1-800-TEACH-12

1-800-832-2412

The Teaching Company®

7405 Alban Station Court, A-106

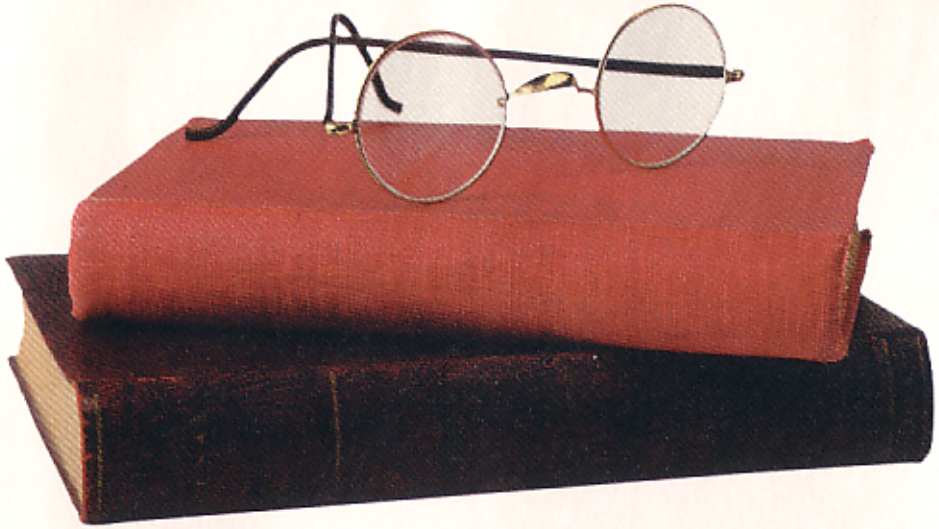
Springfield, VA 22150-2318

The Great Courses on Tape™

Teaching that engages the mind™

The Philosophy of Mind

John R. Searle, D.Phil.



COURSE GUIDE

THE TEACHING COMPANY®

John R. Searle, D. Phil.

University of California, Berkeley

John R. Searle is one of America's leading philosophers. He is the Mills Professor of Philosophy at the University of California, Berkeley, where he has been a member of the faculty since 1959. He attended the University of Wisconsin, where he was awarded a Rhodes Scholarship to Oxford in his junior year. He received his B.A. (1955), M.A., and D. Phil. (1959) degrees from Oxford. He was a lecturer in Oxford from 1956 to 1959, prior to accepting his appointment in Berkeley; and he has held visiting professorships in the United States, England, Canada, Norway, Germany, Italy, France, and Brazil. Professor Searle's books include *Speech Acts*, *Expression and Meaning*, *Intentionality, Minds, Brains and Science, The Rediscovery of the Mind*, and *The Construction of Social Reality*. In 1984, he was the Reith Lecturer on the BBC, and he has appeared frequently on U.S. public television. He has been a Guggenheim Fellow and twice won a Fulbright Award. In 1990 he was President of the American Philosophical Association. He has been a member of the American Academy of Arts and Sciences since 1977. He is currently a council member of the National Endowment for the Humanities. His works have been translated into twenty languages.

©1998 The Teaching Company Limited Partnership

The Philosophy of Mind
Table of Contents

©1998 The Teaching Company Limited Partnership

Course Scope:

In any intellectual era there are certain overriding problems that form the horizon of intellectual life. In our era the leading problem is how to account for our commonsense conception of ourselves as conscious, free, mindful, rational agents in a world consisting entirely of mindless, meaningless, blind, purposeless physical particles in fields of physical force.

This problem is manifest in countless ways. How do the social sciences relate to the natural sciences? What is the nature of mental illness and how does it relate to physical illness? Can we really discover laws of human behavior analogous to the laws of physics? The central problem area, however, lies in the philosophy of mind. What is the nature and structure of the human mind and how does it relate to the rest of reality? The aim of this course is to introduce the student to some of the leading issues in the philosophy of mind and above all to enable the student to think about these problems for himself or herself.

©1998 The Teaching Company Limited Partnership

Lecture One
Dualism, Descartes' Legacy

Scope: To understand contemporary discussions in the philosophy of mind, we need to know something of their ancestry. The modern conception of the philosophy of mind begins with the work of René Descartes (1596-1650), a French philosopher who articulated a crucial metaphysical distinction between two kinds of substances, mental and physical. This view, called "dualism" or sometimes "psycho-physical dualism," survives to the present day. It is, however, subject to decisive objections. Historically, those who reject it typically adopt some version of "monism," the view that there is only one kind of substance in the universe. The two most common monisms are idealism, the view that only mental reality exists, and materialism, the view that only matter exists. Today materialism is dominant in science as well as in philosophy, and the history of the philosophy of mind is in large part an attempt to get a version of materialism which is not subject to decisive objections.

Outline

I. Descartes' Dualism

A. There are two kinds of substances in the world, mental and physical.

1. The essence of the mental is "thinking" (= consciousness).

2. The essence of the physical is extension (= having spatial dimension).

B. The mental and the physical have other distinguishing features.

1. Minds are indivisible (hence indestructible), while bodies are infinitely divisible.

2. Minds are free, while bodies are determined.

3. Minds are known directly, by way of "cogito, ergo sum." Bodies are known indirectly.

C. Six Difficulties with Cartesian Dualism

1. The most important problem is interaction: How can the mind and the body ever interact? How can one causally affect the other?

2. Freedom of the will. If the mind is free, but the body is determined, it looks as if the freedom of the mind makes no difference.

3. Other minds. How is it that I can know that other people have minds, since the only mind to which I have direct access is my own mind?

4. Skepticism in general. If I am locked in my own experiences, how can I ever really know anything of the external world?

5. Sleep. How is it possible that people can be totally unconscious, if a person consists of a mind, and mind is essentially conscious?

©1998 The Teaching Company Limited Partnership

6. Animals. Animals behave as if they had minds, but if so they would have to have immortal souls because minds are indestructible.

D. There have been various attempts to solve these problems within the Cartesian framework. All of these attempts have failed.

Recommended Reading:

Searle: *Minds, Brains and Science,* (Chapter 1)

Descartes: from *Meditations on First Philosophy*, Meditations II and VI. (in Rosenthal, ed.)

Ryle, G. *The Concept of Mind*, Chapter 1, "Descartes' Myth" (in Rosenthal, ed.)

Questions to Consider:

1. What is Cartesian dualism?
2. What sorts of problems does dualism pose for a theory of mind?
3. How do dualists attempt to solve these problems?

©1998 The Teaching Company Limited Partnership

Lecture Two

Alternatives to Dualism: Materialism and Its Discontents

Scope: This lecture discusses the history of doctrines in the philosophy of mind that have constituted a response to Cartesian dualism. It is generally assumed that property dualism encounters many of the same difficulties as substance dualism and for that reason is not acceptable. That leaves monism with two possibilities: idealism and materialism. Idealism was influential up to and through the 19th century, but given the enormous success of the physical sciences, it has not seemed an acceptable option in the middle and later parts of the 20th century. Materialism seems inevitable but unattractive. This lecture is mostly about the recurring difficulties with materialism.

Outline

I. Alternatives to Cartesianism—Property Dualism and Varieties of Monism

- **A.** Property Dualism
  - **1.** Descartes was wrong to think that there are two kinds of substances. But there are indeed two kinds of properties, mental and physical properties.
  - **2.** One and the same body can have both mental and physical properties.
- **B.** Varieties of Monism. Within monism, we need to distinguish between those monists who think everything is mental (idealists), and those who think everything is material (materialists).
- **C.** Behaviorism: Logical and Methodological Behaviorism distinguished.
  - **1.** Methodological behaviorism says we should study behavior as a scientific project in psychology.
  - **2.** Logical behaviorism says that any statement about the mind is equivalent in meaning to a set of statements about behavior.

II. The Failure of Logical Behaviorism

- **A.** Three objections can be raised against logical behaviorism.
  - **1.** There appears to be a kind of circularity. Beliefs can only be analyzed presupposing desires, but desires can only be analyzed presupposing beliefs, for example.
  - **2.** Behaviorism leaves out the causal component in the relation of the mental to the physical.
  - **3.** There are counterexamples of the superactor-superspartan variety. We can imagine someone who acts exactly as if he were in pain, without actually being in pain, and someone who can have a pain without ever manifesting that pain in behavior.

©1998 The Teaching Company Limited Partnership

- **B.** The most important objection arises from common sense. Behaviorism denies what we all know: We all have subjective con-scious mental states; and these are quite different from our behavior.

III. Physicalism, the theory that mental states are identical to brain states, represents an attempt to improve on behaviorism.

- **A.** Type-Type Identity Theories: Every type of mental state is identical with a type of brain state.
- **B.** Weaknesses of Identity Theories
  - **1.** If there really is a contingent identity between the mental and the physical, then there must be two different sets of features to nail down the identity. But that is property dualism. Attempts to answer this were not successful.
  - **2.** Neuronal Chauvinism. It seems too neuronally chauvinistic to suppose that every *type* of mental state must be identical with a certain *type* of neuronal physical state. It seems more natural to suppose that every *token* mental state must be identical with some *token* physical state. So, type-type identify theory led naturally to token-token identity theory.
  - **3.** The common sense objection that we made to behaviorism also applies to type-type identity theory. It seems to leave out the mind.
- **C.** Token-token identity theories have many of the advantages of type-type identity theories without many of the disadvantages. But they raise an unanswered question: What do two physical states have in common if they are different physical states, but the same mental state? The answer to this leads to functionalism: They perform the same function in the life of the organism.

IV. Three influential arguments have been advanced against identity theories of any kind.

- **A.** Thomas Nagel: What it is like to be a bat
- **B.** Frank Jackson: What Mary Knew
- **C.** Saul Kripke: Necessary Identities

V. The functionalist defines mental states in terms of causal relations. Mental states such as beliefs and desires are defined in terms of causal relations between the external input to the system, the internal causal relations among the elements of the system, and the causal output of the system.

- **A.** Advantages of Functionalism: Unlike behaviorism, we now have the causal element in the system. Both beliefs and desires can be explained in terms of causation.
- **B.** Black Box Functionalism and Computer Functionalism Distinguished
  - **1.** Black box functionalism treats the brain as a black box, and it has no theory as to the internal processing.

©1998 The Teaching Company Limited Partnership

2. Computer functionalism says that the internal processing of the box consists in computations. Computer functionalism is the same as Strong Artificial Intelligence.

VI. I summarize the pattern of materialist analyses through behaviorism, functionalism, and computer functionalism. I point out that the general pattern was to treat the problems as a series of technical questions, but each materialistic account seems to have left out some central feature about the mind, such as subjectivity, qualia, or semantic content.

Essential Readings:

Hempel, C: "The Logical Analysis of Psychology" (in Block, ed)

Armstrong, D: "The Nature of Mind" (in Block, ed)

Jackson, F: "What Mary Didn't Know" (in Rosenthal, ed.)

Kripke, S: Excerpts from "Naming and Necessity" (in Rosenthal, ed.)

Nagel, T: "What is it Like to be a Bat?" (in Block, ed)

Putnam, H: "Brains and Behavior" (in Block, ed)

Searle, J.R.: *The Rediscovery of the Mind* Chs. 1 and 2

Supplementary Readings:

Smart, J.C.C.: "Sensations and Brain Processes" (in Rosenthal, ed.)

Block, N: "Troubles with Functionalism" (in Rosenthal, ed.)

Lewis, D: "Mad Pain and Martian Pain" (in Rosenthal, ed.)

Questions to Consider:

1. What is a "type-type" mind-brain identity theory, and how does it differ from a "token-token" mind-brain identity theory?
2. What problems do identity theories, in general, have?
3. What is functionalism? What advantages does functionalism have over physicalism?

©1998 The Teaching Company Limited Partnership

Lecture Three
Strong Artificial Intelligence

Scope: Many people who work in cognitive science and in the philosophy of mind think that the most exciting idea of the past generation, indeed of the past two thousand years, is that the mind is a computer program. Specifically, the idea is that the mind is to the brain as the computer program is to the computer hardware. This view I have baptized "Strong AI," but it is sometimes called "Computer Functionalism." In this lecture I explain the appeal of this view, but I also subject it to a decisive refutation—the "Chinese Room Argument."

Outline

I. The Theoretical Basis of Strong AI
In order to explain the appeal of Strong AI, I have to introduce five somewhat technical notions.

- **A.** Turing Machines
  1. The idea of a Turing machine is an abstract, mathematical notion. For practical purposes, ordinary computers—the kind that you buy in a store—are Turing machines.
  2. The remarkable feature of a Turing machine is that it performs only four operations: Print "0;" erase "1;" print "1," erase "0;" move one square left; move one square right. Modern machines perform these operations at the rate of millions per second.
- **B.** Algorithm—an algorithm is a systematic procedure for solving a problem in a finite number of steps. Computer programs are algorithms.
- **C.** Church's Thesis—this thesis states that any algorithm can be implemented on a Turing machine. For every computable function, there is a Turing machine that can compute that function.
- **D.** Turing's Theorem—this theorem states that there is a Universal Turing machine which can simulate the behavior of any other Turing machine.
- **E.** The Turing Test—this test states that if an expert cannot distinguish the behavior of a machine from that of a human, then the machine has the same cognitive abilities as the human.

II. Strong AI
If we put all these together, we get the idea that the brain is a Universal Turing machine and that human cognitive abilities are computer programs. We test this with the Turing test and come to the conclusion that artificial intelligence is, in principle, capable of creating minds.

III. Refutation of Strong AI

©1998 The Teaching Company Limited Partnership

A. Strong AI and, with it, the Turing test, are subject to a decisive refutation, the Chinese Room Argument.

B. The Chinese Room Argument claims that a monolingual English speaker who is locked in a room with a set of computer rules for answering questions in Chinese would in principle be able to pass the Turing Test, but he would not thereby understand a word of Chinese. If the man doesn't understand Chinese, neither does any digital computer.

IV. Attacks on the Chinese Room. According to the system's reply, the man in the room does not understand Chinese, but the whole room does.

Recommended Reading:

Turing, Alan, "Computing Machinery and Intelligence" (Mind 1950, pp 433-60, reprinted in Anderson, ed)

Searle, John, *Minds, Brains, and Science*, Chapter 2

Questions to Consider:

1. What exactly is Strong AI?
2. Why do many people think it is a correct account of the mind?
3. Briefly summarize the Chinese Room Argument.
4. What is the "system's reply?" What is the answer to it?

©1998 The Teaching Company Limited Partnership

Lecture Four

The Chinese Room Argument and Its Critics

Scope: In this lecture we consider some of the implications of the Chinese Room Argument; we answer the most common arguments against it; and we end with a solution to Descartes' mind-body problem. We begin with the distinction between the problem of consciousness and the problem of intentionality. Many people in AI, even in Strong AI, concede that computers are not conscious, but they think consciousness is unimportant anyway. What matters is intentionality, and computers can have intentionality. One advantage of the Chinese room is that it does not depend on consciousness. It applies to intentionality as well.

Outline

I. Intentionality is defined, and the distinction between consciousness and intentionality clarified.

II. The Chinese Room Argument has the simplicity of an obvious counterexample, but in fact, it has a precise logical structure that can be stated in four steps.

A. Programs are syntactical.

B. Minds have semantic contents.

C. Syntax is not sufficient for semantics.

D. Therefore, programs are not minds.

III. Attacks on the Chinese Room (Continued)

A. There is a large number of attacks on the Chinese Room Argument. There must be over two hundred published attempted refutations. In this lecture, I answer the main types of these arguments.

1. The Robot Reply—If the program were implemented in an actual robot, it would understand Chinese.
2. The Brain Simulator Reply—If the program actually simulated the behavior of a Chinese brain, it would understand Chinese.
3. The "Can't Do It" Reply—In real life, it would be impossible to program a human being so that he would pass the Turing test.
4. The "Wait 'til next year" Reply—Maybe better computer technology will enable us to build thinking computers.
5. The analogy with light and electromagnetism reply
6. The "it's not really computation if done consciously" reply

B. I answer each of these in detail and discuss the implications of the debate for larger issues in the philosophy of mind.

IV. The Solution to the Mind-Body Problem

©1998 The Teaching Company Limited Partnership

A. Brains cause minds.

B. Minds are features of brains.

C. Conclusion: We should treat the mind as a biological problem, as biological as digestion.

Recommended Reading:

Searle, J.R.: "Minds, Brains and Programs" (with commentaries by critics) (in Rosenthal, ed.)

Questions to Consider:

1. The man in the Chinese Room is passing the Turing test in virtue of producing correct output behavior; i.e., presenting the right output symbol for a given input symbol. Does this show that the man understands what the output symbols mean?
2. Is thinking solely a matter of behavioral outputs?
3. How is semantics different from syntax?
4. Can minds have semantic contents that are independent of behavior?

©1998 The Teaching Company Limited Partnership

Lecture Five
Can a Machine Think?

Scope: The Chinese Room Argument was originally intended as a specific refutation of a very specific thesis in the philosophy of mind, Strong AI. However, it raises a very large number of other issues, and in this lecture I try to go through these in a systematic fashion. Could a machine think? Could an artifact think? Could we build an artificial brain, just as we have built artificial hearts? What is the significance of the Chinese Room Argument? Does it really show that "computers can't think?" What exactly of a general philosophical nature is established by the Chinese Room Argument? I begin with the possibility that "connectionist" (parallel distributed processing, neuronal net) architectures might escape the Chinese Room.

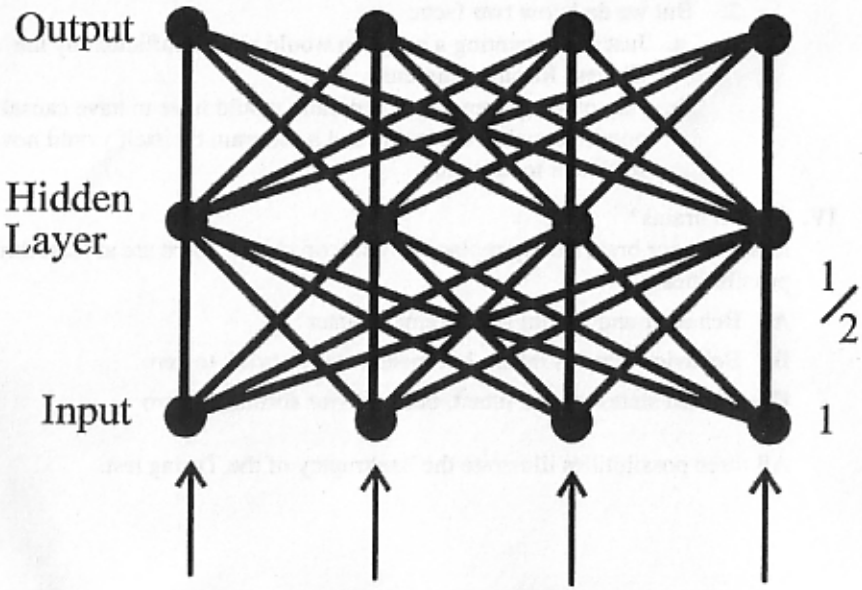
Outline

I. How do connectionist machines differ from von Neumann machines?

A. They operate by interrelated parallel processing, as shown on the diagram.

B. Does this evade the Chinese Room objection?

1. No. By Church's thesis, any computation they can perform can be performed on a Turing machine. Connectionism is faster, but it has no additional computational powers.
2. If we are talking about the physics of specific architectures, then we are doing speculative neurobiology, not computation.



©1998 The Teaching Company Limited Partnership

II. Could a machine think?

A. If by machine we mean a physical system capable of performing functions, then the brain is such a machine, so, of course, machines can think. But that is not what is meant.

B. Could an artifact think?

If you could duplicate the causal powers of the brain, then you would produce a thinking machine. It is, in principle, no more impossible to have an artificial brain than to have an artificial heart.

C. But then, why couldn't a computer think? Why not an artificial brain made of silicon chips?

Answer: Of course, the Chinese Room does not show that Something couldn't both be a computer and be thinking. Indeed, that is what our brains are. We know that our brains can compute, and we can think. Thus it follows that we are thinking computers.

D. But then, what does the Chinese Room show? Does it show that "computers cannot think?" No, what it shows is that implementing the program by itself is not sufficient for thinking. It might both implement and think, but the program is not constitutive of thinking.

III. What would a machine have to have in order to think?

A. To the three premises of the Chinese Room, let's add a fourth: Brains do it; brains cause mental phenomena. From this it follows that any other system would have to duplicate the causal powers of the brain.

B. Yes, but what are those causal powers?

1. We don't know specifically. We believe that it involves neurons and synapses, but we do not know the details.

2. But we do know two facts:

a. Just implementing a program would not be sufficient by the Chinese Room argument.

b. Any other system that could think would have to have causal powers equal to the brain, and a program by itself would not be sufficient to do that.

IV. Silicon brains?

Imagine your brain slowly replaced by silicon chips. There are at least three possibilities:

A. Behavior and mental states remain intact

B. Behavior remains intact, but mental states shrink to zero

C. Mental states remain intact, but behavior shrinks to zero

All three possibilities illustrate the bankruptcy of the Turing test.

©1998 The Teaching Company Limited Partnership

V. Biological Naturalism as a Theory of the Mind.

A. Descartes' questions can be answered by taking the mind seriously as a biological phenomenon.

B. Mental states are caused by and realized in the brain, in the same sense in which digestion is caused by and realized in the stomach and the digestive tract.

VI. Observer-independent and observer-relative features of the world can be distinguished. This is a more important distinction than are the invalid distinctions between mind and matter or machines and nature.

Questions to Consider:

1. If behavior is insufficient to ascribe thought to a system, what is the minimum requirement for the ascription of thought?
2. If we define the brain as a "physical system," is there any reason why we cannot say that machines think?
3. If brains as physical systems have the ability to think, is it possible for other physical systems to have this ability also? What is the necessary feature for a system to have the ability to think?
4. The thesis of biological naturalism is a type of materialist theory. In what way does it differ from the earlier materialist theories?

©1998 The Teaching Company Limited Partnership

Lecture Six

Is the Brain a Digital Computer?

Scope: Earlier we distinguished between Strong AI and Weak AI. Weak AI is the view that cognitive processes like other biological phenomena can be simulated computationally. Strong AI is the view that computational processes are constitutive of mental processes. The Chinese Room Argument refutes Strong AI. But there is an intermediate position between Weak and Strong AI which is immune from the Chinese Room Argument. This position states that mental processes, though not constituted by computation, nonetheless have a computational structure. Mental processes are computational processes operating over the syntactic structure of mental states so construed. On this view the brain is a digital computer, even if there is more to a mind than being a computer program. To have a label, I call this view "cognitivism." This lecture contains a discussion and a refutation of cognitivism.

Outline

I. Cognitivism Defined

- **A.** Cognitivism is the view that every mental state has a syntactical or computational structure, and that mental processes are computational processes operating over the syntactic structure of mental states.
- **B.** This view is not as clear as it could be. But it is widely held in cognitive science. Fodor holds it, for example, and it ties in naturally with the "language of thought" hypothesis.

II. The Evaluation of Cognitivism

- **A.** Cognitivism is not subject to the Chinese Room Argument because it does not say that computation is constitutive of cognition.
- **B.** Nonetheless, it is inadequate as we shall see.

III. The Distinction between Intrinsic and Observer-relative

- **A.** Fundamental to our conception of the natural sciences is that they investigate those features of nature that exist independently of human observers.
  1. Examples of observer-independent features are force, mass, gravity, and photosynthesis
  2. Examples of observer-relative features are money, property, government, and marriage.
- **B.** What about computation? Is it intrinsic or observer-relative?

©1998 The Teaching Company Limited Partnership

IV. On the standard definition, due to Turing, computation is observer-relative.

- **A.** Arguments for Multiple Realizability—the same computation can be realized in different physical structures.
- **B.** But what fact about those physical structures makes them computational? The only fact about them is that they can be assigned a computational interpretation.
- **C.** But then, computation is observer-relative. Except for those few cases of human beings carrying out conscious computations, such as adding 1 + 1 to get 2, computation does not exist in nature. Computation is not *discovered* in nature, it is *assigned*.

V. Consequences for Cognitivism

- **A.** Cognitivism is ill-defined. The question "Is the brain a digital computer?" is ambiguous. It can mean either:
  1. Is the brain intrinsically a digital computer?
     Answer: Nothing is intrinsically a digital computer. Something is a computer only relative to a computational interpretation, or
  2. Can the brain be assigned a computational interpretation?
     Answer: A computational interpretation can be assigned to anything. The answer is trivially, yes.
- **B.** The cognitivists' answer to this is to commit the homunculus fallacy.
  1. They do not take the fallacy seriously because they think it is removed by "recursive decomposition," i.e., by replacing smart homunculi with stupid ones.
  2. But this answer fails. Stupid homunculi are still homunculi.

VI. Comparison with the Chinese Room—the Chinese Room shows that semantics is not intrinsic to syntax. This argument shows that syntax is not intrinsic to physics.

Recommended Readings:

Fodor, J.: "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology." (in Rosenthal, ed.)

Searle, J.R.: *The Rediscovery of the Mind*, Chapter 9.

Questions to Consider:

1. Cognitivism is not affected by the result of the Chinese Room Argument. How does it avoid the Chinese Room Argument?
2. What is the observer-independent/observer-dependent distinction?
3. Is syntax observer-dependent or observer-independent. Why?

©1998 The Teaching Company Limited Partnership

Lecture Seven

Some Solutions to Descartes' Problems

Scope: We have now had a fairly extensive discussion of computational theories of the mind. I have, in passing, tried to present a solution to the mind-body problem: Mental states are caused by brain processes, and mental states are realized in brain structures. We can think of this on analogy with liquidity, solidity, or digestion, for example. In this lecture, I want to show how this solution to the mind-body problem deals with Descartes' problems. We will see that most but not all of his problems have a simple solution.

Outline

I. Descartes had six problems.

A. Interaction

1. Mind-body interaction can be explained in the same way that interaction between higher level and lower level features of other systems are explained. Consider the different levels of description of a piston in a car engine, for example.
2. We are able to give such descriptions because we avoid the assumption that "mind" and "body" name metaphysically exclusive categories.

B. The Problem of Other Minds. We know about other minds just not on the basis of behavior, but on the basis of the causal connections between the environment, behavior, consciousness, and underlying neurobiological mechanisms.

C. Skepticism in General

1. Descartes had a special problem because on his account all we are directly aware of are the contents of our own minds. We have to infer facts about the external world.
2. On my view, we have direct immediate awareness of the external world, and thus one form of Descartes' skepticism is blocked.
3. There are other skeptical possibilities unanswered by my view. How do we know, for example, that we are not brains in vats?

D. How is it possible to have freedom of the will?

1. Descartes does not have a solution to this problem.
2. Neither do I. In a way, my treatment of the problem is worse than Descartes', because we cannot avoid the conviction of our freedom, but yet, on my view, the mind is just a higher level feature of the brain, and the brain is just a physical system like any other.

©1998 The Teaching Company Limited Partnership

E. Sleep. On my view we can be totally unconscious and still be humans. Again, this is a consequence of avoiding the dualists' assumptions. Unconscious states are also possible.

F. Animals

1. It is obvious that at least some animals have consciousness.
2. We do not know how far down the phylogenetic scale consciousness goes.
3. We know of the existence of consciousness in animals not just on the basis of behavior, but on the basis of neurobiology.

G. The Unconscious

1. Descartes thought that unconscious mental states were impossible.
2. In the twentieth century, we have come to think that many if not most of our mental states are unconscious.
3. This view is due to Freud, but the Freudian model has been extended in cognitive science. We will not understand the unconscious until we say more about consciousness.

Recommended Readings:

Searle, J.R.: *Minds, Brains and Science,* Chs. 1 and 10

Searle, J.R.: *Intentionality,* Ch. 10.

Questions to Consider:

1. How can there be causal relations between the mental and the physical?
2. How can we tell whether animals have conscious minds?
3. In what way are the notions of "mind" and "body" not metaphysically exclusive categories?

©1998 The Teaching Company Limited Partnership

Lecture Eight
The Structure of Consciousness

Scope: In this lecture we examine the nature and structure of human consciousness. In the development of our Cartesian tradition of not treating consciousness as part of the natural physical world, there are surprisingly few efforts to examine consciousness on its own terms. If we take it seriously as a biological phenomenon, it seems to me that we need to identify its salient features. This lecture gives a survey of the most prominent ones.

Outline

I. Subjectivity—We need to distinguish ontological subjectivity from epistemic subjectivity. Consciousness has ontological subjectivity.

II. Unity

- **A.** As Kant pointed out, all of our conscious states come to us as part of a single, unified experience and not as a set of discrete experiences.
- **B.** Unity is both vertical and horizontal.
- **C.** Pathological cases such as split brains and Korsakov's syndrome show breakdowns in unity.

III. Intentionality

- **A.** Intentionality is that feature of mental states by which they are directed at or are about objects and states of affairs in the world.
- **B.** Not all conscious states are intentional. Not all intentional states are conscious.

IV. Mood

- **A.** Every conscious state involves some mood or other.
- **B.** Often these moods do not have a name.
- **C.** Moods are not to be confused with emotions.

V. Gestalt Structure. All normal, non-pathological conscious mental states are structured. The gestalt psychologists provide the best evidence for this. We need to distinguish the gestalt structure of particular states, from the general figure-ground structure of consciousness.

VI. The distinction between the center and the periphery: Attention.

- **A.** A remarkable feature of consciousness is that we can shift our attention from one area of our conscious state to another at will.
- **B.** I am not paying equal attention to every element of my conscious field.

©1998 The Teaching Company Limited Partnership

VII. Boundary Conditions. Conscious states come with a sense of their own situatedness. One has a boundary awareness of one's location in space and time as well as of certain other features of one's life and environment.

VIII. The Aspect of Familiarity

- **A.** Conscious states come to us in varying degrees of familiarity.
- **B.** This is a consequence of the aspectual shape of our intentional representations.

IX. Overflow. Conscious states typically refer beyond themselves.

X. Pleasure/Unpleasure. Conscious states are in varying degrees pleasurable or unpleasurable.

XI. Features mistakenly identified as essential to consciousness. In the standard philosophical literature since Descartes, three features of consciousness are generally identified as essential, but I think it is a mistake to see them this way.

- **A.** Self Consciousness. It is often said that all states of consciousness are self-conscious. This seems to me incorrect.
- **B.** Introspection. It is often said that we know of our conscious states by a special epistemic method of introspecting. This also seems to me incorrect.
- **C.** Incorrigibility. It is often said that our knowledge of our conscious states is such that we cannot be mistaken. This also seems to me incorrect.

Recommended Readings:

Searle, J.R.: *The Rediscovery of the Mind*, Chs. 4-6.

Questions to Consider:

1. What are the special features of self-consciousness? What reasons are there for saying that animals might have self consciousness?
2. How does the subjectivity of consciousness raise special problems for the claim that consciousness is a biological phenomenon?

©1998 The Teaching Company Limited Partnership

Lecture Nine

How to Study Consciousness Scientifically

Scope: The most important scientific problem of the present era is one that until recently most scientists did not regard as a suitable topic for scientific investigation at all. It is simply this: How exactly do brain processes cause consciousness? Given our present models of brain functioning, it would be an answer to the question, "How do lower level neuronal firings at the synaptic connections cause all of our subjective experiences?" This is one of those areas of science in which our ability to solve the scientific problem is impeded by a series of philosophical obstacles and misunderstandings. We have in previous lectures begun to see some of these misunderstandings. In this lecture I want to make them fully explicit so that we can remove them. This is one of those areas of science in which progress is impeded by philosophical errors.

There are certain general background assumptions that underlie these specific errors, and I will try to make these assumptions fully explicit as well. As with most philosophical mistakes, once you articulate the problem exactly, you can see its solution.

Outline

I. Three Background Assumptions

A. Residual Cartesian Dualism.
We still tend to think that mental phenomena in general, and consciousness in particular, are not part of the ordinary physical world in which we live.

B. The Distinction between Nature and Machine
Like the mind-body distinction, this was a useful distinction in the seventeenth century that has become an obstacle to progress in the twentieth century. The recent debates about chess-playing computers reveal the sorts of confusions we are making.

C. Many analysts fail to distinguish between those features of reality that are intrinsic or observer-independent, from those that are observer-dependent or observer-relative. It is important to see that consciousness is observer-independent. It is an intrinsic feature of reality.

II. Here are nine philosophical errors that have prevented us from getting progress on this subject matter. I try to state and expose each.

A. Consciousness cannot be defined. We do not have a satisfactory definition.
Answer: We need to distinguish analytical from common-sense definitions. Analytic definitions come at the end, not at the beginning of our investigation. We can get a common-sense definition of consciousness easily at the outset.

B. Consciousness is, by definition, subjective; science is objective, so there can be no science of consciousness.
Answer: We need to distinguish the epistemic sense of the objective-subjective distinction from the ontological sense. Consciousness is ontologically subjective, but that does not prevent an epistemically objective science.

C. We could never explain how the physical causes the mental.
Answer: We know that it happens. Our puzzles are like earlier problems in the history of science such as explaining life and electromagnetism.

D. We need to separate qualia from consciousness and leave the problem of qualia on one side.
Answer: There is no distinction between consciousness and qualia. Conscious states are qualia down to the ground.

E. Epiphenomenalism: Consciousness cannot make a difference to the world.
Answer: Consciousness is no more epiphenomenal than any other higher level features of reality.

F. What is the evolutionary function of consciousness? It plays no role.
Answer: Even our most basic activities, eating, procreating, raising our young, are conscious activities. If anything, the evolutionary role of consciousness is too obvious.

G. The causal relation between brain and consciousness implies dualism.
Answer: This objection confuses event causation with bottom-up causation.

H. Science is by definition reductionistic. A scientific account of consciousness must reduce it to something else.
Answer: We need to distinguish explanatory reductions from eliminative reductions. You cannot eliminate anything that really exists and consciousness really exists.

I. Any scientific account of consciousness must be an information processing account.
Answer: Information processing is observer-relative. Consciousness is intrinsic, observer-independent.

Recommended Readings:

Nagel, T.: "What is it like to be a bat?" *Philosophical Review,* 1974, 83, pp. 435-50. (in Block, ed.)

Searle, J.R.: "The Mystery of Consciousness," in *The New York Review of Books,* 1995, November 2, November 16, Vol. XLII, Number 17, 18.

©1998 The Teaching Company Limited Partnership
©1998 The Teaching Company Limited Partnership

Questions to Consider:

1. How does the subjectivity of consciousness affect the possibility of a science of consciousness?
2. What reasons are there for thinking that consciousness is or is not epiphenomenal?

©1998 The Teaching Company Limited Partnership

Lecture Ten

How the Mind Works: An Introduction to Intentionality

Scope: Most philosophers and cognitive scientists assume that the most important feature of the mind is its intentionality, that feature by which the mind is directed at or about objects and states of affairs in the world. But how can we explain this feature as a biological phenomenon? What is its structure? How do we fit intentionality into the overall structure of the world? Many philosophers think intentionality is really more important than consciousness, and indeed that consciousness can be dispensed with because the only phenomena that we need to explain our behavior are intentionalistic phenomena. Consciousness, on this view, is epiphenomenal.

Outline

I. Relation of Consciousness to Intentionality

- **A.** Unconscious intentional states must be the sort of thing that is accessible to consciousness.
- **B.** Chomsky's Universal Grammar and much of cognitive science violate this constraint.

II. Definition of Intentionality. Intentionality is that capacity of the mind by which it is directed at or about objects and states of affairs in the world. Not all mental states are intentional. For example, there are undirected forms of elation and anxiety. But others, such as belief and desire are always intentional. If you believe, you must believe that something is the case. If you desire, you must desire something.

III. The Structure of Intentionality. The best way to understand intentionality is to see the analogy between intentional states and speech acts. There are four points of connection:

- **A.** Structure. Both have the structure F(p) or S(p).
- **B.** Direction of fit. Both have direction of fit. There are only three directions of fit: mind-to-world, world-to-mind, and null direction of fit
- **C.** Sincerity conditions. Every speech act that has a propositional content is an expression of the sincerity condition which determines the appropriate psychological state.
- **D.** Conditions of satisfaction. Every intentional state with a direction of fit, has conditions of satisfaction. The essence of intentionality is representation of conditions of satisfaction.

©1998 The Teaching Company Limited Partnership

IV. Difficulties with the foregoing analysis; explanations of those states, such as shame and pride, that do not have a direction of fit.

V. Every intentional state only functions in relation to other intentional states. Intentional states do not determine their conditions of satisfaction in isolation.

VI. The Network and The Background. Even the Network is not enough. Every intentional state requires a set of capacities that are not themselves intentional states. There are various arguments for proving this thesis.

Recommended Reading:

Searle, J.R.: *Intentionality*, chs. 1 and 5.

Questions to Consider:

1. What is an intentional mental state and how does it differ from nonintentional mental states such as pain, undirected anxiety, etc.?
2. Intentional states have contents that are propositional in nature. In what sense can these contents be said to be true or false, satisfied or unsatisfied?
3. Intentional states are not self-interpreting. In order to be interpretable, they require a web of other intentional states, i.e., a Network. But even the Network is insufficient for the interpretability of intentional states. What is further required for their interpretability, and why?

©1998 The Teaching Company Limited Partnership

Lecture Eleven

The Structure of Action and Perception

Scope: In this lecture we consider the question of whether the theory of intentionality presented in the previous lecture can be extended to perception and action. We find that it can, but we need to introduce some new notions, particularly the notion of *intentional causation* and the notion of *self-referentiality*, specifically, the *causal self-referentiality* of many intentional phenomena.

Outline

I. The Intentionality of Intentional Action
Can we extend the analysis to intentional action? There are three puzzling counterexamples:

- **A.** Chisolm's counterexample
- **B.** Davidson's counterexample
- **C.** Bennett's counterexample

Our analysis must be able to deal with these examples.

II. A General Theory of Intentional Action

- **A.** The distinction between prior intentions and intentions-in-actions.
- **B.** The causal self-referentiality of each.
- **C.** The bipartite structure of intentional actions. Every action contains both an intention-in-action and a bodily movement or other conditions of satisfaction.
- **D.** How this analysis deals with the counterexamples of "deviant causal chains" provided by Chisolm, Davidson, and Bennett.

III. The Intentional Structure of Perception

- **A.** Perceptual states have conditions of satisfaction, direction of fit, and propositional contents. Examples are given for each.
- **B.** The notion of representation is inadequate to capture the immediacy of perceptual phenomena. The world is not just represented, but it is directly presented to us when we perceive objects, etc.
- **C.** We need the notion of the causal self-referentiality of perception. It is part of the conditions of satisfaction of my perceiving a state of affairs that there is not only the state of affairs there, but its being there causes me to perceive it.
- **D.** The form of intentional causation.

IV. Further Reflections on the Nature of the Background. Intentionality rises to the level of the background abilities.

©1998 The Teaching Company Limited Partnership

Recommended Reading:

Searle, J.R.: *Intentionality*, Chapters 2 and 3.

Questions to Consider:

1. Perceptual states such as vision are intentional states. Thus, they have structural features similar to intentional states such as beliefs and desires, and features that are different from those of beliefs and desires. In what way are the contents of perceptual states different from the contents of beliefs and desires?
2. In what way does the feature of causal self-referentiality avoid some aspects of Cartesian skepticism?

©1998 The Teaching Company Limited Partnership

Lecture Twelve
The Construction of Social Reality

Scope: One of the most remarkable capacities of the human mind is its ability to construct an objective reality of entities that in some sense exist only because we believe them to exist. I am thinking of such things as money, property, marriage, governments, and above all, language. These entities differ from mountains, molecules, and tectonic plates because socially constructed entities exist only relative to human beings, but physical reality is totally independent of human intentionality. This lecture offers a general theory of how such social constructions are possible.

Outline

I. Three Questions about Institutional Reality
- **A.** How do we avoid circularity? Does "money" mean "believed to be money?"
- **B.** Causation. How can the fragile structure of institutional reality function causally?
- **C.** What is the role of language in institutional reality?

II. The Building Blocks of Social Reality. Three notions are necessary to account for the construction of institutional facts:
- **A.** Collective intentionality—humans have the biologically given ability to act and think collectively.
- **B.** The assignment of function—humans are able to assign functions to entities. All functions are observer-relative.
- **C.** Constitutive rules—not all rules regulate. But some constitute the activity that they regulate. These always have the same form: X counts as Y, or X counts as Y in context C.

III. We consider a simple case of the construction of institutional reality.
- **A.** The Parable of the Community Boundary
- **B.** The essential feature of institutional reality is the imposition of function on entities that cannot perform the function in virtue of their physical structure but only in virtue of the collective acceptance of the function.
- **C.** We consider the example of money and the evolution of paper currency in Europe.
- **D.** A General Theory of the Structure of Institutional Facts

IV. The lecture concludes by offering answers to the three questions.

©1998 The Teaching Company Limited Partnership

Recommended Reading:

Searle, J.R.: *The Construction of Social Reality* (Chapters 1-5)

Questions to Consider:

1. Many groups of animals are said to have social facts in virtue of their ability to behave collectively. Humans are said to have a type of social fact, an institutional fact, which exists not solely in virtue of collective behavior, but in virtue of other features that are derived from human mental abilities. What are these features that are necessary for the creation of an institutional fact?

©1998 The Teaching Company Limited Partnership

Bibliography

Anderson, Alan Ross (ed). *Minds and Machines*. Prentice Hall, 1964

Block, Ned (ed). *Readings in Philosophical Psychology*, Vol. 1. Harvard University Press, 1980.

Rosenthal, David (ed). *The Nature of Mind*. Oxford University Press, 1991.

Searle, John. *Intentionality, An Essay in the Philosophy of Mind*. Cambridge University Press 1983.

_____. *Minds, Brains and Science*. Harvard University Press, 1984.

_____. *The Rediscovery of the Mind*. MIT Press, 1992.

_____. *The Construction of Social Reality*. The Free Press, 1995.

_____. "The Mystery of Consciousness" *New York Review of Books*, Nov 2, 1995; Nov 16, 1995, Vol. XLII, Number 17, 18.

©1998 The Teaching Company Limited Partnership