

Handout #4: Panpsychism

1. Nagel's Epistemological and Metaphysical Commitments

Recall the central theses of Nagel's "What is it like to be a bat?"

Nagel's Epistemological Thesis: We do not have a cogent understanding of how we might explain the genesis and nature of phenomenally conscious states of mind. We cannot explain the genesis and nature of phenomenal consciousness using the methods and theories currently employed by cognitive scientists. A satisfying theory of phenomenal conscious must await a conceptual revolution.

Nagel's Metaphysical Thesis: The subjective character or phenomenal quality of experience is not identical with nor reducible to the "physical" (e.g. biological, biochemical, or functional) properties of nervous systems or control systems.

Central Questions: Has Nagel adequately supported these theses? Nagel's epistemological thesis is wholly negative. What kind of positive metaphysical conception of the relation between mental and physical can he articulate to replace the physicalist metaphysics that he rejects?

2. The Argument for Panpsychism

It becomes clear at the outset of "Panpsychism" that Nagel rejects substance dualism but embraces a form of property dualism in its wake.

(1) Property Dualism or Non-Reductive Realism: The subjective qualities of experience or its "phenomenal properties" are non-physical properties of organisms.

"Ordinary mental states like thought, feeling, emotion, sensation or desire are not physical properties of the organism—behavioral, physiological or otherwise—and they are not implied by physical properties alone. Nevertheless they are properties of the organism, since there is no soul, and they are not properties of nothing at all" (1979, 181-2).

Questions: *What function is "implied" performing here?* Is Nagel assuming that an adequate or satisfying explanation of the genesis and nature of phenomenal consciousness would have to *entail* its existence and introspectible character? Is this fair? Does it reveal his commitment to an outdated conception of the nature of science in general? Is this mistaken understanding of scientific understanding responsible for Nagel's property dualism? Consider PGS's evolutionary hypotheses about the evolution of sensory experience and evaluative experience. Does the evidence he utilizes "entail" the existence of sensation or desire in the animals under review? If the relationship between his evidence and his hypotheses (or the premises of his reasoning on the issue and the conclusions he draws from them) is instead "abductive" so that his conclusions are advanced as the most plausible of the coherent conceptualizations of subjectivity that have been proposed, does that justify Nagel in asserting his skeptical epistemological thesis regarding the

prospects of the cognitive sciences in their current form or his metaphysical theses asserting the irreducibility of the mental? (Is property dualism even coherent? Is it a relevant alternative to existing formulations of physicalism or materialism or naturalistic monism?)

Cf. Nagel's claim, "What is the reason to deny that mental properties can be entailed by physical ones?" (What is a "property entailment"? Entailment is a relation between sentences or propositions, right?)

Perhaps property dualism can be reconciled with contemporary cognitive science if we understand these mental properties as in some sense "emergent" phenomena. But Nagel rejects emergence on the road to panpsychism.

(2) Nonemergence: "All properties of a complex system that are not relations between it and something else derive from the properties of its constituents and their effects on each other so combined ... [If we cannot derive] an observed feature of a system...from the properties currently attributed to its constituents...this is [sufficient?] reason to conclude that either the system has further constituents of which we are not yet aware, or the constituents of which we are aware have further properties that we have not yet discovered" (1979, 182)

Question: *What does "derive" mean in this characterization of the rejection of emergent phenomena?* Again, what if no derivation of phenomenality is forthcoming but the "properties attributed to a system's constituents" make it plausible that the system would have an observable property or phenotype? Suppose we can see how, e.g., synaptic strengthening and growth might manifest to a subject or those observing her behavior over time as "memory," even though there is no prospect of anything like a "formal derivation" of the existence of memory from observations of synaptic growth and its biological consequences? I am here thinking of a "formal derivation" as the sort of argument offered in logic and math of conclusions from premises or theorems from axioms utilizing transitions that conform to a privileged set of inference rules. Is this the conception of derivation Nagel is working with here? How does it relate to the looser notion of an "explanatory gap" featured in the work of Joe Levine and Frank Jackson, where that phrase is taken to denote a real explanatory need that PGS is trying to fill? (What is an "explanatory need"? Is this a purely psychological concept? Is adopting a pragmatic conception of "genuine explanatory needs" a necessary first step to abandoning property dualism and embracing the emergence of the mental?)

Notice that Nagel's rejection of emergence still leaves open (as a live possibility for him) a "relational monism" of the sort embraced by Russell which has parallels with the relationalist conception of perceptual experience advanced by Michael Martin. Martin and those sympathetic to his view of perception reject the idea of a "perfect hallucination" which is phenomenologically indistinguishable from the visual experience enjoyed by a subject when she is perceiving veridically. Their idea is that the phenomenal character or subjectivity quality of perceptual experience is itself relational and is constituted by (or emerges from) interactions between the perceiving organism and the phenomena she perceives.

(3) Material Composition: "Any living organism, including a human being, is a complex material system. It consists of a huge number of particles combined in a special way... Anything

whatever, if broken down far enough and rearranged, could be incorporated into a living organism. No constituents besides matter are needed” (1979, 181).

(4) **Panpsychism:** “The basic physical constituents of the universe have [or the universe itself has] mental properties, whether or not they [i.e. these constituents] are parts of living organisms” (1979, 181).

Nagel’s Argument for Panpsychism: (1)+(2)+(3)=(4).

“If the mental properties of an organism are not implied by any physical properties but must derive from properties of an organism’s constituents, then those constituents must have nonphysical properties from which the appearance of mental properties follows when the combination is of the right kind. Since any matter can compose an organism, all matter must have these properties. And since the same matter can be made into different types of organisms with different types of mental life (of which we have encountered only a tiny sample), it must have properties that imply the appearance of different mental phenomena when the matter is combined in different ways. This would amount to a kind of mental chemistry” (1979, 182).

3. The Interpretation of Panpsychism

A. What is meant by “physical”?

[Newly posited] properties are counted as physical if they are discovered by **explanatory inference** from those already in the class. This repeated process starts from a base of familiar, observable, spatio-temporal phenomena and proceeds to take in mass, force, kinetic energy, charge, valence, gravitational and electromagnetic fields, quantum states, anti-particles, strangeness, charm, and whatever physics will bring us next. What the argument claims is that a similar chain of explanatory inference beginning with familiar mental phenomena [i.e. attempts to explain the subjective character of experiences] would lead to general properties of matter that would not be reached along the path of explanatory inference by which physics is extended” (1979, 183).

Questions: *What is meant by “explanatory inference”?* How does this relate to the conceptions of “derivation” and “entailment” that Nagel utilized to characterize the rejection of emergent phenomena and postulate the irreducibly mental character of experience?

Note how Nagel differentiates panpsychism from physicalism in epistemological terms: physicalism is vindicated if we start with observations of objects (e.g. animals and their brains) and end up with a satisfactory theory of phenomenal consciousness even if to explain the actions of animals and their internal processes we need to posit some new fundamental property beyond those we utilize to explain the behavior of inorganic or inanimate phenomena. Panpsychism requires that scientists (or thinking people more generally) enlarge the set of fundamental properties or forces they recognize as a result of their attempts to explain the genesis and nature of those qualitative aspects of experience we identify through introspection.

The difference between panpsychism and dualism rests in the dualistic intuition (not shared by the panpsychist): “the uneasiness that one may feel about the suggestion that mental phenomena should derive from any properties of matter at all” (1979, 183).

Nagel goes on to talk about what it is “legitimate” to infer, which is an epistemic notion, and has interesting points of contrast with “derivation” and “explanatory” inferences. Of course, these ideas are connected, especially if we assume that whenever it is “legitimate” to infer something from our observations: (a) the conclusion we draw explains our observations and (b) allows us to construct a derivation of the observations from the confirmed hypothesis, and (c) therein allows us to construct an informal (non-deductive) “derivation” of the hypothesis from the observations. In the quote, Nagel combines the ideas of explanation and derivation as follows:

“It will never be legitimate to infer, as a theoretical explanation of physical phenomena alone, a property that includes or implies the consciousness of its subject” (1979, 183).

Questions: Why does Nagel think this? Is it because we cannot rule out zombies or because the premises of all legitimate inferences to conclusions about the existence and nature of the mental lives of animals must contain some reference to introspective claims about irreducibly mental properties? Nagel has the latter sort of thing in mind when saying our inferences about other minds “employ concepts understood independently and not introduced through physical theory.” *But don’t the conclusions we draw always involve such (extra-scientific) concepts, even when we are engaging in theoretical physics? After all, it isn’t as though physicists can bracket background logic and background assumptions that go into their basic mastery of natural language when interpreting the results of their experiments.* Which, if either, of Nagel’s conditions on “legitimate” inference to conclusions asserting the existence or character of the mental or phenomenal applies to the inferences PGS makes in drawing his conclusions about the evolution of subjectivity? Which applies to Atkins and Macpherson?

Neutral monism: If sensory observations of matter and introspection on the subjective character of experience each leads us to posit the same set of fundamental properties and forces, we would have to say that the mental and physical are products of some third kind of thing.

Nagel’s a priori rejection of neutral monism: We cannot legitimately infer conclusions about the mind or the subjective character of experience from sensory observations of matter unless we build in mental assumptions or psychophysical principles. But if we make mentalistic assumptions, we are no longer engaging in physics. So neutral monism is impossible.

Task: Think about what PGS is doing or what Birch et al were doing or what cognitive psychologists and animal ethologists do regularly when explaining animal behavior in mental terms and evaluating hypotheses about the conscious perceptions of animals. This isn’t physics, but it *is* science. (Isn’t it?) But it does not involve a “mental chemistry” of the sort Nagel envisions in this essay. Does Nagel’s argument against neutral monism or his argument for panpsychism rely on an overly restrictive view of the physical sciences? Can we do biology without assuming and explaining memories of various sorts? Can we do biology without mentioning consciousness? **Are the cognitive sciences necessarily oriented toward the explanation of non-physical things once we’ve adopted Nagel’s conception of the physical,**

defined, as it is, by inference from sensory observation stripped of all mentalistic assumptions? He seems to accept this consequence on p. 185; citing his articles, “Linguistic and Epistemology” and “Freud’s Anthropomorphism.” **I think this consequence is close to a reductio of Nagel’s conception of the physical.** Shouldn’t he instead try to define the physical in metaphysical terms, rather than defining it, as Nagel here does as whatever we posit in reasoning from our observations of matter or objects we perceive in space?

Nagel’s definition of physicalism: the physical properties and phenomena are those posited in the course of reasoning from sensory observations in a legitimate way while bracketing all psychological or psychophysical assumptions you would otherwise make.

Complaint: This is a bad definition of “physicalism.” See Daniel Stoljar’s work on this subject.

B. Nagel’s Non-reductive Causal Realism

“There will probably be many uniform psych-physical correlations of the form; ‘Whenever an organism is in exactly physical state P it is also in mental state M.’ This may be true of my present total physical and mental states, for example. No doubt more general correlations also exist. On a correlation view that should be enough for M to be causally explained by P. But it is not enough on a stronger view of causation” (1979, 187).

Critique: Even if we adopt an interventionist account of causation of the sort defended by James Woodward at Cal tech. There will be psycho-physical causation and the reverse. Nagel needs to do more than reject a correlationist or “Humean” conception of causation to reject mental causation and causation of sensations and their qualitative character. He needs a strong form of non-reductive causal realism. Instead he returns to the claims he makes in “What it is like to be a bat?” that we simply do not understand how the physical information processing we observe going on in the bat’s nervous system might cause that bat to have experiences with a certain subjective character. This failure to understand is so dire, we should conclude with the panpsychist that there must be some mental character involved in the physical; constituents, of which we are not yet aware, which in combination in the bat’s nervous system, provides it with qualitative experiences and so explains why there is something it’s like to be that bat.

In rehearsing the argument he again defines a metaphysical concept in epistemological and psychological terms. That is, we are told what a qualitative property of experience is or what the subjective character of experience is where this is supposed to be a real, irreducible feature of our experiences. But qualia are defined in terms of a form of understanding that can only be gained through introspection. **Nagel is able to do this by writing of our “understanding of features.”** What we are said to “understand” is not claims nor sentences nor utterances nor diagrams nor representations of any sort. Instead we understand features of our experience and these features can only be understood from one type of point of view. Nagel seems to have in mind the human point of view and the other points of view demarcated by the other species of animal (or animals with the sophistication necessary to be attributed experiences with qualitative character). But this seems arbitrary.

The basic questions: The physicalist can admit that introspection is different from inference from sensory observation. The physicalist can admit that empathetic or projective imagination based (as it is) on experience and perhaps introspection upon experience, differs from abduction from behavior. Why must we go further to define a certain kind or property or phenomena that by definition can only be known directly through introspection or indirectly through empathetic imagination? Can't we have two different ways of knowing without two different kinds of thing to know? If, as I've suggested above, we must combine, introspection and imaginative projection with neurophysiology and other modes of explanation to arrive at a satisfactory account of the mind and the forms of qualitative, consciousness experience enjoyed by minds, would that imply, given Nagel's definitions of the physical and mental in terms of our ways of knowing about them, that we have indeed come to embrace a form of neutral monism; that we have in fact transcended the dichotomy between the physical and mental as urged upon us by pragmatist philosophers of science?