

Précis of *The Conscious Mind**

DAVID J. CHALMERS

University of Arizona

Chapter 1: *Two Concepts of Mind*. I distinguish the phenomenal and psychological (functional) concepts of mind. I argue that every mental state is a phenomenal state, a psychological state, or a hybrid of the two. I discuss the two mind-body problems corresponding to the two concepts of mind, and discuss the various senses of the term “consciousness”.

Chapter 2: *Supervenience and Explanation*. I distinguish varieties of supervenience, especially logical and natural supervenience, where supervening properties covary with base properties across either logically possible (conceivable) or naturally possible worlds. (One can also define a distinct notion of metaphysical supervenience, if one believes that the classes of logically possible and metaphysically possible worlds are distinct; I hold that the two classes coincide.) *Materialism* is the thesis that all positive properties globally metaphysically supervene on microphysical properties. A property is *reductively explainable* in terms of physical properties if it globally logically supervenes on those properties.

I defend the notion of conceptual truth against various objections, and outline a two-dimensional intensional framework for handling *a posteriori* necessity. Every concept has two intensions, a primary intension and a secondary intension. The primary intension delivers a concept's referent in a centered world when the world is considered as actual (i.e., considered as an epistemic possibility); the secondary intension delivers a concept's referent in a world when it is considered as counterfactual. (The primary intension of “water” picks out roughly the “watery stuff” in a world; the secondary intension picks out H₂O.) A statement is *a priori* when it has a necessary primary intension; a statement is necessary when it has a necessary secondary intension. *A posteriori* necessities (such as “water is H₂O”) have a necessary secondary intension but a contingent primary intension. A statement is conceivable (or logically possible) when its primary intension is true in some world; a statement is possible (metaphysically possible) when its secondary intension

* David J. Chalmers, *The Conscious Mind: In Search of a Fundamental Theory* (New York: Oxford University Press, 1996).

is true in some world. So the Kripkean gap between conceivability and possibility is explained at the level of statements, without appealing to a distinction between conceivable and possible worlds. The class of worlds in question is always that of the ideally conceivable (or logically possible) worlds.

I use arguments from conceivability, epistemology, and analysis to argue that *almost* all properties logically supervene on the physical, except perhaps for phenomenal properties and other properties depending on them.

Chapter 3: *Can Consciousness be Reductively Explained?* I use arguments from conceivability, epistemology, and analysis to argue that phenomenal properties do not logically supervene on (are not *a priori* entailed by) physical properties. These include arguments from the conceivability of zombies and inverted spectra, arguments from incomplete knowledge and from epistemic asymmetry, and arguments from the absence of structural/functional analysis. It follows that there is no reductive explanation of consciousness. I apply this to potential explanations in neuroscience, cognitive science, and physics.

Chapter 4: *Naturalistic Dualism*. I argue that phenomenal properties are not necessitated by physical properties, so that materialism is false. The arguments of Chapter 3 are supplemented by an argument that there is no *a posteriori* necessary entailment from physical to phenomenal. I use the two-dimensional analysis of *a posteriori* necessity to argue that no such necessity can save materialism if *a priori* entailment fails. And I argue that there is no reason to believe in “strong” metaphysical necessities that escape the two-dimensional account. I also discuss and argue against the possibility that there is an *a priori* connection that we cannot grasp due to cognitive limitations.

I compare this argument to related arguments for dualism and discuss the position that results. Materialist ontology must be expanded by new fundamental properties (phenomenal or protophenomenal properties) and new fundamental psychophysical laws. The resulting position may be epiphenomenalism, interactionism (although I argue against this), or a Russellian monism on which (proto)phenomenal properties serve as the categorical grounds of basic physical dispositions. I sketch the logical geography of the issues and defend my “naturalistic dualism” from objections.

Chapter 5: *The Paradox of Phenomenal Judgment*. On my position, even if consciousness cannot be physically explained, behavior and functioning can be. So it seems that consciousness is explanatorily (although perhaps not causally) irrelevant to behavior. In particular it is explanatorily irrelevant to claims such as “I am conscious” and related phenomenal judgments (where judgments are defined in functional terms). I call this the “paradox of phenomenal judgment”. I argue that this paradox is counterintuitive but yields no fatal flaws. I address the objections that it implies that we are unable to know about, refer to, or remember our phenomenal states. I argue that these objec-

tions rest on causal theories of knowledge and of reference that we have independent reason to reject in the phenomenal case. Knowledge of and reference to phenomenal states is based on something tighter than a causal relation; it is based on a relation of acquaintance. I discuss the content of phenomenal beliefs and the constitutive relation between experience and phenomenal belief.

Chapter 6: *On the Coherence between Consciousness and Cognition*. I outline the program of searching for a theory of consciousness in terms of psychophysical laws: first nonfundamental, then fundamental laws. Data is hard to come by, but first-person data plus plausibility assumptions in the third-person case can do much work. I discuss some candidates for nonfundamental psychophysical laws, involving the “coherence” between consciousness and “awareness”, a functional state defined in terms of direct availability for global control. I argue that consciousness and awareness covary, as do their structures. I apply this to some methodological questions about the empirical study of consciousness.

Chapter 7: *Absent Qualia, Fading Qualia, Dancing Qualia*. I argue for a principle of organizational invariance: systems with the same fine-grained functional organization will have the same sort of phenomenal states (of natural necessity). I argue against the empirical possibility of “absent” and “inverted” qualia, using thought-experiments involving replacement of neurons by silicon chips. If absent qualia are possible, then “fading” qualia are possible, and if absent or inverted qualia are possible, then “dancing” qualia are possible; but I argue that it is very implausible that fading or dancing qualia are possible.

Chapter 8: *Consciousness and Information: Some Speculation*. I speculate on the form of a fundamental theory of consciousness, suggesting that it may involve a dual-aspect view of information (information realized both physically and phenomenally). I argue that this would cohere well with the other principles, and with the relation between consciousness and phenomenal claims. One version of this view implies a sort of panpsychism; this need not follow, but I argue that it is not as implausible as is often supposed. I suggest that such a metaphysics of information may fit well with the Russellian monism discussed earlier, and discuss some open questions for such a theory.

Chapter 9: *Strong Artificial Intelligence*. I apply previous conclusions to argue that strong artificial intelligence is true: there is a class of programs such that any implementation of a program in that class is conscious. I give an account of computation and implementation to support this conclusion, and I argue against Searle’s “Chinese room” arguments and other objections to strong artificial intelligence.

Chapter 10: *The Interpretation of Quantum Mechanics*. I discuss some relations between a theory of consciousness and the interpretation of quantum

mechanics. After discussing the measurement problem and various interpretations, I defend a version of the Everett interpretation on which wavefunctions never collapse. A major question for this view is why, given that the physical world is superposed, we experience it as discrete. I suggest that this is a question for a theory of consciousness, and argue that the principles I have developed *predict* that there will be discrete experiences in a superposed world. I defend the Everett interpretation from some objections, although I am ultimately agnostic about its truth.