CONSCIOUSNESS AND THE ZOMBIE WITHIN:
A FUNCTIONAL ANALYSIS OF THE BLINDSIGHT EVIDENCE

Ullin T. Place

School of Philosophy,   School of Psychology,

University of Leeds, U.K.   University of Wales Bangor, U.K.

CONTENTS

Abstract

Cowey & Stoerig's (1995) demonstration that the phenomenon of blindsight applies to monkeys with striate cortical lesions in the same way as it does to humans with similar lesions makes it plausible to argue that the behaviour of mammals and probably that of other vertebrates is controlled by two distinct but closely interdependent and interacting systems in the brain which I shall refer to respectively as 'consciousness' and the 'sub-conscious automatic pilot or "zombie" within'.

On this hypothesis, *consciousness* has three functions, (a) that of categorizing any input that is *problematic* in that it is either unexpected or significant relative to the

individual's current or perennial motivational concerns, (b) that of selecting a response appropriate both to the presence of a thing of that kind and to the individual's motivational concerns with respect to it, and (c) that of monitoring the execution of that response. *Conscious/phenomenal experience*, on this view, is the first stage in the process whereby problematic inputs are processed by consciousness. Its function is to modify the figure-ground relations within the central representation of a problematic input until an adequate categorization is selected.

The *sub-conscious automatic pilot* or *"zombie-within"* has two functions (a) that of continuously scanning the total current input and alerting consciousness to any input it identifies as problematic, (b) that of protecting consciousness from overload either by ignoring those non-problematic inputs which require no response or by responding appropriately but automatically to those for which there already exists a well practised skill or other "instinctive" response pattern.

1.   The Evolution of A Theory

In this chapter I develop a theory of consciousness and its unconscious counterpart which I call the "zombie-within". It has its source in two lines of research both of which originated in the 1950s, now more than forty years ago. One of these was an attempt made by the present author in two papers published in the *British Journal of Psychology*, 'The concept of heed' (Place 1954) and 'Is consciousness a brain process?' (Place 1956), to examine the implications for the science of psychology of the work of Wittgenstein (1958; 1953) and Ryle (1949) on the linguistic analysis of what Ryle calls "the logical geography of our ordinary mental concepts". The other was the late Donald Broadbent's (1958) experimental and theoretical analysis of the phenomenon of selective attention in his book *Perception and Communication*.

1.1   Place 'The Concept of Heed' (1954) and 'Is Consciousness A Brain Process?'
       (1956)

In *The Concept of Mind* Ryle (*op. cit.*) shows is that many of our most common psychological verbs, verbs such as 'know', 'believe', 'understand', 'remember', 'expect', 'want' and 'intend', do *not*, as had been traditionally supposed, to processes within the individual of whom they are predicated to which he or she has "privileged access" through the process known as "introspection". These verbs refer to *dispositions* or *performance characteristics* of the individual which are manifested as much in what he or she publicly says and does as in his or her private mental processes. However, the application of these same techniques of linguistic analysis also shows

> an intractable residue of concepts clustering around the notions of consciousness, experience, sensation and mental imagery, where some sort of inner process story is unavoidable. (Place 1956, p. 44)

It was this "intractable residue" to which I was referring when I argued in the same paper that

> the thesis that consciousness is a process in the brain is . . . a reasonable scientific hypothesis, not to be dismissed on logical grounds alone (Place 1956, p. 44)

Central to this concept of consciousness was the idea that the verb phrase 'paying attention to _____' refers to an *internal non-muscular activity* whereby the individual

> exercises a measure of control over the vividness or acuteness of his consciousness of (a) the sensations to which he is susceptible at that moment, or (b) such features of the environment as are impinging on his receptors, without necessarily adjusting his receptor organs or their position in any way. (Place, 1954, p. 244)

In contrast to Ryle who had argued that to pay attention was to perform whatever task one was engaged in at the time with a disposition to succeed in it, I pointed out that

> close attention to his own activity will be of no avail to the unskilled person because he has not learnt to discriminate between the relevant and irrelevant features. On the other hand an acute consciousness of the details of his own activity in relation to the environment may actually detract from the efficiency of performance in the case of an individual who has learnt to make many of the adjustments involved automatically. (Place 1954, p. 247)

Here we have the germ of two ideas which are fundamental to the theory expounded below, (a) the idea that conscious experience is *not*, as it has been too often portrayed by philosophers, a mere passive spectator of what is going on inside and outside the organism, but, when properly focussed, is an integral part of the process whereby the

behaviour of the organism is brought into an adaptive relation to the environmental contingencies, and (b) the idea that in order to perform that function successfully, the implementation of the *tactical details* of a skilled performance must be handed over, as it were, to what I am here calling the "automatic pilot" or "zombie-within" in order to free consciousness to concentrate on those features of the task where important *strategic decisions* are called for.

Although I did not emphasise this point at the time, it will be apparent that the role assigned in this account to consciousness in general and conscious experience in particular is one which has as much application to the control of animal behaviour as it has to that of human beings. What I did not then appreciate is that the other function of consciousness which I emphasised both in 'The concept of heed' and in 'Is consciousness a brain process?', that of enabling the individual to give a verbal description of those aspects of the current situation on which attention is focussed, also has its roots in a mechanism which plays a key role in animal problem-solving. For, as is shown by research on the effects of lesions of the striate cortex in man and monkey ("blindsight"), without conscious experience of the stimuli involved a monkey is unable to categorize and thus recognise either individuals or things of a kind (Humphrey 1974). What I *did* emphasise, particularly in 'Is consciousness a brain process?', was the idea that the remarkable ability of human subjects to give a running commentary on their private experiences, either at the time or shortly thereafter, is a *by-product* of the ability to give a description of and running commentary on that individual's current stimulus environment *in so far as* attention and consciousness are focussed upon it.

1.2   *Broadbent's* Perception and Communication *(1958) and* Decision and Stress
      (1971)

This theory of the functions of attention and consciousness, as I was later to discover (Place 1969), bears a remarkable resemblance to the theory of selective attention expounded by the late Donald Broadbent (1958) in his book *Perception and*

Communication. Basing his conclusions on results obtained from the *dichotic listening experiment* in which conflicting auditory messages are fed by earphones into the two ears, Broadbent introduced the idea that there is a central information processing unit in the brain that is a *"limited capacity channel"* in the sense that it can only process a limited amount of information coming in from the sense organs at any one time. Such a limited capacity channel or LCC requires a *selective attention mechanism* which protects it from overload, partly by excluding aspects of the current total input which are unproblematic and thus do not need to be processed, and partly by holding other inputs that need to be processed in a short term memory store or *"buffer"* until the LCC-entry bottleneck clears.

In *Decision and Stress* Broadbent (1971) introduced a number of modifications to the model he had outlined in the 1958 book. Three are particularly important for our present purpose:

(1) He introduced the term "*state of evidence*" (i.e. "evidence" about the current state of the environment) to refer to the output of the selective attention mechanism and the input into the limited capacity channel, a notion which corresponds to that of "raw" or uninterpreted experience in traditional psychology.

(2) He proposed that the function of the limited capacity channel is to "pigeon-hole" and "categorize" the "evidence" passed through from the selective attention mechanism, where 'to pigeon-hole' is to routinely assign an unproblematic input to its classification and 'to categorize' is either to create a new classification or to extend or otherwise modify an existing classification so as to accommodate an otherwise unclassifiable input. However, as is shown by the "blindsight" evidence described below, the kind of routine automatic behaviour which on the present hypothesis is assigned to the zombie-within requires no conscious experience to supply the "evidence" and no categorization of objects in the affected part of the visual field. Nevertheless, the individual is able to perform routine visually guided tasks, such as reaching for objects and, in the case of the monkey at least, avoiding obstacles. I infer from this that no classification of the input is

necessary for the automatic routine control of behaviour by visual stimuli or those in the other sensory modalities which, on this hypothesis, is mediated by the automatic pilot or zombie-within, and that, therefore, Broadbent's routine "pigeon-holing" of non-problematic inputs does not exist. The only classification of sensory inputs that occurs is the categorization in consciousness of problematic inputs.

(3) He proposed that the parts of the input that are not in the current focus of attention are not filtered out completely, as proposed in 1958, but rather contribute, to a lesser extent than the part that is in the focus, to what Broadbent calls *"the category state"*, the final outcome of the categorization process. Although Broadbent himself does not use that terminology, another way of putting the point would be to say that the input in the focus of attention stands as *figure* to the inputs outside the focus as *ground*.

1.3   Humphrey's *'Vision in a Monkey without Striate Cortex' (1974)*

In his 1974 paper Nicholas Humphrey writes:

> In 1965 Weiskrantz removed the visual striate cortex from an adolescent rhesus monkey, Helen. In the 8 years between the operation and her death in 1973 this monkey slowly recovered the use of her eyes, emerging from virtual sightlessness to a state of visual competence where she was able to move deftly through a room full of obstacles and could reach out and catch a passing fly. (Humphrey, 1974, p. 241)

Nevertheless,

> After years of experience she never showed any signs of recognizing even those objects most familiar to her, whether the object was a carrot, another monkey or myself. (Humphrey, 1974, p. 252).

The full significance of this observation for the theory of consciousness has only become apparent in the light of two subsequent discoveries. The first of these was Weiskrantz's (1986) demonstration that in addition to retaining the same visual abilities (apart possibly from the ability to avoid obstacles which is not demonstrable in a subject with only a partial lesion of the striate cortex) lesions of the striate cortex in man have the effect of completely abolishing visual conscious experience in the affected part of the visual field. The second discovery was Cowey and Stoerig's (1995; 1997; Stoerig & Cowey, 1997) demonstration that lesions of the striate cortex in the monkey have the same effect in

abolishing visual conscious experience of stimuli in the "blind" field as they do in human subjects, despite the fact that the animal can reach for objects in that part of the visual field with almost the same accuracy as for objects in the intact field.

When combined with these subsequent discoveries and interpreted in the light of Broadbent's model, Humphrey's observations show (a) that the function of conscious experience is to provide the "evidence" on which categorization of current inputs is based and without which no categorization of those inputs is possible, (b) that relying only on the sub-cortical visual inputs available to it, the unconscious "automatic pilot" or "zombie-within" can learn by the process of trial-and-error to make many very accurate visual discriminations including the ability to reach for objects and avoid obstacles by sight, and (c) that the behaviour controlled by the unconscious "automatic pilot" or "zombie-within" (e.g., reaching for "unseen" objects and avoiding "unseen" obstacles in the visual field) does *not* require any categorization of the inputs to which system is responding (human subjects describe such responses as "pure guesswork").

1.4 *Weiskrantz's* Blindsight *(1986)*

As already mentioned, Weiskrantz has shown that the effect of lesions of the striate cortex in man is to abolish conscious experience in the affected part of the visual field. It does so, presumably, by depriving the cortex of the "raw material" from which the "evidence" on which categorization of inputs is based. Nevertheless, as we have also seen, patients such as Weiskrantz's subject D.B., show some remarkable visual discrimination abilities, such as the ability to reach for objects in the blind field with considerable accuracy, the phenomenon to which Weiskrantz has given the name "blindsight". For our present purposes, the two most important additional points to emerge from Weiskrantz's study are (a) that the visual discrimination abilities displayed by the blindsight patient always fall short of the ability to judge spontaneously *what kind of a stimulus* has been presented to the "blind" field (thereby confirming Humphrey's observation that without striate cortex categorization of a visual input is impossible), and

(b) that human subjects with striate cortical lesions can only be induced to display the considerable discrimination abilities they retain in the "blind" part of the visual field, by persuading them to *guess* the location of something or which of two specified alternatives was present in a case where they insist that they *"saw nothing"* showing thereby that *without conscious experience of the input* the subject has no way of checking his judgment against "evidence" on which such judgments are normally based)

1.5   Cowey and Stoerig's 'Blindsight in Monkeys' (1995)

In addition to showing that we can use Humphrey's (1974) study as evidence of the effect of completely depriving an organism of its visual conscious experience, Cowey and Stoerig's (1995) paper also provides us with the first conclusive evidence that Descartes was mistaken in thinking that, because only humans have language, because only *they* can describe what their conscious experiences *are like,* conscious experience is an exclusively human phenomenon. It also provides us with a methodology which, when suitably adapted to the species in question, should allow us to demonstrate the "blindsight" phenomenon in other species of mammal, in birds, in other vertebrates, and perhaps even in some invertebrates. If this latter prediction is fulfilled, it will show beyond serious doubt that conscious experience has been present in the brains of free-moving living organisms for a very long time indeed. Even with only the monkey evidence available, the idea, supported by many contemporary philosophers, that conscious experience is a functionless epiphenomenon which appears only with the emergence of *homo sapiens* can no longer be sustained.

1.6   The Ventral and Dorsal Visual Pathways

Recent neuropsychological research on visual agnosias (Farah, 1990; Milner & Goodale, 1995) has drawn attention to the functional significance of an anatomically identified bifurcation within the visual areas of the brain between two "streams" or "pathways", the *ventral stream* and the *dorsal stream.* As originally defined by Ungerleider and Mishkin

(1982), these two pathways bifurcate downstream of the striate or primary visual cortex (V1). The ventral stream travels *via* the extra-striate visual areas (V2-V5) to the *infero-temporal cortex*. The dorsal stream travels upwards to "terminate" in the *posterior parietal cortex*. In other words, the bifurcation between the two pathways lies entirely within the cerebral cortex.

Studies of the behaviour of patients with lesions restricted to one or other of these two pathways show that lesions of the ventral stream result, depending on the site and extent of the lesion, in a variety of functional disorders involving the loss or disturbance of visual conscious experience associated with a loss or disturbance of the ability to recognise objects and the situations in which they occur, conditions such as *prosopagnosia* (loss of the ability to recognise faces) and *simultanagnosia* (loss of the ability to recognise the relations between multiple objects in a visually presented scene). Lesions of the dorsal stream, on the other hand, result in disturbances of the visual control of voluntary movement.
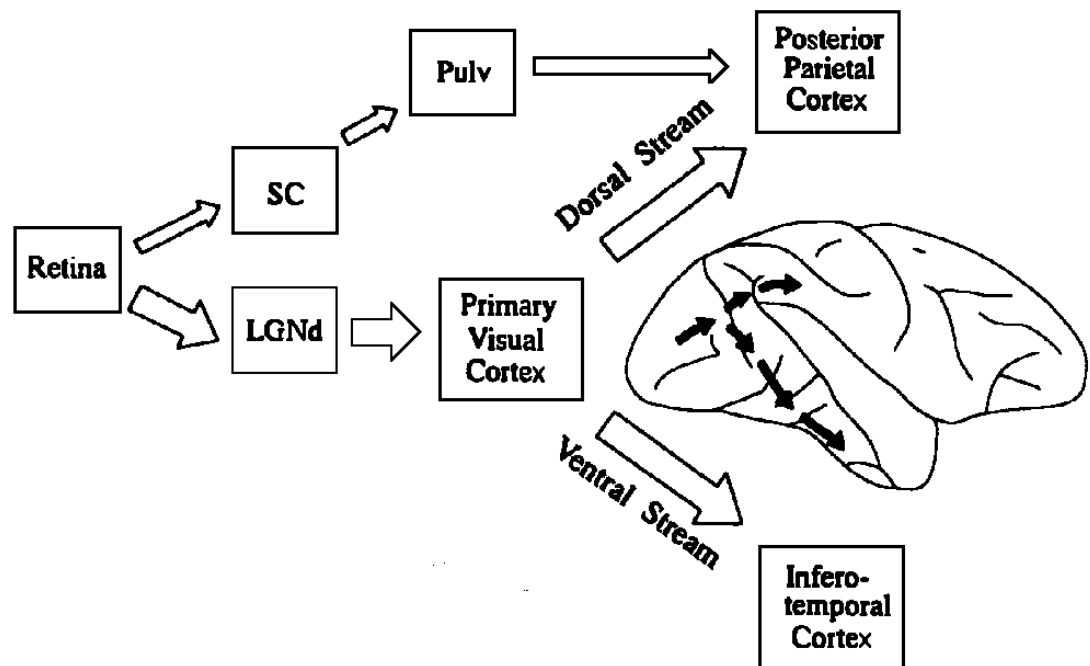

Figure 1: The Ventral and Dorsal Streams (after Milner & Goodale, 1995, Figure 3.1, p. 68). Reprinted by permission of Oxford University Press.

Since the two pathways bifurcate downstream of the striate cortex (V1), they cannot be invoked to explain the phenomena of "blindsight", i.e., the visual functions that survive lesions of V1. However, as is shown on Figure 1, there is *another* pathway converging on the posterior parietal cortex most of which consists of structures lying outside the cortex in the midbrain (*superior colliculus* and *pulvinar*). I call this the "*sub-cortical (S-C) to dorsal*[1] *pathway*". For although it is shown for convenience on Figure 1 above both the dorsal *and* ventral streams properly so-called, in fact, until it reaches its "destination" in the posterior parietal cortex, it is composed of structures (the *superior colliculus* and *pulvinar*) which lie *below* the cortex in the midbrain and thus below both dorsal *and* ventral streams.

As is apparent from Figure 1, identifying the S-C to dorsal pathway gives us a second pair of visual pathways with the same "destinations" as the dorsal and ventral streams (the posterior parietal and infero-temporal cortices respectively), but bifurcating *at the retina* rather than downstream of the primary visual cortex (V1). Balancing the S-C to dorsal pathway is what we may call the "*ventral pathway*" (to distinguish it from the ventral stream which forms part of it) consisting of the *lateral geniculate nucleus*, the *primary visual cortex* (V1), the *ventral stream* (V2-V5) and the *infero-temporal cortex*. The two pathways so-defined differ in two respects:

(1) Apart from the lateral geniculate nucleus, all the structures composing the ventral pathway are in the cortex, whereas all the structures composing the S-C to dorsal pathway, apart from its final "destination", the posterior parietal cortex, are sub-cortical.

(2) Unlike the ventral pathway all of whose component structures apart from its "destination", the infero-temporal cortex, are concerned only with the processing of *visual* information, all the structures composing the S-C to dorsal pathway without exception process information from all sensory modalities.

---

[1] "Dorsal" here only in the sense that, like the dorsal stream properly so-called, it "terminates" in the posterior parietal cortex.

1.7   Recent Work on Attention

Recent work on the phenomenon of selective attention within conceptual analytic philosophy, experimental cognitive psychology and neuroscience has shown that Broadbent's conclusions are in need of considerable modification and elaboration. A recent development which supports the notion that one of the functions of selective attention is to protect consciousness, considered as a limited capacity channel, from overload by alerting it only to those inputs that are problematic for one reason or another comes from conceptual analysis as practised by philosophers in the ordinary language tradition. At a one-day conference on 'Attention and Consciousness' held in the Department of Philosophy, University College, London, on 26th May 1995 the Oxford philosopher Paul Snowdon presented an analysis of the concept of attention in ordinary language in which he drew a distinction between what he calls "Attention-N" ("N" for 'noticing') represented by passive voice expressions such as *Her attention was caught by an unusual ____*   and what he calls "Attention-A" ("A" for 'active') represented by active voice expressions such as *She paid close attention to the colour and shape of the object* or *to what she was doing*. It looks as though Attention-N is an unconscious involuntary mechanism for ensuring that problematic inputs and only such inputs are processed by consciousness, while attention-A is mechanism for ensuring that the focus of consciousness is maintained on a problematic input until it has been adequately categorized. There is evidence to suggest that the involuntary catching of attention by problematic inputs is a midbrain function mediated primarily by the superior colliculus and pulvinar;[2] while the conscious voluntary active holding of the focus onto an input until adequate categorization is achieved is mediated by the posterior parietal cortex.[3]

---

[2] The evidence described in footnotes 9 and 10 (Section 4.2 below) suggests that the function of the superior colliculus is to control the orientation of the relevant sense organs towards the location of a problematic input in environmental space, and that the function of the pulvinar is to control the access of such inputs into consciousness.

[3] See footnote 8 (Section 4.1 below).

From a neuroscientific perspective, Michael Posner (Posner & Petersen, 1990; Posner & Dehaene, 1994) has adduced evidence that in addition to the function which it shares with the intra-cortical dorsal pathway of mediating the visual control of voluntary movement, the S-C to dorsal pathway "(posterior parietal cortex, pulvinar and superior colliculus)" also has an important role in the control of selective attention, constituting, as it does, what he calls (Posner & Dehaene, *op. cit.,* p. 76) "the posterior attention system".

The posterior system would seem to have two functions, (a) that of bringing problematic inputs into the focus of attention in the first place (Snowdon's "Attention-N"), and (b) that of maintaining such inputs in the focus of attention until an adequate categorization is achieved (Snowdon's "Attention-A"). Of these two functions the first involves a mechanism which is necessarily *unconscious* and *involuntary* in the sense that the individual cannot decide or, in the true sense of that word, be instructed to notice things. We often say *Notice this* or *Notice that*; but such instructions only work if they are accompanied either by pointing at or otherwise highlighting the feature in question or by a verbal description of what is to be noticed, thereby creating an expectation of what is to be noticed. In either case the effect of the instruction is to *facilitate* rather than directly *induce* the noticing which remains essentially involuntary. By contrast the function of maintaining the focus of attention on a problematic input, once it has been noticed, until an adequate categorization has been achieved is an activity which is subject to conscious and voluntary control. If, as the evidence seems to suggest, tasks involving conscious and voluntary control are mediated by the cerebral cortex, while those that are unconscious and involuntary are mediated by the mid- and hindbrains, it looks as though Snowdon's Attention-A is mediated by the posterior parietal cortex and his Attention-N by the superior colliculus and pulvinar, with superior colliculus controlling the peripheral aspects (the orientation of the receptor organs) and the pulvinar the ingate into consciousness.

The other attentional system which Posner distinguishes, "the anterior attention system (anterior cingulate and basal ganglia)" would seem, in light of the evidence

adduced by Pashler discussed below, to have the function of initiating and maintaining concentration on the processes of response-selection and response-execution. This type of selective attention, like that which maintains focus of attention on a problematic input until adequate categorization of it is achieved, is under conscious and voluntary control and is to that extent also part of Snowdon's Attention-A.

Finally there is the recognition for which I am personally indebted to Harold Pashler (1991; 1997) that Broadbent's threefold system of *limited capacity channel* with a *"bottleneck"* or *"filter"* which protects the LCC from overloading by restricting input access to it, and a *buffer* which holds prospective inputs until the bottleneck clears has more than one embodiment in the brain. Pashler has shown in his experimental studies of dual-task that the response-selection system is also a limited capacity channel protected by a filter or bottleneck restricting access into it from what Broadbent calls the "category states" generated by *his* limited capacity channel, the input categorization system. A similar bottleneck is to be expected restricting access into the response-execution system. Similar bottlenecks may also exist in the human cerebral cortex to control access to the name-concept selection system (Wernicke's area) and the sentence articulation system (Broca's area). As we have seen, controlling access into the response-selection and response-execution systems and maintaining the focus of attention on these tasks until an appropriate response has been selected and its execution is complete would seem to be functions performed by Posner's anterior attentional system.

2.   The Complementary Functions of Consciousness and the Unconscious Zombie-within

The picture that emerges from the various strands of evidence described above is of two parallel, but complementary and continuously interacting input-to-output transformation systems in the brain which I shall refer to respectively as "consciousness" and the unconscious "automatic pilot" or "zombie-within".

2.1   Consciousness as an Input-output Transformation System

On this hypothesis*, consciousness* is a "*limited capacity channel*" (LCC) or rather a sequence of three such channels which, together with what I call "the *emotion-servo*", have four sequentially ordered functions, (a) the function of *categorizing* on the basis of what Broadbent (1971) calls the "evidence" and which I equate with conscious experience any input that is identified by the zombie-within as *problematic*, in that it is either *unexpected* or *motivationally significant*, i.e., significant relative to the individual's current or perennial motivational concerns (LCC 1), (b) the function of *reacting emotionally* to inputs which have been identified as problematic, both before ("physical" pleasure/pain) and after they have been categorized ("mental" pleasure/pain), thereby ensuring that the subsequent processes of response-selection and response-execution are brought into an adaptive relation to the individual's current and perennial motivational concerns (the Emotion-Servo), (c) the function of *selecting a response* appropriate both to the presence of a thing of that kind and to the individual's motivational concerns with respect to it (LCC 2), and (d) the function of initiating and monitoring the *execution* of the response selected (LCC 3).

The evidence suggests that although much of what goes on is unconscious (in the sense that the details are not available to be described or reported by the human subject), the whole of the cerebral cortex in mammals is devoted to the implementation of consciousness *in this functional sense*. In general it would seem that what the human subject reports are the *outcomes* of the processes of selective attention, categorization, emotional reaction, response-selection and response-execution, rather than the processes themselves. The exceptions to this rule are the process of *sensory conscious experience* which can, to some extent, be described independently of the way it is finally (as opposed to tentatively) categorized, and the thoughts (images and sub-vocal speech) which contribute to, but do not exhaust, the process of response-selection, just as conscious experience of the feedback from the output as it develops contributes to, but does not exhaust, the process of response-execution.

2.2   The Unconscious "Automatic Pilot" or "Zombie-within"

The functions of the *unconscious automatic pilot* or *zombie-within* are (a) that of continuously *scanning* the total current input so as to alert consciousness to any input it identifies as *problematic*, (b) that of *protecting consciousness from overload* either by ignoring those non-problematic inputs which require no response or by responding appropriately, but automatically and *without categorization*, to those for which there already exists a well practised skill or other "instinctive" response pattern.

Like its namesake in popular mythology, the zombie-within is a creature of habit, routine and unquestioning conformity to the instructions it receives from consciousness. Anything out of the ordinary is immediately passed on for processing by consciousness. The one respect in which it differs from the traditional picture of its mythical namesake is in its capacity to learn from experience, limited though that is to the progressive shaping of minor variations in behaviour by their immediate consequences.

The evidence suggests that, with one possible exception, all the functions of the zombie-within are mediated by structures in the midbrain and brainstem. The one possible exception is in the case of the visual functions of reaching for objects and avoiding obstacles which are retained when visual conscious experience is abolished by lesions of the striate cortex (V1), thus occluding the ventral visual pathway and yielding the phenomenon of "blindsight". The dorsal visual pathway which is known to mediate these functions, though composed in the main of midbrain structures such as the superior colliculus and pulvinar, also includes the posterior parietal cortex. It may be, however, that although the functions of reaching for objects and avoiding obstacles do not require *visual* conscious experience and are, to that extent, to be regarded on the present hypothesis as functions of the zombie-within, they *do* require the integration of visual information supplied by the zombie along the S-C to dorsal pathway with conscious experience of the *somaesthetic feedback* from the movements involved as they develop,

and that this integration is the contribution to these functions made by the posterior parietal cortex.

2.3   Interactions between the Two Systems

Although, as the blindsight phenomenon shows, there are other forms interaction between the two systems, the three most important interactions between consciousness and the zombie-within are (a) the action of the zombie in alerting consciousness in general and conscious experience in particular to problematic inputs, (b) the gradual transfer to the zombie-within of stimulus-stimulus expectations and stimulus-response connections formed within consciousness as they become habitual (for PET scan evidence of this process, see Raichle, *et al.* 1994), and (c) the integration of the two systems in a well-developed motor skill where, as the syntactic organization of movement becomes increasingly automatised, i.e., gets taken over by the zombie, the easier it becomes for the mechanisms of selective attention to ensure that consciousness is focussed on those aspects of the task that are crucial from the point of view of effective strategic decision-making and the timely initiation of the selected response.

3.   Modules within Consciousness and the Zombie-within

The multiple functions identified within both consciousness and the zombie-within imply a multiplicity of modules within both systems. Figure 2 shows the arrangement of these modules as I currently construe it. As you will see, the diagram shows the output from the sense organs splitting into two streams, *consciousness* on the left, the *zombie* on the right.
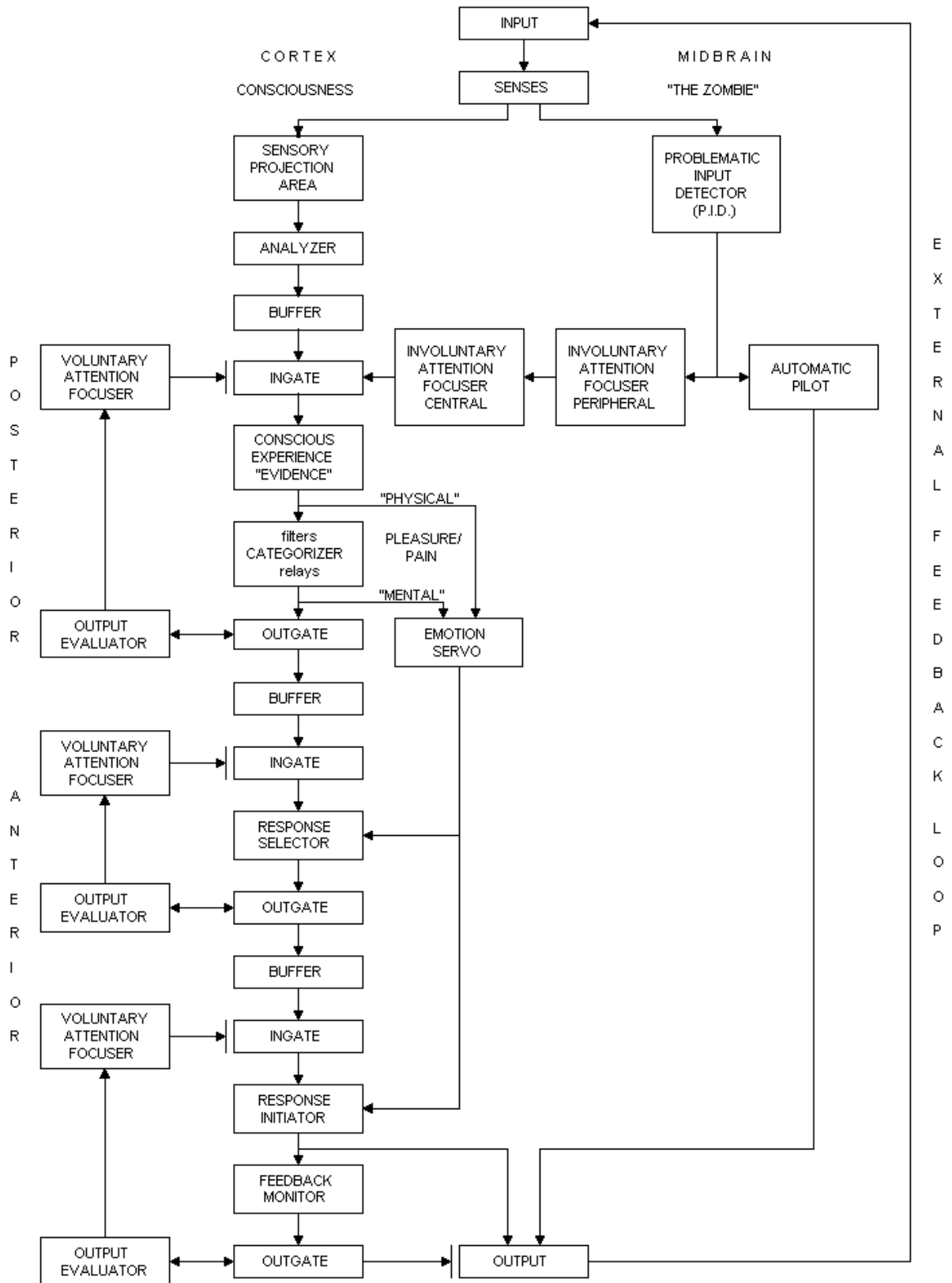
Figure 2: Consciousness and the Zombie-within - Suggested Layout of Modules

3.1   Modules within the Zombie

The zombie is shown as consisting of four functionally defined modules, (a) the
problematic input detector (PID) which separates inputs into problematic and
non-problematic on the basis of relatively coarse criteria of unexpected/expected and
motivationally significant/insignificant, and transmits the former *via* (b) the *involuntary
attention-focuser (peripheral)* which mobilises and directs movements of the head eyes
and body so as to bring the *source* of the problematic input within the range of all
relevant sense organs, (c) the *involuntary attention-focuser (central)* which attracts the
focus of *conscious experience* to that part of the sensorium where the problematic input is
located, while either ignoring or routing non-problematic inputs to output *via* (d) the
automatic pilot.

3.2   Modules within Consciousness

Consciousness, as shown here, consists of three sequentially ordered limited capacity
channels (LCCs), concerned respectively with *input-categorization (perception)*,
response-selection and *response-execution*. The concept of the *limited capacity channel*
comes from Broadbent (1958) whose 1971 book *Decision and Stress* restricts its
application to the process of input categorization. The evidence that there is more than
one such channel in the brain comes from Pashler (1991; 1997). In order to protect it
from overloading, Broadbent's model requires that each LCC be provided with a system
of subordinate modules, including (a) an "ingate" which *controls access* to the LCC, (b) a
buffer or short term memory store in which inputs waiting to obtain access to the LCC
are held until the ingate clears, (c) an *attention-focuser* which maintains the focus of
attention on the task in hand until it is satisfactorily completed, and (d) an *output
evaluator* which checks the candidate outputs of the LCC and allows the
attention-focuser to open the ingate to a new input once the previous information
processing task has been satisfactorily completed, while at the same time opening (e) an

outgate which allows the approved output schema to proceed either directly to the initiation of a response or into the buffer of the next LCC in line.

In the case of the input-categorizing LCC the limited capacity channel is shown as divided into two separate modules, *conscious experience* which, to use Broadbent's (1971) term, provides the "evidence" on which categorization is based and without which the blindsighted subject's judgments become "pure guesswork" and the *categorization* response itself. A similar division is shown within the response-executor LCC between *response initiation* and *feedback control*.

Lying outside this system of three limited capacity channels are two modules, the *ANALYZER* and the *emotion servo*. The existence of the ANALYZER is suggested by the known functions of the visual cortical areas V1-V5. Recordings from individual cells in these areas have revealed what are known as "feature detectors", cells which fire in response to the presence within the current retinal input of various features and patterns which are relevant for the identification of objects and situations in the organism's visual environment. The features which are detected in this way become more and more abstract and involve responding to activity spread over a wider and wider area of the visual field the further they are from V1. Recent research by Steve Luck and Nancy Beach (1996) confirming the Feature Integration Theory (FIT) of Ann Treisman and her colleagues (Treisman, 1988; Treisman & Gelade, 1980; Treisman & Gormican 1988) suggests that the effect of the module here identified as *conscious experience* is to "bind" the "information" provided by these individual feature detectors into a single *Gestalt*, thereby generating the "evidence" on which categorization or interpretation of the input is subsequently based.

Whereas the ANALYZER is construed as a module which precedes and prepares the ground for the process of input-categorization, the *emotion servo* is brought into play by the action of the input-categorization LCC. This activation occurs both before categorization in response to "raw" uninterpreted conscious experience, as in the case of "physical" pleasure and pain, and after categorization, as in the case of "mental" pleasure

and pain. Its function is to provide motivation both for response-selection and for response-execution.

3.3   The Problematic Input Detector (PID)

Fundamental to the system whereby behaviour is controlled by the brain as set out on Figure 2 is the *problematic input detector* (PID). The PID is that part of the zombie-within which determines whether a current input is or is not problematic, alerting consciousness to it if it is, either ignoring it or allowing it to proceed automatically to the selection and execution of a response if it is not.

In order to understand how the PID works two questions need to be answered:

(1) What sorts of input qualify as *problematic*?

(2) How, given that input-categorization does not occur before consciousness has been brought into play, are such inputs *detected* by the zombie?

3.3.1   Varieties of problematic input

An input is problematic if (a) it is *unexpected*, or, (b) if expected, it is *motivationally significant*, i.e., significant relative to the individual's current or perennial motivational concerns. An input is *motivationally significant* if it is (c) something the individual is *searching* or *on the look out* for, (d) a stimulus which is *intrinsically pleasant* or *unpleasant* (i.e., one whose pleasantness or unpleasantness does not depend on how it is categorized or interpreted) e.g., the pleasantness of the sensation of being stroked or the unpleasantness of the sensations of pain and nausea, and (e) a stimulus which has been associated with a motivationally significant past event, such as Plato's lyre that reminds the lover of his beloved.

3.3.2   How different varieties of problematic input are detected

The unexpected. Given the principle of association by contiguity, it can be predicted that an organism will build up a vast number of *stimulus-stimulus expectations* based on observed regularities in the way an input of one type is invariably succeeded by an input of another type. Given a background of such expectations, a second input which differs

21

from that expected on the basis of past experience, given the first input, is going to stand out like a sore thumb.

Objects of search. We may suppose that the PID is sensitized to respond to the objects of search by the *active disposition which initiates the search* and guides it until either the object is found or the search is abandoned.

Intrinsically pleasant and unpleasant stimuli. Although any such effect, if it exists, is concealed by the fact that intrinsically pleasant and unpleasant stimuli immediately attract conscious attention, there is some evidence to suggest that emotional reactions elicited by conscious experience of the stimuli before categorization may also be elicited by the alternative input system which serves the zombie-within. A study by Zihl, Tretter & Singer (1980) cited by Weiskrantz (1986, pp. 125-6) reports "an autonomic electrodermal response . . . to [moving] visual stimuli in the absence of 'seeing'" in a case of blindsight; while Tranel & Damasio (1985), also cited by Weiskrantz (1986, pp. 137-8), "showed that two prosopagnosic patients who failed to recognize familiar faces verbally nevertheless displayed a clear and strong skin conductance response to photographs of familiar faces relative to control faces." This evidence raises the possibility that emotional reactions to intrinsically pleasant and unpleasant stimuli may not be, as it subjectively appears, a response to conscious experience of the stimuli in question. It may rather be *a response to a preconscious input reaching the zombie-within* which, in turn, attracts the focus of conscious attention to the stimulus and emotional reaction as a unitary *Gestalt*. It suggests that these emotional reactions are triggered by a direct connection between the problematic input detector (PID) and the emotion servo. But since its only function would appear to be to ensure a more rapid mobilisation of motivational resources than if it were routed through consciousness, this connection has been omitted from Figure 2 in favour of the phenomenologically more significant contrast between emotional reactions which do and do not depend on the way the experience is conceptualised.

Associations with motivationally significant past events. It is suggested that the function of the *dream imagery* characteristic of REM sleep is to "stamp in" associations between events that have occurred during previous waking period and motivationally significant past events at the expense of motivationally neutral associations formed during the same period. In REM sleep conscious experience is, as it were, being allowed to "freewheel" when decoupled from sensory input, thus leaving it free to generate images, particularly visual ones, whose form is determined only by the new *associative links* formed as a result of the attention-focusing and categorizing of problematic inputs which has taken place during the preceding period of waking, and by the individual's current *emotional preoccupations*.
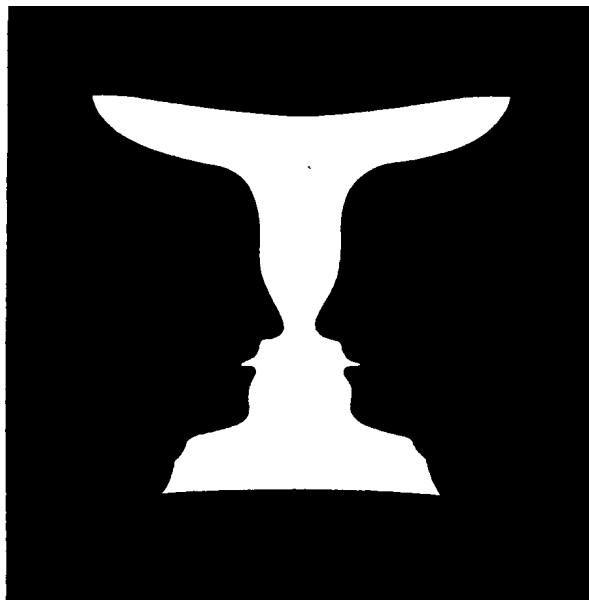

Figure 3: Figure-ground reversal *(After Rubin 1915)*

3.4   Conscious/Phenomenal Experience

Conscious/phenomenal experience, on this view, is the first stage in the process whereby problematic inputs are processed by consciousness. Its function is to provide the "evidence" on which the categorization of problematic inputs is based, by modifying the

figure-ground relations (Figure 3) within the central representation of the input until an adequate categorization is selected.

3.4.1   *Intrinsic figure-ground differentiation* versus *imposed figure-ground organization*

Two kinds of figure-ground relation need to be distinguished. On the one hand there is the *intrinsic figure-ground differentiation* whereby one part of the current input (the figure) stands out from, is more salient and thus catches the attention more readily than, the rest (the ground), simply by virtue of the sharpness and magnitude of the contrast between the two. The other is the *figure-ground organization* properly so-called which is *imposed* on the input from the sensory projection areas of the cortex by the process whereby conscious/phenomenal experience is generated.

The two forms of figure-ground differentiation are connected in that the sharper the intrinsic figure-ground contrast the more *strongly structured* and, therefore, less malleable is the input which is available for moulding by conscious experience. In other words, the larger and simpler the intrinsic contrast between figure and ground (salience) the less room there is for conscious experience to impose a different pattern of figure-ground organization.

3.4.2   *Conscious/phenomenal experience as imposed figure-ground organization*

We have seen that on the present hypothesis it is the output or "*evidence*," as Broadbent (1971) calls it, which is generated by the process of *selective attention* which constitutes the *conscious/phenomenal experience* to which the introspecting subject is responding when she describes what it is like either to receive sensory input from a particular input source in the environment or to imagine being exposed to it. The "luminosity" or "phosphorescence" which is the most striking feature of conscious/phenomenal experience from the standpoint of the introspective observer enables a linguistically competent human to give a running commentary both on the sequence of events in her stimulus environment and her conscious/phenomenal experience of those events at the time and to provide a first hand report on some of them subsequently. It also enables the

organism to check its categorization of a problematic input against the "evidence" on which the categorization is based. Without this check the blindsighted subject loses all confidence in the sometimes remarkably accurate discriminations he is able to make relying solely on the sub-conscious system. Such discriminations, he insists, are "pure guesswork." This lack of confidence may also explain the inability of blindsighted patients to initiate voluntary action based upon their blind field discriminations to which Marcel (1988) has drawn attention.

3.4.3   Mental imagery

Mental imagery is a form of conscious experience which occurs in a variety of different contexts, in dreams, in daydreaming, in the recollection of past events and in planning the future. Phenomenologically it resembles and can sometimes be confused with the kind of sensory conscious experience which provides the "evidence" for the categorization of an input. Moreover, there is evidence from a study (Kosslyn *et al.* 1995) using positron emission tomography (PET) that when a subject forms a visual image of a picture he or she has just been shown, the same pattern of activity develops in all the principal visual areas of the cortex, including V1, as occurs when the subject is looking at the actual picture. This re-instatement of the cortical activity involved in sense perception in the absence of the input otherwise required for its occurrence is undoubtedly the substance behind Hume's much criticised claim

> that all our simple ideas in their first appearance are deriv'd from simple impressions, which are correspondent to them, and which they exactly represent (Hume 1739/1978, p. 4).

Aside from the fact that its form is not determined by *current* sensory input, the principal difference between a mental image and a perceptual experience is its relation to the process of categorization. We have seen that in sense perception conscious experience precedes and provides the "evidence" on which categorization is subsequently based. In the case of a mental image, as Kant (1781/1787/1929, pp. 182-183) demonstrates in developing his concept of the

"schema", the construction of a mental image presupposes a prior categorization or conceptualization of what the image is to be an image of.

In order to account for this kind of control over the process of conscious experience on the present model as laid out on Figure 2, we would have to include a number of "re-entrant" (Edelman, 1987) or "recurrent" (Jordan, 1986) circuits feeding back from the CATEGORIZATION and RESPONSE-SELECTION modules.[4] Not only are such circuits well-attested anatomically, Jordan (*op. cit.*) has shown that, along with the reverberatory circuits (Hebb 1949) required to bridge the gap between the offset of the first stimulus and the onset of the second, such circuits are an essential feature of any neural network that can learn to "expect" or "anticipate" the *second* of two sequentially ordered stimuli on presentation of the *first*. As we have already seen, an extensive repertoire of such expectations is required as a background against which an unexpected input will stand out as figure and thus be referred to consciousness by the zombie-within.

It seems that the recurrent circuits required to account for the generation of mental imagery would need to feedback from the CATEGORIZATION and RESPONSE SELECTION modules to the SENSORY PROJECTION AREAS (such as V1 in the case of a visual image) to ensure that conscious experience is supplied with the necessary "raw material", to the ANALYZER to ensure that it is given the necessary structure, to the relevant INGATE to ensure access into CONSCIOUS EXPERIENCE, and to the VOLUNTARY ATTENTION FOCUSER to ensure the maintenance of the image until it has served whatever purpose it was intended to fulfil. However, in order to avoid too much complication these circuits are not shown on Figure 2. The only recurrent circuits shown are those connecting the OUTGATE of each of the three LCCs to its respective INGATE and the EXTERNAL FEEDBACK LOOP connecting the motor output to sensory input.

---

[4] I am indebted to Pim Haselager of the University of Nijmegen for drawing my attention to the need to emphasize this point.
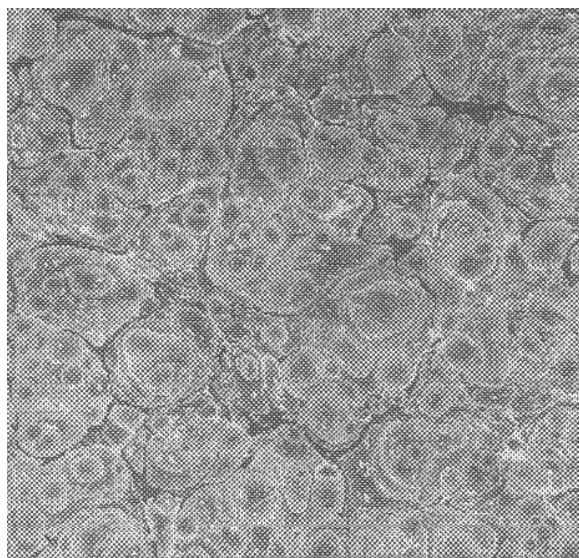
Figure 4: A weakly structured field

It is suggested that in the case of a *mental image* these recurrent circuits impose a pattern of figure-ground organization on a field that is *intrinsically weakly structured* (Figure 4) and does not, therefore, restrict the pattern of organization that can be imposed on it in the way a more salient and strongly structured input would do. This results in a pattern of figure-ground organization which in the extreme case bears no relation to any objective structure in the input source. In the case of vision, the Rorschach (1932/1942) ink blots provide a classic example of a series of such weakly-structured fields which permit and thus promote the formation of a wide variety of such images.[5]

We know from the introspective reports of human subjects that such images occur both in waking consciousness as part of the thought process whereby solutions to problems are generated and as the predominant feature of the dreams that occur during the rapid-eye-movement (REM) phase of sleep. In the latter case there is strong circumstantial evidence for the occurrence of such imagery in the sleep of those

---

[5] The "non-objects" used by Vanni, Revonsuo and Hari (1997) in their experiments are another example.

mammals in which it occurs.[6] Although there is at present no corresponding evidence for the occurrence of mental imagery as an aid to animal problem-solving, it would be surprising if an ability which is almost certainly present during sleep were not exploited for more obviously practical purposes during waking.[7]

3.5   Categorization

Categorization is the process whereby problematic inputs are classified according to the kind of object or situation of whose presence in the organism's stimulus environment the input is a reliable indicator. It is the function of categorization to ensure that the universal, kind or category under which an input is subsumed lines up with what Skinner (1969) calls the "contingencies" operating in the organism's environment. A *contingency* for Skinner is a sequence of events whereby under certain *antecedent* conditions *behaving* in a certain way will have certain predictable *consequences*. By classifying its problematic inputs in a way that enables it to anticipate the consequences of selecting one form of behaviour rather than another, the organism puts itself in a state of readiness to select a successful behavioural strategy appropriate both to the presence of an object or situation of that kind and to whatever may be the organism's current behavioural objectives as and when the occasion for action arises.

3.5.1   Two components of categorization: the input filter and the output relay

We have seen that Broadbent (1971) distinguishes two processes within what we are here calling "categorization", namely, "pigeon-holing" and "categorization" proper. On his view, "pigeon-holing" is simply a matter of slotting an input into a pre-existing category,

---

[6] The inhibition of the skeletal musculature during this phase of sleep makes sense only as a device whose function is to prevent the massive and obviously maladaptive somnambulism which would otherwise occur in response to such imagery.

[7] Tim Shallice (1988; Burgess & Shallice 1996) has drawn my attention to a hypothesis proposed by Schank (1982) which suggests that the original function of mental imagery was to allow the organism to remind itself of the past consequences of the various courses of action suggested by the current stimulus situation as possible solutions to the problem that situation presents.

whereas "categorization" in his sense involves either creating a new category altogether or modifying the boundaries of an existing category so as to fit a new instance. This way of construing the matter is mistaken in so far as it assumes that the organism cannot respond adaptively to an input without specifically classifying it as an encounter with an object or situation of a particular kind. On the present hypothesis, it is only problematic inputs that require classification in this way. Once a behaviour pattern has become habitual, direct input-to-output transformations replace responses mediated by categorization and motivational choice. Evidence for just such replacement of one pattern of brain activity by another as behaviour becomes habitual comes from the recent study by Raichle *et al.* (1994) mentioned above.

The categories which make up an individual's conceptual scheme to one of which every problematic input must be assigned in order for the process of categorization to succeed have two components. One component is the *filter* which selects those inputs which satisfy the entry criteria for the concept in question and rejects those otherwise similar inputs which do not satisfy them. The other component is what we may think of as a *relay* which pre-selects all those behavioural strategies which can be relied on to yield a predictable consequence, when emitted in the presence of an object or situation of that kind. Once this set of behavioural strategies has been pre-selected, a choice is made between them in the light of subsequent environmental conditions and the organism's motivational attitude to the expected consequences of adopting one course of action rather than another. In the case of a linguistically competent human being, an important group of behavioural strategies which are pre-selected in this way are strategies for selecting appropriate words and sentence frames for describing objects and situations of the kind in question.

3.6   The Emotion Servo

29

Skinner's concept of the "three-term contingency" (*antecedent* conditions, *behaviour* called for under those conditions and the *consequences* of so behaving) not only provides a clue to the nature of the concepts or categories the organism uses in classifying its problematic inputs, it is also the key to understanding the operation of what we are calling the "emotion servo." As we have seen, the function of this module is to modulate behaviour in such a way as to bring it into conformity with the organism's motivational objectives. As the contingency unfolds and as its conformity or lack of conformity to those objectives becomes apparent, so the organism's emotional reaction changes. The same sequence of events will evoke a different sequence of emotional reactions depending on the organism's motivational attitude to the anticipated or actual consequences of its behaviour. If the consequence is attractive, anticipating its appearance produces excitement, its actual occurrence, pleasure, its failure to appear when expected, first anger then misery or depression. If the consequence is repulsive or, as Skinner would say, "aversive", anticipating its appearance produces fear or anxiety, its actual occurrence, first anger then misery or depression, its failure to appear when expected, relief.

Each different variety of emotional response is characterized (a) by the type of situation that evokes it, i.e., whether it is *prospective* as in excitement and fear, *retrospective* as in anger, relief and depression, or *focussed on the moment* as in pleasure and disgust, (b) by its position on the *pleasant-unpleasant* dimension, pleasant in the case of excitement, pleasure and relief, unpleasant in fear, disgust and depression, mixed in the case of anger and apathy, (c) by its position on the *arousal* dimension, high in the case of excitement, anger and fear, moderate in pleasure and disgust, low in relief, apathy and depression, and (d) by a characteristic *impulse*, the impulse to sigh in relief, to smile in pleasure, to jump for joy in excitement, to attack in anger, to freeze or run away in fear, to vomit in disgust, to do nothing or punish oneself in depression.

4.   Locating these Modules within the Brain

Evidence from a variety of sources, neurology, electrophysiology, and, most recently, the newly discovered brain-imaging techniques, makes it possible to propose a tentative identification of the modules shown on Figure 2 with specific anatomically defined structures within the brain.

4.1   The Ventral and S-C to Dorsal Pathways

For the most part, I shall confine my remarks on this score to the "upstream" portion of Figure 2, that which precedes the transition within consciousness from CONSCIOUS EXPERIENCE to CATEGORIZATION. For this is the area to which the neurological evidence described in Section 1.6 above relates. That evidence shows that in order to account for the visual functions which survive lesions of primary visual cortex (V1) and the consequent loss of visual conscious experience, the phenomenon known as "blindsight", we must suppose that the residual visual functions are mediated by what we are calling the "sub-cortical (S-C) to dorsal pathway" which proceeds by way of the *superior colliculus* and *pulvinar* to the *posterior parietal cortex*. This S-C to dorsal pathway bifurcates at the retina from the *ventral pathway* consisting of the *lateral geniculate nucleus*, the *primary visual* or *striate cortex (V1)*, and the *extra-striate visual areas (V2-V5)* to the *infero-temporal cortex*. These two pathways, together with the intra-cortical dorsal and ventral streams distinguished by Ungerleider & Mishkin (1982) which bifurcate "downstream" of V1, are shown on Figure 1 (taken from Milner & Goodale, 1995, p. 68).

Whereas all the structures composing the ventral pathway, with the doubtful exception of the infero-temporal cortex,[8] are exclusively visual in function. Those composing the S-C to dorsal pathway (superior colliculus, pulvinar and posterior parietal

---

[8]   There is evidence (Gibson and Maunsell 1997) of cells in IT which respond to cross-modal associations between visual and auditory stimuli in a delayed match-to-sample memory task.

cortex) subserve all sensory modalities. This is consistent with Posner's (Posner & Petersen, 1990; Posner & Dehaene, 1994) hypothesis that the function of these structures, in their capacity as the "posterior attention system", is to control the focus of sensory attention as it switches from one modality to another or concentrates different modalities on the same area of environmental space. This concatenation of evidence allows us, in the case of the visual modality, to identify the bifurcation between consciousness and the zombie-within as shown on Figure 2 with the bifurcation at the retina between a *ventral* pathway consisting of the *lateral geniculate nucleus*, the *primary visual* or *striate cortex* (V1), the *extra-striate visual areas* (V2-V5) and the *infero-temporal cortex* corresponding to the upstream portion of what I am calling "consciousness", and the *S-C to dorsal pathway* consisting of the *superior colliculus* and *pulvinar*, but almost certainly excluding its "destination", the *posterior parietal cortex*, corresponding to the zombie.
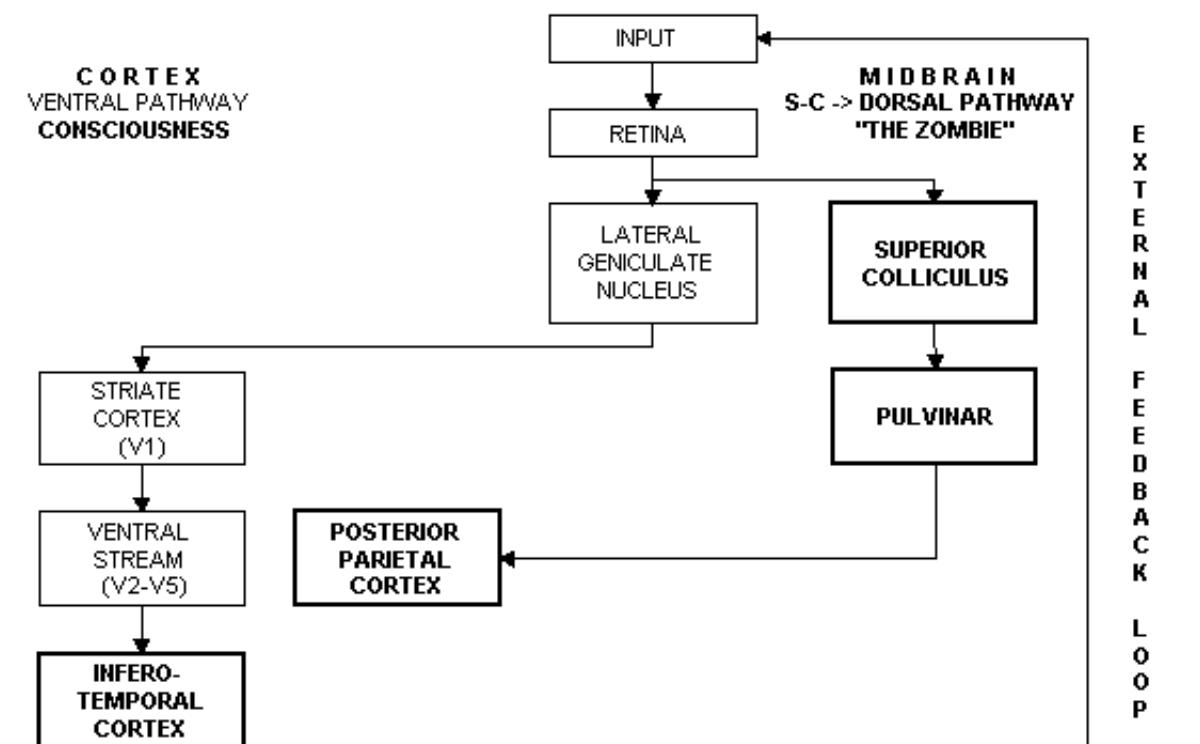

Figure 5: The Dorsal and Ventral Pathways - Multimodal Modules **in Bold**

These relationships are shown Figure 5. It is a re-drawing of Milner & Goodale's diagram (Figure 1 above) which, for the sake of clarity, omits the dorsal stream properly so-called (connecting V1 to the posterior parietal cortex) and is arranged in the same format as the upper part of Figure 2 with the ventral pathway on the left and the S-C to dorsal pathway on the right.

Given the identification of the ventral pathway as the route whereby information is passed from the retina into consciousness and the S-C to dorsal pathway as the route into and through the zombie-within, what are we to say about the one exception, the posterior parietal cortex, which is the terminus of both the S-C to dorsal pathway and the intra-cortical dorsal stream? There would seem to be a connection here between the posterior parietal and the visual functions of reaching for objects (Weiskrantz 1986; Cowey & Stoerig 1995) and avoiding obstacles (Humphrey 1974)[9] which are retained when visual conscious experience is abolished by lesions of the striate cortex (V1), thus occluding the ventral visual pathway and yielding the phenomenon of 'blindsight.' Although it is clear from the performance of human subjects with lesions of the striate cortex ("blindsight") that the functions of reaching for objects and avoiding obstacles do not require *visual* conscious experience of the relevant stimuli and are, to that extent, to be regarded on the present hypothesis as functions mediated by the zombie-within, the fact that the subject in such cases is induced to reach for an object he does not "see" by an appropriate instruction to *guess* where it was shows that conscious experience of *some* kind is involved in the production of such behaviour. A plausible hypothesis would be that reaching and obstacle-avoiding behaviour, though it does not require *visual* conscious experience, *does* require the integration of visual information supplied by the

---

[9] There is at present no evidence that human blindsighted subjects can learn to negotiate obstacles in the absence of visual stimulation routed *via* V1, as Humphrey's monkey subject Helen learned to do. But all the human blindsighted subjects studied thus far have been able to rely on the unaffected portion of the visual field to do this. If, like Helen, they had been compelled to rely on visual information arriving from the retina *via* the sub-cortical route, my guess is that they too would have learned to avoid obstacles in the absence of V1.

zombie along the S-C to dorsal pathway with conscious experience of the *somaesthetic feedback* from the movements involved as they develop, and that this integration is the contribution to these functions made by the posterior parietal cortex. However, some recent evidence (Rossetti, Rode and Boisson 1995; Rossetti 1997) on the somaesthetic counterpart of blindsight which the authors refer to as "numb-sense" shows that a patient (J.A.) with this condition can use his "normal" left hand to point accurately at the location of stimuli applied to the "numb" right hand which he cannot consciously feel. This shows that, provided the blindsight or numb-sense subject can be induced to guess at the location of the target object by pointing at it, successful voluntary movement does not require conscious experience, whether visual or somaesthetic, of the target towards which the movement is directed. What the evidence does not show is that such voluntary movement is possible without conscious experience of the feedback from the movement itself, whether visual, somaesthetic or both.

Although, as this evidence clearly demonstrates, the S-C to dorsal pathway has a *secondary* function in the visual control of voluntary movement, the fact that all its structures process information from *all* sensory modalities, when combined with the brain-imaging and neurophysiological data reviewed by Posner and Dehaene (1994) and the evidence of disorders of attention, such as unilateral neglect, resulting from lesions of these structures, suggest that its *primary* function is to integrate the *involuntary* alerting of conscious attention to problematic inputs from all sensory modalities mediated by the two midbrain structures, the superior colliculus and pulvinar, with the *voluntary* maintenance of the focus of attention on such inputs until an adequate categorization of them is achieved mediated by the posterior parietal cortex. On this hypothesis the posterior parietal is construed as having two functions, (a) a *general* function which is to *maintain* the focus of conscious attention within and between the different sensory modalities (acting on structures such as those in the ventral stream in the case of vision) on inputs to which the focus has been initially attracted by the "zombie" (in the shape of

the superior colliculus and pulvinar) until such time as an adequate categorization of those inputs has been achieved,[10] and (b) a *specific* function which is to control voluntary movement by integrating, through the same mechanism of conscious attention focusing, the visual and somaesthetic feedback from such movements as they develop.

4.2   *Provisional Anatomical Conclusions*

Assuming that this analysis is approximately correct, we are in a position to make some tentative identifications of the modules shown on Figure 2 with some of the actual structures that have been identified anatomically within the brain as laid out on Figures 1 and 5. These tentative identifications are set out on Figure 6 which is a re-working of Figure 2 with the names of the neural structures substituted for the functional descriptions of the modules with which they have been provisionally identified in the preceding discussion in the special case of vision. Thus, in place of the SENSES we have, in the case of the visual modality, the RETINA. In place of the SENSORY PROJECTION AREA we have, in the case of the visual modality, the STRIATE CORTEX. In place of the ANALYZER we have, in the case of the visual modality, V2-V5. In place of the PROBLEMATIC INPUT DETECTOR (P.I.D.) we have, for all modalities, the MIDBRAIN RETICULAR FORMATION.[11] In place of the AUTOMATIC PILOT we have the CEREBELLUM. In place of the INVOLUNTARY ATTENTION FOCUSER - PERIPHERAL we have the SUPERIOR COLLICULUS.[12]

---

[10] Evidence confirming the suggestion that the posterior parietal cortex performs this function is provided by a recent study by Vanni, Revonsuo and Hari (1997) which shows that the magnetic alpha rhythm generated in the parieto-occipital sulcus (POS) is suppressed by object targets to a much greater degree than non-objects, a finding which is readily interpreted as showing the persistence of activity in this area when a stimulus is *not* readily categorized, as compared with the rapid shut down when it is.

[11] As shown by Moruzzi and Magoun (1949).

[12] For the role of the superior colliculus in coordinating, at a pre-conscious level, the position and sensitivity of the different sense organs in relation to particular locations in environmental space see Stein and Meredith (1995). For the role of the superior colliculus in controlling the reflexive orienting response whose absence on the affected side of the body is characteristic of the phenomenon of unilateral neglect see Rafal and Robertson (1995).
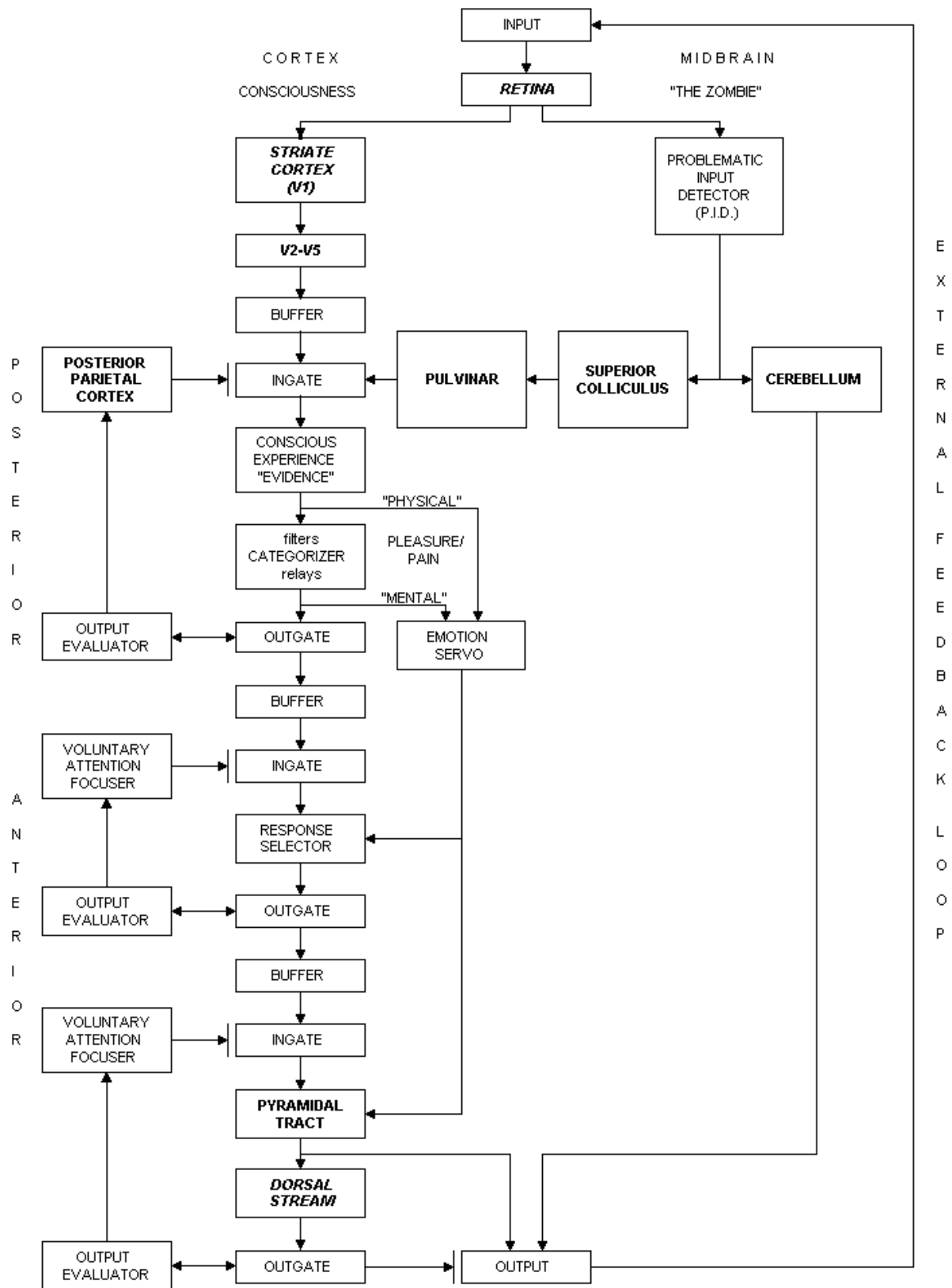
Figure 6: Consciousness and the Zombie-within — Identifications **in Bold**

In place of the INVOLUNTARY ATTENTION FOCUSER - CENTRAL we have the PULVINAR.[13] In place of the VOLUNTARY ATTENTION FOCUSER we have the POSTERIOR PARIETAL CORTEX. Finally, in place of CONSCIOUS EXPERIENCE "EVIDENCE" we have, at least in the case of vision, the INFERO-TEMPORAL CORTEX.[14] You will notice that Figure 6 omits the connection between the PULVINAR and the POSTERIOR PARIETAL CORTEX shown on Figures 1 and 5 and which is needed to explain the visual control of reaching for objects and obstacle avoidance when the relevant parts of V1 have been destroyed ("blindsight"). This has been done in order not to obscure the functionally much more important connection between the PULVINAR and the INGATE controlling access to the INFERO-TEMPORAL CORTEX *alias* CONSCIOUS EXPERIENCE. Further "downstream" the only identification to have emerged at all clearly from the preceding discussion is that between the FEEDBACK MONITOR and the DORSAL STREAM.[15] However, two other

[13] A study by Vanni, Revonsuo and Hari (1997) provides evidence suggesting that the pulvinar is involved in modulating activity in the ventral stream (V2-V5), where object recognition or, as I would think, the preparation of the "evidence" for it occurs and that the effect of such modulation is to "select the next target for ventral processing". If this may be interpreted to mean that the pulvinar controls which parts of the total visual input are currently subject to "ventral processing" and hence in the focus of conscious attention, it supports the suggestion that the function of this structure is to regulate the involuntary attraction of the focus of conscious attention to problematic inputs by processes which are themselves necessarily pre-conscious, i.e., part of what I am calling the "zombie-within".

[14] Evidence confirming this identification is provided by Sheinberg and Logothetis (1997). They showed that in a binocular rivalry experiment 90% of the cells in a monkey's IT respond to whichever of the two rival stimuli is currently in the focus of attention; whereas in V1, V4 and V5 only 20-25% of cells do so. But, apart from that cited in footnote 9 above, there is no evidence of the involvement of other sensory modalities beside the visual in IT. It is, therefore, unlikely that the "unity of consciousness" across sensory modalities which is demanded, as much by functional considerations as by phenomenology, is secured by concentration in a single anatomical location. For this a better candidate is the synchronous firing of cells in different parts of the cortex.

[15] Needless to say, this identification rides roughshod over a number of complexities. The dorsal stream is a body of linked cortical modules connecting the [VISUAL] PROJECTION AREA (V1) with the POSTERIOR PARIETAL CORTEX. It is of similar complexity to that of the ventral stream which appears on Figures 2 and 6 as the [VISUAL] ANALYZER (V2-V5) connecting V1 to the INFERO-TEMPORAL CORTEX. The posterior parietal cortex appears to have two functions (a) the function emphasised on Figures 2 and 6 where it is identified as the VOLUNTARY ATTENTION FOCUSER and the function which it presumably shares with the dorsal stream as a whole and which is the basis for this identification of providing the integration of visual and somaesthetic information required for the FEEDBACK MONITORING of voluntary movement. A further complexity is added by recent evidence (Gallese 1998; Rizzolatti & Arbib 1998) demonstrating the role of the pre-motor cortex (the counterpart of the posterior

identifications have been included on the basis of what has been known for a long time, that between RESPONSE INITIATION and the PYRAMIDAL TRACT and between the EMOTION-SERVO and the HYPOTHALAMUS. Likewise the recent work on "numb-sense" mentioned above will doubtless soon make possible the identification of the somaesthetic counterparts of the purely visual structures shown on Figure 6. It may be that similar identifications can already be suggested for other sensory modalities. If not, future research will doubtless allow us to fill these gaps too.

But there, for the present, I shall let the matter rest. I hope I have said enough to persuade you that we are now in a position to answer the question which has remained unanswered since my (Place, 1956) paper 'Is consciousness a brain process?', namely, 'If consciousness *is* a brain process, which of the various processes in the brain that we now identify neuroanatomically is it?' It turns out in the light of what has been said above that that question is too simplistic. But complicated though it is, I hope I have persuaded you that the rudiments of an answer are within our grasp.

---

parietal on the anterior side of the fissure of Rolando), not only in the visual feedback control of voluntary movement, but in the visual interpretation of the movement of others. As in the case of the role of mental imagery in response selection, to do justice to these complexities on a diagram such as that on Figures 2 and 6 would seriously detract from the sense of a flow of information within consciousness from input to output.

Cowey, Colin Blakemore, and the editors of this volume, Yves Rossetti and Antti Revonsuo.

REFERENCES

Broadbent, D. E. (1958) *Perception and communication*. Oxford: Pergamon.

Broadbent, D. E. (1971) *Decision and stress*. London: Academic Press.

Burgess, P. W. & Shallice, T. (1996) Confabulation and the control of recollection. *Memory*, 4, 359-411.

Cowey, A. & Stoerig, P. (1995) Blindsight in monkeys. *Nature,* 373, 6511, 247-9.

Cowey, A. & Stoerig, P. (1997) Visual detection in monkeys with blindsight. *Neuropsychologia*, 35,   929-939.

Edelman, G. M. (1987) *Neural darwinism: The theory of neuronal group selection*. New York: Basic Books.

Farah, M. J. (1990) *Visual agnosia: Disorders of object recognition and what they tell us about normal vision.* Cambridge, MA: MIT Press.

Gallese, V. (1998) From neurons to meaning: Mirror neurons and social understanding. Paper presented to the Second Annual Conference of the Association for the Scientific Study of Consciousness, Bremen, Germany, June 21st 1998.

Gibson, J. R. & Maunsell, H. R. (1997) Sensory modality specificity of neural activity related to memory in visual cortex. *Journal of Neurophysiology*, 78 (3), 1263-1275.

Hebb, D. O. *The organization of behavior*. New York: John Wiley.

Hume, D. (1739/1978) *A treatise on human nature*. L. A. Selby-Bigge (Ed.), 2nd Edition, P. H. Nidditch (Ed.), Oxford: Clarendon Press.

Humphrey, N. K. (1974) Vision in a monkey without striate cortex: a case study. *Perception*, 3, 241-255.

Jordan, M. I. (1986) Attractor dynamics and parallelism in a connectionist sequential machine. *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Kant, I. (1781/1787/1929) *Kritik der reinen Vernunft*. First and Second Editions, Riga: Hartknoch.   English translation by N. Kemp Smith as *Immanuel Kant's critique of pure reason*,   London: Macmillan.

Kosslyn, S. M., Thompson, W. L., Kim, I. J. & Alpert, N. M. (1995) Topographical representations of mental images in primary visual cortex. *Nature*, 378, 496-498.

Luck, S. J. & Beach, N. J. (1996) Visual attention and the binding problem: a neurophysiological perspective. In R. Wright (Ed.) *Visual attention*. Oxford: Oxford University Press.

Marcel, A. J. (1988) Phenomenal experience and functionalism. In A. J. Marcel & E. Bisiach (Eds.) *Consciousness in contemporary science,* Oxford: Clarendon Press.

Milner, A. D. & Goodale, M. A. (1995) *The visual brain in action*. Oxford: Oxford University Press.

Moruzzi, G. & Magoun, H. W. (1949) Brain stem reticular formation and activation of the EEG. *Electroencephalography and Clinical Neurophysiology*, 1: 455-473.

Pashler, H. E. (1991) Shifting visual attention and selecting motor responses: distinct attentional mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 1023-1040.

Pashler, H. E. (1997) *The psychology of attention*. Cambridge, MA: MIT Press.

Place, U. T. (1954) The concept of heed. *British Journal of Psychology*, 45, 243-255.

Place, U. T. (1956) Is consciousness a brain process? *British Journal of Psychology*, 47, 44-50.

Place, U. T. (1969) Burt on brain and consciousness. *Bulletin of the British Psychological Society*, 22, 285-292.

Posner, M. I. & Dehaene, S. (1994) Attentional networks. *Trends in Neuroscience*, 17, 75-79.

Posner, M. I. & Petersen, S. E. (1990) The attention system of the human brain. *Annual Review of Neuroscience*, 13, 25-42.

Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A.-M. K., Pardo, J. V., Fox, P. T. & Petersen, S. E. (1994) Practice-related changes in human functional anatomy during non-motor learning. *Cerebral Cortex,* 4, 8-26.

Rafal, R. & Robertson, L. (1995) The neurology of visual attention. In M. S. Gazzaniga (Ed.) *The cognitive neurosciences*. Cambridge, MA: MIT Press, ch. 40, pp. 625-648.

Rizzolatti, G. & Arbib, M. A. (1998) Language within our grasp. *Trends in Neuroscience*, 21, 188-194.

Rorschach, H. (1932/1942) *Psychodiagnostik*. Berne: Hans Huber. English translation as *Psychodiagnostics* by P. Lemkau & B. Kronenberg, W. Morganthaler (Ed.). New York: Grune & Stratton.

Rossetti, Y. (1997) Implicit perception in action: short-lived motor representations in space. P. G. Grossenbacher (Ed.) *Advances in consciousness research*. Amsterdam: John Benjamins.

Rossetti, Y., Rode, G. & Boisson, D. (1995) Implicit processing of somesthetic information: a dissociation between Where and How? *Neuroreport*, 6 (3), 506-510.

Rubin, E. (1915) *Synsoplevede Figurer*. Køpenhavn: Gyldendalska.

Ryle, G. (1949) *The concept of mind*. London: Hutchinson.

Schank, R. C. (1982) *Dynamic memory: A theory of reminding and learning in computers and people*. New York: Cambridge University Press.

Shallice, T. (1988) *From neuropsychology to mental structure*. New York: Cambridge University Press.

Sheinberg, D. L. & Logothetis, N. K. (1997) The role of temporal cortical areas in perceptual organization. *Proceedings of the National Academy of Sciences, USA*, 94, 3408-3413.

Skinner, B. F. (1969) *Contingencies of reinforcement: A theoretical analysis*. New York: Appleton-Century-Crofts.

Snowdon, P. (1995) Perception and attention. Paper presented to a one-day conference on 'Attention and Consciousness: Psychological and Philosophical Issues', Department of Philosophy, University College London, 26th May 1995.

Stein, B. E. & Meredith, M. A. (1993) *The merging of the senses*. Cambridge, MA: MIT Press.

Stoerig, P. & Cowey, A. (1997) Blindsight in man and monkey. *Brain*, 120, 535-559.

Tranel, D. & Damasio, A. R. (1985) Knowledge without awareness: an autonomic index of facial recognition by prosopagnosics. *Science*, 228, 1453-1455.

Treisman, A. (1988) Features and objects: The Fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, 40, 201-237.

Treisman, A. & Gelade, G. (1980) A feature integration theory of attention. *Cognitive Psychology*, 12, 97-136.

Treisman, A. & Gormican, S. (1988) Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95, 15-48.

Ungerleider, L. G. & Mishkin, M. (1982) Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.) *Analysis of visual behavior*. Cambridge, MA: MIT Press.

Vanni, S., Revonsuo, A. & Hari, R. (1997) Modulation of the parieto-occipital alpha-rhythm during object-detection. *Journal of Neuroscience*, 17 (18), 7141-7147.

Weiskrantz, L. (1986) *Blindsight*. Oxford: Clarendon Press.

Wittgenstein, L. (1953) *Philosophical investigations*. English Translation by G. E. M. Anscombe. Oxford: Blackwell.

Wittgenstein, L. (1958) *The blue and brown books*. Oxford: Blackwell.

Zihl, J., Tretter, F., & Singer, W. (1980) Phasic electrodermal responses after visual stimulation in the cortically blind hemifield. *Behavior and Brain Research,* 1, 197-203.