We can say heat is molecular motion and that it causes feelings of warmth. We can then analyze heat without paying attention to the subjective character of our experiences of heat. We can do the same with color and other secondary properties by analyzing them as external causes of internal sensations. But if we analyze the sensations as neurological or functional things we have no place to "put" the subjective character in question?

Task: Explain, what, if anything, is wrong with an analysis which defines experience as a representation of some physical phenomenon (e.g. molecular motion or the reflection of light) and then defines the property so represented as a cause of the kind of experience in question. E.g. (a) experiences of red are visual representations (of a certain biochemical sort) of the reflectance of light along certain wavelengths; (b) surfaces are red when they reflect light so as to generate visual representations of them (of a certain biochemical sort).

6. Nagel's Initial Defense of Theses 1 and 3

"If physicalism is to be defended, the phenomenological features must themselves ne given a physical account. But when we examine their subjective character it seems that such a result is impossible. The reason is that every subjective phenomenon is essentially connected with a single point of view, and it seems inevitable that an objective theory will abandon that point of view" (1979, 167).

Response: If we identify sensations with brain states we have the beginnings of an explanation of the "essential connection" between the phenomenal character of these sensations and the point of view of the organism in whose nervous system these sensations are realized: only the organism has these sensations, and if she has something like human introspection, she will be able to identify, classify and report these sensations and no others, because her introspective faculties only give her access to these sensations and no others.

Questions: What about split-brain cases? Should we say, with Bayne and Chalmers, that the left hemisphere has experiences with the phenomenal character of both the left and right visual field but that she can only report on the right visual field? Or should we say that she only has access to the phenomenal character of the experiences she can report and utilize to complete the tasks assigned to her by the experimenters? In footnote 3 Nagel dismisses any essential connection

between self-awareness and phenomenality or subjective character, but his argument is hard to assess because it tendentiously assumes that self-awareness requires human (i.e. sentential) language.

Akins' Objection: Bats don't have a point of view but they have subjective experience. They don't have a point of view because echolocation gives them actionable information but it doesn't represent their prey as three-dimensional objects. It is unlike human vision in this respect. Nagel moves to quickly in equating the having of subjective experience with occupation of a determinate point of view on a world of objects of which you are one.

Birch et al on birds:

> Conscious experience in healthy adult humans is highly unified. You have a single perspective on the world and everything of which you are consciously aware is part of that perspective. All the experiences generated by your brain have a common subject. There are not two or more subjects housed within the same skull.
>
> Psychologists have long been fascinated by pathologies, such as the split-brain syndrome, in which this unity apparently breaks down [36,37]. Subjects who have had the corpus callosum wholly or partially severed sometimes display disunified behaviour when different stimuli are pre- sented to the two halves of the visual field. The same questions can be asked of non-human animals. Birds are particularly interesting in this respect because they are natural split-brains. They have no structure akin to the corpus callosum connecting the two hemispheres of the dorsal pallium, which is homologous to the cortex in mammals [40]. Could every bird be a pair of conscious subjects, intimately cooperating with each other? A similar debate arises with respect to the cerebral ganglia and brachial plexus (a nerve ring around the top of the arms) of the octopus [23,41–43]. These structures are connected, but they have some degree of functional autonomy from each other. Could an octopus have two, or even nine, conscious perspectives on the world? Current evidence does not settle these questions; our aim is only to raise them. (2020, 793).

7. Bats

Bats have experiences. Is it an additional step to say that these experiences have a subjective or phenomenal character? In any event, Nagel assumes that there is something it is like to be a bat. The problem he poses is how we can *know* what it is like to be a bat when bats use echolocation to navigate and we don't know what it is like to echolocate. All we can do is describe our experience and try to imagine how it would differ if we had the physiology of a bat and lived as a bat does.

Questions: *How is this epistemic problem related to the metaphysical "mind-body" problem with which Nagel began his essay?* Might the mind be the brain and might there be fully satisfying biochemical/functional analyses of the subjective character of human experience, even if we never know or even never can know what it is like to be a bat? Could we develop a fully satisfying physicalist analysis of the subjective character of bat experience without knowing what it is like to have that kind of experience? Suppose we can only *introspect* our own experiences and must use our *imaginations* to understand the experiences of other people and animals. Isn't it compatible with this that the experiences we introspect (along with our introspective awareness of them in ourselves and our imaginative appreciation of the experiences of others) can indeed be fully analyzed or described in bio-chemical/functional terms? Wouldn't these differences in the kinds of access we have to the two kinds of mental state in question explain why we can't be acquainted with echolocatory experience, can only dimly imagine what it's like to enjoy experience of this sort, and in this sense we cannot really know what it's like to

enjoy it, even though said experience is fully amenable to scientific analysis or bio-physical reduction?

Footnote 8: In footnote 8 on p. 172, Nagel says, "My point is…not that we canot *know* what it is like to be a bat. I am not raising that epistemological problem. My point is rather that even to form a conception of what it is like to be a bat (and a fortiori to know what it is like to be a bat) one must take up the bat's point of view." But that doesn't affect the objection. How is the difficulty of conceiving what it is like to be a bat, which seems essentially connected to imagining oneself as a bat, supposed to imply some problem for the adequacy of physical theories of the subjective character of echolocatory experience? Mightn't we be bad at imaging having an experience but great at identifying the nature of said experience?

Nagel's Response: The concept of the subjective character of experience or the concept of the particular qualitative character of a particular kind of experience is the concept of something that lacks an objective character.

"It is difficult to understand what could be meant by the objective character of an experience, apart from the particular point of view from which its subject apprehends it…But if experience does not have, in addition to its subjective character, an objective nature that can be apprehended from many different points of view, then how can it be supposed that a Martian investigating my brain might be observing physical processes which were my mental processes (as he might observe physical processes which were bolts of lightning), only from a different point of view?" (1979, 173-4).

See too "Does it really make sense, in other words, to ask what my experiences are really like, as opposed to how they appear to me?" (1979, 178).

Objection: This is question-begging and frankly inconsistent with the admission that the science of bat neurophysiology constrains what it's like to be a bat. Bat experiences must be conceived as having physical and spatial properties to coherently explain why the bat's physiology and behavior set limits on what it might be like for the bat to experience the world and its movements through that world.

Task: Let's look carefully at Akins, MacPherson and Birch et al here to see how objective or scientific investigation of the physiology and (more generally) biopsychology of animal perception might be thought to have determinate implications for what it's like to be the animals with the sensory and cognitive apparatuses so described.

Important here is the view emphasized by Akins that (a) we come to understand experience as representational as we examine its physiological substrate and the role of this substrate in the behavior or life of the animal, and (b) our introspective awareness of our experiences itself leads us to characterize these experiences in largely representational terms (e.g. as visual experience of the colored surface of an object).

Questions: What role does a representational account of experience play in efforts to bridge the explanatory gap that still exists between the cognitive neuroscience of subjective experience and

the phenomenological understanding of experience we gain through introspective reflection upon it? Is representationalism essential to the naturalization of the mind? Might it ultimately hinder progress of this project as radical pragmatists, enactivists and embodied mind theorists have suggested?

8. Extrapolation from Bat Physiology

As I said above, Nagel notes (p. 169) that we can extrapolate from the bat's "structure and behavior" some facts about what it must be like to be a bat. Akins focuses on this admission to pose a tension for Nagel. The problem is that Nagel admits that the physical facts (here facts about the physiology, behavior and environment of the bat) do in fact shed some light on what it is like to be a bat and act as data for acceptable hypotheses about how the bat experiences the world and its movements through the world. So if these facts and further knowledge of this sort *cannot* be adequate to the task, there must be some principled limit to its utility. But how can we know this a priori? Don't we first have to learn everything we can about the bat and only after exhausting this project and finding ourselves with remaining questions, conclude that we cannot know what it is like to be a bat? Perhaps if we had no idea what it must be like to be a bat and had no facts relevant to the determination of this issue, we could then conclude as Nagel does at the outset of his essay, that we can never know or that we can only know what it's like to be a bat by inventing an entirely new science. But once Nagel acknowledges the relevance of what we do know about the bat to his questions about the nature of bat experience, it is hard to see how he can still maintain the existence of certain "in principle" obstacles to adequate understanding.

How physiology impacts the definition of consciousness: Birch et al argue that even after we have distinguish the qualitative character of perceptual experience from other components of consciousness (e.g. self-consciousness or meta-consciosuness) we should still refrain from simply speaking and thinking of levels of consciousness so understood.

"Any measure of p-richness is specific to a sense modality, so we should not refer to a species' overall level of p-richness. … Within a given sense modality, it is possible to resolve p-richness into different components. For example, the richness of visual experience depends on bandwidth (the amount of visual content experienced at any given time), acuity (the number of just-noticeable differences to which the animal is sensitive), and categorisation power (the animal's capacity to sort perceptual properties into high-level categories). Does this make it impossible to develop overall evaluations of p-richness for conscious vision? Not necessarily. If one species outperforms another with respect to all three components, it has richer visual experiences overall. However, if the different components of p-richness are poorly correlated (e.g., because some species have low bandwidth and high acuity, or vice versa), we may decide that cross-species comparisons should use these finer-grained dimensions rather than p-richness" (Birch et al, 2020, 790-1).[1]

Questions: Does this sort of thinking show that a concept of phenomenal consciousness or "what it's like" to perceive or experience something can (and has) been incorporated into cognitive science without the kind of revolution Nagel thinks is needed? Don't the refinements Birch et al.

[1] The have similar things to say about evaluation and the experiential component of preferring one thing to another or experiencing something as good or bad. "Like p-richness, e-richness has more than one component. Rich affect-based decision making takes many inputs into account at once (evaluative bandwidth) and is sensitive to small differences in those inputs (evaluative acuity). If these components turn out to be poorly correlated, we may decide that cross-species comparisons need to use finer-grained dimensions" (Birch et al., 2020, 792).

propose to Nagel's undifferentiated notion of "the subjective character of experience" come from the "bottom up" or from the field rather than a priori from the philosophy seminar room? Does this demonstrate the pragmatists rejection of a principled analytic/synthetic distinction, where defintions are refined and recast as the facts come in from observation and experimentation? Or is Nagel right when he says that reductive approaches to consciousness merely change the subject?

9. Intrinsicality and Footnote 11 on Kripke

"If mental processes are indeed physical processes, then there is something it is like, intrinsically, to undergo certain physical processes" (175).

Question: How is "intrinsically" functioning in this argument? Suppose that we cannot have a good sense of what it is like to be a bat from a neurological description of the neurological processing that is crucial to its echolocatory hunting. Suppose that it is only once we know about the things it hunts and where it hunts them and we get an in-depth account along the lines provided by Akins of the different components of the perceptual processes involved that we begin to get a sense of how things must seem to the bat as it does what it does. Would this show that what it is like to echolocate is not an "intrinsic" character of the neural states of the echolocating bat? Mightn't phenomenal character be complex and relational?

This is also relevant to Nagel's endorsement of Kripke's argument that pain can't be identical to c-fibers firing because we don't identify pain via a description or mode of presentation. I.e. when a scientist identifies a pain via a brain scan she does identify it via a visual mode of presentation (i.e. by viewing the output of the scanner) but when we introspect on our pains the mode of presentation of the pain is just the pain itself. We can discuss whether this is enough to motivate dualism. (I don't think it is.) And it is curious why Nagel thinks Kripke's argument is a good one when Nagel gives an excellent debunking explanation of the dualistic intuitions in what remains of the footnote.

"When we try to imagine a mental state occurring without its associated brain state, we first sympathetically imagine the occurrence of the mental state: that is, we put ourselves into a state that resembles it mentally. At the same time, we attempt perceptually to imagine the nonoccurrence of the associated physical state, by putting ourselves into another state unconnected with the first: one resembling that which we would be in if we perceived the nonoccurrence of the physical state. Where the imagination of physical features is perceptual and the imagination of mental features is sympathetic, it appears to us that we can imagine any experience occurring without its associated brain state, and vice versa. The relation will appear contingent even if it is necessary, because of the independence of the disparate types of imagination" (176, fn11).

(See too PGS's discussion of Nagel's footnote 11, citing Daniel Stoljar.)

But I want to focus here on what Nagel initially says in reaction to Kripke's argument for dualism, "Like Kripke, I find the hypothesis that a certain brain state should necessarily have a certain subjective character incomprehensible without further explanation" (175, fn11).

Explanation: perhaps a neurological structure doesn't have a subjective character "intrinsically." Instead, when this state of the nervous system is operative in the way it is in the life of the animal whose brain it is, it will seem to that animal that such and such is going on in such a such a manner: i.e. that animal will have enjoy an experience with a certain phenomenal character. The problem here is thinking or writing as if phenomenal characters were simple, unanalyzable, intrinsic properties of brain states. (See Dennett on "Quining qualia" where he allows that we have experiences and that there is something it is like to enjoy one, but he rejects qualia if they're defined as intrinsic, simple, etc.)

10. Nagel's Berkeleyean Intuition

Nagel has Berkeley's intuition that "esse est percipi" when it comes to our minds or at least those states of mind that have a (reportable?) subjective character.

"Very little work has been done on the basic question (from which mention of the brain can be entirely omitted) whether any sense can be made of experiences' having an objective character at all. Does it make sense, in other words, to ask what my experiences are really like, as opposed to how they appear to me? We cannot genuinely understand the hypothesis that their nature is captured in a physical description unless we understand the more fundamental idea that they *have* an objective nature (or that objective processes can have a subjective nature)" (1979, 178).

Let's distinguish two things here:

(1) the absence of cases (or even the impossibility of cases) that are introspective analogs of hallucinatory experience. I.e there are cases in which it appears to me as if X but it is not the case that X (or it seems to me that there is an X but there is no X) because my perceptual experience of X is entirely generated by my mind/brain or some external stimulus that in no way resembles or "fits" my perceptual experience. E.g. I can have a visual experience of an apple in front of me when there is none because I am just hallucinating. And in a less extreme case I can have a visual experience of an apple in front of me when there is none because I am seeing a hologram or a two-dimensional drawing of an apple that looks real (etc.) In contrast, it is hard or impossible to describe a case in which it seems to me introspectively that I seem to see an apple or it seems to me introspectively that I am feeling pain, but there is "introspective seeming" is false or inaccurate because I don't seem to see an apple or I am not really in pain. (We need to be subtle here because we surely can mischaracterize or misdescribe the subjective character of our experience.)

But this phenomenon, whatever metaphysical conclusions might be drawn from it, does not entail the conclusion Nagel draws from it: i.e. that our experiences don't have an objective character.

(2) Even if we don't have or even can't have inaccurate or hallucinatory experiences of our experiences, still, if our experiences have functional and material (e.g. spatial) properties (by being in our heads, taking time, playing functional roles, etc.) then they have objective characteristics about which we can be mistaken. The claim that our experiences have no

properties beyond those they seem to have when we reflect upon them introspectively begs the question against the materialist or physicalist and it does not follow from the absence of (or even impossibility of) hallucinatory introspective seemings. Introspection may be incomplete even if it is unlike perception in the manner Nagel here describes.

11. Objective Phenomenology

As Nagel explains in his footnote 11, our understanding of other minds seems to involve simulation or empathy or the "sympathetic imagination." We understand what another person or animal is experiencing by imagining ourselves in his or her situation as best we can and generating a simulacrum of the experiences we (judge that we) would have were we in her situation as we construe it. (Adam Smith's moral psychology is based on this idea.). At the end of "What is it like…" Nagel suggests that the conceptual revolution he has claimed is necessary for an understanding of the subjective character of experiences might be accomplished by constructing a science of experience that eschews imaginative projection or sympathetic imagination for an "objective phenomenology."

"Structural features of perception might be more accessible to objective description, even though something would be left out. And concepts alternative to those we learn in the first person may enable us to arrive at a kind of understanding even of our own experience which is denied us by the very ease of description and lack of distance that subjective concepts afford" (1979, 179).

Question: How does Nagel's idea of an "objective phenomenology" compare to Dennett's idea of a "heterophenomenology" and the more general idea, pursued by Akins and Macpherson, or investigating the physical structure of our sensory and cognitive apparatus and using this to explain and correct our naïve introspective characterizations of what we are experiencing? Consider, for example, the naïve characterizations we give of the unity of our experiences when we compare experience to a "stream of consciousness" or a "train of thoughts" (both phrases are used by Bain and then, following Bain, James). Can we use studies of the cognitive results of brain bisection to correct these reports?

Or consider Aristotle's view that we have exactly five senses, which is the launching point for MacPherson's essay. Did Aristotle ignore the introspective data that might lead one to posit proprioceptive, and somatic senses? (See what MacPherson says about Plato's view that temperature and pain (or tissue damage) are perceived through faculties distinct from touch—a view he reached on introspective grounds rather than examination of human physiology.)

Question: Once we chart the neural substrate of proprioception, somatic awareness, nociception, equilibrioception, and vomeronasal perception (as described by Macpherson),[2] does that provide us with a more objective account of the phenomenal character of the experiences generated by or realized in these substrates?

[2] "Other candidates that have been considered as being additional human senses include senses of hunger, thirst, wet and dry, the weight of objects, fullness of the ladder, suffocation and respiration, sexual appetite, and lactiferousness" (MacPherson, 2011, 126; citing Dallenbach, K. M. (1939). Pain: History and present status. American Journal of Psychology, 52(3), 331–347.)

And how might this physiology-informed investigation of human sensory experience enable us to better answer Nagel's questions about the phenomenal character of echolocatory experience or the experience of the magnetoreceptive perceptions of trout or the infrared perceptions of pit vipers?[3] Look at MacPherson's illustration of the UV light pattern reflected by flowers and the "nectar guide" this provides for bees (2011, 135). Doesn't this co-evolved phenomena shed light (so to speak) on the bee's experience of its "lived world" and how this differs from our experience of objects in that world? More generally, consider how MacPherson's approach to individuating the senses undermines the claim made by Nagel and other defenders of the "explanatory gap" that psychology requires a revolution to account for consciousness (see Nagel's 3rd thesis above). It is not that we have the sciences of the senses on the one hand (e.g. physiology, ecology, etc) and then phenomenology or the study of the conscious or introspectable features of sensory experience on the other. Instead, MacPherson shows how we will probably have to use the phenomenal and representational character of experience to individuate the senses in the first place. See here her questions about whether bees detect UV light with vision and pit vipers see infrared (2011, 134-8). Even questions like "What is an eye?" and "What is a neuron?" depend on a host of criteria and it is far from clear that phenomenological considerations play no role in developing and adjudicating between different answers. Indeed, MacPherson gives them a central (albeit partial) role in her answers to these questions. The science of the mind is an interdisciplinary science, as there is no prospect of investigating (eg) the neurological substrate of various forms of perception *independently* from a study of the ecological features and behavioral repertoires of the animals who possess these nervous systems.

Questions: Don't we need a methodology that integrates phenomenology with the rest of the cognitive sciences? Is Nagel wrong to think that we would need a conceptual revolution to accomplish this end?

[3] MacPherson cites Hughes, H. C. (1999). Sensory exotica. Cambridge, MA: MIT Press.