

Dennett, Consciousness, and the Sorrows of Functionalism

BRUCE MANGAN

Institute of Cognitive Studies, University of California, Berkeley, California 94704

Little is gained, and much lost, by casting an empirical theory of consciousness in a "functionalist" philosophical mold. *Consciousness Explained* is an instructive failure. It resurrects various behaviorist dogmas; it denies consciousness any distinct cognitive ontology; it obliquely adopts many long-standing research positions relating parallel and sequential processing to consciousness, yet denies the core assumption which produced this research; it takes parallel processing ("Multiple Drafts") to be incompatible with educated common-sense views of consciousness (the "Cartesian Theater"), while in fact parallel processing is compatible with some Cartesian Theater views. Contrary to Dennett, the Cartesian Theater does not necessarily imply that contents must fully "arrive" in consciousness at a single, specifiable instant; criticism of the Cartesian Theater based on this attribution is thus without force. And if consciousness is a distinct information-bearing medium, functionalist attempts to "explain" consciousness are inherently inadequate. © 1993 Academic Press, Inc.

The philosopher Daniel Dennett recently proposed (1991) "an empirical theory of consciousness" (p. 17), mixing many of his earlier ideas with some of the central findings of the information processing school of cognitive psychology. As any student of Dennett would suspect, the result is strongly colored by functionalism.

Functionalism is probably the dominant strain in the English-speaking philosophy of mind at present. The unwary might think that a contemporary "functionalist" approach to consciousness would undertake a careful investigation of the actual functional role that consciousness plays in the much larger domain of human cognitive processes. But this is rarely the case in current philosophy, and it is not the main thrust of Dennett's work.

For functionalism is a version of behaviorism.¹ If taken seriously, Dennett's *theoretical* position could well be a disaster for consciousness research. It would undercut the database for the empirical study of consciousness by restricting analysis to inputs and to output "texts" (read stimulus and response) using a method Dennett calls heterophenomenology; in practice this would have us slough off much useful phenomenological information and would encourage habits of mind that all but killed off consciousness research earlier in this century. Furthermore, Dennett's functionalism would have us confound the *least* problematic case of consciousness, that of human beings, with the *most* problematic case, the purported "consciousness" (actual or potential) of computers.

Unfortunately, these consequences are often obscured by Dennett's scattered,

Commentary on D. Dennett. (1991). *Consciousness explained*. Boston: Little, Brown.

¹ See Block (1978) for a discussion of the background of the functionalist/behaviorist relationship.

indirect style of exposition. Typically, he seems to imply one, sometimes plausible position, only to shift ground later on. In this respect, *Consciousness Explained* resembles a kind of intellectual bait-and-switch operation, and in part for this reason it is difficult to be sure of Dennett's precise view on many questions, since they often seem to change as his exposition progresses.

To add to the confusion, Dennett's most fundamental *positive* proposals for understanding consciousness—Multiple Drafts (parallel) and Joycean Machine (serial) modes—while taken from main-line cognitive psychology, are only belatedly and relatively briefly connected with their source. This obscures, but does not diminish, a fundamental tension in *Consciousness Explained*. The current, empirically based, information processing views of consciousness² grew largely from experiments on human beings (not computers claiming to be conscious) and can be traced back to the earliest years of the Cognitive Revolution. But the Cognitive Revolution was based to varying degrees on the rejection of behaviorism, and at least tacitly relied on educated "common sense" assumptions about consciousness—just the sorts of assumptions Dennett attacks in the earlier chapters of his book under the covering rubric "Cartesian Theater."

So there is a paradox beneath the surface of *Consciousness Explained*. Dennett would have us give up intuitions about consciousness that are not only natural and appealing, but have already shown their "cash value" by informing a successful and stable research program—a research program, furthermore, which has already brought out the basic parallel/serial distinction and applied it to consciousness. So the key feature of Dennett's own positive account of consciousness grows from research assumptions he wants to deny.

Dennett's current theory, then, involves three major domains of argument: It (1) attacks various traditional ideas about consciousness, lumped under the term Cartesian Theater; (2) it obliquely sets out some of the core findings of the information processing school of cognitive psychology, especially the parallel/serial contrast as applied to consciousness; and (3) it tries to present these findings in a way that interlards them with functionalist thinking, often using heterophenomenology as the methodological tool for this end.

Functionalism

Dennett generally keeps his functionalism, per se, in the background; it is the Gray Eminence of *Consciousness Explained*. But at one point he does discuss the doctrine directly, bringing out the characteristic computer-oriented slant modern functionalism gives the study of consciousness: "In principle, it makes no difference, the functionalist says, whether a system is made up of organic molecules or silicon, so long as it does the same job." So, for example, in the case of a hypothetical wine-tasting machine.

... if you reproduce the *entire* "functional structure" of the human wine taster's cognitive system (including memory, goals, innate aversions, etc.), you would thereby reproduce *all* the mental properties as well, including the enjoyment, the delight, the savoring that makes wine-drinking something that many of us *appreciate*. Artificial hearts don't have to be made

² See Baars (1988) for an extensive treatment of this approach.

of organic tissue, and neither do organic brains—at least in principle. If all the control functions of a human wine taster's brain can be reproduced in silicon chips, the enjoyment will *ipso facto* be reproduced as well. (p. 31; Dennett's emphasis)

Dennett then tells us that his “book will defend a version of functionalism,” which it certainly does. And for a functionalist like Dennett, it is an article of faith that there can be no real distinction between consciousness in human beings and “consciousness” in certain sorts of computers. (We must wait until p. 281, however, for this to be said directly). This presumption is an absolutely fundamental constraint on Dennett's theory, and one major source of its counterintuitive texture.

Now in one sense this *is* a step beyond behaviorism, since behaviorism simply refused to talk about consciousness. It was “unscientific” and that was that. Dennett's functionalism will allow us to use the word “consciousness,” but only by killing off so many of its standard senses and connotations that the surviving notion of “consciousness” becomes completely counterintuitive. One might think that Dennett's functionalism means that computers, too, can have a phenomenological life. But even this is too commonsensical. In general, as Block (1978) says, functionalism “‘tracks down’ mental states only at the periphery—i.e., through physical, or at least nonmental, specification of inputs and outputs.” Dennett's application of functionalism to the problem of consciousness means that neither neurons nor silicon chips can have a real ‘inner’ life: There simply is no locus of inner experience, of qualia, of feelings of meaning and willing, and so on, as normally understood; no *distinctive* cognitive domain mediates in-coming and out-going neural (or analogous) activity.

Dennett's honesty about the counterintuitive nature of his theory of consciousness is refreshing, if a bit peculiar. He tells us that in contrast with “the alluring simplicities of the traditional view” of consciousness (p. 17), he will offer some “rather startling adjustments . . . that will be as counterintuitive at first to scientists and laypeople alike” (p. 37).

I don't view it as ominous that my theory seems at first strongly at odds with common wisdom. On the contrary, we shouldn't expect a good theory of consciousness to make for comfortable reading. . . . If there were any such theory to be had, we would surely have hit on it by now. The mysteries of the mind have been around for so long, and we have made so little progress on them, that the likelihood is high that some things we agree to be obvious are just not so. (pp. 37–38)

While this makes a virtue of necessity, and is used to justify Dennett's attack on more conventional notions of consciousness, it is still peculiar to *recommend* a theory because it is counterintuitive. Certainly it is easier to come up with an odd theory to cover a set of facts than it is to find an intuitively plausible theory to do the same thing. Usually the ease and simplicity of a theory counts in its favor, not against it.

Nevertheless there is a certain appeal to the proclamation of a radical new theory, and Dennett makes the most of it: His major positive proposal, what Dennett calls The Multiple Drafts model of consciousness, is *initially* and repeatedly said to rest on “novel background assumptions,” (p. 17) and to be “an alternative to the traditional model, which I call the Cartesian Theater. It requires

a quite radical rethinking of the familiar 'stream of consciousness' and is initially deeply counterintuitive" (p. 17).

But many chapters later, after various criticisms of the Cartesian Theater have intervened and his own Multiple Drafts model has been sketched out as a supposed alternative, Dennett mentions, rather briefly, the sources of his "new" Multiple Drafts proposals. On p. 189 we are told in passing (and I believe for the first time) that versions of the Multiple Drafts idea have existed for decades, and Dennett points to one of the oldest war-horses in cognitive modeling, Selfridge's *Pandemonium* (1959). But it is only in chap. 9 that a reader unfamiliar with cognitive psychology and its pervasive interest in parallel processing will finally, although still briefly, learn how greatly indebted Dennett's theory is (functionalism excepted) to this well-established research tradition. And even those bits of research on parallel processing which are discussed in any detail (e.g., a page or so on Anderson's ACT*) are drawn on very largely from computer modeling, not biologically based, literature.

It turns out that when Dennett first said he had revolutionary new ideas to offer with his Multiple Drafts model, he apparently meant that they were largely just new and revolutionary to him. ". . . My theory would have been inconceivable (by me, at least) if it had not borrowed heavily from empirical work in various fields which opened up (to me, at least) new ways of thinking" (pp. 256–257). Dennett then goes on to cite Baars (1988), one of the relatively few direct references to the biologically based approach, as weaving together many strands of prior research into "a version of the emerging consensus that I have gingerly introduced, ignoring some features, and emphasizing others . . ." (p. 257).

So by chap. 6 Dennett has vitiated his own chap. 1 justification for the counterintuitive aspect of his theory. "The alluring simplicities of the traditional view" turn out to have made substantial progress with "the mysteries of the mind" after all, and Dennett now tells us he is merely "introducing" a new consensus in empirical cognitive research.

Certainly no "startling readjustment" in cognitive psychology results from considering parallel processing and its relation to consciousness—although of course there are many different ways to pursue and specify this relation.³ The thesis that parallel, unconscious activity influences the far more limited domain of conscious experience was an integral part of Freud's theories. Within the framework of the Cognitive Revolution, Ulric Neisser brought the parallel/serial contrast to the fore in his classic, experimentally based *Cognitive Psychology* (1967), in effect arguing that consciousness contained both preattentive (parallel) and attentive (serial) aspects, while unconscious processing, including most preattentive activity, was for Neisser parallel. George Mandler, among the first major figures in cognitive psychology to straightforwardly address the place of consciousness in modern cognitive research (the title of his 1975 paper, "Consciousness: respectable, useful and probably necessary" speaks for itself) reinterpreted Neisser's

³ The historically minded can find the germ of the idea of parallel processing in Leibniz (e.g., *New Essays on Human Understanding*, 1769/1981), who held that indistinct "petit perceptions" in consciousness were the result of what today we would call an unconscious and parallel process of extreme complexity.

preattentive processes as much more strictly unconscious and parallel, and equated consciousness with attention. While the equation of consciousness with attention is now the dominant assumption (e.g., Treisman & Gelade, 1980), the notion that consciousness also has a preattentive aspect continues to receive support: for example, the “inattention” experiments of Rock and Gutman (1981). Rumelhart et al. (1987), in perhaps the most prominent connectionist discussion of consciousness, propose that a “vague” and (at least) relatively parallel aspect of experience sometimes precedes the formation of a clear, serial content in experience. And among the connectionists, probably Smolensky (1988) has taken the lead in showing how serial processing can rest on a parallel architecture.

The significant new element in Dennett’s theory, then, is neither the parallel aspect of his Multiple Drafts proposal nor the more specific idea that the serial aspect of conscious activity (the Joycean machine) is founded on a parallel substrate. But Dennett does deserve credit for recognizing how important the parallel/serial distinction has become for the empirical investigation of consciousness and for trying to introduce the point to a wider audience—even if Dennett’s “gingerly” treatment of the empirical work on parallel processing in *Consciousness Explained* makes it a rather oblique and quirky introduction to this research, not to mention the paucity of discussion of the human attention literature, which is, rightly or wrongly, at the heart of current consciousness research.

In any case, the “startling adjustment” Dennett proposes must not be confused with applying parallel/sequential notions to consciousness. Although Dennett begins by treating the Multiple Drafts proposal as if it is a contending alternative to the Cartesian Theater, this is a false opposition on two counts; Multiple Drafts turns out to largely refer to *unconscious* processes, and insofar as Multiple Drafts applies to conscious experience, Multiple Drafts and the Cartesian Theater are not mutually exclusive, but if anything orthogonal to one another (again, excluding functionalist assumptions). We will return to these points later, but first we will concentrate on features of Dennett’s theory that *are* most original and which, in their developed form, do conflict directly with a wide range of common sense notions about consciousness.

Heterophenomenology

Dennett calls heterophenomenology “a giant theoretical step” (p. 73), and, in a sense, he is absolutely right.

Taken at a very general level, and ignoring its functionalist constraints, we can say that heterophenomenology aims to create a phenomenological method *that is expressly designed for the purposes of doing empirical cognitive research*. Obviously the study of consciousness could use a scientifically informed method for directly characterizing its subject matter. But two lines of thought have worked against creating a scientific school of phenomenology, and Dennett tries to circumvent them both.

On one hand, there is the position deriving from Descartes. To counter the skepticism of his age, Descartes maintained that the contents of consciousness give us absolute certainty, in one special sense: Even if an evil demon were to attempt to fool us with hallucinations, the content of our experience would still

be the content of our experience. Experience, *qua* experience, is certain; the accuracy of its "reference" is another matter. The point was developed in this century by Husserl, who maintained that since the contents of consciousness, as such, are "apodictic," incontestable, his Pure Phenomenology is logically prior to natural science.⁴ This doctrine has had many pernicious consequences. Among other things, a scientist looking at the modern phenomenological movement will see consciousness treated as if it inhabits a kind of protected reserve, with its contents fixed absolutely by declaration, and its study exalted as something far beyond mere science. Call this the left wing stance toward consciousness.

On the other hand, there is the right wing stance toward consciousness taken by old style behaviorism. Here the buzzword is not "apodictic" but "incorrigible." And as is often the case, the two extremes turn out to share some common ground. For behaviorism also insists that consciousness cannot be studied by science, and for roughly the same reason: A declaration about one's experience is assumed to end the matter, and thus disputes about the contents of consciousness are thought to be in principle beyond resolution by scientific method.

Now the aim of heterophenomenology is *in part* to break the hold of the incorrigible and apodictic dogmas on consciousness research. In this heterophenomenology appears to be successful, and to this degree Dennett makes a significant step by recognizing that a truly scientific phenomenology is possible and should be attempted. But in the attempt, I believe, Dennett badly miscarries. He pays a great (and quite unnecessary) tax for the particular way he tries to avoid the apodictic/incorrigible presumptions, a tax exacted by his functionalism.

Dennett begins his discussion of heterophenomenology with a set of promises and aims that appear intriguing, plausible, and sensitive to the aims of scientific research. But as the argument unfolds, its import shifts, and the frame Dennett first puts around heterophenomenology turns out to be misleading.

To set the stage for his discussion of heterophenomenology, Dennett points out, very reasonably, that "even if mental events are not among the *data* of science, this does not mean that we cannot study them scientifically. Black holes and genes are not among the data of science, but we have developed good theories about them" (p. 71; Dennett's emphasis). After drawing the parallel with genes and black holes, Dennett turns directly to his notion of heterophenomenology. He tells us what it aims to do, and it is an impressive declaration.

Ignoring all tempting shortcuts . . . here is the *neutral* path leading from objective physical science and its insistence on the third person point of view, to a method of phenomenological description that can (in principle) do justice to the most private and ineffable subjective experiences, while never abandoning the methodological scruples of science. (p. 72; Dennett's emphasis)

Setting aside for the moment what "neutrality" and "justice toward our most private and ineffable experiences" actually turn out to be for Dennett, we can give a first pass characterization of heterophenomenology this way: it proposes to take verbalizations about purported phenomenological states (i.e., observable responses) and treat them as "texts." From the outside, from the scientific third

⁴ See Husserl, *Cartesian Meditations*.

person point of view, we can then use these texts as data for constructing a theory about the inside: about the subject's supposed conscious experiences. But it must be remembered that the heterophenomenological world, the possible phenomenological experiences indicted by the text, is a theoretical domain by definition.

A heterophenomenological assertion about a conscious experience can never claim to be certain; it is as potentially fallible as any assertion made by any scientific theory, but at the same time it is amenable to various kinds of scrutiny and evaluation. There is neither absolute certainty nor inherent intractability in the heterophenomenological method. A text can be used as one source of information to help evaluate hypotheses about conscious states, and we can then take a critical attitude toward any given first person report. Presuming a sufficiently strong background of other findings, we can reinterpret a first person report, revise it, even reject it out of hand.

Now one might think that on Dennett's account we can then say that a heterophenomenological theory refers to real phenomenological states, just as theories about black holes and genes refer to real black holes and to real genes. Of course, we could be wrong about the existence of black holes, and it is possible we could even be wrong about the existence of genes. But past a certain point of confirmation—when it is clear that there is only a very small chance that we are wrong about the existence of an entity like a gene or a black hole—no one (except a doctrinaire anti-realist) is troubled by the standard metamorphosis in which a hypothetical entity posited by a theory becomes for science a perfectly real, existing entity that is interrogated or interpreted by the theory. Thus the natural continuation would be to go from heterophenomenological findings back to a (hopefully) chastened understanding of our actual phenomenology. In other words, real experience is to be *interrogated* by heterophenomenological method.

But this is a move Dennett fights with all his might, for it violates his allegiance to functionalism. Functionalism, like earlier forms of behaviorism, is decidedly uncomfortable with realism (see Block, 1978). Behaviorism and its functionalist variations are more than willing to posit unadulterated existence to their stimulus and response system, or its analogues. But the status of the intervening territory is another matter. Dennett will end up denying that there is any distinctive or fundamental conscious domain residing between in-coming and out-going cognitive activity; for Dennett, assertions about such a conscious domain are nothing but "theoretical fictions." But again, the standard inclination in science, quite independent of the question of consciousness, is to move from very well-confirmed theoretical hypothesis to an ontological conclusion.

By refusing to go along with this standard realist transition, Dennett's heterophenomenology violates not only common sense assumptions about reality in general and the reality of our phenomenological world in particular, but also much common sense scientific practice as well. There is no need to deny realism to satisfy "the methodological scruples of science." Here is another point where Dennett's allegiance to functionalism creates much unnecessary difficulty and counterintuitive squirming. But, again, Dennett seems at first blush to hold a far more moderate position:

People undoubtedly do believe they have mental images, pains, perceptual experiences, and all the rest, and *these* facts—the facts about what people believe, and report when they express their beliefs—are phenomena any scientific theory of mind must account for. We organize our data regarding these phenomena into theorist's fictions. . . . Then the question of whether items thus portrayed exist as real objects, events, and states of the brain—or the soul for that matter—is an empirical matter to investigate. (p. 98; Dennett's emphasis).

But even at this early point, we must distinguish two very different questions, questions constantly confounded by Dennett. Deciding if the data are strong enough to justify moving this or that *particular* content from the status of theorist's fiction to that of existing content *is* an empirical question. Deciding that all phenomenological entities are in principle nothing but theorist's fictions is another matter entirely. Even if we were totally unable to make sense of *some* set of phenomenological claims (Dennett's favorite examples come from perplexities that arise when we consider experience over the range of milliseconds), this hardly justifies the conclusion that, as a general matter, we can *never* move with a clear scientific conscience from heterophenomenology back to real phenomenology, that is, back to real experience.

For Dennett, heterophenomenology's neutrality is not primarily concerned with getting rid of biases that might contaminate our ability to organize or adjudicate this or that *person's* phenomenological claims.⁵ For Dennett's "neutrality" actually refers to the functionalist demand that we must presume strict equality between texts generated by humans and those produced by *any* hypothetical, text-producing entity such as "a zombie, a parrot dressed up in a people suit, or a computer driving a speech synthesizer program" (p. 76). Dennett takes great pains to welcome absolutely hypothetical, nonhuman sources of heterophenomenological information, although these possibilities are of current relevance only to thought experiments. The world is not, in fact, confronted by computers loudly proclaiming that they are conscious and offering detailed heterophenomenological texts to justify the claim. At the moment the consciousness of computers is simply not a scientific problem.

But what the world *is* full of is human beings, constantly making tacit and explicit claims about their experience. For anyone trying to work out an *empirical* theory of consciousness—and this is Dennett's declared aim—human reports about conscious experience are the primary data, and all other sources of data lacking this component (e.g., computer simulations, neurophysiological findings, animal experiments) are decidedly secondary. These secondary data are still, of course, germane, although generally to the degree that they converge with, and give additional support to, the primary database, people's reports about their conscious states.⁶ Yet Dennett displays constant skepticism, if not downright

⁵ This person-centered implication of Dennett's "neutrality" is certainly suggested by Dennett's preliminary discussion of heterophenomenology; in their generally perceptive review of *Consciousness Explained*, for example, Roskies and Wood (1992) take this to be Dennett's point. Here, and in a few other passages of this review, I follow my brief discussion of Roskies and Wood as found in Mangan (1992).

⁶ For a method of "convergent phenomenology" that is expressly designed for the use of cognitive research and avoids both the anti-realist mind set and apodictic/intractable presumptions, see Mangan (1991).

antagonism, toward the primary data in consciousness research, reports about conscious experiences made by human beings.

For in practice, Dennett uses heterophenomenology to cast *general* doubt on the value of using subjective reports for scientific investigation. Affirming this aspect of behaviorist ideology risks doing immense damage by turning us away from the primary database and returning to the dark ages of consciousness research.

Rather than state his anti-realist, behaviorist assumptions directly and argue for them (they are finally mentioned, briefly, after more than 450 pages, in Appendix A), Dennett relies instead on rhetorical devices to imply that there is some kind of inherent folly in taking our reports about direct experience seriously. Early in his exposition, for example, he takes fictive literary texts as the fundamental analogy to heterophenomenological texts—not, significantly, fictitious entities generated by scientific theories. “Some texts, such as novels and short stories, are known—or assumed—to be fictions, but this does not stand in the way of their interpretation” (p. 79). So although fictitious, the heterophenomenological world will still seem to hang together. Again we are told that the reason for maintaining this fictive stance is because “officially we must keep an open mind about whether our apparent subjects are liars, zombies, or parrots dressed up in people suits” (p. 83).

Dennett continues to embrace the world of nonhuman “consciousness,” and his analysis is deflected to accommodate it. So, after a long excursion into the status of fictive entities such as Sherlock Holmes, Mr. Pickwick, and a hero of Dennett’s own devising, Feenoman, Dennett moves on to consider the operations of Shakey, a computer robot that “sees” using the standard manipulation of zeros and ones. Dennett remarks, beginning to anticipate his final position, that “Shakey’s ‘images’ provide an example of how something that really wasn’t an image at all could be the very thing that someone was talking about in the guise of an image” (p. 94).

Finally Dennett’s discussion of heterophenomenology turns to human experience. He begins by reminding us how often people make mistakes, confabulate, and so on. Dennett then briefly considers the classic experiments on internal visualization (e.g., Shepard & Cooper, 1982; Kosslyn, 1980) which are usually taken to show that at least some inner, completely subjective experiences are quite amenable to empirical investigation, and furthermore that in this case the findings are consistent with our more or less common sense intuitions about inner visualization. At this point Dennett makes a begrudging, characteristic remark: “Sometimes the unwitting fictions we subjects create can be shown to be true after all, if we allow some metaphorical slack, for instance . . . recent research on imagery shows that our introspections are not utterly false” (p. 94).

Dennett’s treatment of heterophenomenology is a good example of the expository method used in *Consciousness Explained*. We rarely find a straightforward claim and then a direct, sustained argument for it. In its place we have the often striking use of analogies and other stylistic devices (their aptness is another matter) to create what can be called a presumptive mood. In this case, Dennett creates the presumptive mood that even the simple, experimentally consistent report of an internal visualized rotating object is something of a miracle; even

then it still must be granted “metaphorical slack” and can only rise to the left-handed status of not being “utterly false.”

Here, as is so often the case in *Consciousness Explained*, there has been a substantial shift from what seemed to be the import of an initial claim made by Dennett. Remember when we were promised that heterophenomenology would do justice to our most private and ineffable subjective experiences? Now things have shifted to the point that we are told that even experimentally controlled, quantitative and reliable reports about our inner life (not just stray personal introspections) are to be taken as wildly dubious, although not, perhaps, utterly false. This is not a promising way to develop a scientific phenomenology. The cognitive revolution freed consciousness research; one might uncharitably say Dennett attempts a theoretical counter-revolution, a sort of latter day behaviorist putsch.

Denying Consciousness any Fundamental Status

Even at the end of the book, the precise status of the reality behind the heterophenomenological world is hedged. Dennett seems to think heterophenomenology insulates him from the need to proclaim *any* clear-cut position about the ontology of consciousness. He remarks that one virtue of heterophenomenology is that it is able to evade (his term) what I would call uncomfortable ontological demands: “The indirectness of the heterophenomenological method is precisely a way of evading ill-motivated obligations to “identify” or “reduce” the (putative) entities that inhabit the ontology of subjects” (p. 459). In the end, Dennett will only go this far toward addressing the ontological foundation of heterophenomenology:

In some regards, you could say that my theory identifies conscious experiences with information-bearing events in the brain—since that’s all that’s going on. . . . The question of whether to treat part of the heterophenomenological world of a subject as a useful fiction rather than a somewhat strained truth is not always a question that deserves attention [!]. Are mental images real? There are real data structures in people’s brains that are rather like images—are *they* the mental images you’re asking about. If so, then yes, if not, then no. (p. 459)

Of course it is only through the mediation of a rather complex set of theories that Dennett can grant “information-bearing events in the brain” reality. But Dennett has no trouble converting *these* theoretical fictions into robust ontological assertions. The brain truly exists and “information-bearing events” exist. These Dennett takes to be fully real. But, again, Dennett thinks it is a mistake to grant reality to mental images, or to any other conscious content as normally understood.

This leaves a key point unaddressed. There is certainly nothing in the notion of consciousness in the full, inner, intuitive sense that precludes it from bearing information. An inner image and the “guise” of an inner image can *both* be interpreted as information-bearing events. And if we can assume that consciousness (as intuitively understood) is produced by the brain, then we *can* go on in principle to identify the ontology of subjects. In this way we can move from a hypothetical heterophenomenology back to real phenomenology—and *within the general constraints as set out by Dennett above.*

But this move is not open to Dennett because of a further commitment. For in the last few pages of the book Dennett finally comes clean about a crucial background assumption that clearly helps motivate his squeamish attitude toward the

ontological status of consciousness. This is an assumption that readers, philosophers or not, should have been told about in the first chapter: Over the course of his career, he says,

My fundamental strategy has always been the same: first, to develop an account of content that is *independent of and more fundamental than* consciousness—an account of content that treats equally of all unconscious content-fixation (in brains, in computers, in evolutions “recognition” of properties of selective design)—and second, to build up an account of consciousness on that foundation. First content, then consciousness. . . . This strategy is completely opposite, of course, to the vision of Nagel and Searle, who in their different ways insist on treating consciousness as fundamental. (pp. 457–58; Dennett’s emphasis)

In Dennett, consciousness cannot be allowed fundamental status, and functionalism accommodates this demand. For as we saw above, functionalism views any cognitive process as equivalent to any other, in all important senses, “when it does the same job” (p. 72). For Dennett and functionalism, the *medium* in which cognitive activity takes place is irrelevant.

From his standpoint, Dennett is right to say that he is not trying to “reduce” consciousness to something else. He seems to want to understand content in such a way that it is, as it were, invariant as it passes from locus to locus in the universe. A content is to be understood in such a way that there is nothing distinctive about it in one medium as opposed to another; a content is in essence the same in brains, computers, physical evolution, the external evolution of Dawkins’ “nemes” (chap. 7) and so on. In this sense, then, Dennett is not a reductionist, but might be called a “pass-throughist.”

In the most extreme application of this position, Dennett bravely points out that it allows for a kind of personal immortality. “Current embodiment, though a necessary precondition for your creation, is not necessary for your existence to be prolonged indefinitely.” On the one hand, notions of an eternal soul are wrong, but

if you think of yourself as a center of narrative gravity, on the other hand, your existence depends on the persistence of that narrative . . . which could *theoretically* survive many switches of *medium*, be teleported as readily (in principle) as the evening news, and stored indefinitely as sheer information. (p. 430; Dennett’s emphasis).

This follows the well-known AI point that the medium on which a program is run is quite irrelevant to the information processing structure created by the program itself. The same program, it is said, can be run on a computer made of cogs and wheels, old shoes, beer cans, the population of China, and so on; and of course the same program can be run on different computer hardware. The important thing on this view is the “sheer information,” not the medium in which information happens to be instantiated. Roskies and Wood (1992) complain that this sort of application to consciousness is “disappointing . . . and old hat.” While this seems to me to be an accurate observation, the notions of “information-bearing medium” and “sheer information” do let us use Dennett’s own terms to clarify the intuition he wants to reject. And it is this: When people consider their own conscious contents, they are not only concerned with the sheer information their consciousness bears, *but also with the nature of the particular medium of consciousness, and the particular way that medium bears its information.*

The Cartesian Theater

I believe this way of putting the matter lets us get a bit more precise about the problem Dennett sees lurking in the Cartesian Theater. For if consciousness is a distinct information-bearing medium, then our own direct, full sense of a conscious content would seem to be mingled with the specific character of the medium that bears the content. But this means that from the standpoint of our own first person, conscious experience, we *cannot* move contents into another medium without losing something (i.e., the specific characteristics of the conscious medium), even if functional equivalence is preserved. This, however, violates Dennett's fundamental premise that contents must be completely medium-independent, even to the point that we can contemplate personal immortality in a computer chip. But at first Dennett's objection to the Cartesian Theatre seems to rest on other, broadly motivated scientific concerns.

Dennett initially binds the notion of the Cartesian Theater to "dualism." What's wrong with dualism? Dualism is one of those tricky words that can have many meanings, especially in Dennett. He initially asserts that dualism means seeing consciousness as a mysterious, nonphysical mind stuff, inherently beyond the pale of scientific understanding.

This fundamental antiscientific stance of dualism is, to my mind, its most disqualifying feature, and is the reason why in this book I avoid dualism *at all costs*. It is not that I think I can give a knock-down proof that dualism, in all its forms, is false or incoherent, but that, given the way dualism wallows in mystery, *accepting dualism is giving up*. (p. 37; Dennett's emphasis).

First of all, we saw above that much later in the book Dennett himself shows that substantial progress *has* been made using at least tacitly dualistic, educated common sense assumptions about consciousness. So as a practical matter for scientific research, accepting "dualism" of certain kinds simply does not entail wallowing in mystery and giving up. Psychophysics, for example, began and prospered mightily on an absolutely explicit dualism—working out in precise mathematical terms the relation between the physical magnitude of an external stimulus and its consciously perceived subjective magnitude. With modifications of Fechner's original logarithmic scale (i.e., Stevens' power law, 1975) and with the recognition that context effects also influence psychophysical judgments, psychophysics is dualistic to this day, and probably still possesses the most empirically precise database of all branches of psychology.

Nor is there something inherently unscientific in granting consciousness a distinct physical ontology. Referring to Penrose's *The Emperor's New Mind* (1989), Dennett is willing to concede the possibility that "perhaps some basic enlargement of the ontology of the physical sciences is called for in order to account for the phenomena of consciousness" (p. 36). Dennett does not think that Penrose succeeded in making his particular case, but even Dennett does not object to Penrose's attempt in principle:

It is important to notice that he [Penrose] has been careful *not* to fall into the trap of dualism. What is the difference? Penrose makes it clear that he intends his proposed revolution to make the conscious mind *more* accessible to scientific investigation, not less. (p. 37; Dennett's emphasis)

So at this point, dualism, for Dennett, is not the claim that consciousness occupies its own distinct ontological niche; dualism is equated with views of consciousness that treat consciousness as so radically nonphysical and mysterious that it is beyond the bounds of science. Up to this point, his argument against dualism (as characterized) is both reasonable and balanced. But then Dennett makes one of his shifts: "Scientists and philosophers may have achieved a consensus of sorts in favor of materialism, but as we will see, getting rid of the old dualistic visions is harder than contemporary materialists have thought. Finding suitable replacements for the traditional dualistic images will require some startling readjustments to our habitual ways of thinking. . . ." (p. 37). How does this conclusion follow? Why must we cut out dualistic assumptions root and branch? If materialism can accommodate consciousness, and if (as it turns out) research is making significant progress on the current "dualistic" assumptions, why do we need to substitute a set of deeply counterintuitive assumptions that, furthermore, skew the database for consciousness research and in addition bring with them an ideology that has in fact greatly inhibited consciousness research in the past? And yet by way of this transition, Dennett goes on to argue against various forms of the Cartesian Theater as if it entailed a scientifically unworkable form of "dualism," and that we must presume, in effect, that content is irrelevant to its instantiating medium, i.e., that there is nothing distinctive about the contents of consciousness as such.

The terms "dualism" and "Cartesian Theater" contribute greatly to the presumptive mood coloring Dennett's treatment of the views he wants to supplant. These terms conjure up centuries of misleading associations that *should* be banished and for the most part already have been banished from scientific practice.

But many scientists (and philosophers) still do have the lingering suspicion that if consciousness is ontologically distinct, a huge bifurcation will loom up in nature. This is an unfortunate artifact from the past, and it does persist. However, we do not need to adopt Dennett's radical alternative to get rid of it.

Here the notion of an information-bearing medium comes to our aid. Quite beside neurons, we already know of *many* completely distinct, physically based biological information-bearing media operating within our organism: In the ear alone, the ear drum is a distinct information-bearing medium; the bones of the middle ear are a distinct information-bearing medium; the saline solution in the cochlea is a distinct information-bearing medium; the basilar membrane is a distinct information-bearing medium. Our cognitive apparatus is not *only* made of neurons, although, of course, so far as we know, neurons do the lion's share of cognitive information-bearing. And there are many other physically based, biological information-bearing media that fall outside the standard sense of cognitive activity: DNA, for example, and the immune system.

Taking consciousness to be just one more physically based, biological information-bearing medium hardly mandates "dualism." Recognizing the neuron/consciousness distinction is important for exploring certain interesting and very difficult problems, just as other scientific problems rest on examining the relation between other sets of distinct features of the world. If consciousness is simply one more information-bearing medium among others, we can add it to an already rather long list of media without serious qualms. For in this sense, there is nothing ontologically "special" about neurons *or* consciousness; they are both, as it were,

on the same physical, ontological plane, and no grand dichotomy is mandated by them. So, *contra* Dennett, we can treat consciousness as ontologically distinct without entailing dualism, and still pursue scientific research with a clear conscience.

One problem with Dennett's use of the term Cartesian Theater is that it covers so much territory. At times it refers to very general common sense ideas about consciousness, at times to more refined but still general notions held by many scientists and at least some philosophers (I have called this educated common sense views of consciousness), and at times it refers to this or that specific theoretical position, ranging from Descartes to men like Eccles and Libet. The Cartesian Theater, then, is an extremely generalized, inductive concept and not a theoretical term that Dennett may simply define as he wishes, since its *raison d'être* is to characterize existing views.

Dennett does define a subposition (a classic straw man, as it turns out, Dennett's denial (p. 139) notwithstanding), which he calls "Cartesian materialism."

Cartesian materialism, the view that nobody espouses but almost everybody tends to think in terms of, suggests the following subterranean picture. We know that information moves around the brain, getting processed by various mechanisms in various regions. Our intuitions suggest that our stream of consciousness consists of events occurring in sequence, and that *at any instant every element in that sequence can be classified as either having occurred "in consciousness" or not having occurred "there" yet.* And if this is so, then (it seems) the contentful vehicles of content moving through the brain must be like railroad cars on a track: the order in which they pass by some point will be the order in which they "arrive at" the theater of consciousness and (hence) "become conscious." To determine *where* [Dennett's emphasis] in the brain consciousness happens, trace all the trajectories of information-vehicles, and see what point particular vehicles are passing at *the instant* they become conscious. (p. 144; my emphasis, except as indicated)

On this tacit view of conscious experience attributed by Dennett to otherwise good scientific materialists, consciousness occupies a kind of point-instant, and through it passes a sequence of completely specifiable, discrete contents that become conscious at a certain "instant" and are thus unambiguously present or absent. "Cartesian materialism" understood in this way is then shown to lead to various paradoxes that Dennett explores with gusto throughout chap. 6. For the sake of argument and space, let us assume that the problems Dennett finds in applying Cartesian materialism *as defined* are insurmountable.

But so what? There is a huge chasm between generally held, *explicit* materialist presumptions about consciousness (that it is a distinct cognitive realm, that it is a manifestation of neural processes, that it can stand between in-coming and out-going neural activity) and Dennett's speculation about the tacit view materialists seem to hold about consciousness. Even if Dennett is correct about the existence of these tacit presumptions, and even if they are in fact unworkable, it hardly follows that we must then accept Dennett's radical alternative by default. First we must see if other materialist accounts of consciousness avoid the positions Dennett claims to find in his own manufactured attribution of Cartesian materialism.

We do not have to look far. Even in the passage above, Dennett uses the ubiquitous phrase, "stream of consciousness." In a sense, Dennett has produced

an argument *supporting* the claims of that great Impresario of the Cartesian Theatre, William James. James was a materialist, took consciousness to be quite distinct from its neural substrate, and saw consciousness as sometimes standing between in-coming and out-going neural activity. But James in the *Principles of Psychology* (1890) gave extensive play to the idea that experience is in a sense smeared or spread out, with a halo or “fringe” of experience intermingled with the more discrete contents in consciousness. There is no precise moment when a content is suddenly “there.” It is a complete mistake to take consciousness, for instance, to be “a string of bead-like sensations and images, all separate” (p. 605). Or, we can say, like a string of boxcars. According to James

All our concrete states of mind are representations of objects with some amount of complexity. Part of the complexity is the echo of the objects just past, and in a less degree, perhaps, a foretaste of those just to arrive. . . . The lingerings of the past dropping successively away, and the incoming of the future making up the loss. . . . They give that continuity of consciousness without which it could not be called a stream. (pp. 606–607).

There is a certain indefiniteness *in consciousness* about just what contents are present or absent. And especially relevant to Dennett’s concern with the question of time and experience, James’ extensive phenomenological analysis⁷ led him to conclude that the notion of an instantaneous present *was* a sort of fiction. Phenomenologically, we live in a “specious present.” In contrast, the notion of a “strict present” is “an altogether ideal abstraction” (p. 608).

In short, the practically cognized present is no knife-edge, but a saddle-back, with a certain breadth of its own on which we sit perched, and from which we look in two directions in time. The unit of composition of our perception of time is a *duration*, with a bow and a stern, as it were—a rearward—and a forward-looking end. It is only as parts of this *duration-block* that the *secession* of one end to the other is perceived. (pp. 609–601)

James maintains that the phenomenological sense of “now” is quite elastic, although its objective duration seems to hover around twelve seconds or so (p. 630). Furthermore, James to a degree anticipated the notion of complex, parallel, neural processing underlying the phenomenology, and emerging into it:

With the feeling of the present thing there must at all times mingle the fading echo of all other things which the previous few seconds have supplied. Of, to state it in neural terms, *there is at every moment a cumulation of brain-processes overlapping each other, of which the fainter ones are the dying phases of processes which but shortly previous were active to a maximal degree.* (p. 635)

And later James again speculates that “The feeling of time duration is the immediate effect of . . . an overlapping of brain processes of different phase” (pp. 638–639). And he repeats in even more “parallel” language a few pages later that the cause of our extended sense of the current moment is “probably the *simultaneous* presence of brain-processes of different phase” (p. 642).⁸

⁷ See, especially, his chapters “The Stream of Thought” and “The Perception of Time.”

⁸ See Mangan (1991) for an extensive discussion of James’ phenomenology, especially the function of “fringe” experience in mediating conscious (serial) and nonconscious (parallel) interaction, and the relation of this analysis to connectionism.

The point here is not that James is necessarily right about all of this. In any case, this is not the venue to work out in detail the relation of James' views of consciousness to Dennett's claims. The point is simply that there are other alternatives beyond Dennett's forced choice of "Cartesian materialism" or his own proposals. But this much should already be clear: There does not seem to be anything in the notion of parallel processing and a *generic* notion of "Multiple Drafts" that is inconsistent with the intuition that consciousness is in some sense "its own place." Insofar as Multiple Drafts incorporates the general notions of parallel processing, it is compatible with *either* a Cartesian Theater view of consciousness (which may reflect a core intuition that consciousness is an information-bearing medium) or with Dennett's own proposals.

All in all, it must be said that Dennett has written an intriguing book, perhaps an important book, and, in one sense (for all its bait and switch tactics), a very honest book: It exudes a friendly intellectual *hubris* from its title to its final page. And yet it is balanced by the recognition that "we often learn more from bold mistakes than from cautious equivocation" (p. xi). I suspect a better motto for Dennett's work would be hard to find.

The fruitfulness of an assumption does not, of course, prove it to be true. But many decades of philosophical "behaviorist" attacks on educated common sense notions of consciousness, beginning with Wittgenstein and Ryle, have not produced any positive account of consciousness. By throwing off behaviorism, and taking up an educated common sense view of consciousness, the cognitive revolution became possible, and it quickly developed a substantial body of research that did further our understanding of consciousness. Now some of these findings have attracted the latest in a long line of philosophical behaviorists, Dennett. Given the notion of hybrid vigour, it is possible that Dennett's way of integrating these two schools will produce a new strain of thought that will improve on its progenitors. I myself doubt this and suspect that if successful, Dennett's approach would amount to a behaviorist putsch.

But the value of Dennett's work cannot be judged by its shortcomings alone, and there is no way to know at this point what research program *Consciousness Explained* might further. Unfortunately, the most promising of Dennett's ideas, his attempt to create a phenomenological method designed expressly for cognitive research, was killed in the cradle by Dennett's functionalism. And I have tried to show in a few significant cases how functionalism has damaged some of Dennett's other attempts to work out points of common ground with current cognitive research. Perhaps the most important contribution *Consciousness Explained* will make is to show us just how badly functionalism can skew the empirical investigations of consciousness. Dennett may even have produced the *reductio ad absurdum* of a functionalist analysis of consciousness. But then knowing what to avoid is certainly valuable knowledge.

REFERENCES

- Baars, (1988). *A cognitive theory of consciousness*. New York: Cambridge Univ. Press.
 Block, N. (1978). Troubles with functionalism. Reprinted in N. Block (Ed.), *Readings in the philosophy of psychology*. (1980, Vol. 1). Cambridge, MA: Harvard Univ. Press.

- Dennett, D. (1991). *Consciousness explained*. Boston: Little, Brown.
- Husserl, E. (1960). *Cartesian meditations: An introduction to phenomenology*. Translated by Dorion Cairns. The Hague: Nijhoff.
- James, W. (1890). *The principles of psychology* (Vol. I). New York: Holt.
- Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard Univ. Press.
- Leibniz, G. W. (1981). *New essays on the human understanding*. P. Remant & J. Bennett (Trans). New York: Cambridge Univ. Press.
- Mandler, G. (1975). "Consciousness: Respectable, useful and probably necessary." In R. Solso (Ed.), *Information processing and cognition*. The Loyola Symposium. Erlbaum.
- Mangan, B. B. (1991). *Meaning and the structure of consciousness: An essay in psycho-aesthetics*. Doctoral Dissertation. Berkeley: Univ. of California. Berkeley.
- Mangan, B. B. (1992). *Consciousness unexplained*. Peer Review, *The Sciences*, September/October. New York Academy of Sciences.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton–Century–Crofts.
- Penrose, R. (1989). *The emperor's new mind*. London/New York: Oxford Univ. Press.
- Rock, I., & Gutman, D. (1981). "The effect of inattention on form perception." *Journal of Experimental Psychology: Human Perception and Performance*, 7, 275–285.
- Roskies, A. L., & Wood, C. C. (1992). *A parliament of the mind*. *The Sciences*, May/June. New York Academy of Sciences.
- Rumelhart, E. & McClelland, R., & the PDP Research Group. (1987). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.
- Selfridge, O. (1959). *Pandemonium: A paradigm for learning*. Symposium on the Mechanism of Thought Processes. London: HM Stationery Office.
- Shepard, R. N., & Cooper, L. A. (1982). *Mental images and their transformation*. Cambridge, MA: MIT Press.
- Smolensky, P. (1988). On the proper treatment of connectionism. *The Behavioral and Brain Sciences*.
- Stevens, S. (1975). *Psychophysics*. New York: Wiley.
- Triesman, A., & Gelade, G. (1980). The feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136.