THE FUNCTIONS OF CONSCIOUSNESS AND ITS CONSTITUENT PARTS

Ullin T. Place and Kathleen E. Taylor

[Paper presented to the Annual Meeting of the European Society for Philosophy and Psychology, Oxford, August/September 1995]

1 Introduction: Consciousness is a brain process

In a paper entitled 'The concept of heed' which was published in the *British Journal of Psychology* more than forty years ago, the senior author (UTP) suggested

> "that the logical objections to the statement 'consciousness is a process in the brain' are no greater than the logical objections which might be raised to the statement 'lightning is a motion of electric charges'."

In a better known follow-up paper published in the same journal in 1956, he argued that

> "the thesis that consciousness is a process in the brain is ... a reasonable scientific hypothesis not to be dismissed on logical grounds alone."

The implication of this claim is that its truth or falsity will decided, not by *a priori* argument, but on the empirical evidence of psycho-physical correlation. Forty years ago that empirical evidence did not exist. Now it does.

In order to identify the brain process or system of brain processes in which consciousness consists, we must first find out what biological function it subserves. And what we need to look for in order to settle *that* question is a neurological condition whose effect is to knock out consciousness in a particular part of the sensorium. This in itself will not be sufficient to tell us which process in the brain consciousness *is*, because we know enough about how the brain works to know that a function as complex as this will not be restricted to the part of the brain whose destruction causes the function to disappear. What it will do, however, is allow us to observe what the organism which lacks this bit of consciousness is and is not able to do.

2 "Blindsight"

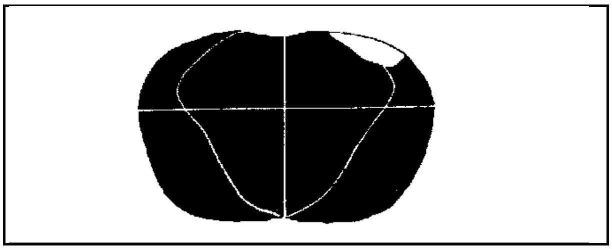
Over the past twenty years a neurological condition that answers to this specification has been identified. This is the condition known as "blindsight" since it was so named by its discoverer, Larry Weiskrantz. We will assume that everyone knows (a) that blindsight is a phenomenon which results from lesions of the striate or primary visual cortex, (b) that its principal effect is to abolish conscious experience in the affected part of the visual field, and (c) that, despite the absence of conscious experience, significant residual discrimination abilities can be demonstrated in the blind field. However, because of the way the nerve fibres from the retina of the eye project onto the striate cortex, tumours which affect only the striate cortex and whose surgical removal produces the phenomenon seldom affect both halves of the visual field. In all cases that have been investigated hitherto, the blind field is on one side only and occupies only a part of that side of the visual field. Consequently, in going about his daily life the patient can always move his head and eyes in such a way as to bring objects that attract attention into that part of the visual field that is unaffected by the lesion. This means that the only way to study what the blindsighted patient can and cannot do, relying only on his blind field, to instruct him to use his intact field to fixate a target, thereby preventing the eye-movements which would otherwise bring any stimuli presented into the range the intact part of the field, and allowing the effect of stimuli presented only to the blind field to be observed. Under these highly artificial conditions the only responses, apart from those of turning the eyes towards or reaching for an object, which have been investigated, are the verbal responses of reporting the presence and guessing the nature of the stimuli presented.

Because of these limitations, in order to take the matter further, we need to study the behaviour of an animal, preferably one of a species closely related to man in whom the *whole* of the striate cortex has been surgically removed. If, as we think, we are entitled to suppose that such an operation would have the effect of completely abolishing the animal's visual consciousness,[1] by studying what such a creature can and cannot use its eyes to do, we should be able to discover precisely what having consciousness adds to the discriminative abilities which the organism possesses in its absence.

---

[1] An elegant experimental demonstration that lesions of the striate cortex produces the same loss of visual phenomenal experience in monkeys as it does in humans has recently been reported by Cowey and Stoerig (1995).

Now it so happens that in 1965 Weiskrantz performed just such an operation on an adolescent female rhesus monkey called Helen. Over the next eight years until her death in 1973, Helen's behaviour was studied intensively by Nicholas Humphrey who reported his observations in a paper in *Perception* in 1974. Unfortunately, as was discovered on autopsy carried out by Alan Cowey after her death and as illustrated by this slide which comes from Nick Humphrey's paper, a small part of the striate cortex corresponding to an area in the upper right periphery of the visual field, was left intact after the operation. However, since there was no evidence in her behaviour of head and eye movements tending to bring objects into the part of the visual field where some striate cortical cells and hence, presumably, some visual consciousness remained, Humphrey was able to conclude



Helen's visual fields showing the probable extent of the cortical field defect

"For my part, I believe the balance of probability is that none of Helen's vision was mediated by the spared striate cortex."

Assuming that we can accept Humphrey's reassurance on this point, we can use his observations as to what Helen was and was not able to use her eyes to do after the operation as evidence for what visual functions do and do not require consciousness. The results of this study are clear cut. Although for some months after the operation she behaved as if totally blind, after a time many visual functions began to recover either spontaneously or as a consequence of specific training by the experimenter. Like the blindsighted patients she showed a remarkable ability to orient towards and reach for objects which, by analogy with the human cases, she could not "see." She also showed a remarkable ability to use her eyes to avoid obstacles in her path when moving through them. On the other hand she never showed the slightest sign of recovering her ability to recognise either a particular individual or the kind of object or situation she was confronted with. As Humphrey puts it:

"After years of experience she never showed any signs of recognising even those objects most familiar to her, whether the object was a carrot, another monkey, or myself".

Combining this evidence with that from studies of human blindsighted patients and other more general considerations, allows us to draw the following conclusions about the functions of consciousness as a whole, about the different parts of consciousness and the functions of those parts and about how it relates to those other functions that lie outside consciousness and which are retained when consciousness is absent.

3 The function of consciousness, defined as that which is absent from the affected part of the visual field in cases of blindsight, is to enable the organism to identify the kind or category of object and situation which is the source of any sensory input which is problematic, either because it is unexpected or because it is significant relative to the organism's current or perennial motivational concerns. Classifying the input source in this way allows the organism to select a behavioural strategy which is appropriate both to the nature of the situation with which it is currently confronted and its motives and objectives with respect to that situation.

4 Four sub-systems forming the functional anatomy of consciousness

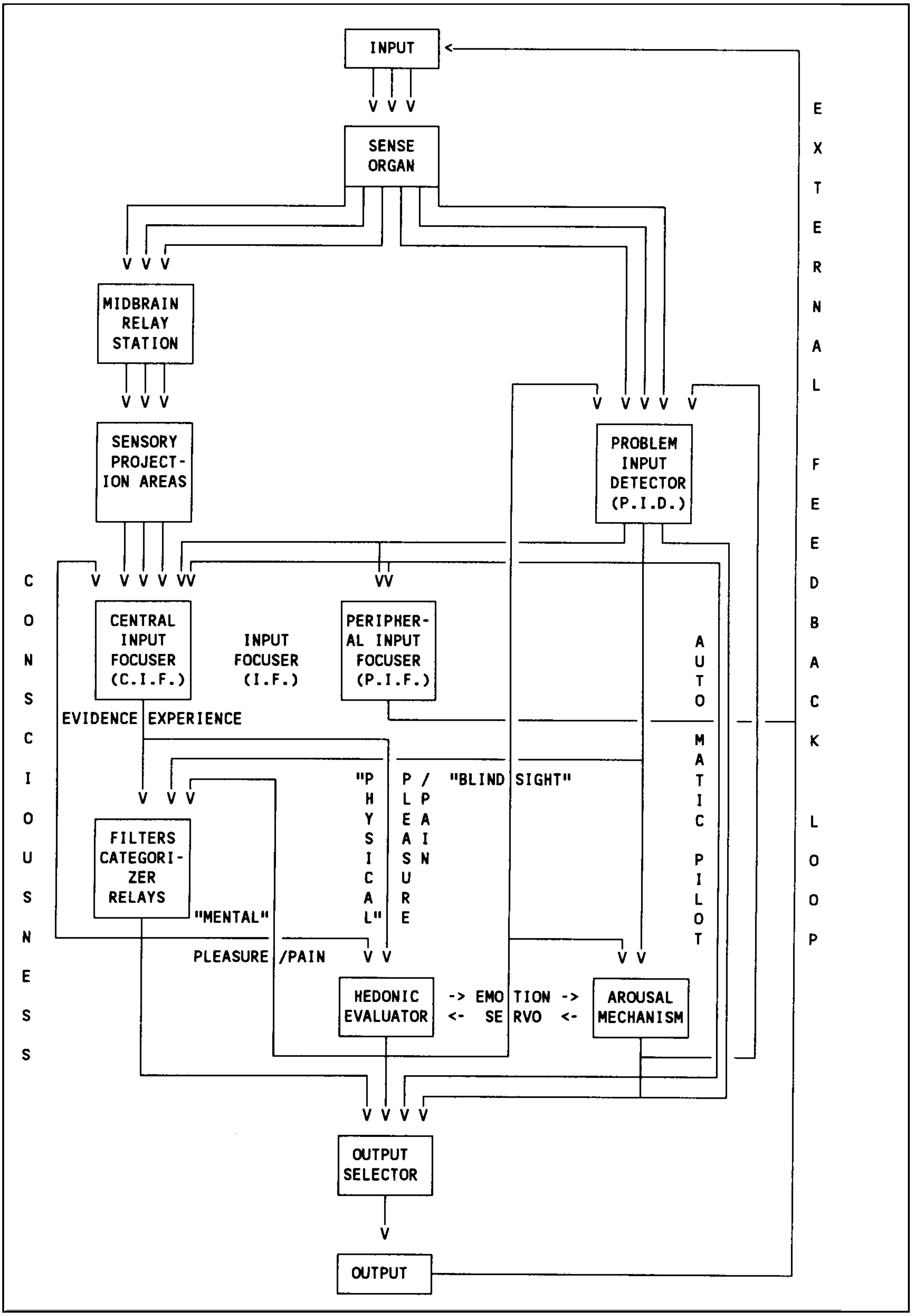
The evidence suggests that the execution of this function involves four sub-systems or modules each subserving a different sub-function within the overall function of detecting problematic inputs classifying them according to the nature of their source and selecting an appropriate response:

(a) A subconscious *problematic input detector* (PID) whose function is to scan the total current input for anything problematic and attract the focus of attention to any input thus detected.

(b) An *input focuser* (IF) whose function is to concentrate the sensory and input-processing resources of the organism in such a way as to maximise the probability that the input source will be correctly classified (categorized).

(c) A *categorizer* whose function is to select an appropriate universal, category or concept under which the problematic input source is subsumed.

(d) An *emotional servo* whose primary function is to modulate the level of activity and general direction of the system as a whole in such a way as to meet any threat to the survival of the individual or its progeny or any opportunity to enhance its chances of surviving and reproducing itself, but which is also



The functional anatomy of consciousness

sensitive to the outcome or anticipated outcome of any behavioural project with which the individual is motivationally involved. This system also has the function of allowing the organism to recuperate when no such threat or opportunity is present.

Conceived in this way, the functional anatomy of consciousness can be laid out in the form of a block diagram as shown on this slide which has been developed in collaboration with Kathleen Taylor. This is our version of what Martin Davies calls "a consciousness box".

5 Interpreting blindsight in terms of the functional anatomy of consciousness

The phenomenon of blindsight is to be explained on this hypothesis by supposing that, at least in the case of vision, the input from the senses into the brain splits into two channels before it reaches the midbrain. One channel constitutes the input which is scanned for problematic features by the subconscious PID. This sub-conscious input channel also takes over the control of output selection from consciousness as behaviour becomes habitual and automatic as a consequence of repetition. The other input channel is relayed to the various sensory projection areas of the cortex which, in the case of vision, is the striate cortex. The output of the various sensory projection areas provides the raw material from which the brain constructs the *central focus of attention* which constitutes the "evidence", as the late Donald Broadbent calls it in his 1971 book *Decision and Stress*, on which the categorization of the input source is based. This central focus of attention is the *phenomenal experience* which, in the absence of the raw material normally supplied by striate cortex, disappears from the affected part of the visual field in cases of blindsight.

In blindsight, the subconscious input channel is still intact. So are the categorization and emotional response systems. Moreover, there have to be connections between the subconscious input channel and the conscious output selection process based on categorisation and emotional evaluation in order to engineer the transition from conscious selection of behaviour to the automatic control of habitual behaviour by the subconscious input channel. It is therefore, not surprising to find that once an individual has recovered from the initial trauma of a striate cortex lesion, sensory inputs to the blind field can still influence categorization and, to a lesser extent, emotional response through the subconscious input channel. Such influence, however, is necessarily limited to those aspects of the input to which the subconscious system needs to be sensitive in order to perform its two functions, that of detecting problematic inputs and that of regulating behaviour when it is "on automatic pilot". Not only does such control fall a long way short of giving the blindsighted individual the ability to categorize the source of inputs in the affected part of the visual field, in the absence of conscious experience the brain is deprived of the ability to check its categorization of the source of an input against the "evidence" on which that categorization is based. In other words, categorization in blindsight, though sometimes remarkably accurate, becomes "pure guesswork."

6 Detecting problematic inputs

The evidence also allows us to be more specific about how the functions of the different sub-systems or modules involved in consciousness are implemented. In the case of the subconscious PID, there are two issues that confront us -

(i) the nature of the scanning mechanism which is constantly searching the whole of the current sensory input for anything problematic, and

(ii) how problematic inputs come to "stand out" from the unproblematic majority.

With regard to the first of these, a number of rhythmical patterns of neural excitation have been detected in the brain over the years, which have been interpreted as evidence for some such scanning mechanism. Such a mechanism would detect something unusual or problematic as a disruption of the normal rhythmical pattern as it swept over the surface of the brain. Evidence which has been interpreted in this way includes the *alpha rhythm* observed on the EEG during periods of drowsiness and the more recently discovered *40 Hz neuronal oscillations* which sweep the thalamus and cortex from front to back.

With regard to the second issue, given the principle of association by contiguity which has a history as principle of learning going back to Plato, it can be predicted that an organism will rapidly build up a large 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 a large number of such expectations, a second input which differs from that expected on the basis of past experience, given the first input, is going to stand out like a sore thumb. That this is the mechanism responsible for the attraction of the focus of attention to the unexpected cannot be doubted. What is more difficult to understand is how attention becomes attracted to inputs that are *motivationally significant* for the organism without necessarily being unexpected. Three types of such motivationally significant inputs need to be distinguished:

(a) objects of search, particulars or things of a kind which the individual is "on the look out for",

(b) what ethologists call "innate releasing stimuli" such as the secondary sexual characteristics of the opposite sex, or the features of an infant of the same or different species, and

(c) stimuli which are associated with motivationally significant past events, such as Plato's lyre that reminds the lover of his beloved.

Of these 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. The innate releasers are, presumably, to be explained in terms of innately determined synaptic weights within the various neural networks of which the "grey matter" of the brain consists. A theory of how dream imagery makes stimuli associated with motivationally significant past events stand out against a background of motivationally neutral associations will be presented below.

7 Two components of the input focuser: the peripheral and the central

Although in the intact organism it functions as a single integrated process, the input focuser has two sub-components, a *peripheral component* and a *central component*. The peripheral component consists in muscular movements of the trunk, limbs, head and receptor organs in such a way as to bring any problematic input source within the sensitive range of any receptor organ which may be relevant to securing a successful identification of the kind of thing it is. In the case of sensory modalities such as vision and somatic sensation, peripheral movements also have the effect of focusing stimulation from the problematic input source on the most sensitive part of the receptor surface, the tips of the fingers in the case of somatic sensation in primates, the fovea of the eye in the case of vision. In the case of vision moreover, the response of accommodation has the function of altering the focal length of the lens of the eye so as to ensure that a sharp image of the problematic input source is projected onto the fovea, while that of convergence ensures the overlap of the images on the two retinas at the point where visual attention is focused.

The central component of the attention-focusing mechanism (central input focuser - CIF) is illustrated in the case of the auditory modality by the well known "cocktail party effect", extensively studied by Broadbent and his co-workers in the "dichotic listening" experiment in which two different auditory messages are fed simultaneously into the two ears by means of ear-phones. Before categorization of a problematic feature embedded in a total input array can take place, the activity in that part of the array which is problematic must be accentuated so that it stands out as *figure* with respect to the rest of the array as *ground.* Moreover, if the figure which emerges does not immediately suggest an interpretation, the internal field must be re-organized until a recognizable figure is obtained whose categorization makes sense in relation to the current behavioural context. An example from the visual modality which illustrates the way in which different distributions of an array between figure and ground yield different interpretations is given on this slide which will be familiar to most of you. Here the effect of treating the white part of the array as figure relative to the black as ground is to give the appearance of a kind of chalice; while treating the black part as figure relative to the white as ground gives us two human faces looking at one another. You can see it either as the one or as the other. What you cannot do is see it as both simultaneously.

Figure-ground reversal

8 Intrinsic figure-ground differentiation v. imposed figure-ground organization

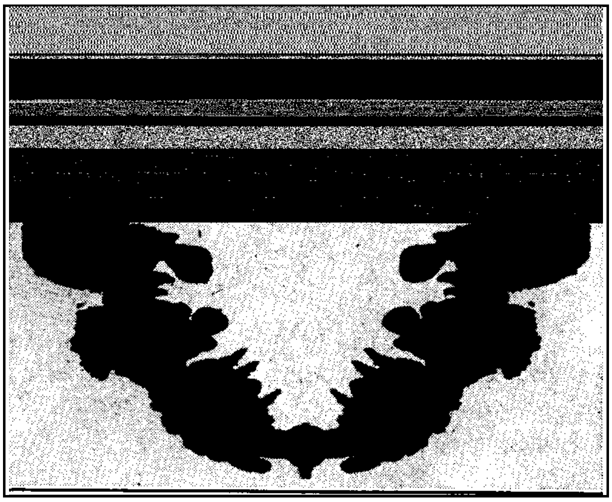
Two kinds of figure-ground relation need to be distinguished. On the one hand there is the *intrinsic figure-ground differentiation* which determines the salience of an input before the central input focuser (CIF) has got to work and thus the ease with which attention is 'caught' by the action of the PID. 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 CIF.

The two forms of figure-ground differentiation, though they become separated in the case of vision by lesions of the visual projection area, the striate cortex, are intimately connected in that the sharper the intrinsic figure-ground contrast the more strongly structured and, therefore, less malleable is the input (whether in vision or any other modality) which is available for moulding by the CIF. In other words, the larger and simpler the intrinsic contrast between figure and ground (salience) the less room there is for the CIF to impose a different pattern of figure-ground organization.

9 Phenomenal experience as imposed figure-ground organization

We have seen that on the present hypothesis it is the output or "evidence", as Broadbent calls it, which is generated by the central input focuser (CIF) which constitutes the 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 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 experience of those events at the time, together with 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.

10 Mental imagery



An ink blot similar to those used in the Rorschach Test

In the case of a mental image, a pattern of figure-ground organization is imposed on a field that is intrinsically weakly structured 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 ink blots, which, for the benefit of any not familiar with them, look something like that on this slide, 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.

We know from the introspective reports of human subjects that 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 mammals in which it occurs. 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.

11 Categorization and contingencies

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 B.F. Skinner 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*. Only by classifying its problematic inputs in a way that enables it to anticipate the consequences of selecting one form of behaviour rather than another can the organism ensure the subsequent selection of a successful behavioural strategy appropriate to whatever may be the organism's current behavioural objectives when the occasion for action arises.

12 Dimensions of emotional response

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 type of emotional responses is characterized

(a) by the type of situation that evokes it,

(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,

(d) by a distinctive variety of impulsive behaviour that is evoked, sighing in the case of relief, smiling in the case of pleasure, jumping for joy in excitement, attacking in anger, freezing/running away in fear, vomiting in disgust, lack of response and intra-punitive (self-punishing) behaviour in depression.

13 The function of dream imagery

Although there is no evidence directly relating to dreaming either in Humphrey's study of the rhesus monkey Helen or in Weiskrantz's study of 'blindsight' in human patients with lesions of the striate cortex, there is an explanation of the function of the dream imagery reported by human subjects on waking from "rapid eye-movement" (REM) sleep which interprets it as the link between the two levels at which the emotional motivational system operates, the conscious (CIF) level which must be activated before motivational evaluation of the input can occur and the sub-conscious (PID) level which controls the attraction of attention to motivationally significant inputs. In REM sleep the CIF is being allowed, as it were, to 'freewheel' when decoupled from all but the most unexpected and motivationally significant inputs. As is shown by the rapid eye-movements from which this phase of sleep derives its name, both the peripheral and the central components of the mechanism are involved; but it is only the central component represented by the dream imagery which on the present hypothesis has a specific function during sleep. Decoupled from sensory input and with, in the case of the visual modality, a wholly unstructured input being transmitted to the striate cortex from the sub-cortical structures which regulate the level of arousal within the conscious system, the central component of the mechanism is free to generate images, particularly visual ones, whose form is determined only

(a) 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

(b) by the organism's current motivational pre-occupations.

Earlier in this paper, when discussing the different varieties of problematic input to which the focus of attention gets directed when they are detected by the PID, we drew attention to the way the focus of attention is attracted to inputs which stand out as unexpected against a background of expectations of how one input type will succeed another, based on repeated experience of such sequences in the past history of the individual in question. What remains to be to be explained is how inputs which have been associated with motivationally significant events in the individual's past history come to stand out against a background of motivationally neutral associations, thereby allowing the PID to attract the focus of attention to them, even when they are not unexpected. This in our view is the function of the dream imagery which is the outstanding feature of REM sleep both in humans and in those other species of living organism in which it occurs. By rehearsing recently established associations under conditions where motivational/emotional factors are free to control the form of the images and the way they succeed one another, those associations which are motivationally significant will be selectively 'stamped in' at the expense of those that are not, in such a way that the types of input involved will attract attention when they recur in the waking state.

References

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

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

Cowey, A. (1974) Atrophy of retinal ganglion cells after removal of striate cortex in a rhesus monkey. *Perception* 3:257-260.

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

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

Llinás, R. R. and Grace, A.A. (1989) Intrinsic 40 Hz oscillatory properties of layer IV neurons in guinea pig cerebral cortex *in vitro*. *Society of Neuroscience Abstracts* 15:660.

Place, U. T. (1954) The concept of heed. *British Journal of Psychology* 45:234-255.

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

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

Skinner, B.F. (1938) *The behavior of organisms*. Appleton-Century-Crofts.

Skinner, B.F. (1969) *Contingencies of Reinforcement*. Appleton-Century-Crofts.

Steriade, M., CurróDossi, R., Paré, D. and Oakson, G. (1991) Fast oscillations (20-40 Hz) in thalamocortical systems and their potentiation by mesopontine cholinergic nuclei in the cat. *Proceedings of the National Academy of Science of the U.S.A.* 88:4396-4400.

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