Consciousness and Its Implications

Professor Daniel N. Robinson

THE TEACHING COMPANY ®

Daniel N. Robinson, Ph.D.

Philosophy Faculty, Oxford University
Distinguished Professor, Emeritus, Georgetown University

Daniel N. Robinson is a member of the philosophy faculty at Oxford University, where he has lectured annually since 1991. He is also Distinguished Professor,
Emeritus at Georgetown University, on whose faculty he served for 30 years. He was formerly Adjunct Professor of Psychology at Columbia University.

Professor Robinson earned his Ph.D. in neuropsychology from City University of New York. Prior to taking his position at Georgetown, he held positions at Amherst College, Princeton University, and Columbia University.

Professor Robinson is past president of two divisions of the American Psychological Association: the Division of History of Psychology and the Division of Theoretical and Philosophical Psychology.

Professor Robinson's publications cover an unusually wide range of disciplines, including law, philosophy of mind, brain sciences, psychology, moral philosophy, American history, and ancient history. He is former editor of the *Journal of Theoretical and Philosophical Psychology*. He is also the author or editor of more than 40 books, including *Praise and Blame: Moral Realism and Its Application*, *Wild Beasts & Idle Humours: The Insanity Defense from Antiquity to the Present*, *An Intellectual History of Psychology*, *The Mind: An Oxford Reader*, and *Aristotle's Psychology*.

In 2001, Professor Robinson received the Lifetime Achievement Award from the Division of History of Psychology of the American Psychological Association and the Distinguished Contribution Award from the Division of Theoretical and Philosophical Psychology of the American Psychological Association.

Table of Contents
Consciousness and Its Implications

Consciousness and Its Implications

Scope:

The subject of consciousness is among the most vexing in both philosophy and science, and no less tractable in psychology, where the conceptual problems are often neglected. As a "state," consciousness seems resistant to translation into physical terms and measurements, though its dependence on a healthy nervous system appears to be as close to a "cause-effect" relationship as any in the natural sciences.

The aim and scope of these 12 lectures must be modest, for the subject is as vast as that of human and animal awareness. What I hope to convey may be distilled into four main points: First, that consciousness and mental life are *sui generis*; they are not "like" anything else. They are not like anything that is material or physical and seem to require for their fuller understanding a science not yet available, if ever available. Second, what distinguishes consciousness (and the term presupposes consciousness *of* something) from all else is its *phenomenology*—there is something it is like to be "conscious" that is different from all other facts of nature. Third, conscious awareness is a power possessed by the normal percipient, including non-human percipients. This power is such that much that impinges on the sense organs is filtered out and sometimes only the weakest but the most "meaningful" of occurrences gains entrance. Fourth, such powers vary over the course of a lifetime, are subject to disease and defect, and thus, lead to questions of profound ethical consequence.

Here, then, is a topic in which science, philosophy, medicine, and ethics are merged, the result being issues at once intriguing and unsettling.

Lecture One
Zombies

Scope: In this course, we will attempt to unravel the nature of consciousness, its provenance, and its function. We begin with an examination of the concept of the zombie, which functions effectively as a physical entity *without* consciousness. If a system can solve problems and process information without consciousness, of what value is consciousness? The question of ethics is raised if we consider that entities without consciousness cannot be judged for their actions. Could such an entity strive for moral improvement? The subject of consciousness is vast and varied and, as a philosophical problem, far from an easy solution.

Outline

I. Our core questions in this course on consciousness are: What is it? How does it come about? What is it for?

A. Popular speech is rife with references to consciousness. We talk about being "half conscious" or "unconscious" of something; the act of "daydreaming" reflects by contrast on a vividly conscious life; the patient in the emergency room is suffering a "loss of consciousness."

B. Zombies, the "walking dead," accomplish what they do without consciousness.

1. Philosophical zombies are different from the Hollywood version.

2. They are created to test certain notions we have about the essence of mental life and the properties that life must have to qualify as "consciously" lived.

C. Some years ago, Güven Güzeldere summarized various ways of configuring such entities and then understanding their nature.

1. One might make a device that is indistinguishable from conscious human beings in the way it behaves, though its internal machinery would be nothing like our own (a *behavioral* zombie).

2. A better "fit" than the behavioral zombie is the *functional* zombie, which does and says what we do and say; its underlying systems function as ours do but do not include anything by way of consciousness, let alone self-consciousness.

3. The third kind of zombie, the *identical* zombie, has an anatomy fully identical with that of a human.

D. These three types of zombie capture the various ways philosophers have attempted to dissolve the seeming mystery of consciousness.

1. One solution to the problem of consciousness is behavioristic: X is properly regarded as "conscious" to the extent that its behavior is relevantly like that of anything that is regarded as being conscious.

2. Other philosophers might use more stringent criteria: Not only must there be behavioral similarities of the right sort, but these must come about in the right sort of way.

3. The behavior must express underlying physiological processes of just the sort that underlie our own actions and speech.

4. To the extent that the device functions the way we do, we are permitted to regard it as conscious in the relevant sense.

5. But this entity nonetheless has no consciousness as we understand that state; physical foundations are unable to account for the consciousness itself.

E. If not physical properties, what other properties are there? Consider a system entirely physical but nonetheless conscious. How does a system entirely and solely physical come to have this defining mental property?

F. Some philosophers reject the very idea of zombies in any terms that would settle issues in philosophy of mind.

1. Nigel Thomas, in his essay entitled "Zombie Killer," argued that the very foundational premise on which zombie examples are constructed is defective.

2. If we accept what is now widely endorsed by scientists and philosophers—that our own conscious experiences are nothing more than the result of our own brain and bodily functions—then zombies are, indeed, problematical.
3. But if zombies are a conceptual possibility, then functionalism must be false.

G. In his essay **"**Does Consciousness Exist?" William James (1842–1910) argued that the function of "knowing," if not explained by "consciousness," must still be explained by *some* concept.

II. We recognize the difference between the mere registration of an event and our knowledge of it: our *knowing* it—which raises the question of the value of direct awareness.

A. Direct awareness is not a requirement for learning or memory. A computer's memory can take on whole megabytes of new information, while being unaware of the function it serves.

B. So what is consciousness good for? This begs the question of whether the best account of those defining features of human nature are to be understood within an essentially evolutionary context.

C. The more refined conceptions of evolutionary influences do not require that every property be useful.
1. Consciousness, it might be argued, is an epiphenomenon of evolution—like an architectural spandrel, it is not functional.
2. The fully adapted being is akin to the zombie, but, owing to the formation of that being, consciousness appears as a byproduct.
3. Thus, consciousness may be regarded as a product of evolution, but with no embarrassment to the theory of evolution, just in case consciousness has no real function.

D. To be a zombie may be like sleepwalking.
1. To be unaware implies having no conception of consciousness and self-consciousness in others.
2. What form of interaction might be possible among such entities? Could there be crime and punishment, moral improvement, an aesthetic dimension to life, and so on?
3. None of us would claim to be willing to live this sort of life because it is not as rich a life as ours.

E. What of such "dissociative" disorders as "multiple personalities" that might have someone conscious of being someone else, or of stages of infancy at which there is consciousness but no basis for its personalization, or sleep, or dream-states, or of *autoscopy* and near-death experiences?

III. The word *consciousness* did not take on its current meaning in English until late in the 17th century. John Locke's *An Essay Concerning Human Understanding* appeared in 1690 and made use of the term in a new way.

A. Prior to Locke's interpretation of the term, Thomas Hobbes said that what is "conscious" is just what is understood in common.

B. In Locke's work, consciousness takes on its private character, its contents found through introspection, by a mind able to examine its own content and that of no other.

C. The word is problematical, and matters seem to become even more unsettled when we add a "self" to it.

Essential Reading:

Güzeldere, Güven, "Three Ways of Being a Zombie." Presented at the University of Arizona conference *Toward a Science of Consciousness*, April 8–13, 1996, Tucson, Arizona.

James, William, "Does Consciousness Exist?" *Journal of Philosophy, Psychology, and Scientific Methods* 1 (1904): 477–491.

Supplementary Reading:

Hobbes, Thomas, *Leviathan* (1651), book 1, chapter 7, p. 31.

Locke, John, *An Essay Concerning Human Understanding,* book 2, chapter 1, section 19, p. 115.

Questions to Consider:

1. Since zombies seem to be able to do so much without consciousness, what might be the effect of consciousness should they possess it while doing just these things?
2. Can there really be zombies? Can there be a "caretaker" zombie who is nonetheless "unconscious"?

Lecture Two
Self-Consciousness

Scope: In this lecture, we consider the proposition that conscious life is grounded in the real essence of mind and, as such, is somehow insulated from the changes that might otherwise be brought about by mere "matter in motion." If the constituents of our own bodies continuously change, can we still retain an identifiable "self"? We again turn to the British empiricist John Locke (1632–1704) for his contribution to the issue.

Outline

I. What is the relationship between consciousness and knowledge?

- **A.** It is possible to imagine a zombie called George that responds to his name.
  - **1.** What is unimaginable is his *knowing himself* to be George, because, presumably, a zombie does not "know" anything. To know something is to be conscious of it and zombies are not conscious.
  - **2.** Consequently, they are not self-conscious.
- **B.** The claim "I am conscious of a rabbit in the garden" is different from the claim "I know there is a rabbit in the garden."
- **C.** There is a difference between being *conscious of* and being *conscious*.
  - **1.** The former is always subject to error.
  - **2.** Being conscious or aware is to be the possible subject of an experience, i.e., the *self*.
  - **3.** In the older phenomenological literature, there is the common assumption that consciousness invariably includes a "self."
  - **4.** For there to be knowledge, motives, desires, and beliefs, there must be consciousness, but these features of mental life are often in operation without the actor reflecting on all of them.

II. With *reflective* consciousness, the focus is on *self* as the subject or source, but who or what is "self"?

- **A.** We could argue that an old ship, *Old Faithful*, that has been extensively rebuilt, is no longer *Old Faithful*. Since none of the constituents of our bodies remains constant over time, it might be argued that there is no *self* as such, for everything about us physically changes from moment to moment.
- **B.** Against all this is a venerable philosophical conception of entities that retain their identity over time and independently of physical changes.
  - **1.** On this understanding, a thing is what it is essentially, even if a number of so-called accidental changes are imposed on it.
  - **2.** Intellectuals of the 17th century, especially Newton and Galileo, held that the last word on what really matters is to be provided by science.
  - **3.** Thomas Hobbes (1588–1679) in his *De Corpore* of 1642 held that two distinguishable entities cannot be the same and that there was continuity of the person throughout the seasons of bodily change, given one condition.
- **C.** Hobbes set the stage for John Locke's analysis of the issue.
  - **1.** In his *Essay Concerning Human Understanding*, Locke develops the distinction between "real" and "nominal" essences.
  - **2.** What we come to regard as the "essence" of something arises from our own tendency to classify it in a manner that is convenient by our own lights.
  - **3.** What will give Locke or anyone else that continuing identity that might have been incorrectly regarded as one's "real essence" is no more than that on which the various habits and dispositions of the mind settle.
  - **4.** It is the manner in which human intelligence perceives and uses entities that determines their "nominal essence."
  - **5.** Locke gave the problem of personal identity its modern formulation.

6. Faithful to Newtonian science, Locke regarded the real essence of a thing to be beyond the power of sense, a congeries of sub-microscopic particles held together by gravitational forces but perceived in ways that generate such *nominal* characterizations as "Fellow of the Royal Society" or "a rational animal."
7. These characterizations arise from conventional discourse, the contingencies of culture and context, the nuances of perception, memory, and mental life.
8. Locke accepts that there is a real essence, but he rejects the claim that this is given to us in our observations of a thing.

D. Locke reduces personal identity itself to the merely contingent contents of consciousness: It is hostage to the limitations of memory and the vagaries of experience.

1. Locke's famous example of the prince and cobbler, the contents of whose respective consciousness were switched, tells us that each man will be "the same man" but not "the same person."
2. The "real essence" of neither is disclosed by such nominal features as a princely bearing or special skill in turning leather.

E. How does this all play out in the matter of self-consciousness and that vexing *self* of self-consciousness?

1. Older philosophical schools tend to identify self with soul, namely, with an enduring feature of the individual, immune to changes that otherwise alter the conditions of the body.
2. On this understanding, behind all the changes that might take place in the life of John Smith, there is an essential John Smith that remains unchanged throughout the seasons of life.
3. If Locke's metaphysical mission is one of "Newtonianizing" the mind, then the first step must be the elimination of such a worrisome, nonphysical item as the "essential" self and the seemingly dual nature of reality—partly physical and partly something else.

III. Locke is considered one of the fathers of British empiricism, which reduces all that is factually knowable to what is observable or subject to perception.

A. If a personal identity is analogous to *Old Faithful*, then to the extent to which various old planks of memory continue to be present in consciousness, the personal identity of the individual is preserved.

B. Locke's cobbler and prince will each be not merely the subject of experiences but will experience these as *his own* and, thus, as experiences that cannot be had in just this way by any other.

1. The amnesiac is not lacking in self-consciousness, though he may have lost his Lockean "personal identity."
2. Locke knew this; he granted that we know ourselves to be the subjects of our experience *intuitively* and not as a result of systematic observation.

IV. William James proposed that there were different senses of "self," including the "material" self, the "social" self, and the "spiritual" self.

A. Considering the spiritual self at all is a reflective process.

B. Locke thought our knowledge of God and of ourselves was "intuitive," whereas our knowledge of the necessary truths of geometry was "demonstrative" in that a formal argument is required to demonstrate the truth of the conclusions.

C. Apart from these special cases, all of our factual knowledge comes, on Locke's account, from experience.

V. It is unclear that intuitive knowledge of one's self must be included among the necessary starting points in examining and defining consciousness.

A. Zombies are mindless. If consciousness is not necessary to account for seemingly mental achievements, why would self-consciousness be requisite?

B. Moreover, is it really the case that a person always knows *as a necessary truth and immediately* that the thoughts occupying consciousness are his own?

C. If "normal" knowledge of consciousness depends on the brain's health and proper functioning, isn't it *contingent* rather than *necessary*?

Essential Reading:

James, William, *Principles of Psychology*, pp. 283–296.

Robinson, D. N., ed., *The Mind (Oxford Readers)*.

Supplementary Reading:

Bealer, G., "Self-consciousness," *Philosophical Review* 106: 69–117.

Questions to Consider:

1. In what sense might someone be conscious but not aware of himself or herself as the conscious entity?
2. If everything in the body (and brain) changes, minute-by-minute, how can our "self-consciousness" be explained in physical terms?

Lecture Three
The "Problem" of Consciousness

Scope: In this lecture, we look at the perplexing relationship between the immaterial and the physical. We ask what it is about consciousness that would concern a physicist, and we address the claim that "physics is complete." We discover what Aristotle had to say about "real being," substance, and causality and raise the question of how the physical world interacts with a mental world not reducible to anything physical.

Outline

I. The core problem in the specialty of philosophy of mind is the problem of consciousness. Yet in our day-to-day affairs, we have no "problem" with consciousness; we scarcely think about it.

- **A.** Today, consciousness is an issue that tests both philosophers and theoretical physicists.
- **B.** Yet 18th-century leaders of thought were persuaded that no problem was inaccessible to the methods and perspective of the new physics.
  - **1.** Pierre-Simon Laplace (1749–1827) summed up the confidence of his age when he claimed that if we had complete knowledge of the position and velocity of every particle in the universe, the direct application of Newton's laws would allow us to predict perfectly every future event.
  - **2.** Laplace recognized that such complete knowledge was beyond the reach of human powers, but his bottom line, shared by a number of scientists today, was that physics is complete, even if our comprehension of its completeness is forever imperfect.
- **C.** This idea is ancient. The Greek atomists Leucippus and his student Democritus believed all of reality is exhausted by invisibly small physical particles and the spaces between them.

II. What does it mean to say that physics is complete?

- **A.** First, it adopts the metaphysical position—physicalism—that the only and ultimate reality is physical.
- **B.** The term *metaphysics* originated with 1st- and 2nd-century A.D. commentators on works by Aristotle (384–322 B.C.). They referred to the treatise Aristotle wrote after his treatise on natural science as "*meta ta phusica*"— "after the treatise on nature."
- **C.** Aristotle's "*meta ta phusica*" is the pioneering work in metaphysics, and its influence is nothing less than current.
  - **1.** Aristotle declared it an inquiry into what really is, combined with a critical examination of the grounds on which we may claim to know this or anything else.
  - **2.** The desire to know is not satisfied merely through sensory experiences. We know about things when we understand what brings them about and how they are related to other things.
- **D.** In Book I, Aristotle distinguishes among the different senses of "causation."
  - **1.** Taking the example of an ordinary object, a coffee cup, one "cause" of the coffee cup is the matter of which it is composed—its *material* cause.
  - **2.** But the material is insufficient to explain the causality associated with the cup. There must also be a *formal* cause for the cup to have the right shape.
  - **3.** To produce the right shape, the cup must have an *efficient* cause, which explains how the cup came to have the shape it has.
  - **4.** How did the makers of the cup know how to produce the cup? They had a plan to work from—what Aristotle calls the *final* cause.
- **E.** Aristotle's *Metaphysics* focuses on the issue of *real being*, and on this question, the limits of sense-based knowledge become evident.
- **F.** Physicalism regards all really existing things to be substantially physical. This invokes the notion of "substance."

G. For Aristotle, the substance of a thing is "that which is peculiar to it, which does not belong to anything else."

1. We can say that Mary has a pain, not that a pain has Mary.

2. Mary is not a predicate qualifying some other subject.

H. If the phrase "physics is complete" means that any and every entity to which real predicates are applicable is itself a physical substance, and if all predication in reality includes only physical properties, then "the problem of consciousness" is just one more problem to be solved by physicists.

I. If, however, consciousness is just a "code word" for an entity whose substantial nature is self-reflecting mental life, itself not reducible to anything physical, then the "problem of consciousness" is beyond the reach of physical analysis and physics is *not* complete.

J. In his *Metaphysics*, Aristotle addresses this issue with his categorization of real entities.

1. *Sensible* entities are those whose being is readily established through perception and can be explained by science.

2. Another class of entities comes under the heading of "unmovable"—changeless over time and in their essence.

III. We might postulate that "consciousness" is just that immoveable "substance" within the framework of which all change and all spatio-temporal affairs unfold and gain their real existence.

A. George Berkeley (1685–1753), early in the 18th century, concluded that the real, material world required for its subsistence a representation in *mind*.

B. On Berkeley's account, what is anything if not a set of perceptible properties? What is anything if not a representation in some consciousness?

C. The question at this point is ontological: What has "real being"; what really exists?

D. If physics is complete, we are committed to a monistic ontology of monistic materialism.

E. The problem of consciousness is one of discovering the manner in which entirely physical things and combinations of things come to generate a physical state or condition that we call "consciousness."

F. If, with George Berkeley, we find the most telling arguments being those that deny the independent reality of matter and require of the seemingly material world the foundational realty that is *mind*, then we retain a monistic ontology, but in this case, we would call it *monistic idealism*, real existence now being in the form of *idea*.

G. The more commonsense position to which we tend is that of dualism: There really is a physical world independent of us and a mental world of consciousness and its contents.

1. If we adopt the position of dualism, however, we come up against the problem of explaining what kind of "stuff" this mental stuff is.

2. If, as our commonsense ontology requires, it is immaterial, then we have the daunting question of just how an immaterial "desire to raise my arm" leads to my arm being raised.

H. Somewhere in this mix of questions and answers there seem to be assumptions that have not had the benefit of serious challenge.

1. As with the medieval "problem" of witches, we often try to solve a problem by ignoring the real problem and developing a sound solution to a very different problem.

2. As we shall see in subsequent lectures, there are a number of candidate problems, many candidate solutions, and much candidate evidence in the search for an explanation of consciousness.

Essential Reading:

Block, N., O. Flanagan, and G. Güzeldere, eds., *The Nature of Consciousness: Philosophical Debates*.

Chalmers, D. J., *The Conscious Mind: In Search of a Fundamental Theory*.

Supplementary Reading:

Flanagan, O. J., *Consciousness Reconsidered*.

Robinson, D. N., ed., *The Mind: (Oxford Readers)*.

Sartre, Jean-Paul, *Being and Nothingness*, p. 20.

Questions to Consider:

1. Earlier ages of philosophy reveal no concern about consciousness being a "problem." What modern developments have made it problematical?
2. If consciousness really is a "problem," where should we look for a solution? And why "there," instead of somewhere else?

Lecture Four
The Explanatory Gap

Scope: This lecture discusses the so-called *explanatory gap* that is inherent in the problem of consciousness. Philosophy of mind favors *foundational* explanations: The problem is seen as a gap between the dynamics of the nervous system and the nature of consciousness itself. Can causal relationships be established between neural events and conscious life? Some deny the existence of an explanatory gap at all. In the end, must we resign ourselves to the idea that this is just one of life's elusive facts?

Outline

I. If the major problem in philosophy of mind is the problem of consciousness, then it is so because of the so-called *explanatory gap*.

- **A.** Employment of the term *explanatory gap* suggests that consciousness resists location within the otherwise totally natural or physical domain and that were we to have the right sort of explanation, we would be able to move consciousness from a liminal location to the secure precincts of physical things.
- **B.** If the premise of an explanatory gap is accepted, then criteria must be developed to judge the quality of an explanation and its seeming validity.
- **C.** How do we explain the success of a trip to the Moon? One reasonable answer would be to claim that macro-level laws of physics must be correct because the mission successfully went to and returned from the Moon, based on the proposition that the laws of physics are correct.

II. The big gap in philosophy of mind has been seen as the gap between dynamics of the nervous system and the nature of consciousness itself.

- **A.** The majority of today's foremost philosophers would probably expect that gap to be filled by some sort of causal law.
- **B.** For example, Jack is awakened by his alarm clock and says, "I'm going to be late for work." He seems to have moved from the state of unconsciousness to a state of consciousness. How can we determine what it was that moved him from one state to the next?
  1. Common sense tells us that it was the alarm clock that awoke Jack.
  2. But perhaps a more meaningful explanation might be given by establishing causal relationships between neural events and conscious life. Brain events unfailingly arouse one from sleep to consciousness, even without a loud sound.
  3. Explanations based on brain function are more *foundational* than those based on the ordinary experiences of daily life. To know about brain function is to know what restores consciousness. To know about loud sounds is to know no such thing.
  4. One basis on which to judge the quality of explanations is to consider as better the one that is applicable over a wider range of instances; that is, the better explanation is the more foundational one.
- **C.** But it is not enough to say that Jack was restored to consciousness because of activity in his brain. We need to be more specific.
  1. Yet even at a more molar level (neural units within the cranium), we still cannot find any intelligible means of connection between the mental state (i.e., consciousness) and some internal activity involving one set of structures but not some other set of structures.
  2. Not every relationship is a causal relationship, nor does correlation imply causality.

III. There seems to be something of an explanatory gap between any number of paired relationships, where we sense that the first is somehow responsible for the second, but we cannot figure out how this responsibility is best understood.

- **A.** Isaac Newton (1642–1727) accounted for the fall of objects toward the center of the Earth through gravity. But to cite gravity as the cause does not explain just how gravity has this effect.

B. David Hume (1711–1776), a leading figure in the Scottish Enlightenment of the 18th century, continues to exert an influence on the question of causation.

1. Hume insisted that there is nothing in the external world that presents itself as a cause. We do not see, hear, or touch "causes."

2. Hume claimed that all such causal attributions are based on experience, specifically, on what he called the "constant conjunction" of events.

3. The explanatory gap existing between brain events and conscious experience is cut from the same cloth as the explanatory gap existing between the mass of objects and their behavior in free fall.

C. A contemporary philosopher, Michael Tye, argues that the notion of a "gap" is illusory. He reaches this conclusion by a complex analysis of how various concepts are formed and how they function at the level of our understanding of things.

1. Take the visual experience of "red."

2. According to physicalism, "red" = brain state "B."

3. But, as Tye notes, this strikes us as unconvincing, for describing brain state "B" will be nothing like the experience of "red."

4. Tye points to a basic confusion between the sense of a term and its reference.

5. When we think of what it is that the left side of the equation is referring to, we just think in terms of phenomenal experience; we more or less "picture" red or redness, and in so doing, we absorb the left side of the equation into the conceptual realm of things "felt" or "experienced."

6. Tye would have us accept the thesis that the actual references of our phenomenological expressions just constitute states of the brain; so-called *conscious* states simply *are* states of the brain under a different set of conceptual categories.

D. If the explanatory gap really exists, we still must ask whether it is a measure of where we have arrived in what is finally a work in progress or whether it promises to remain just one of those eternally elusive facts of mental life.

E. To assume that the explanatory gap is at once real and ineliminable in principle is to withdraw from this very framework at least one feature of reality, namely, our consciousness—which carries with it all the qualities that resist translation into ergs, volts, pints, and grams.

Essential Reading:

Levine, J., *Purple Haze: The Puzzle of Consciousness*.

McGinn, C., *The Mysterious Flame: Conscious Minds in a Material World*.

Supplementary Reading:

Tye, M., "Phenomenal Consciousness: The Explanatory Gap as a Cognitive Illusion," *Mind* 108: 705–725.

Questions to Consider:

1. What should any good explanation of a common occurrence include?

2. What are the grounds on which we regard an explanation as *explanatory?*

3. Where, within science and life, are there no "explanatory gaps"?

Lecture Five
Mental Causation

Scope: While causal relationships might be plausibly defended in attempts to close the explanatory gap, they pose problems. How can a mental activity cause a physical activity? A physicalistic solution is unsatisfactory because mental life is not reducible into physicalistic terms and concepts. It can be argued that matters of causation are principally matters of science, not philosophy. Moreover, much of daily life does not lend itself to causal explanations.

Outline

I. Causal relationships are serious contenders to close the explanatory gap in the relationship between mental and physical events, but they face serious problems.

A. It is customary to expect in any causal relationship that the antecedent-consequent events (causes and effects) will be of a kindred type.

1. Examples such as billiard balls colliding demonstrate action that takes place between two physical bodies.

2. Can the same be said of mental causation? How can a decision I take to make a glass of iced tea make parts of my body move?

3. At the commonsense level, and in the ordinary affairs of life, we take for granted that the plans and purposes shaped within our own mental life causally bring about actions capable of realizing those very plans and purposes.

4. Considered philosophically, however, this state of affairs is highly perplexing. And considered scientifically, it seems virtually impossible.

B. One solution to this problem is to adopt a form of physicalism and insist that decisions, judgments, plans, and purposes are simply code words for events in the brain, but this approach is flawed.

1. If we were to recast all statements that we have made about our decisions, judgments, plans, and purposes in physicalistic language, our statements would be unintelligible.

2. If we were to try to match up all of the characteristics of physical events with all of the characteristics of mental events, nothing would match.

C. Many philosophers have coupled a physicalist explanation with another explanation to satisfy the inherent problems of the physicalist approach.

1. Donald Davidson (1917–2003) defended the thesis that he labeled *anomalous monism*.

2. Davidson's monistic ontology proposed that there is only one kind of stuff in reality: physical stuff; yet mental life is not reducible or directly translatable into physicalistic terms and concepts.

3. The "anomaly" is that, on Davidson's account, mental events and processes *do* bring about physical events in just the sense of strict causation.

4. To do this, they must be physical but of such a nature as to preclude reduction to the terms of the physical sciences.

5. This is *anomalous*, for we are required to retain a dualistic vocabulary even as we acknowledge a monistic reality—hence, *anomalous monism*.

6. The most obvious criticism of this view is that it permits the very mental properties that make the causal model problematical.

7. If Davidson were willing to permit mental properties and, at the same time, regard them as irreducible to physical properties, this begs the question: How did these mental properties get caused by the physical properties of the brain?

8. Davidson resisted the very notion of properties and the conventional view that one thing causes another through some sort of causal mechanism that requires identification.

9. Instead, he took the position that where there is causation, one thing causes another by just causing it!

10. His skepticism regarding "properties" is echoed by other philosophers who have argued that "properties" are simply ensembles of causes. For example, to say that water is a "universal solvent" is to say that it *causes* things to dissolve.

II. To answer the question **"**What is a mental property?" we need to identify mental entities.

A. The most obvious mental entities are experiences and thoughts.

1. It is not clear how we should convert the experience or perception of an object into a mental property.

2. In terms of pain, although we can say that an activity in relevant areas of the brain causally brings about a specific mental property, namely, intensity of pain, we still have the problem of accounting for how it is that the experience (sensation) of pain causes a bodily reaction, such as moving our hand away from fire.

3. Moreover, many of our self-protective reactions take place at the level of reflexes and require no consciousness whatever, let alone a definite "experience."

B. Thoughts are philosophically referred to as "intentional objects," meaning the tendency of all mental acts toward an object.

1. This *intentionality* is of central importance to our consideration of mental causation.

2. The philosopher/psychologist Franz Brentano (1838–1917) pointed out that whether the activity is that of perceiving, desiring, remembering, feeling, knowing, or intending, it is always *about* something.

3. Brentano drew a sharp line between the mental and the physical. The latter is never *about* something.

4. How can any physical feature or property give rise to the "aboutness" of thoughts, feelings, and so on? How do mental events, which are about something, causally bring about physical events lacking in this very property?

C. Some philosophers have suggested that we should content ourselves with ordinary forms of explanation.

1. The philosopher Georges Rey suggests that problems be divided into those "peculiar to the mind" and those outside the mind.

2. This fits the notion of customary explanations as containing the "causal" conditions within themselves.

3. But the central task of any discipline, once it has mapped out its territory, is to arrive at a settled position on just how far its explanatory resources are likely to take it.

4. In philosophy of mind, to raise the question of mental causation moves the philosophical question to a scientific question.

5. Is it any surprise that the philosopher asking what is, at base, a scientific question, soon discovers the inability of philosophical modes of explanation to settle it?

III. It can be argued that questions of causation are, in principle and always, scientific questions. Furthermore, much of what we do in our daily lives is not best understood in causal terms.

A. Much of what we do in the ordinary affairs of life we have a reason for doing, reason being understood in a fairly broad sense—the person acting has some end or goal in mind.

1. The primary antecedent conditions for actions are reasons.

2. When viewed at a metaphysically safe distance, it would seem quite hard to ask how a reason for acting causes the action that is intelligibly related to it.

3. To act for a reason is to engage the larger world. The action is tied to an understanding of the surrounding environment and what it affords by way of possibilities for acting.

4. Having the same reason but finding oneself in a radically different environment might call for radically different courses of action.

5. Thus, causal accounts, drawn from the much more orderly world of physical objects in motion, prove to be far too limited when we attempt to bring them into the world of real life as actually lived.

B. Are causes and reasons fundamentally different?

1. Mustn't reasons be a species of cause if reasons are able to bring actions about?

2. To require this connection is to adopt the very physicalistic perspective that a "reasons" account challenges.

3. As of now, we are strongly inclined to regard all instances of motion and activity as candidates for explanations based on physical causes. That position could be called metaphysically "correct," but it could also be false.
4. Despite the tremendous growth of knowledge over recent decades, the problem of mental causation, nonetheless, is pretty much where it was in the time of the ancient Greek philosophers.

Essential Reading:

Heil, John, and Alfred Mele, eds., *Mental Causation*.

Supplementary Reading:

Kim, J., "Mental Causation and Consciousness: The Two Mind-Body Problems for the Physicalist," in *Physicalism and Its Discontents*.

Questions to Consider:

1. What is a "cause"?
2. We enter a room, flip a switch, and light floods the room. What would enter into a complete "causal" account of this occurrence?

Lecture Six
Other Minds

Scope: The problem of how we know there are other minds than our own is part of the epistemological problem of how it is that I know anything. It is more than a linguistic problem, as we shall see in this lecture, when we explore the positions taken by Ludwig Wittgenstein and the Scottish philosopher Thomas Reid, who countered skepticism with a pragmatic approach to the problem of other minds.

Outline

I. At the heart of a certain form of skepticism, known as *solipsism*, is the question: How do I know there are any minds other than my own?

- **A.** A solipsist is prepared to make no claim other than the fact of his own existence and mental life. The problem of other minds, then, is just the problem of answering the solipsist.
- **B.** In one sense, the problem of other minds is cut from the same cloth as all fundamental problems in epistemology: Basically, the problem of other minds is just part of the problem of how it is that I know anything or can claim that I know something.
  - **1.** I can perceive that I have a mind, but I cannot perceive any mind other than my own.
  - **2.** Even if an apparatus existed that could show us an image of another mind, another mind in a conscious state associated with pain, for example, we would have no knowledge of the pain as actually felt.
- **C.** Ludwig Wittgenstein (1889–1951) held that language may not be a reliable guide to correctly interpreting the statements of another, as any two persons may associate two different meanings with a single word.

II. But the problem of other minds is surely not merely a linguistic problem.

- **A.** The problem of other minds may be seen to arise once we attach all of our knowledge claims to direct perception.
- **B.** It is an evidentiary problem, in the sense that it becomes "problematical" only when we use the wrong sort of evidence.
- **C.** If direct perception constitutes the only justification for claiming to know anything, then, in fact, whole realms of what we regard as the "known" become quite obscure.
  - **1.** I do not directly perceive such laws as the laws of the internal combustion engine; my belief in them is inferential.
  - **2.** If I were a scientist, my belief would be based on the developed conception of the lawfulness of nature itself, namely, that if such laws were not persistent over time, the very coherence of the cosmos would be dissolved.
  - **3.** Little of such abstract reasoning is available at the level of direct perception.
- **D.** It could be argued, however, that to talk of the "coherence of the cosmos" simply reflects another of our prejudices.
  - **1.** A "coherent cosmos" can arguably be thought of as just one of an infinitely large number of pictures that might be drawn to capture the nature of reality.
  - **2.** This sort of question is at the heart of a contemporary issue within philosophy of science that is usually categorized as "realism versus antirealism."
  - **3.** Do our scientific laws really express "reality"?

III. The *pragmatic* ground is one particular ground of justification for knowledge claims of any sort.

- **A.** A pragmatic ground belongs to a class of assumptions or dispositions absent which the very conduct of life would become nearly impossible.
- **B.** The 18th-century Scottish philosopher Thomas Reid (1710–1796) advanced commonsense arguments against the seemingly impregnable position of the skeptic.

1. Reid held that we do not enter the world with utterly blank slates by way of our psychological or mental resources; instinct and intuition are natural endowments, supplemented by perception and learning.
2. For Reid, if the rich resources of perception are to serve their own purposes, they must not be nullified by the philosophical pretensions of the skeptic.

C. On the problem of other minds, Reid argued that while we cannot directly perceive any other mind external to our own, we do perceive the behavior of others.

1. On Reid's understanding, the possibility of any form of social interaction presupposes certain instinctual patterns of behavior.
2. The meaning of these patterns will be understood by members of that species, even across species.

D. To briefly summarize Reid's position on language, we begin with John Locke's theory, which we can take as the conventional position on the manner in which words come to have meaning.

1. According to Locke, and many others, a pattern of sound becomes meaningful as speech when language users accept it as such.
2. Reid was perhaps the first to recognize that this account will not work. In order for there to be agreements, there must be some language in place by which to establish them.
3. Reid distinguished between what he called *natural* language and the language we speak, which he dubbed *artificial* language.
4. Natural language refers to body language and intonation and is recognizable across species. In the prelinguistic world of primitive man, it served as a support for the development of artificial language.
5. As with other natural endowments, there is no rational justification for judging the signs and symbols of natural language.

E. Reid held the problem of other minds to be analogous to his concept of language: When other creatures express themselves in a manner comparable to our own natural expressions of our emotions and desires, we are led insensibly to the belief that they, too, have a mental life that includes just these emotions and desires.

IV. To the extent that Reid's argument is from analogy, it faces philosophical and conceptual objections.

A. Norman Malcolm (1911–1990) argued that because an individual's own mental states and behavior are confined to one person, this affords no grounds for a generalization to any other person.

1. Malcolm's argument, however, can be seen as inadequate.
2. There may be a confusion as to just what is being generalized.
3. The generalization does not proceed from one mind to all other minds but from a very large number of correlated mental states and behavioral expressions.
4. In the course of a day, each one of us has thousands of correlated mental states and behavioral expressions on which to ground our "generalizations."

B. What Reid refers to as natural language is very close to Wittgenstein's notion of natural expressions; certain natural expressions constitute a warrant for believing that someone else is happy, in pain, and so on.

C. Clearly, our recognition of other minds constitutes the basis of one of the most important ingredients of social life: empathy.

D. Both our conduct toward others and what we expect from them depend to a considerable extent on our ability to project ourselves into their situation.

E. We may tentatively conclude that nature fits both humans and nonhumans with certain basic tendencies and dispositions such that we are able to adjust to significant environmental impositions without having to learn how to cope with everything.

Essential Reading:

Avramides, A., *Other Minds*.

Questions to Consider:

1. Your best friend happens to be a state-of-the-art robot. How can you tell?
2. Some psychiatric patients present themselves as having multiple personalities. Is this evidence of multiple "minds" in one body?

Lecture Seven
Physicalism Refined

Scope: We have seen some of the difficulties associated with explaining mental events on the basis of physical evidence. In this lecture, we examine two alternative theories, the *identity* theory, which does not accept that there are uniquely mental events, and the *supervenience* theory, which requires that a person cannot move from one mental state to another without moving from one physical state to another.

Outline

I. Two alternative perspectives to explaining mental events on the basis of physical evidence are the *identity* theory and the *supervenience* theory.

- **A.** The counterintuitive character of the identity theory presents difficulties, while its particular strength lies in the fact that it does not aim to explain causal relationships between mental and physical events; it does not accept the proposition that there are, in fact, bona fide and uniquely mental events.
  - **1.** Although there are various forms of the identity theory, they all proceed from an ontological position according to which there is only one kind of entity in reality, namely, physical reality.
  - **2.** Thus, all forms of the identity theory adopt *monistic materialism* (or *monistic physicalism*), a term that pays deference to the fact that not everything physical has mass, for example, an electric charge.
  - **3.** Monistic physicalism asserts that whatever has real existence is physical.
  - **4.** Monistic physicalism explains immaterial entities with examples such as lightning, which is, in actuality, an electrical discharge.
- **B.** Philosophers use an identity of *reference* to distinguish from an identity of *meaning*.
  - **1.** The terms *Morning Star* and *Evening Star* refer to the same entity. Although two sentences containing these terms may well have different meanings, they have identical referents.
  - **2.** Similarly, it is not that electrical discharges cause lightning. Electrical discharges *are* the lightning; both terms refer to the same phenomenon.
- **C.** The identity theory can be illustrated thus: When Samantha says she has a toothache, she is referring not to a mental state but to a brain state.
  - **1.** One argument to this effect has been advanced by the philosopher J. J. C. Smart (b. 1920) who, because he regards it as unbelievable, rejects an ontology that requires two radically different forms of reality, one physical and the other—who knows what?
  - **2.** The identity theory is parsimonious in that it reduces reality from a two-substance to a single-substance affair.

II. Is the identity theory sound?

- **A.** In its usual form, the identity theory asserts that mental states and processes, properly understood, are actually states and processes taking place in the brain.
  - **1.** On this account, there is actually only one set of states and processes, namely, neurophysiological states and processes, not *mental* states as such.
  - **2.** The identity theory establishes the mental as physical.
  - **3.** The identity theory asserts that there is not a relationship between two distinct entities, only one entity for which we seem to have two modes of expression.
  - **4.** This can be summarized by saying that M is identical to P, where M = mental states and P = physical states.
- **B.** Considering the soundness of the theory requires considering the nature of identity relations in general.
  - **1.** The philosopher Gottfried von Leibniz (1646–1716) advanced a criterion for testing such alleged identities.

2. Leibniz's criterion of the identity of the indiscernibles proposes that distinguishing the mental from the physical could be ascertained by substituting a statement about a given mental state with one about a specific brain state.
3. For example, defenders of the identity thesis would say that there are not two ontologically distinct events to explain a toothache, one being pain and the other being neurophysiological events in the brain.
4. Applying Leibniz's criterion would show that the toothache pain is just a brain process—just one process.
5. But some have argued that there is something that might be truthfully said about pain that could never be truthfully applied to statements about brain processes without a fundamental alteration in the truth of the statement itself.

C. Other objections to the identity theory include the position that however mental states and physical states are to be understood, it surely is not the case that, if they are identical, they are so *necessarily*.
1. We can imagine a world in which there are mental events without there being any physical events or a world in which mental events and physical events simply occur in parallel.
2. Obviously, the identity thesis cannot plausibly maintain that mental events just are physical events and are so *necessarily*.

D. Defenders of the identity theory get around this difficulty by arguing that the identity in question is not one of necessity but one of *contingency*.
1. Of all the things that mental events conceivably might have been, it is just contingently the case that they are physical events.
2. That the specific person having a toothache is Samantha is said, here, to exemplify a *contingent identity*.
3. Yet Venus is itself, and necessarily itself, no matter how many distinct names (e.g., Morning Star and Evening Star) are assigned to it. That is a necessary truth, not a contingent fact.
4. We accept that, of all the names a baby might have been given, Samantha was the contingent choice, and thus, we agree that the woman with the toothache is contingently named Samantha. But we do not agree that, of all the women Samantha might have been, she is just contingently herself!

E. Saul Kripke (b. 1940) in his *Naming and Necessity*, published in 1980, distinguished between what he termed *rigid designators* and *proper names*.
1. What Kripke argued is that objects picked out by proper names are never necessarily identical to any set of properties or conditions.
2. Rigid designators, on the other hand, are what they are *in all possible worlds*.
3. An entity is itself necessarily.
4. Therefore a particular mental state remains what it is in all possible worlds, whereas the brain states that would serve as candidate substitutes are not necessarily what they are in all possible worlds.

F. It is probable that every version of the identity theory will finally be jettisoned for several reasons.
1. Suppose it is the case that there is nothing in the domain we call the "mental" that is anything but a collection of physical events and processes within the cranium and the balance of the body.
2. Then, one conclusion to be drawn from the theory, if it is true, is that it is not consequential.

III. The theory of *supervenience* owes its currency to Donald Davidson's essay "Mental Events."

A. According to Davidson, mental characteristics have a relationship with physical characteristics that is dependent—or supervenient—on them.
1. In its usual form, the theory does not begin by denying mental states.
2. Defenders of supervenience theories are willing to retain the ordinary language used to refer to various states and properties associated with psychological life, as well as to refer to various common items, such as tables.
3. In the case of a wooden table, for example, the wood supervenes on something more fundamental, namely, molecules, and the molecules supervene on atoms, and so on.

4. We understand that there are not two ontologically distinct realms: one with a dining room table and another with atoms. We accept that the table is constituted of materials, some of them invisible, but absent these, there would be no table.
5. A table cannot change its shape without an alteration in the physical properties on which the visible and palpable property of shape supervenes.
6. In terms of mental states, two persons cannot be different in the pains that they feel without the underlying neurophysiological processes being different; a person cannot move from one mental state to another without moving from one physical state to another.

B. The problem with supervenience theory is that in every application of it where it seems sound, the properties in question do not match up with anything that makes the "mind-body problem" a problem.

Essential Reading:

Armstrong, D. M., *A Materialist Theory of the Mind*.

Kim, J. ed., *Supervenience and Mind: Selected Philosophical Essays*.

Poland, J., *Physicalism: The Philosophical Foundations*.

Questions to Consider:

1. If physicalism is true, what difference would it make?
2. If physicalism is true, what happens to the concept of personal responsibility?

Lecture Eight
Consciousness and Physics

Scope: It has been argued that the real problem with physicalism is that we do not know enough about matter itself. The explanation of the phenomenon of mental life may demand a physical science beyond our current reach. In this lecture, we look at arguments from the laws of thermodynamics and quantum physics in pursuit of a solution to the question of the unification of the physical and the mental.

Outline

I. Radical versions of physicalism fail at the level of our intuitions.

A. We rely on intuitions, where neither the force of logic nor the evidence of sense supports one account at the expense of all other accounts of complex phenomena.

B. Thoughts are not like things, and feelings have no shape, which is sufficient for most persons to view the physicalist agenda as excessively optimistic.

C. More than one philosopher has suggested that such a judgment is premature.

1. Galen Strawson (b. 1952), in his book *Mental Reality*, argues that the real problem with physicalism is that we simply do not know enough about matter itself.

2. On an agnostic materialism view, the perplexing nature of the mind-body problem stems from unreasonable faith in our understanding of the fundamental nature of matter.

3. Strawson remains committed to monistic physicalism, but he remains neutral, or as he puts it "agnostic," as to just what the physics of it all will prove to be.

4. It is important to keep in mind that physics at the micro level is still a fairly young subject.

II. There are reasons to doubt that our current understanding of physics is up to the task of explaining conscious life, while the possibility remains that physics—especially quantum physics—may yet be able to do so.

A. The Oxford mathematician and theorist Roger Penrose (b. 1931) sets limits on the extent to which contemporary physical science may be viewed as promising in relation to the problem of consciousness.

B. Penrose concludes that the phenomenon of mental life may require a physical science not yet at hand.

1. Kurt Gödel developed a mathematical theorem known as the *incompleteness theorem*, which holds that a mathematical system of sufficient complexity to include an arithmetic is logically incomplete; arguments external to it must be imported in order to render it complete.

2. In light of Gödel's theorem, Penrose concludes that the proper model of human cognition is not computational, for, if the essential nature of human thought were computational, the limitations imposed by Gödel's incompleteness theorem would be manifest.

3. We are able to reflect on our own problem-solving maneuvers without importing into the system some set of axioms, otherwise unknown to us, in order for us to make sense of what we are thinking.

4. To say that we do not need such an external problem-solving apparatus is to say, among other things, that the essential character of human consciousness is *non-algorithmic*.

5. Penrose concludes that mental life is simply not explicable in terms that physics now offers us.

C. Many think that the physics we have at least rules out alternatives to physicalism, if we consider the first and second laws of thermodynamics.

1. The first law establishes that the internal energy of a system is equal to the heat added to the system minus the work done by the system.

2. To insist that the term *energy* is reserved solely to physical forms is simply to assert physicalism, not to establish its adequacy; there could be some sort of *mental* energy. The first law of thermodynamics is neutral as regards the form of energy.

3. The second law asserts that in any closed system, if anything physical is brought about, it is at the cost of energy in the system.

4. Another conception of the law is expressed in terms of *entropy*; unaided by external influences, entropy (or "disorder") tends to increase. To oppose this, work must be done and must be supplied from sources external to the system.
5. From this, we can see what troubles the physicalist about the idea that mental events bring about bodily events. To have such an effect, the mental events would have to supply energy, and this would be a measurable property of the system.

D. The thermodynamics argument is less than convincing.

1. In purely formal terms, the thermodynamics equations are neutral as regards the nature of the units on each side of the equal sign.
2. Furthermore, thermodynamics laws apply to closed systems in a state of thermodynamic equilibrium, and defenders of mental causation can argue that at least some psycho-physical transactions find the system neither closed nor isolated.
3. In the real world, systems are permeable to all sorts of external influences, giving ample opportunity for anti-entropic outcomes.
4. It is sometimes argued that evolution itself violates the entropy version of the thermodynamic laws, for, through evolution, there is a progressive increase in complexity and organization.

III. The prospect has been raised that consciousness and quantum physics are in some bizarre way interdependent.

A. One of the most influential schools of quantum physics, the so-called *Copenhagen school*, has argued that the results of experiments at the micro level can be understood only in terms of the influence of the act of observation itself.

1. Niels Bohr's model of the atom restricts the locations that electrons can occupy when moving from one orbital plane to another; they are entirely probabilistic, as are the energy-level transitions themselves.
2. Further theoretical developments, driven by experimental results, present a microcosm in which the particles occupy multiple states and multiple locations at the same time. The very act of measurement "collapses" these to a single state, namely, a now specifiable single location and state.
3. David Bohm (1917–1992) furthered the speculations of Eugene Wigner (1902–1995) on the connection between quantum physics and consciousness, applying the theoretical and observational aspects of quantum physics directly to the functions of the brain and to the problem of consciousness.
4. Bohm, with his student Yakir Aharanov, showed that an electromagnetic field could have effects in spatial regions otherwise fully shielded (the so-called *Aharanov-Bohm effect*, which violates the core canons of classical physics).
5. Bohm later advanced the thesis that the brain at the micro (quantum) level is an informational system such that the mental and the material merge.

B. Suppose the Copenhagen interpretation is correct, namely, that it actually is the case that only by way of observation and measurement, by way of the introduction of some element or derived product of consciousness, that quantum uncertainties "collapse" into determinate states.

1. The influence that consciousness or its derivatives might have is not by way of contact but by way of that dimensionless entity, information.
2. The very nature of consciousness remains still unaddressed.

C. The essence of quantum phenomena is statistical.

1. What quantum physics overturned, at least at the level of micro- physics, is the classical world of absolute space, absolute mass, and absolute motion.
2. The essence of normal, adult mental life, however, is the capacity for logical modes of analysis and argument; modes of comprehending the broad divide that separates all that is "probable" from that which is necessary.
3. The world of quantum physics is a world of pure contingency, while the world of formal logic is one of pure certainty.

4. There appears to be some sort of "modal" mismatch between the probabilistic nature of quantum reality and the necessities attaching to the conscious construction of such abstractions as mathematics and logic.
5. Quantum effects, owing to their statistical nature, become more accessible at the level of observation in the form of averages arising from innumerably large events at the level of particles.
6. The brain possesses billions of neurons and comparably innumerable connections among them. At any given time, we might consider the functioning brain as the averaged outcome of all these small activities.

Essential Reading:

Albert, D., *Quantum Mechanics and Experience*.

Bohm, D., and B. J. Hiley, *The Undivided Universe*, chapter 15.

Lockwood, M., *Mind, Brain, and the Quantum: The Compound 'I'*.

Supplementary Reading:

Heisenberg, W., *Physics and Philosophy*.

Questions to Consider:

1. Quantum events are inherently probabilistic. Are they a good model for mental events?
2. If reality is finally a composite of mindless quanta, how could I "know"?

Lecture Nine
Qualia and the "Mary" Problem

Scope: This lecture pursues the problem of physicalism through the examination of qualia, as framed by the philosopher Frank Jackson's famous "Mary" problem. We find that having scientific knowledge of natural phenomena is the not the same as having experience of such phenomena. Likewise, qualia are not successfully explained by reference to functional states.

Outline

I. The problem of *qualia* was framed by the philosopher Frank Jackson (b. 1943).

- **A.** In Jackson's analogy, Mary is a scientist who is forced to investigate the world from a black-and-white room via a black-and-white television monitor.
  - **1.** She knows all about the physics of light and colors.
  - **2.** When she is released from her black-and-white world, she finds that her previous knowledge was incomplete, although she possessed all the necessary physical information.
- **B.** *Qualia* refer to qualities that enter into our experience of things; for example, roundness and redness stand as qualia.
- **C.** Jackson argued that if physicalism were adequate in all respects, then to know everything about the physics of the situation is to know everything.
  - **1.** If Mary knows everything about the physics of light, then presumably she knows everything about light.
  - **2.** But if she knows everything about light, how is it that on entering the world of color for the first time, she now knows something that heretofore she did not know?
  - **3.** Jackson argued that there is a kind of knowledge delivered to us by way of our experience that cannot be delivered any other way.
- **D.** This is not a new challenge.
  - **1.** Johann Wolfgang von Goethe (1749–1832) published his *Theory of Color* in 1810, in which he contended that Newton's theory of light explains everything except what we actually see.
  - **2.** For Goethe, no matter how complete our understanding of the physics of light might be, that understanding alone cannot predict or explain the actual experiences we have of light and vision.

II. Are qualia the same for two different persons, and is there any method by which we can answer such a question?

- **A.** We could set up an experiment to test whether two persons see the same color, but it would tell us nothing about what might be called the "inner experience" of each of these persons.
- **B.** There is no reason to assume that identical physiological responses in two different persons unfailingly yield identical experiences.
- **C.** Nonetheless, there is a general tendency to regard the brain's record as somehow more valid or revealing than the mere utterances of actual persons.

III. Regarding the "Mary" problem, other philosophers have raised critical questions.

- **A.** Daniel Dennett (b. 1942) proposed that Mary is either lacking in relevant knowledge or has not made proper use of it.
  - **1.** Assuming that Mary really does know everything, Dennett argues, Mary would have to know how the deep structures of the brain respond to information from the peripheral sense organs such that these brain mechanisms give rise to qualia.
  - **2.** Even if Mary knew everything about the relationship between color experience and the physiology of the nervous system, absent reported qualia by actual persons about their own experiences, there would be no basis upon which to make any sense out of whatever data scientists had extracted from the brain.

B. We must keep in mind that the brain is constantly active, the visual system responding incessantly to events that reach the retina.

1. If this is going to provide the basis upon which to judge what an individual is seeing, it will never be enough to know only the activity going on in the nervous system.

2. There will have to be some way of tying particular events in the nervous system to the qualia reported by each individual.

3. It will be the qualia reported by an individual that provide the relevant data if the events in the nervous system are to explain *anything* about the experience.

4. Accordingly, it will not be the brain sciences, or brain sciences alone, that will settle the "Mary" problem.

C. Perhaps Jackson missed a major point in setting up Mary in a color-impoverished environment.

1. As the philosopher Paul Churchland (b. 1942) noted, conditions such as this will prevent the normal development of those mechanisms responsible for the processing of color information.

2. In light of these deficits, Mary cannot be said to know everything about the effect of light, for part of what goes into "everything" includes processes she is lacking.

3. Jackson's initial thesis is based on the assumption that Mary's visual system is normal, which suggests Jackson was just missing a significant piece of scientific information.

4. Churchland's rejoinder includes the fact that relevant experiences are necessary for us to learn about color in any way at all.

D. We could modify the "Mary" problem and have Mary exposed to poison ivy, to which she is allergic.

1. Mary knows that those who are allergic to poison ivy develop a rash and an itch.

2. It is not clear, however, that although she knew everything about this allergy before suffering from it herself, the actual itchiness presented her with something she could have claimed to "know" before the fact.

3. A description of an experience is different from the experience itself.

E. Jackson's "Mary" problem fails to distinguish between knowing *that* and knowing *how*.

1. I can read a book on all that there is to know about riding a unicycle, but I still do not know how to ride a unicycle.

2. But does this distinction between knowing *how* and knowing *that* actually get to the heart of the issue?

IV. Given the complexity of the nervous system, it is tempting to believe that all that hardware is surely able to bring about qualia.

A. The qualia problem, however, is related to, but different from, questions about the factors that seem to be necessary for their occurrence.

1. Ned Block (b. 1942) proposed the "Chinese nation" thought experiment, the point of which is that qualia such as pain are not successfully explained by reference to functional states.

2. Similarly, Gottfried Leibniz proposed a machine he called a "mill" that could think, feel, and perceive, but an examination of its works would just reveal parts, not anything that would explain perception.

3. In 1974, Thomas Nagel (b. 1937) questioned what it is like to be an entity that has experiences of a certain kind; he argued, "An organism has conscious mental states if and only if there is something that it is to *be* that organism."

B. Experiences have a transforming effect, different from the effects of knowledge.

V. The problem of consciousness comes to be a problem because of certain core assumptions about physicalism that we are strongly inclined to make.

A. While we may not be able to accept the idea that physics is complete, we tend increasingly to accept the idea that problems at the level of fact can be settled scientifically, if they can be settled at all.

B. Consciousness seems to be the one fact in which all other facts are located, and this could be a key to a fuller understanding of the problem.

1. Our awareness of consciousness is likely be heightened by our ability to compare times when we lose consciousness (through sleep, illness, and so on) against times when we regain it.
2. Nevertheless, the number of times we lose consciousness in sleep has not significantly deepened our understanding of consciousness.

Essential Reading:

Block, N., O. Flanagan, and G. Güzeldere, eds., *The Nature of Consciousness: Philosophical Debates*.

Jackson, F., "What Mary Didn't Know," *Journal of Philosophy* 83: 291–295.

Nagel, T., "What Is It Like to Be a Bat?" *Philosophical Review* 4: 435–450.

Supplementary Reading:

Carruthers, P., *Phenomenal Consciousness: A Naturalistic Theory*.

Kirk, R., *Raw Feeling: A Philosophical Account of the Essence of Consciousness*.

Questions to Consider:

1. Did Mary *really* learn something when she first saw colors?
2. How does the "Mary" problem relate to the problem of "other minds"?

Lecture Ten
Do Computers Play Chess?

Scope: Whether a machine is capable of thinking is a question that evokes the problem of other minds and one that found responses in the work of Alan Turing and John Searle. The latter's "Chinese room" analogy seeks to refute the relevance of computational power to the question of consciousness and intelligent behavior. Ludwig Wittgenstein posited that games and rules are cultural artifacts. Playing a game by the rules is not the same as getting the gist of a rule.

Outline

I. In 1997, the world's greatest chess player, Garry Kasparov (b. 1963), was defeated in a chess match with an IBM computer named "Deep Blue."

- **A.** The IBM computer was capable of processing 200 million chess positions per second, while Kasparov could process three or four positions per second.
- **B.** Moreover, the computer, unlike Kasparov, was not subject to physical phenomena, such as emotion and stress.
- **C.** Since chess strikes many as the quintessential activity of an intelligent being, Deep Blue offers an example of a machine outsmarting a human being.
- **D.** Kasparov, like Deep Blue, was the beneficiary of very special circuitry, a very special brain.
- **E.** If Kasparov's brain qualified him as an intelligent being, why would we deny Deep Blue the same status?

II. To appreciate the burden of this question, we turn to Alan Turing's 1950 article entitled "Computing Machinery and Intelligence."

- **A.** Alan Turing (1912–1954) imagined an "imitation game" designed to show that a machine, properly programmed to answer questions, will be indistinguishable from an intelligent human being.
- **B.** With stunning prescience, Turing claimed that one day the notion of machines thinking would be commonly accepted.
- **C.** Turing was contradicted in his own time by a Professor Jefferson who held that machines could not be considered to have intellectual qualities unless they could create from feelings and emotions.
- **D.** Turing refers to this objection as "the argument from consciousness." We cannot impute intelligence to a machine unless it is conscious of its own achievements.
  1. Turing examines the outcome of taking this argument to its extreme, which is essentially the solipsist point of view.
  2. Turing seems to be tracking the "problem of other minds" from a commonsense position: When something behaves and judges as we do, we assume it does so with the same psychological resources we need to do the same.
  3. If it were necessary to have direct knowledge of the consciousness attending the actions, then we could vouch only for ourselves as having minds—the solipsist point of view.
  4. Turing's devices accomplish what they do in a manner that seems similar to our own accomplishments, namely, by following established procedures and providing correct answers.
- **E.** In his "Chinese room" analogy, John Searle (b. 1932) critiques this perspective.
  1. Persons ignorant of the Chinese language participate in an experiment in which they follow written instructions on how to arrange Chinese ideograms in a sequence that makes no sense to them but can be read by those who understand Chinese.
  2. What Searle seeks to establish here is the manner in which a seemingly mental achievement is gained by a process in which there is no comprehension whatever; it mimics the functioning of computers.
  3. Searle's "Chinese room" analogy tries to show the irrelevance of computational power to the question of consciousness and intelligent behavior.

4. To accept the "Chinese room" analogy as establishing this is to accept that, in principle, the achievements in the field of artificial intelligence would leave untouched all the interesting questions about consciousness and understanding.

F. Thomas Nagel's question, "What is it like to be a bat?" infers that to be like something is to be the subject of some experience.

1. There is no basis on which to impute any of Kasparov's emotions, fatigue, and so on to Deep Blue.
2. It is not solipsism to deny consciousness to matter as long as we grant it to those like ourselves.
3. With all due respect to Alan Turing, whatever Deep Blue is doing, it is surely not thinking.

III. If Deep Blue is not thinking, is it playing a game? To answer this, we must establish just what it is that makes any activity a game.

A. Any attempt to provide an exhaustive definition of what is a game is doomed to failure.

B. Ludwig Wittgenstein argued that games are recognized for what they are only by certain persons and under certain conditions.

1. A game might be understood as an activity that expresses certain cultural forms and norms.
2. The difference between a street fight and a boxing match, for example, is recognized for what it is only by those understanding various conventions, instructed within a given culture, and having a developed conception of what constitutes a "match."

C. But games are not only cultural artifacts; they involve rules.

1. There is a distinction between the actual physical constraints that guarantee a given course of action and following a rule.
2. A restaurant's "no smoking" rule would still be violated if we lit a cigarette and let it burn without actually inhaling it.

D. Do computers follow the rules of chess?

1. Those who follow rules get the gist of the rules.
2. Following a rule is not a built-in causal connection but, rather, a conventional one.
3. A given rule will apply across a virtually limitless number of instances such that no one could find them all in advance.
4. Computers do not "get the gist" of the rules of chess. Deep Blue was programmed to have its chess pieces move according to the rules of chess.

Essential Reading:

Penrose, R., *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*.

Searle, J. R., *The Rediscovery of the Mind*.

Supplementary Reading:

Turing, A. M., "Computing Machinery and Intelligence," *Mind* 49: 433–460.

Questions to Consider:

1. Must there be consciousness for there to be "play"?
2. What are the necessary features of an activity that qualify it as a "game"?
3. Well, *do* computers play chess?

Lecture Eleven
Autism, Obsession, and Compulsion

Scope: Knowing what happens when the functions of consciousness are defective may help us learn more about the nature of consciousness. In order to adapt to any particular environment, a normal sensory system resorts to active or passive filtration. This does not happen in cases of autism and other neurotic disorders. Moreover, those with such disorders cannot come to know what it is to be like someone else through conscious awareness and the integrative achievements of the mind.

Outline

I. Psychologists refer to our ability to continue a conversation when background noise is very loud as the *cocktail party effect*, and this includes the ability to control what we listen to when different sets of sound are present at the same time.

- **A.** To be conscious is to be *aware of something*, and the process by which we are able to direct our awareness is *attention*.
- **B.** Under normal circumstances, there are two principal means by which events in the external world are denied access to consciousness; both involve filtering.
  - **1.** One means is fixed and based on the operating characteristics of a particular sensory system.
  - **2.** The second means by which events in the external world are barred from entry into conscious awareness is that of *active* filtering, sometimes called *selective* attention (the cocktail party effect being a good example).
  - **3.** In the absence of such filtering, it would be impossible to adapt to any environment.
- **C.** The relationship between consciousness and awareness is direct, as is the relationship between awareness and attention.
- **D.** If we cannot get outside our own consciousness in order to study it—in the same way that fish are not likely to *discover* water, for it is the abiding condition of their lives—we might nevertheless learn more about the nature of consciousness by seeing what happens when its functions are defective.

II. There has been a surge in research in recent decades of childhood conditions that affect the functions of consciousness.

- **A.** Two such conditions fall in the category of autism spectrum disorders (ASD) and include autism proper and Asperger's syndrome.
  - **1.** Asperger's syndrome is a mild form of autism disorder syndrome, and sufferers often have an elevated sensitivity to noise or loud sounds.
  - **2.** Those who suffer from autism spectrum disorders also have symptoms that seem closely tied to the failure to direct and maintain conscious awareness.
  - **3.** Such children are slower in learning to interpret what others think and feel.
- **B.** Another such disorder is attention deficit hyperactivity disorder (ADHD).
- **C.** There are now many skeptics who are inclined to think that quite normal behavioral problems are being classified as one or another of these abnormal conditions.
- **D.** What is common across cases of autism is the obsessional nature of the controlling thoughts; in this regard, autism and obsessional neuroses have much in common.
  - **1.** Persons suffering obsessional disorders engage in repetitive behavior that is organized and, at a certain level, meaningful, but finally serves no purpose.
  - **2.** These activities commonly include such obsessions as the repeated checking of doors, locks, lights, and so on; the repetitive washing of clothes; or the verbal repetition, sometimes for hours, of a single telephone number.
- **E.** The flattening of emotion, so common in autism, ties directly in with the diminished capacity for empathy.
  - **1.** Where the Asperger syndrome generates auditory defensiveness, the autistic syndrome seems to generate emotional defensiveness.

2. The field of consciousness narrows as a consequence, and once narrowed, the correctives of reality have no means of access.
3. The French psychologist Pierre Janet (1859–1947) theorized that neurosis at its base is the narrowing of the field of consciousness.
4. On this account, neurosis is a disorder at the level of information processing.

F. Another tragic case was that suffered by the classical musician Clive Waring, who lost portions of his brain associated with memory.

1. Waring, while in all respects conscious, alert, and aware, was unable to retain any memory of events that had taken place as little as half a minute earlier.
2. Waring's search for his own mind was moment to moment, with nothing enduring.

III. Conscious life and awareness are not restricted to information processing in the neutral sense of information.

A. We are able to integrate what we hear in the voice and see on the face of another with that provided by the balance of the context.

B. Owing to certain normal and natural dispositions, the result of this integration is in the form of empathy and sympathy; we come to know what it is like to be someone else.

1. This is an achievement of conscious awareness and the integrative achievements of mind, one that makes possible an understanding of lives different from our own.
2. Try to imagine the quality of life available to those unable to perform these integrations and reach states of awareness.

C. It is never, however, a simple task to uncover what it is like to be someone else.

1. Few have done as much to alert us to this as Dr. Oliver Sacks (b. 1933), whose book *The Man Who Mistook His Wife for a Hat* is based on clinical investigations of patients suffering from right-hemisphere brain defects.
2. One of those patients displayed the full range of autistic symptoms yet could draw with amazing accuracy and in a manner that revealed a comprehension of the feelings of others.

D. We all engage in behaviors that to some degree could be labeled autistic, compulsive, or obsessional, but that does not mean that we know what it is like to be autistic or neurotic.

E. There are many other clinical conditions that display aspects of conscious life and awareness otherwise obscured by normal life.

1. These include dissociative disorders, which might involve large gaps in memory or the adoption of multiple personalities.
2. Another version—depersonalization disorder—finds the patient seeing himself at a distance and reflecting on his own behavior as if he were a disinterested witness.
3. These conditions are especially intriguing because the patients have consciousness and are aware.
4. Then, too, there are many altered states of consciousness, some caused by the use and misuse of drugs and alcohol.

Essential Reading:

Sacks, O., *The Man Who Mistook His Wife for a Hat*.

Supplementary Reading:

Baron-Cohen, S. *Mindblindness: An Essay on Autism and Theory of Mind*.

Questions to Consider:

1. Autism is characterized by greatly "flattened" emotions. Are emotions a necessary aspect of *mental* life?
2. How would you reply to someone who said that computers have minds but are "autistic" at the level of interpersonal and emotional life?

Lecture Twelve
Consciousness and the End of Mental Life

Scope: In this lecture, we consider the case of comatose patients in relation to the nature of consciousness, self-consciousness, and *epistemic justification.* Such issues raise the vexing question of criteria used to analyze mental states and the equally critical questions of ethics, rights, and duties in relation to those incapable of protecting their own rights. We end this course with the consideration that it is conscious life that defines our humanity.

Outline

I. The conventional doctrine of neurology and neuroanatomy claims that neurons in the central nervous system are not regenerative, but that doctrine is being challenged by recent research in genetics and developmental neurobiology, as well as by cases of coma patients who have recovered.

A. Many such patients actually resume a life not unlike the one they had left behind before the trauma, but these cases are the exception to the rule.

B. When the cells in the traumatized area of the brain die, deficits ensue, some of which can be overcome by rehabilitation, and some of which cannot.

1. The degree of recovery tends to be greater in early childhood.

2. The more extensive the involvement of the cerebral cortext, the greater the effects on consciousness and the various powers that depend on it: focused attention, problem-solving, general awareness, and cognitive modes of represenation.

3. These powers may be severely degraded—even go undetected by the most careful means—in a patient who is otherwise seemingly "awake" and physiologically competent.

4. Life-sustaining processes can continue to support the body's biological requirements even when cortical functions have all but ceased.

5. Even with all the data that modern machines can generate, confusion and doubt will prevail regarding a coma patient's mental state, as in the famous case of Terri Schiavo.

6. In one systematic study, investigators found that 17 of 40 persistent vegetative state (PVS) patients were misdiagnosed, and one-third displayed recovery during the period of the research.

II. Cases of comatose patients relate to matters growing out of the central issues raised in our earlier lectures. Consider the doctors' dilemma of deciding whether or not such a patient has some inner mental life.

A. Thomas Nagel's question— "What is like to be a bat?"—hearkens back to an argument developed by Aristotle in the 4th century B.C.

B. For Aristotle, the dividing line between plant and animal life was established by sensation.

C. Pulling together Nagel's and Aristotle's understandings, we might say that mental states, at least minimally, are states in which something is sensed.

1. A generalization of this criterion is the proposition that any creature that is the subject of sensation is in a mental state.

2. Moreover, any creature that is potentially the subject of sensation is potentially the possessor of a mental state, from which we can infer that to be the potential subject of sensation is to be *potentially* conscious and to be the actual subject of sensation is to *be* conscious.

D. Does putting Nagel and Aristotle together in this way make for a sound proposition?

1. Any number of sensations are present as we dream, but wouldn't it be contradictory to state that a dreamer is conscious?

2. There seem to be important differences between states of consciousness and states of self-consciousness.

3. What distinguishes dream states from waking states is not consciousness per se but what in metaphysics is sometimes referred to as *epistemic justification*.

4. What we believe about happenings as we dream is without the quality and character of justification enjoyed by things experienced in the waking state; in the waking state, we can justify our reported experiences by sharing them—intersubjective agreement is possible.
5. This issue is relevant to issues involving comatose patients, at least those who, on recovery, claim to have overheard people talking while they were seemingly unconscious.

III. Let us rehearse the conventional wisdom that our discussion of coma patients raises on key questions in ethics and moral thought.

A. Beginning with our understanding of obligations and duties, we have the duty not to exploit or unjustly benefit from the vulnerabilities of others.

B. When we consider our duty relating to the rights of those who are incapable of protecting their own rights, we must confront the nature and limits of the duties of the caretaker.

C. These rights and duties do not apply to those who are "brain dead," though the criteria for "brain death" are controversial.

D. Another pertinent question is: Just how much by way of biological function must a creature have to enjoy some measure of respect?

E. While advances in medicine and personal health have resulted in a more active and flourishing older population, decline in old age—including intellectual decline caused by such conditions as Alzheimer's disease—attacks what we take to be the very essence of our humanity: our mental life.

F. People make living wills to guard against the preservation of their lives should they succumb to a mentally debilitating disease. On what criteria do they base their judgment of a life not worth living?

1. How do caretakers decide when to honor a living will or take lethal action?
2. The "slippery slope" argument proposes that we hold back lethal action in the clearest of cases, lest we take lethal action in less clear cases.
3. In the case of animals, it is only in the philosophy seminar that serious doubts are raised as to whether animals have consciousness.
4. We take it to be a measure of civilization that those least able to care for themselves are the beneficiaries of the solicitude of others.
5. We care not only about creatures that are subject to pain and suffering but about our own character.

G. Our humanity seems to begin and end with conscious life. It is how we use all the rest that serves as a final judgment on whether that consciousness was a gift or a test that we have failed.

Essential Reading:

Jennett, B., *The Vegetative State: Medical Facts, Ethical and Legal Dilemmas*.

Sacks, O., *The Man Who Mistook His Wife for a Hat*.

Supplementary Reading:

Andrews, Keith, Lesley Murphy, Ros Munday, and Clare Littlewood, "Misdiagnosis of the Vegetative State: Retrospective Study in a Rehabilitation Unit," *British Medical Journal* 313 (1996): 13–16.

Horner, Philip J., and James H. Gage, "Regenerating the Damaged Central Nervous System," *Nature* 401 (2000): 1.

Questions to Consider:

1. It is said that the law has a compelling interest in the protection of life. Does this interest extend to a determination of when and whether one might end one's own life?
2. To say that my life is "mine" is to say something different from saying that it is a possession. How is the difference to be understood?

Glossary

Aharanov-Bohm effect: Refers to effects produced by an electromagnetic field in regions otherwise fully shielded.

Algorithm: A "recipe" of sorts for accomplishing specific tasks or solving specific problems.

Alzheimer's disease: An age-related disease whose most devastating symptom is dementia. It affects about five percent of those between 65 and 75 but as many as half of those over 85.

Anomalous monism: Donald Davidson's term referring to the inability to translate mental terms directly into physical terms, even while accepting that all mental states refer to what at base are physical states.

Anti-realism: A thesis in philosophy of science that rejects the view of scientific laws as veritable pictures or exact reflections of reality.

Artificial intelligence: The seemingly intelligent behavior of computational systems powerful enough to solve complex problems.

Artificial language: Thomas Reid's term for any developed language of words and grammar.

Asperger's syndrome: One of the autism-related disorders, milder than autism proper.

Attention deficit hyperactivity disorders (ADHD): Term covering a range of symptoms that include impulsive behavior and inability to maintain focused attention.

Auditory defensiveness: A symptom of autism such that the patient becomes defensively unresponsive to disturbing sounds.

Autism: A pervasive developmental disorder involving language, behavior, and social interactions, the symptoms appearing early in childhood and often associated with a degree of mental retardation.

Autism spectrum disorders: A group of disorders that includes autism and Asperger's disease.

Autoscopy: The seeming ability to see oneself from a position external to one's actual location; e.g., "out-of-body" experiences.

Brain death: A condition of the brain satisfying widely adopted criteria within neurology and including absence of recordable electrical activity.

"Chinese nation" analogy: Ned Block's fanciful analogy to the nervous system, here, with each citizen of China equipped to function as a neuron. Each citizen is able to enter a complex network of information via phone calls that mimic perfectly the signals generated in an individual's nervous system and are associated with the sensation of pain, although no one in China thus experiences pain.

"Chinese room": John Searle's famous room analogy in which persons ignorant of Chinese nonetheless arrange Chinese ideograms according to rules. The result is a message meaningful to those who know the Chinese language but not to those who assemble the message.

Cocktail party effect: The ability to maintain focused attention on one's own conversations in a setting in which loud noise is "filtered" out.

Compulsion: A neurotic repetition of stereotypic behaviors.

Deep Blue: The IBM computer that defeated Garry Kasparov.

Demonstrative: A mode of argument leading to conclusions seen to be both certain and logically necessary.

Depersonalization disorder: A form of mental illness in which one feels detached from one's own body and experiences.

Dichotic: A mode of presentation of sounds in which each ear receives a separate input.

Dissociative disorder: Exemplified by "multiple personality" and often related to traumatic experiences in early childhood.

Dualism: The thesis that there are two basic and different constituents of reality, one physical and one mental.

Efficient cause: Typically, the event immediately preceding an effect and having a measurable, physical influence, such as one billiard ball hitting another.

Empirical: Accessible to direct or aided observation.

Enlightenment: The 18th-century epoch of heightened intellectual and revolutionary undertakings, impelled by unstinting confidence in science.

Entropy: The tendency toward disorder. In thermodynamics, the byproduct of work in the form of heat; in cosmology, the forces that oppose order and coherence.

Epistemology: The branch of metaphysics devoted to a critical appraisal of knowledge claims and the modes of inquiry and explanation on which such claims are based.

Essentialism: The thesis that specific "kinds" of things (apples, persons, copper) are what they are owing to certain core and defining properties.

Explanatory gap: The "gap" in an otherwise completely physical account of reality; the gap arising from the resistance of mental phenomena to physicalistic explanations.

Final cause: In Aristotle's sense, the "that for the sake of which" other causal modalities are recruited; the goal or end toward which actions tend.

Formal cause: In Aristotle's sense, the feature that makes a thing the sort of thing it is; the difference between a collection of bricks and a house is that the latter has the requisite form of a house.

Functionalism: The thesis that the right understanding and explanation of any process is in terms of the functions it serves; two systems are relevantly the same when they perform the same functions.

Gödel's theorem: The proof that any mathematical system of sufficient power to include arithmetic depends on axioms that cannot be proved within the system itself. All such systems are, in this sense, "incomplete."

Identity of indiscernibles: A version of Leibniz's law of identity; X and Y are identical to the extent that they cannot be discerned separately.

Identity thesis: An attempted solution to the mind-body problem based on the thesis that events in the brain do not cause mental events, but that the latter are simply the brain events themselves.

Imitation game: Alan Turing's term for the game used to show that a concealed machine, properly programmed to answer questions, will be indistinguishable from an intelligent human being.

Incorrigibility: Not subject to correction; our statements about our own pains and perceptions are said to be "incorrigible" in this sense.

Intentionality: A term indebted to Franz Brentano and capturing the "aboutness" of mental events. Thoughts, beliefs, and desires are "about" their objects in a way that distinguishes such mental events in principle from anything that is merely physical.

Intuitive: A mode of knowing that results in immediate and certain knowledge; for example, that up is not down, black is not white, and so on.

Laplace's demon: That super mind postulated (as impossible) by Laplace that, equipped with knowledge of the position and momentum of every particle of physical reality and the laws of physics, would unerringly predict all future events.

Law of contradiction: It is not possible for something to be and simultaneously not be.

Leibniz's law: See **Identity of indiscernibles**.

"Mary" problem: A thought experiment devised by Frank Jackson in an attempt to show that qualia exist and to refute the physicalist solution to the mind-body problem.

Material cause: In Aristotle's sense, the matter that is necessary for something to be the subject of other causal influences.

Mental causation: Thoughts, desires, and motives bringing about musculo-skeletal events.

Metaphysics: The branch of philosophy addressed to the question of *being* as such; the ultimate contents of reality.

Monistic idealism: The thesis that the ultimate reality is not physical but mental and that the physical world, to the extent that it exists at all, does so in a totally mind-dependent fashion.

Monistic materialism: The contents of reality include only material kinds of "stuff."

Natural language: In Thomas Reid's philosophy especially, the classification of innate tendencies toward facial expressions, posture, vocalization, and so forth as "natural" means of communication.

Nominal essence: John Locke's term for the properties of reality that we assign to things in the form of names; a reality reflecting not the core physicality of things but our conventional ways of knowing them. See **Real essence**.

Obsession: A neurotic state of mind in which an idea or theme cannot be abandoned.

Ontology: The philosophical specialty devoted to the question "What is there?"

Other minds: An issue in philosophy arising from the fact that we have direct access only to our own thoughts, desires, and beliefs and, therefore, that our assumption that there is any mind other than our own is groundless.

Persistent vegetative state (PVS): A term in clinical neurology referring to patients in an enduring comatose state, with ostensible signs of mental life but with these being purely reflexive.

Personal identity: The conventional term in philosophy referring to the continuity of one's "self" over time and amidst all the biological and social changes to which one is subject.

Philosophy of mind: The branch of philosophy devoted to an examination of the "mental" and its relation to the balance of reality.

Physicalism: A thesis committed to the position that all reality is, at base, physical.

Pragmatism: A standard less of "truth" than of the grounds on which claimed truth finds its ultimate test.

Qualia: The subjective properties of things as experienced, for example, tastes, colors, sounds.

Quantum physics: The physics of the ultimately small and the forces governing their behavior.

Real essence: Locke's term for what is the ultimate reality of a thing; its invisible but physical substrate having nothing in common with that "nominal essence" assigned to it by convention.

Realism: In philosophy of science, the thesis that the laws of science are accurate and direct depictions of reality as such.

Rigid designator: Saul Kripke's term for an ontological status that obtains "in all possible worlds." *Pain* is a rigid designator in that, wherever it is and under any set of descriptions or contexts, pain is pain.

Romantic rebellion: A term referring to the aesthetic reaction against the claims and methods of science.

Solipsism: The thesis that all reality, known as it is solely through one's own mental representations, exists only in these representations.

Somnambulism: Sleep-walking.

Spandrel: In architecture, a feature not intended to function in the way it does but brought about by the arrangement of other structures having a definite function. Within evolutionary theory, the metaphor of the spandrel is intended to convey the appearance of something not itself "selected" but arising from features that were.

Substance: In Aristotle's phrase, "that which is peculiar to it, which does not belong to anything else … substance means that which is not predicable of a subject."

Supervenience: A term covering the sequential dependencies from the most basic to the most apparent; thus, tables supervene on wood; wood supervenes on molecules; molecules on atoms, and so forth. On this account, the mental allegedly supervenes on the physical.

Zombies: In philosophy of mind, a term for the "undead"; philosophical zombies are resourceful, though lacking all consciousness and self-awareness.

Biographical Notes

Aristotle (384–322 B.C.): Ancient Greek philosopher and scientist whose concepts form the basis for much of the Western intellectual tradition.

George Berkeley (1685–1753): A master of optics and mathematics and defender of a form of *idealism* designed to defeat the skeptical implications of materialism. Late in life, he would be appointed bishop of Cloyne.

Ned Block (1942–): Harvard trained, former Chair of Philosophy at the Massachusetts Institute of Technology (MIT), and now Professor of Philosophy at New York University (NYU). Opposed the Turing-test approach to establishing the nature of intelligent behavior; an important contributor to issues involving computational models of human cognition.

David Bohm (1917–1992): A contributor to the Manhattan Project, an associate of Albert Einstein, and a major figure in the development of quantum physics. The charged political climate of the 1950s resulted in his self-exile from the United States.

Franz Brentano (1838–1917): Devoted his early scholarship to the philosophy of Aristotle and was later ordained as a Roman Catholic priest. He abandoned the priesthood over the issue of papal infallibility. His *Psychology from an Empirical Standpoint* (1874) developed the concept of "intentionality" as it would figure in philosophy of mind.

Paul Churchland (1942–): Professor of Philosophy at the University of California, San Diego, Churchland is the quintessential physicalist in the matter of mind and mental life, ardently advocating nothing less than a "neurophilosophy."

Donald Davidson (1917–2003): Among the most influential figures in contemporary philosophy, known especially for his analysis of causal explanation and the putative distinction between "reasons" and "causes" as accounts of significant human actions. His academic life was spent at Stanford (1951–1967), Princeton (1967–1970), the Rockefeller University (1970–1976), the University of Chicago (1976–1981), and from 1981 until his death, Berkeley.

Daniel Dennett (1942–): University Professor and Austin B. Fletcher Professor of Philosophy and Director of the Center for Cognitive Studies at Tufts University. Dennett is one of the half-dozen most influential writers and theorists in the field of cognitive neuroscience.

Kurt Gödel (1906–1978): Czech-born and one of the leading logicians of the modern age, publishing his incompleteness theorem at the age of 25. For many years, a Fellow of Princeton's Institute for Advanced Study.

Thomas Hobbes (1588–1679): Educated in classics at Oxford; a royalist during the Cromwell years and, thus, exiled in Holland. Impressed by the works and research of Galileo, Hobbes developed political and moral theories based on physical-mechanistic principles, and he examined the nature of political organization within a mechanistic framework.

David Hume (1711–1776): Arguably, the most influential philosopher in the English language. He did much to absorb traditional problems in epistemology and ethics into an essentially psychological framework, arguing against the objectivity of morals and the proposition that "truth" is ever independent of the mental processes devoted to its discovery.

Edmund Husserl (1859–1938): German philosopher, student of Brentano, and recognized father of phenomenology. His treatise on ideas (*Ideen*) in 1913 drew attention to the difference between the act of consciousness and the phenomena toward which it is directed.

Frank Jackson (1943–): Currently Distinguished Professor at the Australian National University, where he has also served as Professor of Philosophy and head of the Philosophy Program, as well as Director of the University's Institute of Advanced Studies.

William James (1842–1910): American philosopher and psychologist who led the movements in pragmaticism and functionalism. His functionalistic approach to the mind-body problem is seen by some as the culmination of late-19th-century thought on the subject.

Garry Kasparov (1963–): Russian-born grandmaster and, at the time of the match with Deep Blue, widely regarded as perhaps the greatest chess player of all time.

Saul Kripke (1940–): Influential philosopher of language; Professor of Philosophy at Princeton and at the City University of New York; offers formidable arguments against mind-brain identity theories.

Pierre-Simon Laplace (1749–1827): Precocious in his mastery of mathematics, Laplace would find a tutor and supporter in d'Alembert in Paris, where, over a course of years, he contributed significantly to several branches of mathematics, especially differential equations and probability theory.

Gottfried Leibniz (1646–1716): German philosopher and logician, inventor (independently of Newton) of differential and integral calculus, and a significant contributor to the widest range of intellectual and philosophical issues.

John Locke (1632–1704): Oxford educated, a father of British empiricism, a major influence on the philosophical and political thought of the modern world, and a talented physician.

Thomas Nagel (1937–): Born in Belgrade, earned the B. Phil. from Oxford and Ph.D. from Harvard. Currently Professor of Philosophy and of Law at New York University.

Roger Penrose (1931–): Fellow of Balliol College, Oxford, knighted for his contributions to mathematics and science; brother of Jonathan Penrose, a chess grandmaster. His arguments against the possibility of today's physics to explain consciousness are profound, influential, and controversial.

Thomas Reid (1710–1796): Father of Scottish "Common Sense" philosophy; best known for his anti-skeptical arguments against David Hume and as a staunch defender of the methods of Bacon and Newton in approaching the nature of mind and mental life.

Oliver Sacks (1933–): London-born and Oxford-trained neurologist, Sacks has written with cogency and poignancy on the complexities of brain function as revealed in the neurology clinic.

Jean-Paul Sartre (1905–1980): Earned a doctorate in philosophy and became a leader of the existentialist school of philosophy. His essays and fiction did much to underscore the separation between an indifferent world external to consciousness and the lived life afforded by this recognition.

John Searle (1932–): U.S.-born Oxford Rhodes Scholar and Oxford-trained philosopher; Slusser Professor of Philosophy at Berkeley; a major figure in philosophy of mind and philosophy of language.

J. J. C. Smart (1920–): Scots-Australian, Oxford-educated philosopher and member of the philosophy faculty of the Australian National University; a significant figure in philosophy of mind, defending a form of mind-brain "identity."

Alan Turing (1912–1954): British mathematician, one of the fathers of today's computer science, and one of the great "code breakers" of World War II. His original contributions to mathematics were foundational for today's information sciences.

Johann Wolfgang von Goethe (1749–1832): Genius of Romantic intellectualism, his philosophical contributions laid the foundation for later developments in German idealism and phenomenology; acute commentator on scientific matters and early advocate of evolutionary principles.

Eugene Wigner (1902–1995): Winner of the 1963 Nobel Prize in Physics; born in Budapest, Hungary, he became a U.S. citizen in 1937 and was, for many years, Thomas D. Jones Professor of Mathematical Physics at Princeton University.

Ludwig Wittgenstein (1889–1951): A genius of the modern world, Austrian- born to one of the wealthiest of Viennese families. His years at Cambridge with Bertrand Russell found him challenging the received wisdom of traditional philosophy, his major tool of criticism being a special form of linguistic analysis.

Bibliography

Essential Reading:

Albert, D. *Quantum Mechanics and Experience*. Cambridge: Harvard University Press, 1992. To quote the book jacket: "Ever since physics first penetrated the atom, what it found there has stood as a radical and unanswered challenge to many of our most cherished conceptions of nature."

Armstrong, D. M. *A Materialist Theory of the Mind*. London: Routledge, 1968. This is a now "classic" defense of philosophical materialism and a criticism of alternative perspectives.

Avramides, A. *Other Minds*. London: Routledge, 2001. Perhaps the best recent work on the subject. The author carefully sets forth the problem of other minds, various attempts at a solution, and a cogent "commonsense" defense.

Block, N., O. Flanagan, and G. Güzeldere, eds. *The Nature of Consciousness: Philosophical Debates*. Cambridge: MIT Press, 1997. A number of the major issues covered in these lectures are addressed in the chapters of this book.

Bohm, D., and B. J. Hiley. *The Undivided Universe*. London: Routledge, 1993. The authors "undivide" the universe by attempting to account for consciousness and the "mental" in a manner compatible with quantum physics.

Chalmers, D. J. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press, 1996. David Chalmers provides an extensive bibliography and thoughtful discussions of central issues in philosophy of mind. It is perhaps the best introduction to the subject.

Güzeldere, Güven "Three Ways of Being a Zombie." Presented at the University of Arizona conference, *Toward a Science of Consciousness*, April 8–13, 1996, Tucson, AZ. www.conferencerecording.com/conflists/tsc96.htm. The author reviews various types of zombie: behavioristic, functionalistic, and physical clone. The assets and liabilities of each model are discussed in relation to various philosophies of mind.

Heil, John, and Alfred Mele, eds. *Mental Causation*. Oxford: Clarendon Press, 1993. Chapters in this volume clearly examine and assess the difficulties encountered in attempts to explain mental-physical causal relationships.

Hurley, S. *Consciousness in Action*. Cambridge: Harvard University Press, 1998. This major work by Susan Hurley is quite technical in places but draws attention to a number of the subtler aspects of the problem of consciousness.

Jackson, F. "What Mary Didn't Know." *Journal of Philosophy* 83 (1986): 291–295. This is the article by Jackson that more or less put the ball in play.

James, William. "Does Consciousness Exist?" *Journal of Philosophy, Psychology, and Scientific Methods* 1 (1904): 477–491. (Easily obtained either on the Internet or in anthologies of William James's major works). James in this essay rejects "consciousness" as an entity and defends it as a function—the function of "knowing."

———. *Principles of Psychology* (1890). Cambridge: Harvard University Press, 1983. Perhaps the finest and surely the most well written introduction to the subject.

Jennett, B. *The Vegetative State: Medical Facts, Ethical and Legal Dilemmas*. New York: Cambridge University Press, 2002. A landmark exploration of the state of wakeful unconsciousness; the author is a noted neurosurgeon.

Kim, J., ed. *Supervenience and Mind: Selected Philosophical Essays*. Cambridge: Cambridge University Press, 1993. Kim has been the leading defender of supervenience theory as a solution to the mind-body problem. This volume includes his most influential work.

Levine, J. *Purple Haze: The Puzzle of Consciousness*. New York: Oxford University Press, 2001. Levine discusses the "explanatory gap," underscoring the peculiar nature of consciousness in an otherwise mindless world of physical objects.

Lockwood, M. *Mind, Brain, and the Quantum: The Compound 'I'*. New York: Oxford University Press, 1989. Michael Lockwood offers an original application of quantum physics to philosophy of mind. It is the best book on this subject.

McGinn, C. *The Mysterious Flame: Conscious Minds in a Material World*. New York: Basic Books, 1999. Entirely accessible to a general audience, this book regards consciousness as simply being beyond the reach of scientific or philosophical explanation. The book is all the more compelling because its author is an accomplished philosopher.

Nagel, T. "What Is It Like to Be a Bat?" *Philosophical Review* 4 (1974): 435–450. Reprinted in *Mortal Questions* (Cambridge University Press, 1979). This proved to be one of the seminal articles on the core issue of consciousness. It is among the most cited articles in contemporary philosophy.

Penrose, R. *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics.* Oxford: Oxford University Press, 1989. Penrose here develops the strongest arguments against computational models of mental life.

Poland, J. *Physicalism: The Philosophical Foundations*. Oxford: Clarendon Press, 1994. A good, clear, and generous exposition of physicalism.

Robinson, D. N., ed. *The Mind (Oxford Readers)*. New York: Oxford University Press, 1998. Dozens of original articles appear in this anthology, with extensive introductions to the major sections.

Sacks, O. *The Man Who Mistook His Wife for a Hat.* New York: Simon & Schuster, 1985. A handful of astonishing cases from the neurology clinic and a gentle reminder of the complexity of mental life as actually lived.

Searle, J. R. *The Rediscovery of the Mind*. Cambridge: MIT Press, 1992. In this work, John Searle challenges any number of computational approaches and presents his own neuropsychological position on mind and mental life.

Strawson, G. *Mental Reality*. Cambridge: MIT Press, 1994. This is a vigorous defense of a radically materialistic philosophy of mind.

Supplementary Reading:

Andrews, Keith, Lesley Murphy, Ros Munday, and Clare Littlewood. "Misdiagnosis of the Vegetative State: Retrospective Study in a Rehabilitation Unit." *British Medical Journal*, 313 (1996): 13–16. This research leads to the conclusion that, in the diagnosis of persistent vegetative states (PVS), errors are quite common, even on the part of otherwise well-trained and experienced physicians.

Armstrong, D. M., and N. Malcolm. *Consciousness and Causality: A Debate on the Nature of Mind*. Oxford: Blackwell, 1984. David Armstrong takes a committed materialist position on consciousness. Norman Malcolm, in the patrimony of Wittgenstein, offers constructive challenges.

Baron-Cohen, S. *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge: MIT Press, 1997. Here, the reader confronts the phenomena arising from autism and the bearing of such clinical conditions on our understanding of mental life.

Bealer, G. "Self-consciousness." *Philosophical Review* 106 (1997): 69–117. This is a closely argued account of the nature of self-consciousness and the problems it presents to philosophy of mind.

Carruthers, P. *Phenomenal Consciousness: A Naturalistic Theory*. New York: Cambridge University Press, 2000. The author attempts to absorb qualia into a larger naturalistic framework, rendering a scientific approach to the problem at least plausible.

Catalano, J. *Thinking Matter: Consciousness from Aristotle to Putnam and Sartre*. New York: Routledge, 2000. As the subtitle suggests, philosophers from ancient Greece to our own time have wrestled with the question of how a material entity comes to have conscious life.

Crick, F., and C. Koch. "Why Neuroscience May Be Able to Explain Consciousness." *Scientific American* 273 (6): 84–85 (1995). Francis Crick is a Nobel laureate. In this article, he and Koch offer what might be regarded as an especially optimistic view on the extent to which science will explain consciousness.

Dennett, D. C. "The Unimagined Preposterousness of Zombies." *Journal of Consciousness Studies* 2 (1995): 322–326. Whether or not there can be zombies is a serious question. Dennett lays down interesting skeptical arguments.

Flanagan, O. J. *Consciousness Reconsidered*. Cambridge: MIT Press, 1992. Flanagan leaves no doubt about his own position: Consciousness is real, plays an important causal role, and resides in the brain.

Foss, J. *Science and the Riddle of Consciousness: A Solution*. Amsterdam: Kluwer, 2000. Yet another attempt to demystify consciousness and render it accessible to scientific research and theory.

Güzeldere, Güven. "Varieties of Zombiehood." *Journal of Consciousness Studies* 2 (1995): 326–333. In this article, the various candidate zombies are described, and their place within philosophical discourse is nicely reviewed.

Harnad, S. "Why and How We Are Not Zombies." *Journal of Consciousness Studies* 1 (1994): 164–167. Another serious if not fully successful critique of the "zombie" issue within philosophy of mind.

Heisenberg, W. *Physics and Philosophy*. New York: Harper and Row, 1958. Here, one of the great scientists of the 20th century considers the philosophical problems faced by modern physics and the implications to be drawn from these problems.

Hobbes, Thomas. *Leviathan* (1651). New York: Touchstone, 1997. One of the most important works of philosophy, this book was one of the earliest attempts to understand politics and society in their modern form. Hobbes's discussions on the nature of the individual are most pertinent to this course.

Horner, Philip J., and James H. Gage. "Regenerating the Damaged Central Nervous System." *Nature* 401 (2000). Recent evidence is offered here of the ability of central nervous system neurons to regenerate after injury.

Kim, J. "Mental Causation and Consciousness: The Two Mind-Body Problems for the Physicalist." In *Physicalism and Its Discontents*, edited by C. Gillett and B. Loewer. Cambridge: Cambridge University Press, 2001. An in-depth discussion of the "relationship between mind and matter," focused on mental causation and consciousness.

Kirk, R. *Raw Feeling: A Philosophical Account of the Essence of Consciousness*. New York: Oxford University Press, 1994. Kirk is interesting in that he is prepared to accept the mental but not as fundamental. He defines the physical as whatever physics finally includes as physical theory.

Locke, John. *An Essay Concerning Human Understanding.* Edited by P. Nidditch. Oxford: Clarendon Press, 1975. One of the classic works in philosophy of mind; a forceful defense of empirical theories of mind.

Lycan, W. G. *Consciousness and Experience*. Cambridge: MIT Press, 1996. A strong advocate of the "representational" theory of brain function, concluding that, "once representation itself is (eventually) understood, then not only consciousness in our present sense but subjectivity, qualia, 'what it's like,' and every other aspect of the mental will be explicable in terms of representation together with the underlying, functionally organized neurophysiology."

Nelkin, N. *Consciousness and the Origins of Thought*. Cambridge: Cambridge University Press, 1996. Defended here is the view of consciousness as an instance of "higher order" thought, consciousness as the system's reflections on its own operations.

O'Shaughnessy, B. *Consciousness and the World*. New York: Oxford University Press, 2000. The author says in his introduction, "Consciousness has from the start an appointment in the concrete, with the World in its ultimate physical form"; he then proceeds to integrate consciousness with perception and attention within the framework of adaptation.

Perry, J. *Knowledge, Possibility, and Consciousness*. Cambridge: MIT Press, 2001. A defender of physicalism, Perry advances a theory of "antecedent physicalism'' and deploys it against various challenges to physicalism, including the zombie argument.

Robinson, H., ed. *Objections to Physicalism*. New York: Oxford University Press, 1993. Very good essays on the central postulates of physicalism and the challenge to them arising from mental life.

Sartre, Jean-Paul. *Being and Nothingness.* Translated by H. E. Barnes. New York: Philosophical Library, 1956. This is the "bible" of existentialism and a somewhat poetic precis on the limits of understanding.

Seager, W. E. *Theories of Consciousness: An Introduction and Assessment*. London: Routledge, 1999. This is a very useful introductory treatise. Some philosophical sophistication is required, but the presentations are sound and thoughtful.

Siewert, C. *The Significance of Consciousness*. Princeton: Princeton University Press, 1998. The author takes consciousness seriously. He challenges theories that equate consciousness with a functional role or with information processing. He does not underestimate the challenges posed by qualia and by the authority of first-person reports. He usefully includes clinical findings.

Turing, A. M. "Computing Machinery and Intelligence." *Mind* 49 (1950): 433–460. This is the "classic" defense of artificial intelligence by one of the true pioneers.

Tye, M. "Phenomenal Consciousness: The Explanatory Gap as a Cognitive Illusion." *Mind* 108 (1999): 705–725. Tye argues that there actually is no explanatory gap between consciousness and the physical once our conceptual house is put in order.

