Minds, Brains and Science. by John Searle
Review by: William J. Rapaport
*Noûs*, Vol. 22, No. 4 (Dec., 1988), pp. 585-609
Published by: Wiley
Stable URL: http://www.jstor.org/stable/2215459

# CRITICAL REVIEWS

John Searle, *Minds, Brains and Science* (Cambridge, Massachusetts: Harvard University Press, 1984), 107pp.

WILLIAM J. RAPAPORT
STATE UNIVERSITY OF NEW YORK AT BUFFALO

## TO THINK OR NOT TO THINK

### 1. INTRODUCTION

John Searle's *Minds, Brains and Science* is an elegantly written examination of a number of major philosophical issues: the mind-body problem, the nature of human action, philosophy of social science, and the problem of free will. The clarity and informality of the exposition stems from the book's origins: it is based directly on Searle's 1984 Reith lectures, which were originally aired on the BBC and later published (in slightly different form) in *The Listener* (Searle 1984a-f). I find it difficult to imagine a lecture series of equal intellectual depth on U.S. radio. Fortunately for those of us on this side of the Atlantic, however, the lectures are now available in book form. As befits an attempt to make "the results and methods of modern analytic philosophy . . . available to a much wider audience" (p. 7), there are no footnotes, though there is a short bibliography and an index. The main focus of the book (Chapters 1-3) is the mind-body problem, which will also be the focus of my review; the discussions of action, social science, and free will (Chapters 4-6) are interesting, but will only be covered briefly here.

Searle begins by citing three "themes" that underlie his lectures:

(Theme 1) We know very little about the functioning of the human brain, despite the work of neurophysiologists, of Freud, and of research in artificial intelligence (p. 8).

(Theme 2) The conscious mind is a biological phenomenon, despite our "inherited cultural resistance" to this (p. 10).

(Theme 3) The terms 'mind' and 'science' are not well-defined enough to help us understand these themes (p. 10).

Now, one of the nicer features of the book is Searle's way with words, so I hope that I will be forgiven if I cannot find any better way of presenting some of Searle's ideas than by quoting him. Concerning Theme 1, he points out that

> the most remarkable thing about brain functioning is simply this. All of the enormous variety of inputs that the brain receives . . . are converted into one common medium: variable rates of neuron firing . . . [which] produce all of the variety of our mental life. (p. 9)

This suggests a number of interesting questions: Could this common medium serve as a biological analogue of Fodor's mentalese? Could it serve as a common language for what AI researchers call "knowledge representation"? Is Searle's "remarkable thing" perhaps a reductionistic thesis to the effect that, at bottom, all mental phenomena are physico-biological? Searle's answer, in terms of *causation* and *realization*, leads to Theme 2, which I think is almost—but not quite—right. By way of contrast, I shall sketch an alternative view of the mind as an *abstract* notion *implemented* in the physico-biological brain.

As for Theme 3, Searle gives a rough definition of 'mind' as "the sequences of thoughts, feelings and experiences, whether conscious or unconscious, that go to make up our mental life" (pp. 10-11). Ignoring the apparent circularity, the core of this definition is mind considered as sequences of thoughts, feelings, and experiences. So, if the mind is biological (as Theme 2 states), then these sequences are sequences of, say, neuron firings. Now, assuming that the sequences of thoughts, feelings, and experiences are at least *correlated* in some way with such neuron-firing sequences, we can ask whether they are *reducible* to them, or whether they are, instead, independently describable (perhaps in a computational manner) and merely *implemented* in the neuron firings. I think that Searle wants to have it both ways in *Minds, Brains and Science*, and I want to suggest that he misses a third level of description. His notion of "causation", for instance, is not exactly the ordinary notion of causation, nor is it the converse of "realization", and this obviates many of his conclusions. Moreover, current research by cognitive scientists is beginning to provide a more precise definition of 'mind' in computational terms with which the abstract view of mind is consistent. As we shall see, however, Searle thinks that cognitivism is wrong.

I have praised Searle's stylistic clarity; but I found his discussion and arguments frustrating. In this review, I shall explore Searle's presentation of the mind-body problem in some detail, followed by a briefer presentation of his discussion of action, social science, and the will. (I'll give the six original lecture titles, too; they're far more fun than the chapter titles.)

## 2. MINDS, COMPUTERS, AND COGNITIVE SCIENCE

### 2.1. The Mind-Body Problem ("A Froth on Reality")

Searle's main question in the first chapter is this:

> We think of ourselves as *conscious, free, mindful, rational* agents in a world that science tells us consists entirely of mindless, meaningless physical particles. . . . How can a mechanical universe contain intentionalistic human beings—that is, human beings that can represent the world to themselves? (p. 13)

Good question! Note that a similar question can be asked of computers: How could a mechanical computer (or a running computer program) represent the world to itself? Cognitive scientists who accept the computational metaphor, as well as researchers in the sub-area of AI called 'knowledge representation', assume that computers (or programs) *can* do this. So, will Searle's answer to his question apply to computers? Knowing Searle, we expect that it won't; so one thing to keep in mind is: Why not?

Searle announces his belief "that the mind-body problem has a rather simple solution" (p. 14), though he admits in Chapter 2 that he only provides the "outlines of a solution" (p. 28). He raises a preliminary question, consistent with Theme 2, which hints at the solution as well as at the reason that the solution won't apply to computers: "Why does the mind seem more mysterious than *other* biological phenomena?" (p. 14; italics added).

Before presenting his own analysis of mental phenomena, Searle dismisses other analyses in a typically frustrating passage:

> most of the recently fashionable materialist conceptions of the mind—such as behaviourism, functionalism, and physicalism—end up denying . . . that there are any such things as minds as we ordinarily think of them. That is, they deny that we do really *intrinsically* have subjective, conscious, mental states and that they are as real and as irreducible as anything else. . . . (p. 15)

Whether one agrees with this depends on how one "ordinarily thinks of" the mind. But I find these three theories strange bedfellows in any case. In particular, functionalism (which need not be materialist; cf. Fodor 1981) specifically allows, if not the mind as a substance, then at least the mind as mental states and processes that are *not* reducible to, but are implemented in, physical substance.

Searle's own analysis begins by considering "four features of mental phenomena" that need to be reconciled with the scientific conception of the world (again, I rely on Searle's picturesque language):

(Feature 1) *Consciousness*: "the central fact of specifically human existence" (p. 16); "How . . . could this grey and white gook inside my skull be conscious?" (p. 15).

(Feature 2) *Intentionality*: "how . . . can atoms in the void represent anything?" (p. 16).

(Feature 3) *Subjectivity*: the privacy of our mental states.

(Feature 4) *Mental causation*:. Descartes's interactionist problem—do "thoughts . . . wrap themselves around the axons or shake the dendrites or sneak inside the cell wall and attack the cell nucleus?" (p. 17).

To achieve the reconciliation, Searle presents two theses:

(Thesis 1) "[A]ll mental phenomena . . . are caused by processes going on in the brain" (p. 18).

(Thesis 2) "[M]ental phenomena just are features of the brain" (p. 19).

Thesis 2 is easy to reconcile with Thesis 1, if brain states and processes are "features" of the brain, for surely brain processes can physically cause other brain processes. But Thesis 1 seems to be the opposite side of the coin of Feature 4: After all, do neuron firings "wrap themselves around" my mental experiences or "shake" my feelings or "sneak inside" my thoughts? Descartes's problem is a hard one because nothing short of the identity thesis seems able to resolve it. So Thesis 1 cannot be *quite* right—unless, of course, the sort of causation involved is not ordinary, purely physical, causation. Searle claims that it is; but, as I have argued elsewhere, it isn't (Rapaport 1985, 1986a). Rather, the notion of causation that makes Thesis 1 true and consistent with both functionalism as well as the computational theory of abstract data types and their implementations (cf. Berztiss & Thatte 1983) is one that is more like Jaegwon Kim's "Cambridge dependency" or—even more to the point—Hector-Neri Castañeda's "consubstantiation" (Kim 1974; Castañeda 1972: 13ff; 1975a: 145f). Let me expand on this a bit.

Searle distinguishes, here as well as in his earlier writings on these topics, between two levels of physical phenomena. For example, solidity is a *macro*-level phenomenon that he says is "caused by" the *micro*-level phenomenon of the lattice structure of molecules. Therefore, according to Searle, solidity is—again in his terms—"realized in" that lattice structure (pp. 20-22). Similarly, the mind is a macro-level phenomenon "caused by" and "realized in" the micro-level phenomena of the brain.

But this relationship is *not* a causal one in any sense. It is much more like the relationship between two descriptions of the same event, or—reifying the descriptions—it is the relationship between two intensionally distinct but co-extensional event-levels (or event-"guises"). A correct picture of what is going on requires a third, *intermediate*, level. At the top-, or macro-level, there is the *abstract notion or description* of some phenomenon (say, solidity or some mental phenomenon) considered independently of lower levels. This is much like the computer-science notion of an abstract data type. At the bottom-, or micro-level, there is some physical *implementing medium*, considered independently of higher levels. And, in the middle, there is the *implemented phenomenon*. A useful analogy is this: A character in a play, say, Hamlet, corresponds to the abstract level; an actor, say, Olivier, corresponds to the bottom-level implementing medium; and

Olivier-playing-Hamlet corresponds to the intermediate level—the implemented Hamlet.

Searle's "realization" is a relation between the *top*- and bottom-levels, but his "causation" is a relation between the *intermediate* and bottom levels—and real, physical causation is a relation among events at the bottom-level only. For instance, consider Searle's analysis of an internal combustion engine in terms of causation and realization (Searle 1983; this analysis is slightly more elaborate than his analysis of solidity in *Minds, Brains and Science*). As shown in Figures 1 and 2, Searle holds that (1) bottom-level phenomena (such as electron movements) "cause and realize" top-level phenomena (such as a rise in temperature), (2) some bottom-level phenomena "cause" other bottom-level phenomena (e.g., electron movements cause hydrocarbon oxidation), and (3) some top-level phenomena "cause" other top-level phenomena (e.g., a rise in temperature causes an explosion). But this is not detailed enough. It needs to be augmented by the intermediate level and the following relationships (as shown in Figure 3):

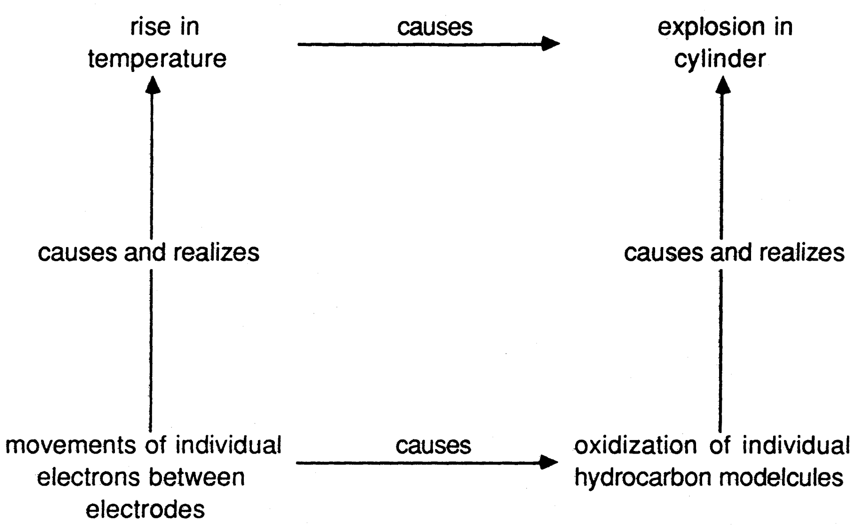


Figure 1.

Searle's analysis of his "causation" and "realization" relationships in an internal combustion engine (adapted from Searle 1983: 269).

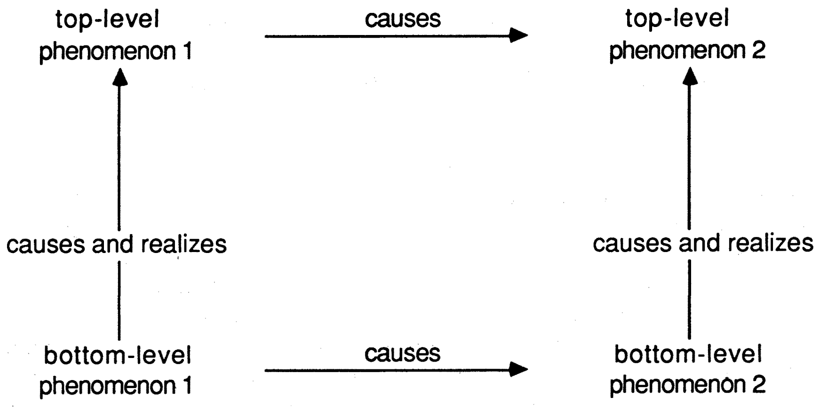


Figure 2. Searle's general analysis of causation and realization

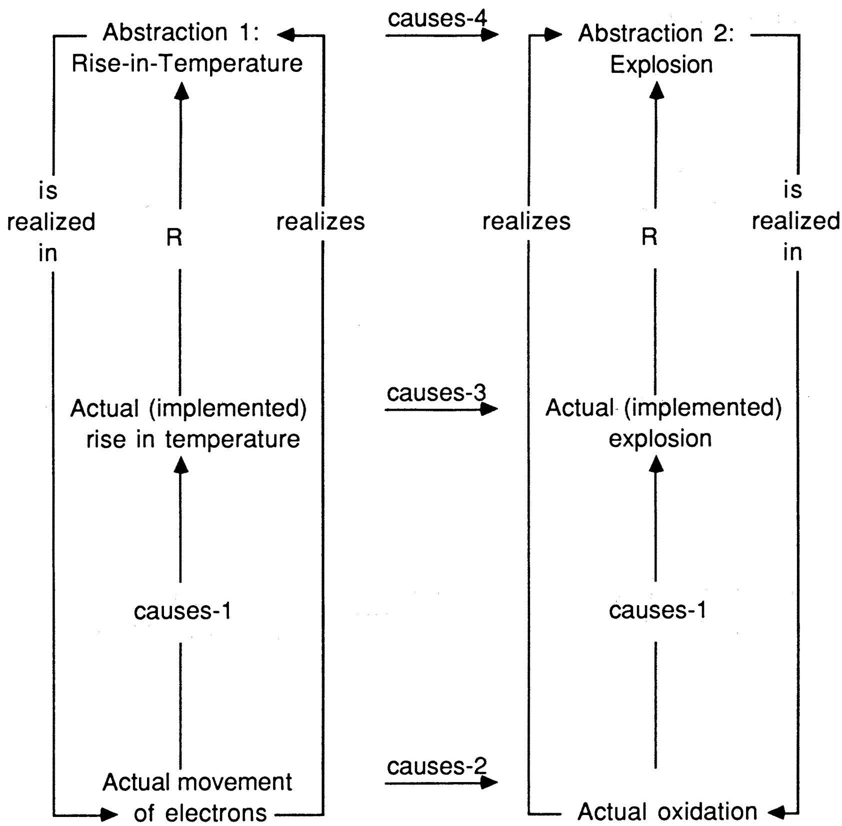


Figure 3.

The causation/realization analysis of a combustion engine using the intermediate level.

- a certain kind of relationship (causal, according to Searle) between the bottom-level phenomena and the intermediate-level phenomena—call it "causation-1";
- a certain (arguably distinct) kind of causal relationship among bottom-level phenomena—call it "causation-2";
- a certain (arguably distinct) kind of causal relationship among intermediate-level phenomena, paralleling causation-2—call it "causation-3";
- a certain (arguably distinct) kind of relationship (arguably causal) among top-level phenomena, paralleling causation-1 and causation-2—call it "causation-4";
- a certain kind of relationship between intermediate- and top-level phenomena—call it "R"; and
- the relationship of realization (or implementation) between bottom- and top-level phenomena.

How are causation-1, causation-2, causation-3, and causation-4 related, and what is R? I have proposed the following (in Rapaport 1985):

- R is the relation of *being an instance of* (i.e., falling under a species).
- Causation-1 can be defined in terms of realization and R-converse. As I mentioned earlier, it could be Kim's Cambridge dependency, although I favor interpreting it as Castañeda's relation of consubstantiation. The main point, however, is that *causation-1 is not a kind of (ordinary, physical) causation at all.*
- Causation-2, on the other hand, *is* ordinary, physical causation.
- Causation-3 is probably best defined in terms of causation-1 and causation-2, although it could be taken as an instance of causation-4 (i.e., causation-3 stands in relation R to causation-4).

Before I offer an interpretation of causation-4, note that the intermediate-level phenomenon is *produced* when a bottom-level device realizes (implements) a top-level abstraction. A general theory of the relationships of causation (Searlean and otherwise), realization (or implementation), and bottom-, intermediate-, and top-level phenomena is presented in Figure 4 and can be described (rather less compactly than in the figure) as follows:

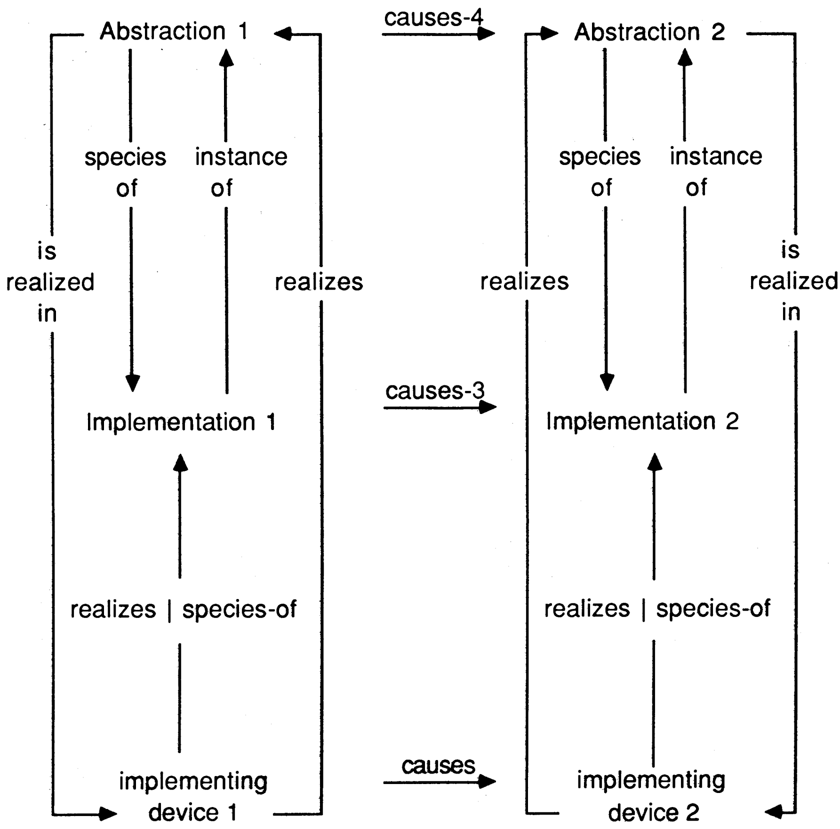


Figure 4.

The general causation/realization analysis using the intermediate level.

First, something like the familiar Aristotelian genus-species relationships hold between the intermediate- and top-level descriptions:

- An Abstraction (i.e., the abstract notion or description of some phenomenon) is a species of an Implementation (i.e., the implemented phenomenon); conversely, an Implementation is an instance of an Abstraction.

Second, realization (implementation) is a relation between bottom- and top-level phenomena:[1]

- An implementing device realizes an Abstraction; conversely, an Abstraction is realized in an implementing device.

Third, *ordinary, physical causation is properly located only at the implementation level,* since that is the level of physical events:

- One implementing device can (physically) cause events in itself or in another implementing device.

It follows that:

- An implementing device realizes a species of an Implementation.
- An implemented phenomenon (i.e., an Implementation) is an instance of that which is realized in an implementing device.
- Implementation-1 *causes-3* events in Implementation-2 iff Implementation-1 is an instance of an Abstraction, which is realized in an implementing-device-1, which, in turn, causes events in an implementing-device-2, which realizes a species of Implementation-2. (The two Implementations might be the same, and the two implementing devices might be the same.)
- Abstraction-1 *causes-4* events in Abstraction-2 iff Abstraction-1 is realized in an implementing-device-1, which causes events in an implementing-device-2, which realizes Abstraction-2. (The two Abstractions might be the same, and the two implementing devices might be the same.)

Alternatively, causation-4 could be taken as primitive—it is, in some sense, the Abstraction *Causation*—and physical causation could be defined as an implementation of it.

But let us return to Searle. How does he solve the puzzles raised by his four features of mental phenomena?

(Feature 1): Consciousness can arise from neuron firings in the same way—whatever that is—that life arises from complex biochemical processes (cf. pp. 23-24).

Searle gives a useful description of how the conscious sensation of pain arises from nerve stimulations (pp. 18-19).

(Feature 2): Atoms can have intentionality in the same way that certain biological processes can produce thirst, which is also intentional.

I find this somewhat less satisfactory, since thirst is not a case of *representational* intentionality. That is, I don't see how an explanation of the intentionality of thirst (considered as the desire to drink, or as being directed to some liquid) will help us to understand *semantic* cases of intentionality. Thirst, like pain, seems to lie more on the "body" side of the mind-body problem. Although thirst and pain are important and interesting cases of mental phenomena, I do not see them to be either relevant or paradigm cases in the present context. Searle's point, I think, is that his causation/ realization relation explains how physical things can be intentional. As I have argued, I don't think he's gotten the relations right; but I'm not sure that my suggestion of the intermediate level accounts for it, either. Rather, I don't think that intentionality can be explained along these lines at all. Of course, if Searle's point here and in the previous solution is only that these are mysteries, but not *more* so than other scientific mysteries (p. 24), then I can (grudgingly) agree with him.

(Feature 3): This is handled rather abruptly: "the existence of subjectivity is an objective scientific fact like any other" (p. 25).

(Feature 4): "[B]ecause mental states are features of the brain, they have two levels of description—a higher level in mental terms, and a lower level in physiological terms" (p. 26).

So, the lower-, or bottom-level description is used to explain mental causation. As I've indicated, this is *almost* right, but the intermediate level is needed to complete the story.

Searle sums up the chapter with his solution to the mind-body problem: *Naive physicalism* ("all that exists in the world are physical particles with their properties and relations" (pp. 26-27)) *is consistent with naive mentalism* ("mental phenomena really exist" (p. 27) and can be conscious, intentional, subjective, and causal), *and both are true*. I agree but go one step further: where Searle holds that the mind is realized in the brain, I hold that mental phenomena can be implemented in many different media, including computers—which brings us to the next chapter.

### 2.2 Can Computers Think? ("Beer Cans and Meat Machines")

In answering the question whether computers can think, Searle contrasts two views. There is the view that he put forth in Chapter 1: "Mental processes are caused by the *behaviour* of elements of the brain . . . [and] are realised in the *structure* that is made up of those elements" (p. 28; italics added). Not that in this more careful formulation, what causes mental processes (namely, behavior) is not the same thing that realizes them (namely, structure). This view is constrasted with "[t]he prevailing view in philosophy, psychology, and artificial intelligence . . . [which] emphasises the analogies between the functioning of the human brain and the functioning of digital computers" (p. 28).

These analogies underlie so-called strong AI. But, considering the intermediate level proposed above, these two views are not inconsistent, *unless* what Searle means is that mental processes are caused by and realized in *only* brain processes. But why should he think that? How *can* he think it? For, if there is a way to describe mental processes independently of the implementing (biological) medium, then they can be implemented in some other (non-biological) medium. Possibly there are some mental processes that cannot be independently described in this way (e.g., pain—though even here it's possible to describe the general functioning of pain*like* phenomena). But are *all* mental processes incapable of such a description? This is (roughly) the claim of Hubert Dreyfus (1979): that human mentality must be humanly embodied. But surely it is possible that non-human mentality could be non-humanly embodied. This raises the question whether such "mentality" would *be* mentality. The answer must surely be 'yes', if mentality is taken to be abstractly describable, and if these two kinds of mentality (human and non-human) are taken to be different *implementations* of it.

There is another way in which the answer might be 'yes', though this takes a bit more arguing: It is often suggested that a simulation of a phenomenon is not an instance of the phenomenon being simulated. For example, simulated hurricanes are not real hurricanes. After all, as people often point out, simulated hurricanes don't get you wet. I think it is wrong to suppose that this shows that simulated Xs aren't Xs. Simulated hurricanes won't get *you* wet, but they *will* get a simulated you simulatedly wet; if they didn't, they wouldn't be very good simulations of hurricanes. The proper way to look at it is that both simulated and real hurricanes are implementations of an abstract notion of hurricane; in this way, they can both legitimately be said to be hurricanes. But suppose that it is the case, *in general*, that simulated Xs aren't Xs. Still, there might be *some* values of X for which simulated Xs *are* Xs. In particular, simulated mentality seems to me to be a good candidate for such an X. If I have a conversation with a computer that passes the Turing Test, it might very well be the case that I could learn something from it. To us the hurricane metaphor, it might indeed get me "wet": it might give me information in much the same way that a Xerox copy of Searle's book can give me the same information that an actual copy would. Such a copy, to perhaps all but certain book-collectors, is the book itself. Clearly, this takes a more subtle argument than I can offer here, but it is one that is worth pursuing. (Cf. Rapaport 1986b.)

Let us return to the earlier point, that machine and human mentality are not the same but might be different implementations of abstract mentality. Does this show that strong AI is wrongheaded? Perhaps; but only if the AI is *very* strong, i.e., that *human* intelligence is its goal. Now, I am not (yet) convinced that this cannot be achieved. Any theory of human mentality that is sufficiently precise to be correct might thereby be sufficiently precise to be a computational—hence, programmable—theory. (Searle disagrees; this is the topic of Chapter 3—see Sect. 2.3.3, below.) The only exception would be if there were some non-analytic, non-formalizable "residue" (as Dreyfus believes). But there is no reason (yet) to think that there will be. In any case, as a philosopher *and* an AI researcher, I am more interested in "computational philosophy" than in "computational psychology". That is, I am more interested in computational theories of *how mentality is possible* independently of any implementing medium than

I am in computational theories of what *human* mentality is and how in fact it is implemented.

Searle, of course, thinks that strong AI *is* wrongheaded, since he takes it to entail a proposition that he claims is false: "that there is nothing essentially biological about the human mind" (p. 28). But this is not a consequence of strong AI, for the physical implementing medium *is* essential to *human* minds and *human* intelligence. According to strong AI, "if you made a computer out of old beer cans powered by windmills; if it had the right program, it would have to have a mind" (pp. 28-29). Searle takes this as a reductio of strong AI. I, on the other hand, think that it is true, though, of course, such a contraption wouldn't be a *human* mind. For Searle, the implementing medium is essential for mentality, not merely for human mentality:

> Suppose we designed a machine that was molecule-for-molecule indistinguishable from a human being. . . . [I]f you can duplicate the causes, you can presumably duplicate the effects. (p. 35)

Therefore, "a man-made machine [could] think" (p. 35). Note that this is not inconsistent with the claim about beer cans and windmills, since the molecule-for-molecule machine might be biological. But why only "molecule-for-molecule"? Why not something that was *structurally* alike at the macro-level but *physically* different at the micro-level? What is it about the beer-can-and-windmill contraption that *prevents* it from having a mind, according to Searle? Answering this is the aim of Chapter 2, and Searle's answer depends only on "the very definition of a digital computer", namely, "that its operations can be specified purely formally . . . in terms of abstract symbols . . . . But the symbols have no . . . semantic content; they are not about anything" (pp. 30-31). That is, they are purely syntactic, whereas "Our internal mental states, by definition, have certain sorts of contents. . . . [T]he mind has more than a syntax, it has a semantics" (p. 31).

I think that there are two separate, but related, arguments here, which are at the heart of Searle's celebrated Chinese-Room Argument. The one we have been looking at is an argument from biology:

Computer programs are non-biological.
Mentality is biological.
∴ No non-biological program can exhibit biological mentality.

This is the argument Searle must make to explain why the beer-can-and-windmill brain, which is non-biological, can't exhibit mentality, which *is* biological. But non-biological computer programs can exhibit *non*-biological mentality. That is, my reply to Searle's argument from biology is the three-level theory of causation and implementation that I offered in response to his two-level theory of "causation" and realization.[2]

The second argument is an argument from semantics:

Computer programs are purely syntactic.
Mentality is semantic.

"[S]yntax alone is not sufficient for semantics" (p. 34).
∴ No computer program can exhibit mentality.

The third premise is the crucial one. Searle calls it a logical truth; I believe that it is empirically false. Of course, it all depends on what's meant by 'semantics'.

Searle defends these arguments by appeal to the Chinese-Room Argument. Suppose that you, who understand no written Chinese, are locked in a room and supplied with an algorithm for taking Chinese input (which, to you, are mere formal symbols) and producing appropriate Chinese output (more formal symbols, to you) in such a way that native Chinese speakers outside the room cannot distinguish your input-output behavior from that of another native Chinese speaker. Then

> there is no way *you* could learn any Chinese simply by manipulating these formal symbols. . . . *you* behave exactly as if *you* understood Chinese, but all the same *you* don't understand a word of Chinese. (pp. 32-33; italics added.)

Unfortunately, the version of the argument Searle presents here is not as carefully stated as in the original version (Searle 1980), and it is easily open to attack by what he calls the systems reply (cf. p. 34). For it is not *you*, the manipulator of the symbols (the "interpreter", in computer jargon)—who "learns" anything or "behaves exactly as if" you understood Chinese. Rather, what understands Chinese is you-the-interpreter *plus* the algorithm and whatever supplementary data structures are in the room. This becomes evident from the allegedly parallel case in which you, still inside the room, get *English* input and, without the benefit of an explicit algorithm (i.e., relying only on English being your native tongue), output appropriate English. But this case is *not* parallel. For a parallel argument, we would need to imagine an English *room*, with an algorithm for *English* that must be explicitly followed. Here Searle is right: Of *course* you don't need such an explicit algorithm. But that's only because you've *internalized* ("compiled") your algorithm for English.[3]

So let us move on to the second step of the Chinese-Room Argument, which is that the program cannot understand because it lacks semantics. However, it *doesn't* lack semantics. What does it mean to *have* semantics? It could mean one of two things: (1) The program must be able to associate its internal formal symbols with external objects (call this "external semantics") *or* (2) the program must be able to associate its internal formal symbols with *other* internal formal symbols (call this "internal semantics"). How do *we* do (1)? I don't know—but, *if we do*, I see no reason why a computer couldn't do it either. (This argument may be unsatisfying, but it is at least as strong as Searle's arguments in Chapter 1.) However, I don't think that we make these associations; here my representationalist sympathies come to fore.[4] The closest we come to doing (1) is by doing (2); by associating one set (better, one interconnected network) of internal symbols with another. The first network of symbols might be linguistic, while the second might be the internal representations of external sensory input (visual, auditory, tactile, olfactory, or a combination of these). That is, a linguistic string *by itself* has no meaning; but a string *in a context*—a network of other strings, linked to direct causal representations of the world—*does* have a meaning. (Incidentally, Searle's answer to the question of what *content* is (p. 31) is consistent with this.) Now, these correlations among internal symbols are methodologically solipsistic, but they are also all that's needed for the appropriate semantics.

In the parallel English-Room Argument, "You understand the questions in English because they are expressed in symbols whose meanings are known to you" (p. 33). But *how* are they known? I repeat: *either* by direct links to the external world (and if it's possible for *us* to make these links, it should also be possible for the computer) *or* by internal links to other internal symbols. And there must be such internal links *anyway*: any direct link to external objects is made because there are internal representations of external objects, which are produced by causal links to these external objects (perhaps in the manner of Sayre 1986), and then there must be internal links among the various internal symbols.

Searle calls this the "robot reply", and he answers it as follows:

> Suppose the robot picks up a hamburger and this triggers the symbol for hamburger to [be input to the program] . . . . As long as all I have is the symbol with no knowledge of its causes or how it got there, I have no way of knowing what it means. The causal interactions between the robot and the rest of the world are irrelevant unless . . . [they] are represented in some mind or other. (p. 35)

And, Searle continues, this cannot be done only syntactically. But why not? Correlation among symbols is a form of "mere" symbol manipulation. And why couldn't the robot have symbols representing the source of other symbols? Again, how do *people* do it? (The robot reply and the network view of semantics are elaborated on in Shapiro & Rapaport 1986, 1987 and Rappaport 1988a.)

At the end of Chapter 2, Searle puts all his arguments together (pp. 38-41; my comments follow in brackets):

(1) "Brains cause minds." [But, as I suggested in the last section, this is *not* ordinary, physical causation; since Searle takes it to be such, I reject (1).]

(2) "Syntax is not sufficient for semantics." [I have suggested in this section that, for the purposes of understanding, it *is* sufficient.]

(3) "Computer programs are entirely defined by their formal, or syntactical, structure." [This *may* be okay—there are, however, theories of denotational and operational program semantics.]

(4) "Minds have mental contents; specifically, they have semantic contents." [The acceptability of this depends on whether 'semantic contents' is to be taken "internally" or "externally".]

∴ (Conclusion 1) "No computer program by itself is sufficient to give a system a mind." (From (2), (3), and (4).) [This is the argument from semantics. But if (2) is false, there is no reason to accept the argument.]

∴ (Conclusion 2) "The way that brain functions cause minds cannot be solely in virtue of running a computer program." (From (1) and (Conclusion 1).) [But (1) is false, and (Conclusion 1) is not supported. Moreover, if 'cause' is taken as I have urged, the negation of this conclusion is plausible.]

So what else *does* help brain functions cause minds? "[T]heir biology matters" (p.40):

∴ (Conclusion 3) "Anything else that caused minds would have to have causal powers at least equivalent to those of the brain." (From (1).) [This is acceptable. But what *are* those causal powers? Searle says they are biological, but I think that at best they are not even causal. (Cf. Rapaport 1985 for a demonstration that this argument begs the question.)]

∴ (Conclusion 4) "For any artefact . . . which had mental states equivalent to human mental states, the implementation of a computer program would not by itself be sufficient." (From (Conclusion 1) and (Conclusion 3).) [Since (Conclusion 1) is unsupported, so is this.]

### 2.3. Cognitive Science ("Grandmother Knew Best")

In this chapter, Searle argues against cognitivism. He begins by claiming that

> we have commonsense explanations of people's behaviour in mental terms . . . . And we suppose that there must also be a neurophysiological sort of explanation of people's behaviour in terms of processes in their brains. The trouble is that the first . . . works well enough in practice, but is not scientific; whereas the second is certainly scientific, but we have no idea how to make it work in practice. (p. 42)

But the first *can* (or, at least, might) be scientific:[5] Recent AI research is trying to formalize such "commonsense" explanations (cf. Hobbs & Moore 1985). Moreover, any *computational* theory of the mind would provide precisely such explanations and would surely be scientific. So, one way of looking at this statement of Searle's is as a challenge to the very possibility of a cognitive *science* based on the computational metaphor. Indeed, Searle denies the alleged dichotomy between the commonsense and neurophysiological explanations, *and* he denies that the cognitive-science approach is correct (pp. 42, 43). He offers four reasons why people might think that cognitivism is true:

(1) There are *psychological reasons*, such as reaction-time experiments and Chomsky's theories.

(2) There is an argument about *rule-following*: Computers and humans function in a similar way, since computers follow rules when processing information and humans do so when they think.

(3) Our brains internalize a *theory* underlying mental abilities.

(4) The *computer/program relationship* is a good model for understanding the brain/mind relationship.

I can agree with Searle that (2) and (3) are not *by themselves* the strongest reasons for cognitivism. Rather, I take (4) to be the best such reason (as discussed in the previous sections). But I think that Searle's objections to (1)-(3) are based on some misunderstandings. Let me begin with (1), the psychological reasons.

#### 2.3.1 Reaction-Time Experiments

Of the two kinds of psychological evidence for supporting that cognitivism is true,

> The first comes from reaction-time experiments, that is, experiments which show that different intellectual tasks take different amounts of time for people to perform. The idea here is that if the differences in the amount of time that people take are parallel to the differences in the time a computer would take, then that is at least evidence that the human system is working on the same principles as a computer. (pp. 44-45)

It seems to me that Searle has this backwards. First, what reaction-time experiments show is *not* that humans work like computers. Rather, what they show is that there *are* mental processes taking place, contrary to behaviorism—which only pays attention to the extensional, input-output relationship. That is, reaction-time experiments show that there are internal processes mediating the input and the output. Second, if the time differences for people do parallel those for a computer, that tells us more about the computer program than it does about humans. It tells us that the *computer* program might be a good model (or theory) for the *human* psychological process. (Of course, that tells us something about humans, too.) Third, the actual time differences are irrelevant, because they depend on the computer architecture, and everyone agrees that that's radically different from human brain architecture. Only *relative* time differences *might* count, if the kind of mechanisms that we program match the kind that are being simulated.[6]

#### 2.3.2 Rule-Following

Against (2), Searle argues that the rule-following is different for computers and for humans. Computers "only act in accord with certain formal procedures" (p. 47), whereas humans are "guided by the . . . content or the meaning of the rule'' (p. 46). That is, human rules are intensional—two rules might both equally well *describe* my behavior, but I *follow* only one of them (p. 46)—whereas either rule would be an acceptable instruction for the computer. Nicholas Goodman (1987) argues along similar lines: the Church-Turing thesis is not a reduction of the notion of ''algorithm'' to that of, say, Turing-machine program, since an algorithm is an intensional entity that contains as an essential component a description of the problem that it is designed for, but no such description forms part of the Turing-machine program.

There is clearly a kernel of an important truth in Searle's and Goodman's observations, but I don't think that either of them refute cognitivism. First, a computer might *in fact* follow (in whatever sense) one rule rather than some other rule that is behaviorally (i.e., input-output; i.e., extensionally) equivalent. Second, the rule actually followed, but not the equivalent one, might be linked into the system's internal network of rules and representations (the system's ''knowldege base'', or ''mental model'', or ''discourse representation structure''). Third, if there *were* two such rules *both* of which were so linked in the *computer's* knowledge base, then such rules might also be linked in the *human's* mind. In that case, what counts for the human is being *aware* of following one of the rules rather than the other—what counts is having some information (''knowledge'') about what the rules ''say'' and what they are for. Can computers be thus ''aware'' of one rule and not of another, or be ''aware'' of the purpose of the rules? I see no reason why not. Recent work in belief systems has shown how a computer can have both true and false beliefs about itself and its behavior: it can fail to be aware of what truly causes it to behave as it does; conversely, it can be aware of the rules that it actually follows. (For surveys, cf. Halpern 1986 and Rapaport 1988b, cf. also Rapaport & Shapiro 1984, Rapaport 1986c.) Moreover, by expressing the rules and the system's ''beliefs'' about the rules in an epistemic or doxastic language, it can have information abut the purpose of the rules. After all, isn't that how we humans have such information?

Searle has an answer of sorts to this (reminiscent of Fred Dretske's discussion of calculators in his recent APA Western Division presidential address; Dretske 1985), which harks back to the syntax/semantics issue of the previous chapter:

> the notion of information-processing embodies a similar massive confusion. . . . The sense in which I do information-processing when I think is the sense in which I am . . . engaged in certain mental processes. But in that sense of information-processing, the calculator does not do information-processing, since it does not have any mental processes . . . It simply mimics, or simulates the formal features of mental processes that I have . . . . In adding 6 and 3, it doesn't know that the numeral '6' stands for the number six . . . . (p. 48)

But how do *we* know that '6' stands for six? One answer, as I have suggested, is that we have a network of arithmetic concepts. But a computer can have precisely such a network of its own. Does it have *our* concept of six? As long as it has our theory of arithmetic (e.g., Peano's axioms), there is no way to force it to have "our" (the intended) interpretation of the symbols, any more than *your* concept necessarily matches mine. All that counts is that we can count with them. (Cf. Rappaport 1988a.)

### 2.3.3. Theories

One frequently hears it said that (strong) AI programs are theories about psychological behavior, theories that are precise down to the last detail, hence implementable, and that any computer running such a program thereby exhibits that psychological behavior. AI researchers test computational psychological theories by bringing them to life, literally. But, according to Searle, there are no such theories:

> a much simpler hypothesis would be that the physiological structure of the brain constrains [meaningful behavior] . . . without the intervention of an intermediate level of rules or theories. (p. 51)

However, the point of the theory talk in cognitive science is that a formal description of those physiological contraints *is* a theory, not that a theory *intervenes*. And if the description is computational, then it can be programmed. In that sense, the program just *is* a theory of mental behavior. Perhaps Searle's point is actually the Gibsonian ecological point that sometimes the brain simply *does* rather than *calculates* (p. 52). But even if this is true for *some* mental processes, there is no reason to think that it is true for all of them.

Finally, Searle reminds us to

> Think of . . . mental processes as biological . . . . Now there will be two gross levels of description in the causal account of how vision [say] takes place in animals. There will be first a level of the neurophysiology . . . . Now in addition to that, there will also be a mental level of description. . . . But in addition to the level of the neurophysiology, and the level of intentionality, we don't need to suppose there is another level; a level of digital computational processes. (p. 54)

This is another frustrating passage, since it suggests that Searle misses the whole point of the computational metaphor in cognitive science, namely, that the level of intentionality *is* the level of computational processes. The working hypothesis of cognitive science (the computational metaphor) is that intentional descriptions are computational descriptions. They are (a) implementable in different physiological media, because (b) they are computational in form.

## 3. ACTION, SOCIAL SCIENCE, AND THE WILL

In the final three chapters, Searle applies his solution of the mind-body problem to three other major philosophical topics. I shall confine myself to brief summaries.

### 3.1. The Structure of Action ("A Walk to Patagonia")

Searle's goal in Chapter 4 is to produce a theory of human action "consistent with his account of the mind-body problem and [his] rejection of artificial intelligence" (p. 57). He presents four features of actions, three features of intentional states, and eleven principles of the structure of human action. I will content myself here with brief comments on these.

#### 3.1.1. Features of Actions

(1) An action cannot be identified with physical movements (p. 57).

This is because one set of physical movements might "continue" many different actions, and one action can be "performed by" different movements. This is suggestive: Is the relationship of physical movements to actions the "causation/realizaton" relationship of Chapter 1 (or perhaps the abstraction/implementation relationship)? I found Searle unclear on this point.

(2) "[I]n acting, what I am doing depends in large part on what I think I am doing" (p. 58).

That is, acting is intenTional and requires intenSional language to express it. Recall the discussion of rules and the Church-Turing thesis, above: perhaps actions are to algorithms as the movements that "implement" them are to Turing-machine programs.

(3) The actor "is in a special position to know what he is doing" (p. 58).

(4) Our ability "to identify and explain the behaviour of ourselves and other people . . . rests on our mastery of a certain set of principles . . . [which] are themselves part of the actions" (pp. 58-59).

Now, earlier, in Searle's arguments against cognitivism, we were told that two rules might equally well describe my behavior (p. 46), so presumably these principles enable us to distinguish between two such rules. If human actions can be explained as implementations of abstract actions, then the existence of co-extensive, but intensionally distinct, descriptions of the actions might be implementation side effects.

#### 3.1.2. Features of Intentional States

(1) Intentional states have a content and a mental type (pp.60-61).

For example, according to Searle, the contents of 'I believe that I will leave the room' and 'I want to leave the room' are the same, but their mental types are different. However, it is arguable that even their contents are different: the content of the former is a proposition ("I will leave the room"), whereas the content of the latter is what Castañeda (1975b) calls a practition ("I to leave the room").

(2) Content and type relate an intentional state to the world, and an intentional state determines "conditions of satisfaction" (pp. 60-61).

A state is satisfied iff the world matches its content.

(3) Intentional states can cause things to happen.

The notion of causation here is "intentional causation": "the mind brings about the very state of affairs that it has been thinking about" (p. 61). That is, it produces the match between world and content.

### 3.1.3. Principles of the Structure of Human Action

(Principle 1) "Actions . . . consist of . . . a mental component and a physical component" (p. 63).

Again, this suggests that the mental component is an abstract description of the action and the physical component is an implementation of it.

(Principle 2) "The mental component is an intention" (p. 63).

That is, the mental component is *about* something. "If successful, the mental component causes . . . and . . . represents the physical component" (p. 63), where:

(Principle 3) "The kind of causation which is essential to both the structure of action and the explanation of action is intentional causation" (p. 64).

What is needed to complete the picture is an explication of "intentional" causation. However, besides Principle 2, all Searle says in *Minds, Brains and Science* is that it is *not* "a matter of regularities or covering laws or constant conjunctions" (pp. 64-65). But then why call it "causation"? Searle owes the reader of this book a better analysis of this sort of relation.[7] It doesn't seem to be causation-1, which goes in the other direction. I suggest that it is simply implementation. Or, perhaps, it is the composition of causation-4 and implementation.

(Principle 4) "[T]here is a . . . distinction between . . . actions which are premeditated . . . and . . . actions which are spontaneous" (p. 65).

The former are the sort of action involved in what AI researchers call planning and what philosophers call practical reasoning. The latter are the sort of action involved in normal conversation.

(Principle 5) "[T]he formation of prior intentions ["formed before the performance of an action"] is, at least generally, the result of practical reasoning" (p. 65).

One consequence that Searle draws from Principles 1-5 is that "the preferred description of an action is determined by the intention in action" (p. 66), that is, "the intentions we have while . . . performing an action" (p. 65).

(Principle 6) "[T]he content that causes behaviour by way of intentional causation must be identical with the content in the explanation of the behaviour" (p. 67).

Of course, if the former content is a practition and the latter is a proposition, they cannot be identical.

(Principle 7) "Any intentional state . . . [only determines its conditions of satisfaction] as part of a network of other intentional states" (p. 68).

This, too, is suggestive, especially in light of my earlier remarks about syntax vs. semantics, since, I would say, the same is true for semantics: the meaning of an expression is given by its location in a network of other expressions.

(Principle 8) "The . . . network of intentionality only functions against a background of human capacities that are not themselves mental states" (p. 68).

Similarly, meanings can be grounded in connections to the external world, but those connections are mediated by internal representations.

### 3.2. Prospects for the Social Sciences ("A Changing Reality")

In Chapter 5, Searle considers why the sciences of human behavior have not been as successful as such natural sciences as physics and chemistry. The answer, he says, depends on "certain radical differences between human behaviour and the phenomena studied in the natural sciences" (p. 71).

The deductive-nomological method of explanation in the natural sciences "is quite worthless" for the social sciences, "even if we had [the appropriate] laws" (pp. 71-72). This is justified by a thought experiment: Suppose that you do A for intentional reason R (e.g., a belief) and that there is a general theory that says that all people like you do A. Now, why did *you* do A? Because the theory says so? But the theory has only *predictive* powers; it is itself in need of explanation. So it must be that you did A because of R; R has *explanatory* power.

Here is Searle's argument in full about the nature of social science. Suppose that there is a social science that is as good as a natural science and that it predicts a revolution. Now, the revolution is both a social event and a physical event. Therefore,

> the laws that predict the revolution will have to make the same predictions at the level of the revolutions and their participants that the laws of molecule movements make at the level of the physical particles. (p. 75)

But they can't, because:

> The phenomena . . . that we pick out with concepts like . . . revolution . . . are not grounded systematically in the behaviour of elements at the more basic level . . . . (p. 77)

That is, social phenomena cannot be reduced to physical phenomena the way that some physical phenomena (e.g., temperature) can be reduced to (or identified with) other physical phenomena (e.g., mean molecular kinetic energy). Why not? Because of "the mental character of social phenomena" (p. 79): because social, but not physical, phenomena are such that the thoughts that people have about the phenomena are part of the phenomena (p. 78)—again, compare the discussion of rules, above. As is the case with actions, a social phenomenon can be physically realized in so many different ways "that there can't be any systematic connections between the physical and the social or mental properties of the phenomenon" (p. 78).

This appears to be inconsistent with Searle's *non*-reductionistic analysis of temperature shown in Figure 1. But I think it also misses the point that the very fact that mental (if not social) phenomena *can* be described in an intensional and computational language makes *anything* that implements that language be intentional. That is, no "bridge principles" (p. 82) are needed to link a mental description to a physical one. If the physical one implements the mental one, it just *is* mental. This is not the 'is' of identity, but the 'is' of implementation.

### 3.3. The Freedom of the Will

Chapter 6 is the least satisfying of the chapters, due to Searle's intellectual honesty in admitting that

> when it comes to the question of freedom and determinism, I am . . . unable to reconcile the two. . . . [T]he problem of the freedom of the will . . . has not been solved . . . and . . . is likely to stay with us. (p. 86)

Nevertheless, as in each of the preceding chapters, there is a nice discussion of the problems. (Readers may be interested in Watts 1984, a complementary panel discussion on free will that Searle participated in with Colin Blakemore, Richard Gregory, and Colin McGinn. It was part of the original Reith Lectures, but not reprinted in *Minds, Brains and Science.*)

## 4. CONCLUSION

As I indicated at the beginning of this review, *Minds, Brains and Science* is interesting and well-written: Searle discusses a number of the more interesting, difficult, and important issues in contemporary metaphysics and epistemology in an appealing and unifying way. But, as I have tried to indicate in the body of this review, his arguments are extremely frustrating and, I think, flawed, though definitely worth tangling with. Nonetheless, I wholeheartedly recommend the book to anyone interested in some of the major issues of contemporary philosophy.[8]

## NOTES

[1]There appears to have been very little formal analysis of realization or implementation. One analysis takes it to be that of semantic interpretation (Tenenbaum & Augenstein 1981: 1, 6, 45). In Rapaport 1985, 1986a, I argue that it is *not* instantiation, since, *inter alia*, from the point of view of the theory of abstract data types, two data types can implement each other.

[2]Searle has responded to this point as follows (private communication):

> The point about biology . . . can be stated as follows:
>
> 1. As a matter of empirical fact, brain processes cause consciousness and other mental phenomena.
>
> 2. It follows trivially from 1. that any other system capable of causing mental phenomena would have to have causal powers at least equivalent to the brain.
>
> 3. It is an empirical question not to be settled on *a priori* grounds whether or not any other given system has these causal powers. It is an empirical question whether beer cans, silicon chips, green slime, or any other hunk of junk has consciousness.
>
> I think it is preposterous to suppose that beer cans either are or could become conscious, but that is an empirical claim on my part. It is a claim which [Rapaport] denies, but that is his problem.
>
> 4. Any other system which duplicated the causal powers of the brain in producing consciousness and other mental phenomena could not do it solely in virtue of instantiating a computer program (for reasons I spell out at some length).

I agree with statements 1 and 2, with two caveats: (a) At best, it is *human* brain processes that cause *human* mental phenomena. (b) I do not think that the relationship is causal; rather, it is one of implementation. As for statement 3, given my interpretation of the relationship as one of implementation, the "empirical" question becomes: Does a given system satisfy the abstract description of consciousness (or whatever)? Looked at this way, the "empirical" issue becomes a behavioral (input-output) issue *plus* an issue about the nature of the mediating processes (depending on whether one demands "strong" or "weak" AI). In the case of beer cans, what seems preposterous to Searle seems obvious to me.

[3]I have become more convinced of the systems reply now than ever before, since I recently found myself in a sort of Chinese room when I acquired a new microcomputer. To learn its operating system, I carefully followed the directions for "opening windows" and "editing". I followed them perfectly, and the system treated me as if I had been a knowledgeable user. But I didn't understand one bit of what I was doing. My gosh, I thought; could Searle be right after all? Fortunately for my argument, the answer is 'no'; because when I reached the point where I no longer needed to check every step of the instructions, when opening windows and editing became "second nature", I *did* understand. What was the difference? In the later case, I had *memorized* the instructions; I had *internalized* them. So the systems response to Searle *is* (partially) correct. It is not the *person* alone in the room who understands, but it is the entire system (person + algorithm + room) that understands. Or, at least, if there is any understanding going on at all, it is to be located there, rather than in the person alone.

[4]Representationalism has the potential for a curiously paradoxical situation. Suppose that a representational theory of mind is successfully implemented in a computer program that passes a Turing test. It might yet turn out to be the case that Dennett (1978) is right when he says that the data structures postulated by intentional-stance theories need not have any neurophysiological correlates. For example, suppose that the human brain works like a "distributed connectionist machine" that is not representational. In that case, one might say that computers have minds, but humans don't! (On connectionism, cf., e.g., the articles in *Cognitive Science 9.1* (January-March 1985).)

[5]I intend 'scientific' here to be taken in Searle's sense, whatever that is. Perhaps it is (as Randall R. Dipert has suggested to me): "unified with explanations in the physical sciences"

[6]Searle has responded to these remarks as follows (private communication):

> [Rapaport] says that I have the reaction time experiments "backward." But this is not in fact my argument. This is an argument given on behalf of cognitivism that I am attacking. And the argument that is given for cognitivism is exactly the way I represent it.

Unfortunately, the informality of the book's lack of references works against Searle here, so I (and several cognitive scientists I have consulted) have been unable to determine who has given such an argument.

[7]Searle does give an account of intentional causation in *Intentionality* (1983). As he states in a reply to my remarks (private communication),

> The cases I give of intentional causation are obviously causal in the ordinary sense of *making something happen*. [Rapaport] is bothered by my claim that it is "not a matter of regularities or covering laws or constant conjunctions." But those are precisely the mistaken features of the traditional Humean definition of causation that I am attacking at some length. It is just begging the question to assume that they give us a correct account of causation.

In my view, however, the relationship is not causal at all, but implementational.

[8]This material is based upon work supported by the National Science Foundation under Grant Nos. IST-8504713 and IRI-8610517 and by SUNY Buffalo Research Development Fund Award No. 150-8537-G. Some of it was presented at colloquia in the Departments of Philosophy and of Computer Science at SUNY Buffalo. I am grateful to Kenneth F. Barber, Gail A. Bruder, John Case, Hector-Neri Castañeda, Joyce H. Daniels, Randall R. Dipert, Judith F. Duchan, Newton Garver, Julie Gerhardt, Nicolas D. Goodman, Shoshana L. Hardt, Richard LaCroix, Hector J. Levesque, Anthony S. Maida, Anthony Ralston, Raymond Reiter, Andrea K. Schnall, John R. Searle, Erwin M. Segal, Stuart C. Shapiro, and Zeno G. Swijtink for discussions and comments on these issues and on earlier versions of this review.

## References

Berztiss, Alfs T., & Thatte, Satish (1983), "Specification and Implementation of Abstract Data Types," in M.C. Yovits (ed.), *Advances in Computers*, Vol. 22 (New York: Academic Press): 295-353.

Castañneda, Hector-Neri (1972), "Thinking and the Structure of the World," *Philosophia 4* (1974): 3-40; reprinted in 1975 in *Critica 6* (1972): 43-86.

Castañeda, Hector-Neri (1975a), "Identity and Sameness," *Philosophia 5*: 121-50.

Castañeda, Hector-Neri (1975b), *Thinking and Doing: The Philosophical Foundations of Institutions* (Dordrecht: D. Reidel).

Dennett, Daniel C., "Intentional Systems," in D. C. Dennett, *Brainstorms: Philosophical Essays on Mind and Psychology* (Cambridge, MA: MIT Press): 3-22.

Dretske, Fred (1985), "Machines and the Mental," *Proc. and Addresses American Philosophical Assoc. 59*: 23-33.

Dreyfus, Hubert L. (1979), *What Computers Can't Do: The Limits of Artifical Intelligence* (New York: Harper & Row, rev. ed.).

Fodor, Jerry A. (1981), "The Mind-Body Problem," *Scientific American 244* (January 1981): 114-23.

Goodman, Nicolas D. (1987), "Intensions, Church's Thesis, and the Formalization of Mathematics," *Notre Dame Journal of Formal Logic 28*: 473-489.

Halpern, Joseph Y. (ed. 1986), *Theoretical Aspects of Reasoning about Knowledge: Proceedings of the 1986 Conference* (Los Altos, CA: Morgan Kaufmann).

Hobbs, Jerry R., & Moore, Robert C. (eds.) (1985), *Formal Theories of the Commonsense World* (Norwood, NJ: Ablex).

Kim, Jaegwon (1974), "Noncausal Connections," *Noûs 8*: 41-52.

Rapaport, William J. (1985), "Machine Understanding and Data Abstraction in Searle's Chinese Room," *Proc. 7th Annual Meeting Cognitive Science Society* (Hillsdale, NJ: Lawrence Erlbaum Associates): 341-45.

Rapaport, William J. (1986a), "Philosophy, Artificial Intelligence, and the Chinese-Room Argument," *Abacus 3* (Summer 1986): 6-17.

Rapaport, William J. (1986b), "Searle's Experiments with Thought," *Philosophy of Science 53*: 271-279.

Rapaport, William J. (1986c), "Logical Foundations for Belief Representation," *Cognitive Science 10*: 371-422.

Rapaport, William J. (1987), "Belief Systems," in S. C. Shapiro (ed.), *Encyclopedia of Artificial Intelligence* (New York: John Wiley): 63-73.

Rapaport, William J. (1988a), "Syntactic Semantics: Foundations of Computational Natural-Language Understanding" in J.H. Fetzer (ed.), *Aspects of Artificial Intelligence* (Dordrecht: Kluwer): 81-131.

Rapaport, William J. (1988b), Review of Halpern 1986, *Journal of Symbolic Logic 53*: 660-670.

Rapaport, William J., & Shapiro, Stuart C. (1984), "Quasi-Indexical Reference in Propositional Semantic Newtorks," *Proc. 10th Int'l. Conf. Computational Linguistics* (COLING-84; Stanford University) (Morristown, NJ: Assoc. for Computational Linguistics): 65-70.
Sayre, Kenneth M. (1986), "Intentionality and Information Processing: An Alternative Model for Cognitive Science," *Behavioral and Brain Sciences 9*: 121-165.
Searle, John R. (1980), "Minds, Brains, and Programs," *Behavioral and Brain Sciences 3*: 417-57.
Searle, John R. (1983), *Intentionality: An Essay in the Philosophy of Mind* (Cambridge: Cambridge University Press).
Searle, John R. (1984a), "1984 Reith Lectures—Minds, Brains and Science: 1. A froth on reality," *The Listener 112* (8 November 1984): 14-16.
Searle, John R. (1984b), "1984 Reith Lectures—Minds, Brains and Science: 2. Beer cans and meat machines," *The Listener 112* (15 November 1984): 14-16.
Searle, John R. (1984c), "1984 Reith Lectures—Minds Brains and Science: 3. Grandmother knew best," *The Listener 112* (22 November 1984): 16-18.
Searle, John R. (1984d), "1984 Reith Lectures: 4. A walk to Patagonia," *The Listener 112* (29 November 1984): 8, 10-11.
Searle, John R. (1984e), "The Reith Lectures: 5. A changing reality," *The Listener 112* (6 December 1984): 15-18.
Searle, John R. (1984f), "1984 Reith Lectures: 6. The freedom of the will," *The Listener 112* (13 December 1984): 10-12.
Shapiro, Stuart C., & Rapaport, William J. (1986), "SNePS Considered as a Fully Intensional Propositional Semantic Newtork" *Proc. 5th National Conference on Artificial Intelligence* (AAAI-86; Philadelphia), Vol. 1 (Los Altos, CA: Morgan Kaufmann): 278-83.
Shapiro, Stuart C., & Rapaport, William J. (1985). "SNePS Considered as a Fully Intensional Propositional Semantic Network," in G. McCalla & N. Cercone (eds.), *The Knowledge Frontier* (New York: Springer-Verlag,): 262-315.
Tenenbaum, Aaron M., & Augenstein, Moshe J. (1981), *Data Structures Using Pascal* (Englewood Cliffs, NJ: Prentice-Hall).
Watts, Geoff (moderator) (1984), "The Reith Lecturers 1984 Discussion: Are minds really free?", *The Listener 112* (20-27 December 1984): 46-47.

# Daniel C. Dennett, *Elbow Room: the varieties of free will worth wanting* (Cambridge, Mass.: The M.I.T. Press, 1984), x + 200 pp., $8.95.

PETER VAN INWAGEN
SYRACUSE UNIVERSITY

The announced purpose of this book is to dispel the confusions and anxieties by means of which philosophers persuade one another that there is such a thing as "the problem of free will." For that is what they do: If they