SELF OR NO-SELF? CONVERGING PERSPECTIVES FROM NEUROPSYCHOLOGY AND MYSTICISM

by Brian L. Lancaster

Abstract. The nature of self is examined in relation to psychological observations which reveal some form of dissociation of knowledge from consciousness. Such dissociations are apparent in cases of blindsight, and amnesic patients displaying implicit memory effects, among others. While amnesic patients, for example, are unable consciously to recall material previously presented, such material does influence subsequent physiological and psychological processes. Thus, it is not the memories themselves that have been lost, but the ability to make conscious connection to them. In attempting to account for such observations, theoreticians generally have posited some kind of "consciousness system" that may become dissociated from brain modules dealing with specific processing.

It is argued here that a view of self along the lines of the Buddhist concepts of no-self and the conditioned nature of "I" introduces a more parsimonious perspective on the neuropsychological data. A theory of the nature of self is presented that constitutes a synthesis between key ideas drawn from Buddhist and other mainly mystical traditions and the scientific observations in psychology. Central to this theory is the role that the left hemisphere's interpreter (Gazzaniga 1985; 1988a; 1988b) plays in constructing a unified "I." This "I" is, in effect, a hypothesis that the mind generates to introduce some coherence into otherwise fragmentary mental elements. Although it appears to be the causal focus of the individual's behavior and experience, it is in fact a retrospective construction and not a true causal structure of the mind. This theoretical view is discussed in relation to various meanings of the term *consciousness* and also in relation to the relevant neuropsychological cases.

Keywords: blindsight; Buddhist psychology; consciousness; implicit memory; mind-brain; self.

Brian L. Lancaster is Senior Lecturer in Psychology, School of Health Sciences, Liverpool John Moores University, 15–21 Webster Street, Liverpool L3 2ET, England.

[*Zygon*, vol. 28, no. 4 (December 1993).]
 ISSN 0591-2385

In his study of mysticism, F.C. Happold asserts that "mysticism has its fount in . . . a consciousness of a *beyond,* . . . of an *unseen* over and above the seen" (Happold [1963] 1970, 18). Indeed, union with the divine—the ultimate "beyond"—is frequently held to be the goal of mysticism. The rich imagery associated with mystical experiences may then be interpreted as the form such a "beyond" takes as it is apprehended by the finite mind.

Were this interest in the "beyond" the only feature of mysticism, there would be little room for a discussion of the common ground with recent developments in neuropsychology. However, the general heritage of the mystical traditions also includes much material concerned with the nature of self and the personal ego or "I."[1] Such material serves as a springboard and guide to the novice's own introspections, which form a crucial part of the psychological dimension in the mystic path. Thus, for example, the contemporary Hindu sage Sri Ramana Maharshi advises that "The only inquiry leading to Self-realization is seeking the source of the 'I' with in-turned mind" (cited in Wilber 1977, 321).

In general, the Jewish and Christian mystical traditions do not engage in extensive analysis of the nature of "I." "I" is simply viewed as an obstacle to be overcome on the path towards the mystic's goal. Mystical texts from these traditions give the individual plenty of advice as to how to attenuate "I" through, among others, the renunciation of desires, cultivation of the fear of God, detachment from the world of the senses, and various meditative devices. Unlike the Sufic and, more especially, the Eastern traditions, however, there is little development within the mystical literature of a detailed psychology of self in terms of the minutiae of "I."

Buddhism in particular, and Hinduism to a lesser extent, has developed a rich literature concerned with the manner in which "I" comes into existence, its role in the individual's psychology, and its relation to consciousness. These ideas have impacted on Western philosophical and psychological thought, both directly, through translation and dissemination, and indirectly, through the work of such groups such as the theosophists and such "teachers" as the Caucasian philosopher-mystic Gurdjieff. The Gurdjieffian system in particular is highly psychological in approach (Tart [1986] 1988). It appears that Gurdjieff distilled his philosophy mainly through contact with Tibetan Buddhism and Sufism (Moore 1991). Perhaps in deference to the contemporary Western outlook, he seems to have brought the psychological features of these traditions to particular prominence in his own mystical philosophy. I will accordingly draw on ideas from this system as well as from more traditional Buddhist

sources in this article. My intention is to consider what seem to be the major features of "I" as understood by these mystical traditions and to assess their relevance in relation to developments in our understanding of brain function.

Scientific data are invariably interpreted in relation to some generally accepted paradigm. In the case of psychology, such a paradigm revolves around our general intuitive grasp of the nature of the person and of self. A fundamental tenet of mysticism, however, is that such an everyday intuition of self is misleading. Buddhist teaching, for example, focuses on the notion that there is actually no coherent personal self. The realization of such a notion would not be expected to come about without extensive training in the art of introspection through meditative and other practices. The gist of the present article is that the contemporary "everyday" view of self and, in particular, of the sense of "I" is an impediment to a more parsimonious understanding of recent developments in psychology than that generally promulgated in the relevant literature.

The Psychological and Neuropsychological Background

The relevant psychological literature is concerned with situations in which knowledge or action may be said to be dissociated from normal conscious awareness. Such terms as *knowledge* and *conscious awareness* are difficult to define with precision. For the moment, however, these terms will be used in a general way, much as the "lay person" would, I think, understand them. Thus, if subjects in an experimental situation can respond accurately to some form of test regarding material previously presented to them, then they are understood as having knowledge of that material. If they are unaware of the source of that knowledge, being unable to introspect on the original material, then that knowledge is effectively dissociated from conscious awareness. These terms will be scrutinized more fully below in the introduction of the mystical perspective and again in the final section, which considers neuropsychological dissociations from a more theoretical standpoint.

Blindsight. As an example of my usage of these terms, consider the patient described by Paillard, Michel, and Stelmach (1983). This patient had an extreme insensitivity to touch on the right side of her body, attributable to a defect in her left parietal lobe. In the study by Paillard and his colleagues, one of them touched her affected right arm while she was blindfolded. As expected, she displayed no

awareness of being touched even when considerable pressure was applied. The patient could, however, to her own considerable surprise, point to the approximate location on which she had been touched on her right arm. The dissociation of knowledge from conscious awareness in this case is well captured by the patient's own words: "But I don't understand that! You put something here. I don't feel anything and yet I go there with my finger. . . . How does that happen?" (Paillard, Michel, and Stelmach 1983, 550).

This study has been viewed as analogous to a series of studies demonstrating what has come to be known as *blindsight* (Weiskrantz, Warrington, Sanders, and Marshall 1974. See Cowey and Stoerig 1992; Weiskrantz 1986, 1990 for reviews). Weiskrantz and his colleagues (1974) coined this term to refer to the abilities of their patient DB. DB had had a major part of his right striate cortex surgically removed in an attempt to alleviate severe migraine attacks attributable to a malformation in that brain region. As a consequence of this surgery, DB was "blind" in the greater part of his left visual field. (The striate cortex is the major brain region receiving input from the retinae via the lateral geniculate nuclei.)

The researchers demonstrated that despite this subjective blindness, DB was able to display some degree of knowledge for stimuli presented in the blind field. Thus, again to the patient's own surprise, he could accurately reach towards the location of a light and discriminate between an *X* and an *O* and between lines of vertical, horizontal, or diagonal orientation. In all cases DB viewed his responses as mere guesses.

Marcel (1988) reports evidence of more complex levels of processing for visual stimuli in the blind field of such patients. He has observed, for example, how a patient forced to reach for objects in the blind field will make preparatory adjustments of arm, wrist, and fingers appropriate to the object's shape, orientation, size, location, and distance. All of these properties must accordingly have been adequately processed despite the lack of visual awareness.

The core characteristic of these demonstrations of blindsight is that patients' visual abilities are unaccompanied by any visual awareness. In interpreting these observations, most authors focus on the visual pathways to the midbrain and the subcortical regions of the brain which are still intact in these patients. These pathways represent some 10 percent of the retinal output. Presumably, the residual visual functions observed in blindsight patients are subserved by brain regions reached by these spared pathways. It seems reasonable to infer further that conscious awareness of visual input in normal individuals is brought about in some way which is dependent on

the major pathway to the striate cortex. The deeper implications of blindsight for our understanding of consciousness will be considered in the final section of this article.

Prosopagnosia and Amnesia. Recent reviews (Schacter, McAndrews, and Moscovitch 1988; Milner and Rugg 1992) have drawn attention to a number of other neuropsychological conditions in which parallels may meaningfully be drawn with the dissociation of knowledge from conscious awareness apparent in blindsight. Weiskrantz is hardly exaggerating when he refers to these as a "virtual epidemic of dissociations" (Weiskrantz 1992, 2), and he rightly emphasizes their centrality to current speculation regarding the mechanisms of conscious awareness. I will give examples of two such conditions—prosopagnosia and amnesia.

Prosopagnosic patients have a specific impairment in the ability to recognize faces. Thus, when shown faces which would normally be immediately recognized (e.g., those of family members or famous media figures), these patients typically score at chance level on tests of recognition. More subtle testing, however, has revealed signs of spared memory for faces in some patients, but in a manner which is outside of conscious awareness. Thus, for example, familiar faces presented to prosopagnosic patients give rise to a different pattern of skin conductance responses (Tranel and Damasio 1985), eye movement patterns (Rizzo, Hurtig, and Damasio 1987), and evoked brain potentials (Renault, Signoret, Debruille, Breton, and Bolgert 1989) by comparison with those recorded in response to unfamiliar faces. Psychological tests of a behavioral nature have also demonstrated covert recognition of faces in prosopagnosic patients (e.g. De Haan, Young and Newcombe 1987).

Analogous effects have been demonstrated in amnesic patients. Graf and Schacter (1985) have coined the phrase *implicit memory* to refer to the kind of memory spared in these patients. Thus, "implicit memory is revealed when previous experiences facilitate performance on a task that does not require conscious or intentional recollection of those experiences" (Schacter 1987, 501). Implicit memory is contrasted with *explicit memory* which "is revealed when performance on a task requires conscious recollection of previous experiences" (Schacter 1987, 501).

Studies demonstrating implicit memory effects in amnesic patients have been extensively reviewed (Richardson-Klavehn and Bjork 1988; Roediger 1990; Schacter 1987; Shimamura 1986; see also Lewandowsky, Dunn and Kirsner 1989). A study by Graf, Squire and Mandler (1984) serves well as an illustration of the basic

paradigm and pattern of results. Amnesic patients were presented with lists of words and subsequently tested for their memory of the words. The patients were impaired relative to controls on tests of free recall, recognition, and cued recall. However, a test of word completion revealed normal performance. The cued recall and word completion tests both employed word "stems" comprising the first three letters of words initially presented. The difference between the two tests lay in the instructions to subjects. In the cued recall tests, subjects were instructed to use the stems to help them *remember* the words ("Try to think of a word from the [lists] with the same beginning letters"), whereas the word completion test avoided any reference to the memory aspect of the study, asking subjects simply to write "the first word that comes to mind" to complete the stem. The difference in patients' performance across these two conditions indicates the spared memory.

Certain amnesic patients do not show all the implicit phenomena to the extent that other amnesic patients do. Similarly, not all prosopagnosic patients display covert recognition of faces (Newcombe, Young, and De Haan 1989). Such negative effects are presumably explicable in terms of the complexity of underlying brain damage. Patients displaying similar gross psychological impediments may not be uniform in their extent of injury or in the specific cognitive processes affected. Thus, for example, Newcombe and her colleagues suggest that prosopagnosia in which covert effects are not apparent may be attributable to damage affecting specifically perceptual aspects of face processing. This would contrast with those cases in which damage interferes more specifically with the memory dimension in face recognition and does give rise to covert effects.

The presence of implicit, or covert, effects in at least a large proportion of cases indicates that our models of memory need to offer some kind of explanation. It has furthermore been established that implicit phenomena occur in normal subjects. For example, demonstrations of the psychological effects of subliminal inputs bear similar characteristics to those of implicit memory in amnesics, and we may infer that a common mechanism is at work (Schacter 1987). The related priming effects, extensively observed in normal subjects in the absence of conscious memory, similarly point to a generalized memory capacity dissociated from consciousness.

The CAS and the Executive System. It has been extensively argued over recent years that the brain is modular in function (Fodor 1983; Gardner 1983; Gazzaniga 1985; Mountcastle 1978; Shallice 1988). Some have suggested that the various phenomena indicating a

dissociation of knowledge from consciousness may be conceptualized by positing a separate system for consciousness in some form—perhaps global or supervisory in nature. It is argued that a module may continue to function perfectly adequately even though it may have become in some way cut off from the brain system responsible for consciousness.

In a recent version of such a theory, Daniel Schacter (1989; 1990) posits a Conscious Awareness System (CAS) which is distinct from, but normally interacts with, the modules that process information pertaining to specific cognitive domains. Thus, a prosopagnosic patient demonstrating implicit memory for faces may be viewed as suffering from some form of damage which has severed the connections between the module(s) dealing with face processing and the CAS. In Schacter's model, the CAS sends outputs to another system, the Executive System, which is considered to be responsible for intentional retrieval of memory. Thus, amnesic patients are not only unconscious of their spared memory, but also unable to consciously initiate retrieval.

Schacter's model is useful to the extent that it appears to account for most relevant observations of implicit phenomena. He also proposes some tentative neuroanatomical bases for the CAS and the Executive System, thus introducing a stimulating additional level of explanation to that which is only cognitive. However, modelling the deficit in amnesia does not necessarily explain it. Indeed, as Schacter himself remarks, "to postulate that conscious awareness depends on a specific mechanism in no way explains how consciousness is achieved or exactly what it is" (Schacter 1990, 369).

It is with a view to advancing the debate into the difficult areas raised by these latter questions that I turn now to a consideration of the contribution that may be brought from mystical literature.

On "I"—Perspectives from Mysticism

As a scientific discipline, psychology is wary of probing the nature of consciousness. Schacter's approach is representative of many models in cognitive psychology and neuropsychology in its attempt to conceptualize relevant processes and leave consciousness as some kind of "added ingredient"—an extra box on a flow diagram. It is assumed that the explanatory power of the rest of the model is not affected by any uncertainty over the real nature of the consciousness box in the flow diagram. It is this assumption that I regard as questionable.

While there is a sense in which consciousness is certainly a mystery

and beyond the lens of psychological science, there is a second sense in our use of the term which is more directly available to psychological analysis. Consciousness in the first sense might be conceptualized as the backdrop to mental action. It is a largely indefinable quality which somehow underlies the reality of mind. I regard it as indefinable since it cannot be reduced to other common categories such as process or structure. It cannot be described objectively since one cannot stand outside it. It is nevertheless the very heart of experience—the "Within" to things, as Teilhard de Chardin puts it (Teilhard de Chardin [1955] 1959). Some might refer to this as "pure consciousness" or simply "awareness." It may be a property of the biosphere as a whole (Deikman 1973) or even a fundamental property of the universe (Bohm 1980; Schrödinger 1964). More prosaically, it may be an emergent property of brain processing (Sperry 1969). There are many candidates by way of explanation but few answers! This form of consciousness, which I shall refer to as Consciousness I, is the business of metaphysics, not psychology.

I follow Edoardo Bisiach in identifying a different meaning of the term consciousness (which he refers to as C_2), namely "the access of parts or processes of a system to other of its parts or processes" (Bisiach 1988, 103). In particular, I characterize what I shall call Consciousness II as arising when "I" accesses, or is in some way connected to, other mental events. Thus, when I say that I am conscious of my pen, it is Consciousness II that is being described, and it arises by virtue of the connection effected between "I" and those brain processes concerned with the pen. What exactly "I" is, is the subject of further discussion in the final section of this article. For the moment, it should be noted that it is viewed as a model of self which includes not only bodily, but also other dimensions of our experience of personhood.[2]

To return to models such as that of Schacter, depicting consciousness I as somehow beyond the explanatory power of the model is clearly necessary. Consciousness II, however, is of the essence of what the model attempts to address. The hallmark of the amnesic patient displaying implicit memory effects, for example, is that the key memory images are not available to "I"—the patient has no consciousness of them. The question is, Why not? What is it about memory and self that can lead to the kinds of dissociations of knowledge from consciousness (that is, specifically Consciousness II) discussed in the previous section? I think the answer to these questions lies in an analysis of the nature of self, rather than in further refinements of our models of types of memory system.

Weiskrantz suggests that "to 'remember' is to enable one to

compare past with present, to reflect, to link separate past events, to order them, and to do so *in relation to one's self as a coherent 'thing'* " (Weiskrantz 1988, 193, italics mine). Thus, he would hold the amnesic patient to have difficulty in relating the memory images to self, much as I have suggested above (substituting "I" for "self"). But *is* one's self a coherent "thing," as he posits? If it is not, as an examination of mystical literature suggests, where does the difference between the normal individual and the amnesic or blindsight patient really lie? The question will not go away simply by putting Consciousness II into a box. It isn't the icing on the cake; it's the cake itself!

According to Serge-Christophe Kolm, the first two words of the Buddhist credo, *ye dhamma,* mean both "I follow the doctrine" and "I am only [composed of] simple elements" (Kolm 1986, 255). This latter meaning conveys the essence of the notion of "no-self." The sense of "I" is clearly an all-pervading human characteristic and cannot be denied. However, the sense that it is an actual thing in itself—an unchanging essence of the person—is, according to Buddhism, an illusion. "I" represents a conditioned attachment to objects of the senses, thoughts and feelings. As objects are perceived, this sense of "I" attaches to them, thereby giving them immediate personal meaning. Given, however, that there is no core to self, no "I" in itself, each element with its "I" connection is essentially disparate. As Krishnia Venkata Ramanan writes, "In truth the self that is the object of the notion of 'I' is a complex of the functions of elements that appear and disappear, but the ignorant hold to the self as a simple, substantial entity. This is an error" (Venkata Ramanan [1966] 1987, 57–58).

Derek Parfit (1987) refers to this perspective on self as the *bundle theory,* in contradistinction to the more traditional *ego theory* of Western psychology. We find the concept argued particularly strongly in the teaching of Gurdjieff. Ouspensky, one of the principal transmitters of Gurdjieff's ideas, expresses the point as follows: *"Man has no permanent and unchangeable I.* Every thought, every mood, every desire, every sensation, says "I". . . . Man has no individual I. But there are, instead, hundreds and thousands of separate small I's, very often entirely unknown to one another. . . . *Man is a plurality*" (Ouspensky 1950, 59).

The individual is thus characterized by a kind of multiplicity of "I." Furthermore, just as there is no unity of self in any given moment, so there can be no continuity of "I" over time. The conviction we have that a continuous core of "I" lies at the heart of our experience is simply an illusion. The sense of "I," as indeed the sense

of any object perceived by the mind, arises as a conditioned pattern imposed upon an ever-changing set of elements. The apparent continuity is in the conditioning, not in the elements themselves. Central to such conditioning is one's personal name since it encourages one to believe that the self designated by that name is continuous from moment to moment, just as is the case with the name itself. For the Buddhist, of course, such a view of self is illusory.

In Buddhist thought, the person—and the sense of "I"—is in the category of a "derived name." It derives its being from the causal factors to which it relates, but it has no independent existence. It therefore follows that we are deluded when we consider ourselves (that is, "I") as the instigator of action. Parfit quotes the words of the Buddha: "O Brethren, actions do exist, and also their consequences, but the person that acts does not" (Parfit 1987, 21). Clinging to the sense of "I" as the instigator of action merely distracts one from the actions themselves.

For Buddhism, a favored image to depict the mind is that of an excited troupe of monkeys. Each monkey seems to be incessantly chattering with apparent oblivion to its neighbors and mindlessly flitting from one branch to the next. So too with the untrained mind, which effectively consists of separate parts articulating disparate desires. Maurice Nicol, a pupil of Gurdjieff, describes the mind's incessant chattering as conversations emanating from the many distinct "I's" within us: "Have you ever listened to your "I's" talking in you? Often "I's" carry on a long conversation, but you do not observe it. You think it is you talking to yourself" (Nicoll [1956] 1984, 1721). These "conversations" are themselves reinforcements of each "I"s' precarious identity. By articulating its own viewpoint, it solidifies its claim to represent the mind of which it is actually only a part. However, as we have already seen, such an "I" lacks continuity. In a subsequent moment it may be toppled as another "I" asserts its role on the mind's center stage. The result is a kind of *self-commentary*—an incessant chattering of what are in effect subpersonalities (Rowan 1990). Any continuity within this situation lies only on the surface of the commentary. Our delusion lies in an acceptance of this surface level as emanating from a true self. In reality the surface of the commentary only masks the fragmentary nature of the "actors" (subpersonalities) themselves.

Such a perspective is described in these traditions along with instructions as to how one might overcome the state of delusion; how one might begin to introduce a level of discipline into the mind. However, my concern in this article is only with the model of mind being articulated. Does this model advance our ability

to explain the kinds of neuropsychological phenomena discussed earlier?

Before attempting an answer to this question in the next section, it would perhaps be useful to summarize the key points of the view of the mind held by Buddhism and allied traditions. First and foremost, there is *no coherent self.* Clearly then, there can be *no continuity of self.* The sense of "I" arises as a momentary mental concept having a conditioned attachment to other current mental processes. There is therefore not a single "I" but a *multiplicity of "I".* A kind of *self-commentary* masks this multiplicity. Finally, just as there is no self as a perceiver of the world, so there is no self initiating actions in the world. Actions are initiated automatically.

TOWARDS A SYNTHESIS

In an attempt to draw together the neuropsychological and mystical material reviewed in the previous sections, it is fruitful in the first place to conceptualize "I" as a mental model (Blackmore 1986; Oatley 1988). Such a formulation may be said to find its root in psychological literature in William James's argument that "I" is a *thought.* As such it has no intrinsic, unchanging permanence. The apparent continuity of "I" comes about because each successive "I-thought" includes the previous one in its make-up.

We may conceive of the "I" as a model generated by the brain, giving some degree of coherence to current brain activity related to external and internal inputs. As Susan Blackmore puts it, "My model of self in the world, my model of reality, 'I,' is a vast, intricate construction, forever trying to model everything; and forever at the mercy of changes in input" (Blackmore 1986, 83). In line with the mystical perspective, we may accept that this model has no intrinsic permanence. It is a model which is constructed one moment and reconstructed the next—as Blackmore suggests, "forever at the mercy of changes in input." In relation to the neuropsychological data, we may simply assume that the brain damage giving rise to the conditions discussed earlier somehow precludes key information from being incorporated into the current construction of "I."

This, however, begs as many questions as it answers. How is "I" constructed? Why may some material, as in the case of amnesics or blindsight patients, be excluded from incorporation in "I"? In what way does the illusion of the continuity of "I" arise? In attempting to answer the first of these questions, I posit that information is stored in memory together with a representation of the "I" which arose at the time that the information was registered. I call this stored

representation of "I" an "I-tag" (Lancaster 1991). The "I" experienced in any given moment itself derives principally from all those "I-tags" currently activated.[3] The manner in which this is achieved will be discussed below.

In exploring these ideas, it is necessary to be cognizant of the highly interactive and dynamic nature of the memory system. In any moment many stored representations are being activated by the current stimulus array, not to mention other ongoing thoughts. Each such representation not only contributes its "I-tag" to the current "I" model, but is also itself continually updated since it is envisaged that the current "I" becomes the new "I-tag" attaching to that representation.

While the foregoing may sound somewhat complex, or even bizarre, it is predicated on a few quite straightforward premises. These may be summarized as follows.

1. Currently active cognitive/neural models are presumed to automatically activate related memory traces.
2. Currently active models are presumed to automatically generate memory traces of themselves.
3. All models active at the same time are presumed to generate memory traces having some associative relation to one another.

Given that "I" is a cognitive/neural model, it follows from premise 2 that it will generate a memory trace. Premise 3 implies that this trace will bear a strong associative connection to the traces of other models current at the same time. My term "I-tag" simply reflects these two points—that it is a memory of that moment's "I" model, and that it is strongly associated with contemporaneous mental activity.

Premise 1 underlies the notion of "I" being constructed from currently active "I-tags." As current sensory activity generates neural models (e.g., arising through feature coding), related memory traces are activated. Each of these delivers its "I-tag," which is then incorporated into the current model of self. Clearly, in any moment a highly complex array of "I-tags" is generated. And from moment to moment this array will constantly change. This situation is reminiscent of the multiplicity of "I" and the absence of any continuity in "I" discussed in the previous section.

I envisage these "I-tags" to be constellated into broad groupings in the individual's memory structure. Such a grouping would be structured around features of "I" common to diverse situations. To the extent, however, that discontinuities in "I" across different circumstances were pronounced, parallel discontinuities in the organization of "I-tags" would be established. Such discontinuities I

hypothesize to underlie the fragmentation of the individual into subpersonalities. The organization of "I-tags" into groupings presumably introduces a degree of coherence into the situation, reducing what would otherwise be the chaos of a constant flux of "I-tags." Nevertheless, such coherence would be relative only.

The Interpreter Module. I hypothesize that a final stage of coherence is engendered by a specialized module that Michael Gazzaniga (1985; 1988a; 1988b) considers to operate in the brain's left hemisphere (in the majority of right-handers). He has termed this module the *interpreter.* Gazzaniga views the brain as consisting of functional modules as described earlier. The interpreter's role is such that it "considers all the outputs of the functional modules as soon as they are made and immediately constructs a hypothesis as to why particular actions occurred" (Gazzaniga, 1988b, 219).

In Gazzaniga's view, the hypothesis generated by the interpreter becomes the conscious belief of the subject. When, for example, the command *walk* was flashed to the right hemisphere of a commissurotomy patient, the patient got up and began walking. When asked why he was walking away from the testing area, the patient replied, "I'm going into the house to get a Coke." For Gazzaniga, this statement reflects the cause of the walking behavior as construed by the interpreter. The patient himself *believes* his thirst to be the cause of his walking. The interpreter thus compulsively fabricates a causal structure within the individual's experience. In this case it is unable to include the command word *walk,* which is restricted to the right hemisphere on account of the commissurotomy.

"I" as Hypothesis. The delusion associated with such mental assignation of causes is redolent of the Buddhist perspective on the mind discussed earlier. In bringing Gazzaniga's view of the interpreter in line with Buddhist thought, we may postulate that the same process as that generating a causal explanation of behavior is involved in generating "I." *"I" itself is a hypothesis;* it is a model generated to bring unification to the disparate output of functional brain modules. It is a hypothesis of an ordered causal relationship pertaining between the profusion of "I-tags" activated from moment to moment. Put simply, this hypothesis *is* the unified "I" which is central to our experience. It is not that the interpreter is simply binding "I-tags" together. Rather, it synthesizes an ongoing narrative out of events, whose very coherence is founded on this concept of a unified "I" as actor.

I am proposing that the incoherence that would be associated with

the shifting mix of "I-tags" is overcome because the interpreter generates a single "I" as the subjective focus of experience. As the mystical view asserts, "I" exists only as a conditioned frame of reference for other mind events. I further suggest that the "self-commentary" discussed in the previous section is itself the work of the interpreter. The interpreter in effect weaves a web of explanation—primarily verbal in nature—to unify a person's experience. This last point raises issues concerning the relation of language to self, which I have discussed elsewhere (Lancaster 1991).

CONCLUSION

In summary, it is proposed that "I-tags" provide the matrix from which the sense of a unified "I" is generated by the interpreter. "I" is thus a retrospective construction forever keyed into one's past experience of images represented by present mental activity. In other words, as soon as "I" am, I am no longer in the present moment.

Consciousness, in the Consciousness II sense, depends upon relationships to the "I" constructed by the interpreter. Events whose mental representations are incorporated into this "I" model are those events of which the subject is conscious.

We may now see how this "I-tag" theory advances our understanding of those dissociations of knowledge from consciousness discussed earlier. In common with other theoreticians, I envisage a breakdown between those modules handling specific task-related information and the "consciousness system." I go further than the others by exploring how this consciousness system may be involved. My foregoing discussion indicates that "I-tags" and the construction of "I" by the interpreter underlie consciousness of relevant material in a crucial way. The kind of dissociations discussed earlier may be attributed to a breakdown either in the interpreter's access to relevant information or in the associative bonding between memory traces and appropriate "I-tags."

The first kind of problem I consider to underlie blindsight, for example. Whatever residual visual processing does take place, it is presumably unavailable to the interpreter. The visual information is therefore never incorporated in the construction of "I"—it remains "unconscious." Weiskrantz (1986; 1988) has convincingly argued that the problem in blindsight lies in the operation of a *commentary system.* In order for us to be "aware" of visual information, such information has to have been included in the ongoing schematizations which characterize the commentary system. The commentary system is viewed by Weiskrantz as being activated "when we

characteristically are able to carry out cognitive operations upon current or stored events" (Weiskrantz 1986, 170). In the terms advanced in the present article, I consider these "cognitive operations" to be those involved in generating a coherent focus for conscious experience, i.e., operations directed to generating "I" as discussed above.

Amnesics displaying implicit memory effects I would explain in terms of the second kind of problem mentioned above—a disturbance in the associative bonding between memory traces and appropriate "I-tags." I envisage that explicit recall comes about when the current "I" establishes connection with the target material's "I-tag." If I wish to remember where I put my keys, I effectively attempt to connect to the "I" that lost them. If I can effect this connection, then the memory "pops into" my consciousness. This theory proposes that it is not only memories of events/objects that require consolidation, but the "I-tags" associated with them. The more coherence there is between successive constructions of "I," the more effective will be the individual's "conscious" memory.

To take the example of the prosopagnosic patient showing implicit memory effects, I postulate some form of breakdown in the storage/consolidation of "I-tags" associated with faces. The memories of faces themselves are, in fact, intact.

It should not be surprising that in setting out to explore the nature of self, I have introduced a concept—that of "I-tags"—the full understanding of which may come about only when memory itself is understood. Self and memory are intimately interrelated. They are "merely two sides of the same fact, or two different modes of viewing the same fact" (John Stuart Mill, cited in James [1890] 1950, 356). I believe that any difficulty in understanding my concept of "I-tags" is in principle no different from that involved in understanding the nature of any memory. What form does the physical representation of psychological events (the engram) take, and how is it that this physical structure encodes nonphysical experiences? These are questions to which contemporary psychology has only the vaguest of answers. I may be able to specify some neurons that respond to light of a given wavelength, for example, but how does a neuronal network (if that is what the engram consists of) give rise to the memory of a rose's redness or beauty? I see no reason why the remembered beauty of the rose is any different in principle from the remembered "I-ness" which is integral to the experience.

As Mary Warnock argues, "to count as a memory a cognitive experience, or thought, must contain the conviction that I myself was the person involved in the remembered scene. The image, if there is one, must be labelled not only 'this belongs to the past' but

also 'it belongs to *my* past'" (Warnock 1987, 59). Whatever its cerebral basis, I regard the "I-tag" as such a label. The theory enunciated in this article has explored the intricate relationship between the label as a memory event and the sense of self as an ongoing construct.

As far as brain mechanisms are concerned, I think that the kinds of properties associated with hippocampal processes in memory are of particular interest with respect to "I-tags." Many aspects of memory, including short-term memory and long-term priming effects, appear to be little affected by damage to the hippocampus. But where circumstances require representations of a cognitive nature and/or conscious retrieval, its presence becomes critical (Hintzman 1990). I would argue that these are precisely the kinds of circumstances in which the "I-tag" mechanism is required.

Teyler and DiScenna (1986) have proposed that when memories are laid down in the cerebral cortex, an index is established in the hippocampus. The hippocampal record indexes the cortical location of the relevant memory. I envisage the "I-tag" as playing much the same role as this hippocampal index. Indeed, it could be argued that for humans, the "I-tag" *is* the index. In other words, there may be some evolutionary continuity between the self-referencing system in lower mammals and that in ourselves. The difference in this regard between myself and a rat is perhaps only that I am conscious of the self-reference to particular memories (i.e., I am *self-conscious*). The hippocampal index even for a rat may itself well be a representation of self in a more primitive form. Lacking an interpreter, the rat surely has no "I," so the index could not consist of "I-tags." Nevertheless, the index may be presumed to include the rat itself as the central reference point within whatever system of coordinates (e.g., spatial) employed.

The major points which arose through my consideration of the mystical perspective have all found expression in the theory enunciated here. The doctrine of "no-self" seems to me particularly apposite to the data with which neuropsychology is currently grappling. The "I-tag" theory presents itself, I believe, as a fruitful synthesis of mystical and scientific traditions. As an articulation of the way in which the sense of "I" is devoid of real substance or continuity, it may be seen as a commentary on Buddhist and similar doctrines. To the extent that it draws on data arrived at through scientific observation, especially of neurological patients, it is a theory that sits within the Western scientific tradition. Just as it has arisen through an attempted integration of introspective and

scientific observation, so its status will be subject to these same two branches of substantiation.

The theory will stand or fall on its ability to generate hypotheses amenable to experimentation. But it should, I feel, also be evaluated in relation to our introspections, particularly where these have been sharpened by the kind of mental discipline that mystical traditions ideally develop. The key concepts—those of "I-tags," the multiplicity of "I," the role of the interpreter, etc.—are not mere intellectual abstractions, but aspects of that which is central to the experience of every one of us—the self. Ultimately, the "I-tag" theory, if accepted, would change the way we view our very selves. And that is perhaps where its strongest affinity with the mystical traditions really lies.

NOTES

1. Throughout the present article the term *I* is used in preference to *ego*. Two factors underlie this preference. First, differing authors have introduced differing shades of meaning to the term *ego* over the years, and it consequently tends to convey a somewhat imprecise meaning today. Secondly, *I* is a more direct term for English-speaking readers. *Ego* has acquired the connotation of something cold and removed, whereas *I* conveys the intended meaning, namely the immediate personal sense of the experiencing self.

2. I am not intending to imply that these two different uses of the term *consciousness* reflect a difference in the fundamental sense of what consciousness actually is, for that would be to engage in the metaphysical issues, which are beyond the scope of the present article. I wish only to draw attention, for purposes of clarification, to the distinction in our usage of the terms. The difference between Consciousness I and Consciousness II arises specifically with respect to the role that "I" plays in the situation.

The distinction is central to the mystical perspective on Mind as some kind of fundamental reality in the universe which transcends the individual mind. As Ken Wilber notes, "the Level of Mind, being pure consciousness, is never conscious of itself, and so is unconscious" (Wilber 1977, 147). To *be conscious of* requires an experiencing "I," that is, an identification with the object's relation to one's personal, subjective world. In the mystical transcending of subjectivity, the mystic may "consciously live as Mind" (Wilber 1977, 146), but she or he is not conscious *of* Mind. There is no paradox in these statements so long as the distinction between the two uses of the term *consciousness* is clearly understood.

3. There is a circularity in this argument concerning the nature of "I." What "I" am in the present derives from what "I" was in comparable situations in the past. Such a formulation relates to what might properly be described as the *habitual* sense of "I." A complete description of the sense of self would need to include some consideration of the manner in which new elements become incorporated within "I" (particularly with regard to its development in childhood). Such a brief is beyond the scope of the present article, concerned as it is with the kinds of neuropsychological cases discussed earlier.

REFERENCES

Bisiach, Edoardo. 1988. "The (Haunted) Brain and Consciousness." In *Consciousness in Contemporary Science,* ed. Anthony J. Marcel and Edoardo Bisiach, 101-20. Oxford: Clarendon Press.

Blackmore, Susan. 1986. "Who am I? Changing Models of Reality in Meditation." In *Beyond Therapy: The Impact of Eastern Religions on Psychological Theory and Practice,* ed. Guy Claxton, 71-85. London: Wisdom Publications.

Bohm, David. 1980. *Wholeness and the Implicate Order.* London: Routledge and Kegan Paul.

Cowey, Alan, and Petra Stoerig. 1992. "Reflections on Blindsight." In *Neuropsychology of Consciousness,* ed. A. David Milner and Michael Rugg, 11–38. London: Academic Press.

De Haan, Edward H. F., Andy Young, and Freda Newcombe. 1987. "Face Recognition without Awareness." *Cognitive Neuropsychology* 4: 385–415.

Deikman, Arthur J. 1973. "The Meaning of Everything." In *The Nature of Human Consciousness: A Book of Readings,* ed. Robert E. Ornstein, 317–26. San Francisco: Freeman.

Fodor, Jerry A. 1983. *The Modularity of Mind.* Cambridge, Mass.: MIT Press.

Gardner, Howard. 1983. *Frames of Mind: The Theory of Multiple Intelligences.* New York: Basic Books.

Gazzaniga, Michael S. 1985. *The Social Brain.* New York: Basic Books.

———. 1988a. "The Dynamics of Cerebral Specialization and Modular Interactions." In *Thought without Language,* ed. Lawrence Weiskrantz, 430–50. Oxford: Oxford Univ. Press.

———. 1988b. "Brain Modularity: Towards a Philosophy of Conscious Experience." In *Consciousness in Contemporary Science,* ed. Anthony J. Marcel and Edoardo Bisiach, 218–38. Oxford: Clarendon Press.

Graf, Peter, and Daniel L. Schacter. 1985. "Implicit and Explicit Memory for New Associations in Normal and Amnesic Subjects." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11: 501–18.

Graf, Peter, Larry R. Squire, and George Mandler. 1984. "The Information that Amnesic Patients Do Not Forget." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10: 164–78.

Happold, F. C. [1963] 1970. *Mysticism: A Study and an Anthology.* Rev. ed. Harmondsworth, Middlesex, UK: Penguin Books.

Hintzman, Douglas L. 1990. "Human Learning and Memory: Connections and Dissociations." *Annual Review of Psychology* 40: 103–39.

James, William. [1890] 1950. *The Principles of Psychology.* Vol. 1. New York: Dover.

Kolm, Serge-Christophe. 1986. "The Buddhist Theory of 'No-Self.' " In *The Multiple Self,* ed. by Jon Elster, 233–65. Cambridge: Cambridge Univ. Press.

Lancaster, Brian. 1991. *Mind, Brain and Human Potential: The Quest for an Understanding of Self.* Shaftesbury, Dorset, UK: Element Books.

Lewandowsky, Stephen, John C. Dunn, and Kim Kirsner, eds. 1989. *Implicit Memory: Theoretical Issues.* Hillsdale, N.J.: Erlbaum.

Marcel, Anthony J. 1988. "Phenomenal Experience and Functionalism." In *Consciousness in Contemporary Science,* ed. Anthony J. Marcel and Edoardo Bisiach, 121–58. Oxford: Clarendon Press.

Milner, A. David and Michael Rugg, eds. 1992. *The Neuropsychology of Consciousness.* London: Academic Press.

Moore, James. 1991. *Gurdjieff: The Anatomy of a Myth.* Shaftesbury, Dorset, UK: Element Books.

Mountcastle, Vernon B. 1978. "An Organizing Principle for Cerebral Function: The Unit Module and the Distributed System." In *The Mindful Brain: Cortical Organization and the Group-Selective Theory of Higher Brain Function,* ed. Gerald M. Edelman and Vernon B. Mountcastle, 7–50. Cambridge: MIT Press.

Newcombe, Freda, Andy W. Young, and Edward H. F. De Haan. 1989. "Prosopagnosia and Object Agnosia without Covert Recognition." *Neuropsychologia* 27: 179–91.

Nicoll, Maurice. [1956] 1984. *Psychological Commentaries on the Teaching of Gurdjieff and Ouspensky.* Vol. 5. Boulder, Colorado: Shambhala.

Oatley, Keith. 1988. "On Changing One's Mind: A Possible Function of Consciousness." In *Consciousness in Contemporary Science,* ed. Anthony J. Marcel and Edoardo Bisiach, 369–89. Oxford: Clarendon Press.

Ouspensky, P. D. 1950. *In Search of the Miraculous: Fragments of an Unknown Teaching.* London: Routledge and Kegan Paul.

Paillard, Jacques, Francois Michel, and George Stelmach. 1983. "Localization without Content: A Tactile Analogue of 'blindsight.' " *Archives of Neurology* 40: 548-51.

Parfit, Derek. 1987. "Divided Minds and the Nature of Persons." In *Mindwaves: Thoughts on Intelligence, Identity and Consciousness,* ed. Colin Blakemore and Susan Greenfield, 19-28. Oxford: Basil Blackwell.

Renault, Bernard, Jean-Louis Signoret, Bruno Debruille, Françoise Breton, and Francis Bolgert. 1989. "Brain Potentials Reveal Covert Facial Recognition in Prosopagnosia." *Neuropsychologia* 27: 905-12.

Richardson-Klavehn, Alan, and Robert A. Bjork. 1988. "Measures of Memory." *Annual Review of Psychology* 39: 475-543.

Rizzo, Mathew, Richard Hurtig, and Antonio Damasio. 1987. "The Role of Scanpaths in Facial Recognition and Learning." *Annals of Neurology* 22: 41-45.

Roediger, Henry L. III. 1990. "Implicit Memory: Retention without Remembering." *American Psychologist* 45: 1043-56.

Rowan, John. 1990. *Subpersonalities: The People Inside Us.* London: Routledge.

Schacter, Daniel L. 1987. "Implicit Memory: History and Current Status." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13: 501-18.

———. 1989. "On the Relation between Memory and Consciousness: Dissociable Interactions and Conscious Experience." In *Varieties of Memory and Consciousness: Essays in Honor of Endel Tulving,* ed. Henry L. Roediger and Fergus I. M. Craik, 355-89. Hillsdale, N.J.: Erlbaum.

———. 1990. "Towards a Neuropsychology of Awareness: Implicit Knowledge and Anosognosia." *Journal of Clinical and Experimental Neuropsychology* 12: 155-78.

Schacter, Daniel L., Mary P. McAndrews, and Morris Moscovitch. 1988. "Access to Consciousness: Dissociations between Implicit and Explicit Knowledge in Neuropsychological Syndromes." In *Thought Without Language,* ed. Lawrence Weiskrantz, 242-78. Oxford: Oxford Univ. Press.

Schrödinger, Erwin. 1964. *My View of the World.* Cambridge: Cambridge Univ. Press.

Shallice, Tim. 1988. *From Neuropsychology to Mental Structure.* Cambridge: Cambridge Univ. Press.

Shimamura, Arthur P. 1986. "Priming Effects in Amnesia: Evidence for a Dissociable Memory Function." *Quarterly Journal of Experimental Psychology* 38A: 619-44.

Sperry, Roger W. 1969. "A Modified Concept of Consciousness." *Psychological Review* 76: 532-36.

Tart, Charles T. [1986] 1988. *Waking Up: Overcoming the Obstacles to Human Potential.* Shaftesbury, Dorset, UK: Element Books.

Teilhard de Chardin, Pierre. [1955] 1959. The Phenomenon of Man. Translated by Bernard Wall. London: Collins.

Teyler, Timothy J., and Pascal DiScenna. 1986. "The Hippocampal Memory Indexing Theory." *Behavioral Neuroscience* 100: 147-54.

Tranel, Daniel, and Antonio R. Damasio. 1985. "Knowledge without Awareness: An Autonomic Index of Facial Recognition by Prosopagnosics." *Science* 228: 1453-54.

Venkata Ramanan, Krishnia. [1966] 1987. *Nagarjuna's Philosophy as Presented in the Maha-Prajnaparamita-Sastra.* Delhi: Motilal Banarsidass.

Warnock, Mary. 1987. *Memory.* London: Faber and Faber.

Weiskrantz, Lawrence. 1986. *Blindsight: A Case Study and Implications.* Oxford: Clarendon Press.

———. 1988. "Some Contributions of Neuropsychology of Vision and Memory to the Problem of Consciousness." In *Consciousness in Contemporary Science,* ed. Anthony J. Marcel and Edoardo Bisiach, 183-99. Oxford: Clarendon Press.

———. 1990. "Outlooks for Blindsight: Explicit Methodologies for Implicit Processes." *Proceedings of the Royal Society* B239: 247-78.

———. 1992. "Introduction: Dissociated Issues." In *Neuropsychology of Consciousness*, ed. A. David Milner and Michael Rugg, 1-10. London: Academic Press.

Weiskrantz, Lawrence, Elizabeth K. Warrington, M.D. Sanders, and J. Marshall. 1974. "Visual Capacity in the Hemianopic Field Following a Restricted Cortical Ablation." *Brain* 97: 709-28.

Wilber, Ken. 1977. *The Spectrum of Consciousness*. Wheaton, Ill.: Theosophical Publishing House, Quest Books.