

## **Applying continuous modelling to consciousness**

Kathleen Taylor, D.Phil.

University Laboratory of Physiology, Parks Road, Oxford, OX1 3PT, U.K.

Affiliation: University of Oxford  
Tel: 01865 272540  
Fax: 01865 272469 (departmental fax)  
E-mail: kathleen.taylor@physiol.ox.ac.uk

### **Abstract**

Much of neuroscience is currently dominated by an information processing metaphor which is largely conceptualized in discrete terms. An alternative metaphor conceptualizes information flow as continuous. A qualitative set of hypotheses based on this metaphor, the energy model, is described here. This model considers information transfer in terms of the flow of an abstract variable, energy, between points in a field comprising the extent of the nervous system. Although extremely simple, it generates some intriguing consequences. In particular, it provides a useful way in which to look at consciousness. Traditional problems of consciousness, such as qualia and the unity of consciousness, are briefly addressed, and outlines are sketched of the answers given by the energy model.

## **Introduction**

Despite many attempts in recent years, e.g. (Cai, 1991; Lazarev, 1992; Llinas and Pare, 1991; Niedermeyer, 1997), there is still no generally accepted theory of many large-scale neural phenomena, such as sleep, the EEG and consciousness. One reason for this is that the currently dominant information processing approach is essentially functional in emphasis, e.g. (Guigon *et al.*, 1994), and determining the functional significance of large-scale phenomena has proved unexpectedly problematic. A second reason is that treating neurons as discrete information processors when modelling large-scale phenomena leads to a combinatorial explosion of parameters, an insoluble problem given current computing power.

In other sciences, numerous researchers are beginning to apply continuous modelling techniques to phenomena long considered discrete. In brain science too, continuous modelling is being applied to large-scale neuronal phenomena. Here I present a very simple qualitative model, the energy model, which despite its simplicity makes a number of testable predictions. Many other more complex quantitative approaches exist in the literature (for example Liley *et al.*, 1998; Wright, 1997). However, what the energy model lacks in complexity it gains in interpretability and breadth of scope. It has intriguing implications when applied to consciousness, including a potential answer to the functional problem of what consciousness is for.

I should note at this point that the approach taken here is neuroscientific, not philosophical. The unashamedly neuro-physicalist approach taken in this paper undoubtedly glosses over many philosophical subtleties. For example, I use the term  $\text{mental object}=\text{throughout}$  to refer to qualia and any other contents of consciousness, without attempting to discriminate between types (for the reason that according to the model here presented, they are all instantiated in the same way, as patterns of brain activity). Nevertheless I have attempted to sketch ways in which the energy model may be applied to such philosophical issues. It seems

to me that the model has a potential for such application, and that this in itself makes it worthy of further consideration.

### **The Energy Model**

The traditional view of neurons holds them to be discrete entities performing computational functions. However, they can also be conceptualized as points in a continuous energy field (Taylor, 2000). Fields are defined over some specified domain, here the nervous system. Much of my discussion will focus on the more restricted example of activity within cerebral cortex.

Energy, which is an abstract variable related to the transmission of information between neurons, has the following properties. Firstly, it flows from inputs to outputs via the synaptic connections (weights) between neurons. That is, energy enters the field via the inputs (sensory transducer structures such as the retina) at a given time, and leaves via the outputs (motor transducer structures such as the muscle endplate) to terminate in behaviour, which dissipates energy in interactions with the environment. Secondly, the rate of energy flow is dependent on the strength of the synaptic weights, while the amount of energy flowing into the system is dependent on how active the inputs are. Thirdly, the amount of energy at a given point in the field at a given time reflects the neuronal activation at that point and time.

### **The concept of energy**

I should state to begin with that by 'energy' I do not mean the intimate biochemical details of synaptic function. 'Energy' is an approximation: instead of treating synaptic transmission as discrete, and facing the consequent combinatorial explosion, I treat it as continuous. Undoubtedly more complex hypotheses would more accurately reflect synaptic behaviour. However, I feel one should begin at the beginning, with the simplest model capable of generating interesting predictions.

The terms ‘information flow’ and ‘information processing’ are used widely in cognitive neuroscience, while in neuroimaging studies talk of ‘neuronal activation’ is commonplace. Precise definitions of these terms are hard to achieve, and they may often be used with different referents. For example, a PET and a MEG study may both discuss their findings in terms of ‘neuronal activation’, even though one is measuring levels of radioactive oxygen and the other electromagnetic changes. Nevertheless, there is consensus that the terms ‘information’ and ‘neuronal activation’ do have some meaning, that there is some quantity which is stimulated in cortex by sensory input, transmitted through cortex and in some sense dissipated by motor output. The energy model is an attempt to capture this intuition. The currently dominant information processing metaphor is usually expressed in discrete terms, envisaging each neuron as an individual processor, each cortical area or functional network as a ‘module’, and so on. Yet within this intellectual environment the term ‘information flow’ is often used. All the energy model does is take that term literally.

In physics, it is difficult to say what energy actually is. Rather, it is defined by its relationships to other variables. I have tried to do the same with my ‘energy’, relating it to other variables (for example synaptic strengths, cortical distances), stating that it is closely linked to ‘neuronal activation’, and so on. The reason for using the term ‘energy’ is that the behaviours being modelled seem most closely fitted by the concept of an energy field.

#### On what parameters does energy flow depend?

This can be illustrated by considering the simplest case: energy flow between two points A and B. The assumption made is that a connection between two points in the field can be treated as a cylindrical medium through which energy flows (see Figure 1). The rate of energy flow per unit time from point A to point B is proportional to the strength of the weight (the synapse) between A and B: the larger the weight, the more energy can flow along it in a given time, *ceteris paribus* (all things being equal). The rate of flow is also proportional to the difference in energy between the two points: the larger this gradient between A

and B, the faster the flow of energy between them, *ceteris paribus*. The flow rate is inversely proportional to the distance between A and B: longer distances mean less energy is transferred from one point to the other in a given time.

How fast energy flows also depends on the type of neuronal 'material' through which the energy is flowing, and on the amount of 'weight leakage' (removal or addition of energy within the process of transfer along the weight) which occurs as energy flows from one point to another. For simplicity, I have assumed that the type of material is the same throughout the nervous system, and that no weight leakage occurs.

### Weights

For simplicity, the weight which connects points A and B is proposed to be adjusted by some biologically plausible mechanism such as Hebbian learning (Obermayer *et al.*, 1995; Pennartz, 1997). This is assumed to occur over a longer timescale than that of synaptic transmission. The weight change at a given time therefore depends on the amount of energy in A and B at that time. High energies at A and B lead to greater weight change if maintained over time than briefly high inputs. In the two-point closed system discussed above, maintained high energy levels, due to high levels of input over time to both A and B, would lead to uncontrolled weight increases.

However, our simple example does not in reality exist in isolation (it is not a closed system). Energy is transmitted on to other points with a speed which increases as the strength of the weight increases. Repeated co-activation of inputs strengthens the weights and shortens the time taken for energy to flow from A to B, and from these points to others. So stronger points will actually be active (have energy flowing through them) for less time than weaker ones. Consequently the degree of weight change will be less. As the activation time decreases to the point where it is less than the timescale of weight change, weight change will tend to zero. Thus weights will settle towards fixed values by a process of gradual optimization, avoiding the uncontrolled increases typical of

pure=Hebbian learning in simulations. Numerous biological means of achieving such settling=have been proposed: neuronal plasticity is likely to be regulated by normalizing mechanisms (Carandini *et al.*, 1997) such as receptor up- and down-regulation and dendritic growth.

Over time, this tendency of frequently used synapses to settle=will allow energy to flow quickly through them and into other connections. In a sense, a point= function= is to balance the energy leaving it against the amount flowing into it at a given time. Learning is used to adjust the weights to maximise the rate and smoothness of energy flow from one point to another, and to minimise the amount of energy lingering at a given point. The degree of strengthening between two points A and B will tend to decrease as the distance between them increases: *ceteris paribus*, if energy takes longer to get from A to B then the time period over which A and B are both highly active, and the weight between them can therefore strengthen, is likely to be less.

### Outputs

Energy leaves the system via a range of outputs and is dissipated in patterns of behaviour (here broadly defined to include not only muscular control but neuroendocrine and neuroimmune effects). Many behaviour patterns are mutually exclusive. Two such patterns cannot be activated simultaneously, and yet both may be possible responses to a given sensory stimulus. In the oculomotor system, competition appears to occur between representations of conflicting eye movements; neurons in the frontal eye fields send excitatory projections to neurons in the superior colliculus which encode the same eye movement parameters, but inhibit neurons encoding different parameters (Schlag Rey *et al.*, 1992). Similar competitive mechanisms may also occur in other areas.

This competition effectively sets a threshold on motor output: any potential output must be strongly enough activated to overcome inhibition from any other potential outputs active at the same time. This corresponds to a localised build-up of energy in the pathways encoding a particular output representation, relative to

other pathways encoding competing representations. When the threshold is reached, behaviour is generated, rapidly draining energy from the system. The longer the energy takes to reach threshold, the longer the organism's reaction time.

### Inhibition

Inhibitory connections are usually modelled as having a negative sign. On the energy model, the effect of this is to reverse the direction of flow. Consider two points A and B, with A more active than B. If the weight between them is excitatory, energy will flow from A to B. If the weight is inhibitory, energy will flow from B to A. This has a blocking effect: it prevents energy from A reaching B and continuing from B to other points. Instead the energy which reaches A must find alternative routes.

### Energy and input probability

The flow of energy effectively acts to recode likelihood in terms of time. A given input must pass through a complex matrix which weights it according to the organism's prior experience. Low probability inputs will tend to take considerably longer to stimulate an output than more familiar (high probability) inputs, because the synaptic weights which process them are less well-established and therefore slower to transmit energy.

Energy flow also depends on the complexity of the initial input, and its intensity. Energy from more complex inputs will tend to stimulate an output more slowly than energy from simpler inputs; since more complex inputs tend to activate more pathways, more conflict resolution is required. Similarly, less intense inputs have lower energy, so take longer to reach threshold. High-intensity inputs, whether low or high probability, will stimulate a large flow of energy, so they will be likely to reach output before weaker inputs activating the same pathways. However, the organism's past experience will still influence the speed with which

a high-intensity or complex input stimulates an output behaviour. These hypotheses are testable.

### **Prefrontal Cortex**

This is interesting when we come to consider the structure of cortex. Neuroanatomical studies suggest that the primary sensory areas, such as primary visual cortex (V1), can feed activation to motor (output) areas of cortex via a fast and relatively direct route. For vision, a major contributor to this sensorimotor transformation is parietal cortex (Elkington *et al.*, 1992). The information it processes is primarily, although not exclusively, concerned with motion and the spatial location of the stimulus. Visual information is also transmitted via a more indirect, and somewhat slower, route, involving areas of cortex in the temporal lobe. These pathways, the classic *what*-and *where*-divisions of Ungerleider and Mishkin (Ungerleider and Mishkin, 1982), are thought to interact at multiple levels.

These are not, however, the only routes from sensory input to motor output. Indeed, calling them pathways makes the process of information transfer sound more clear-cut and spatially limited than in fact it is. As activation flows from V1, for example, it is carried by other routes as well as those directly processing sensorimotor transformations. MEG and ERP studies show complex, spreading patterns of activation for even the simplest stimuli (Kelso *et al.*, 1998). A major recipient of information from *what*=(temporal lobe) and *where*=(parietal lobe) areas, as well as from the insular and cingulate cortices, is the prefrontal lobe (Kawamura and Naito, 1984; Barbas, 1988).

Prefrontal cortex (PFC) can be thought of as acting as an interface between sensorimotor function and more cognitive functions such as memory. Its many subdivisions receive input from, and send output to, distributed networks throughout cortex (Knight *et al.*, 1999). Thus it can be considered to function in parallel with stimulus-driven input-output processing (Buchel and Friston, 1997),

allowing the incorporation into such processing of complex contextual information, and monitoring cortical activity. Prefrontal outputs are thought to be both inhibitory and excitatory (Knight *et al.*, 1999). Targets of prefrontal connections are not restricted to cortex, but incorporate many important subcortical areas and pathways including the thalamus, superior colliculus and brainstem (Leichnetz and Gonzalo Ruiz, 1996).

Weaker, more complex or less familiar inputs are more likely to activate prefrontal cortex

*Ceteris paribus*, low probability inputs will generally be processed more slowly than more familiar inputs, and will likely result in a more complex pattern of activation. They will be more likely to activate the PFC prior to a motor response. This is because, on the energy model, energy is a limited resource; for a given stimulus, if less energy flows down one pathway (e.g. because that pathway is weaker), more is available to flow down alternative routes. A low probability input will therefore tend to activate any given synapse weakly, but to activate a greater number of synapses, compared with a more familiar input of equal intensity.

Put another way, energy from high probability inputs will flow mostly through fast, frequently used pathways between input and output, and will tend to reach the output before it reaches prefrontal areas. Energy from low probability inputs will flow along paths to output, and to prefrontal areas, much more evenly. It is thus more likely that a low probability - unfamiliar - input will activate prefrontal areas, prior to a motor response, than a high probability input. Infrequent visual stimuli have been shown to activate PFC transiently (McCarthy *et al.*, 1997). Since, on the energy model, the activation of an output motor response ‘drains’ the energy out of the nervous system, energy from the input stimulus is thereafter no longer available to activate the PFC.

The same pattern holds for strong and weak inputs (of equivalent familiarity), and for complex and simple input patterns. The amount of energy provided by a

stimulus is linked to the time that input takes to be processed. In addition, the faster an input is processed, the less its activation of additional pathways ('path leakage') during processing. Strong inputs will therefore tend to involve less path leakage than weaker inputs. In neuroanatomical terms, they will be less likely to activate the PFC prior to a motor response.

### Prefrontal inhibition can cause localised energy build-ups

It is also worth noting that on the energy model inhibition tends to block energy flow. Energy will therefore tend to build up (pool) in the vicinity of the inhibited unit. Since the main direction of energy flow, up to the first motor response, is generally presumed to be from sensory to motor (posterior to anterior), this inhomogeneity will tend to be greater on the sensory side of the inhibition. Because of path leakage, energy will tend to flow away into sensory areas connected with the area in which the pooling is occurring. Phenomenologically, this evokes the mind's great tendency to drift onto other ideas associated with (triggered by) the original thought. Focusing consciousness on a particular mental object, such as a thought - equivalent, in the model, to the localised buildup of energy - is hard to maintain.

### Function of the PFC

Given that the PFC appears to have a large number of inhibitory outputs, and given the wide range of areas from which it receives and to which it projects, it is likely that a major function of this area could be to control the flow of energy by generating or suppressing local inhomogeneities. This control would not, of course, be *de novo* - that is simply to give the homunculus a new residence in the prefrontal lobe. Rather, prefrontal control of energy flow in other areas of cortex would be driven by the inputs received by the prefrontal lobe at a given time. These inputs would be from sensory areas, but also from more cognitive (higher level) areas stimulated by the current sensory input.

How is such control exerted in the developed brain? Consider the example of a novel and complex visual stimulus S, which initially activates representations of many possible targets and eye movements in the parietal cortex and the frontal eye fields respectively (Goldberg and Bruce, 1990; Platt and Glimcher, 1997). On the energy model, energy provided to the field by S will flow into any pathways previously used by energy from stimuli with similar features (for example the same colour or location in space). As well as the sensorimotor ('where') pathways encoding target and eye movement representations, these pathways may include object analysis ('what') temporal lobe pathways which encode notable features of S (for example large size or bright colour), and hippocampal/limbic pathways which encode previously formed associations between stimuli like S and other sensory, cognitive or emotional stimuli. S will not stimulate an immediate motor output, removing energy from the field, because of the conflict between multiple potential outputs. Some energy will therefore have time to reach the PFC from pathways encoding representations of features of S or its associations.

The PFC can act as an interface between stored knowledge (representations of which are 'retrieved' by the energy flowing through them from S) and the sensorimotor pathways activated by S. Energy reaching the PFC from object analysis areas, for example, which would otherwise have dispersed, can be transmitted preferentially to a chosen sensorimotor representation, boosting its activity sufficiently to generate a motor output. Which representation is chosen will depend on how connections between the PFC and other areas have been formed by experience. If the PFC operates on the same competitive principles found in other areas of cortex such as the frontal eye fields (Schlag Rey *et al.*, 1992), one would expect to find that excitatory connections between a point in the PFC and pathways encoding a given sensorimotor representation would be balanced by inhibitory connections between that point and pathways encoding competing representations. This is empirically testable in primates if not in humans.

The PFC allows energy to flow between networks encoding information about a stimulus S and those encoding sensorimotor transformations of S. This coordinating role facilitates the strengthening of connections between these networks. Even excluding links via subcortical regions, neuroanatomical studies in monkey suggest that parietal and temporal lobe areas are highly interconnected (Jouve *et al.*, 1998), so such connections are likely to be direct as well as to run via the PFC. Furthermore, the direct connections will be over shorter distances, given the relative position of the parietal, temporal and prefrontal lobes. Since as discussed earlier shorter distances imply greater weight change, this means that over time, as stimulus S becomes familiar, the direct weights will strengthen faster than those via the PFC, leading to decreased activation of prefrontal areas. Prefrontal activity thus tends to be self-limiting. In a sense, the function of the PFC is to minimise the amount of energy passing through it by maximising the speed of energy transfer from input to output.

The PFC can thus be thought of as a large number of associative networks, each linking fragments of cognitive and sensory information about particular stimuli together to provide an experiential filter through which all but the strongest or most familiar inputs pass. To begin with, this filter will be broad and non-specific; it is refined by experience. Developmentally, the PFC appears to be the youngest part of cortex: it is thought to mature until adulthood in humans (Rosenberg and Lewis, 1995; Woo *et al.*, 1997). As the filter develops it becomes more able to manipulate the flow of energy, channelling it into restricted pathways (for example by inhibiting other options), and thereby strengthening those channels much more efficiently. This allows learning to be faster and more precisely specialized than if no such manipulation occurred, which facilitates the further refinement of the filter. As the filter refines, one would expect to see corresponding differentiation of the underlying tissue, and indeed the adult PFC of both monkeys and humans is not homogenous, either in its cytoarchitectonics (Rajkowska and Goldman-Rakic, 1995a; Rajkowska and Goldman-Rakic, 1995b), neurotransmitter distributions (Gebhard *et al.*, 1995; Goldman-Rakic *et al.*, 1990), or connectivity (Carmichael and Price, 1994; Cavada *et al.*, 1995).

## The PFC and executive control

According to the energy model, the PFC acts to minimise its own activity by optimising flow from inputs to outputs. It does this by providing an ‘experiential filter’ which allows stored information to contribute to the process of resolving conflict between potential outputs. I speculate that this activity is equivalent to ‘executive function’. This is perhaps the most unsettling aspect of the energy model. At any given moment, energy is pouring into the nervous system, and hence into cortex, from a huge variety of inputs. If the input patterns are familiar they can be processed (energy transmitted to the relevant outputs) fast; no executive processes, thought to be mediated by prefrontal areas (Weinberger, 1993), need be engaged. Even if prefrontal areas are engaged, however, there is no specific decision-making module, no seat of the will, in which a rational comparison of all possible options takes place. Energy from the input-processing networks simply flows into connected areas of the PFC (flushing out any lurking homunculi en route) and from the PFC most speedily to whichever networks experience has selected to have the strongest links with the PFC. One implication of this is that once the organism’s history has set up its brain pathways, there is no way in which the organism can decide to change those pathways by itself. Some input stimulus (whether from the body or the external environment) must be the trigger which evokes the prefrontal activation essential for such change to occur.

Having set out the energy model, let us consider how it re-interprets consciousness.

## **Consciousness**

Philosophers and neuroscientists working on this topic have tended to distinguish two aspects or types of consciousness (Block, 1995; Weiskrantz *et al.*, 1995). Firstly, there is a generalized state of awareness: the stream—or flow—of experience. In the human domain, this type of consciousness is used to discriminate between robots, zombies and unconscious or brain-dead humans on

the one hand, and awake humans on the other. Secondly, there is a much more specialized state of "internal monitoring" in which the conscious subject seems to "check" his internal states ("tapping into the stream of consciousness") and can report on or otherwise utilize their content (Levine, 1990). This ability, sometimes called self-consciousness, is often thought of as being restricted to humans, though the debate on this point is lively. In a sense, the former is a potential for being conscious, the latter an actualization of this potential. Nomenclature varies. I will use the terms "awareness" and "monitoring" respectively. Of course definitions vary also; consciousness is highly controversial. It is not clear to what extent awareness and monitoring are separate. The functions and interrelationship of these two types of consciousness have not been satisfactorily defined. What does the energy model say about them and their relationship?

#### Awareness and monitoring

Cortical activity is defined by the behaviour of an energy field with time. It is the structure of the field, on the energy model, which constitutes awareness. Both the postulated energy field and awareness are continuous, highly variable, and susceptible to changes in both external inputs (Bottini *et al.*, 1995) and internal inputs from one part of the field to another - for example from the ascending reticular activating system, thought to control alertness (Mesulam, 1981), and from the body, via the hypothalamus and cingulate (Vogt and Gabriel, 1993). In addition, changes in awareness are correlated with changes in the electrochemical activity of the cortex, as EEG and other studies tell us (Menon *et al.*, 1997; Schellberg *et al.*, 1990). On the energy model, energy represents information flow in cortex. The assumption is that the electrochemical activity at a synapse reflects information flow (however measured) across that synapse; in which case awareness will also be related to information flow, that is to energy.

As energy flows through the nervous system from input to output, flow may be more or less smooth. If the inputs reaching cortex are sufficiently complex or weak, PFC areas may be stimulated before the motor response is initiated (and

energy drained away). Depending on what associations are triggered by the energy flow, the PFC may manipulate flow to cause inhomogeneities in some areas. Prefrontal inhibition can be direct (via a GABAergic<sup>1</sup> connection), or by utilizing local inhibitory networks (of GABAergic interneurons). It has been suggested that such local inhibitory networks can implement a winner-takes-all policy over time, such that smaller activations of units within the network quickly tend to zero, eventually leaving only the initial maximum active (Taylor and Stein, 1999). This implies that inhibition may act to enhance already active areas of the energy field, while suppressing flow through other, less active areas. Again, this has a phenomenological correlate: being conscious of something means not being conscious of very many other things. I propose that as the focus of consciousness on a single mental object becomes more intense, more energy pools in the network concerned with processing that object, and less is available to flow through other networks. Monitoring can be thought of as the effect of the PFC on energy flow. In its extreme form, this can reduce the energy field to having a peak of activity in a single neural pathway - although because of path leakage, as described above, this state of affairs can rarely be maintained for very long. In passing, I speculate that states of consciousness such as those found in meditation - in which the mind is described as empty - are reverse phenomena, in that energy is fairly evenly distributed through the cortical part of the energy field, with no large inhomogeneities. If correct, this predicts an absence of prefrontal activation during such altered states.

In cognitive terms, monitoring is equivalent to having attention highly focussed on activity in the area where pooling occurs. When energy flows smoothly and swiftly from input to output, attention (pooling) does not occur to the same extent. Behaviours thus generated are referred to as automatized. The energy model conceptualises an automatic movement as one in which the input-output pathways are so well-established that there is little leakage of energy into alternative routes (path leakage). There is therefore little competition between alternative input-output possibilities to be resolved before the movement is

---

<sup>1</sup>γ-amino-butyric acid is the main inhibitory neurotransmitter in cortex.

generated. If consciousness is focussed on such behaviours during their execution, the result is disruption of the flow, and hence, of the smooth execution of the behaviour. By contrast, complex, non-automatic movements will generate much more path leakage, since the pathways are by definition less well established. That is, they will activate many more pathways, including those concerned with contextual information (e.g. in cingulate and prefrontal regions). This gives rise to many possible energy flows competing for control of the motor output. Here it may be advantageous to focus consciousness on one flow at the expense of others; it speeds up decision-making.

The usefulness of monitoring is therefore threefold. Firstly, as described above, it may facilitate faster and more precise learning by restricting the number of paths through which energy can flow. Secondly, it may provide a mechanism for strengthening one potential pathway over another, using the collated experience of the organism to do so. This may help reduce competition for outputs (conflict resolution). Such a mechanism allows past experience efficiently to modulate sensorimotor networks. Thirdly, monitoring has proved invaluable for humans, because the PFC can be refined until its filter is capable of discriminating individual, complex and highly abstract thoughts. This process of extreme refinement is associated with the development and use of language.

### **Implications of the energy model**

#### **Consciousness is time-dependent**

One of the implications of the energy model is that for any given input state, the organism will only be conscious of that input if there is enough time between input and output for an inhomogeneity to build up. If the input-output pathways are so well established that energy flows smoothly through the field, the organism's reaction time will be fast, and it will not be conscious of either the input or the decision to initiate a response. This is the case for highly-automatized movements such as walking (in adults, at least). In other words, the model predicts that the likelihood of an organism becoming aware of a given stimulus

will correlate negatively with the reaction time of a motor response to that stimulus. The reverse also holds: during periods of high activity, an organism is likely to be less conscious of stimuli than during rest or low activity.

### Consciousness and memory

By providing a more differentiated energy field, monitoring may also improve the efficiency of memory storage, since a clear and distinct pattern of neural activity is presumably easier to encode than a broad swathe of undifferentiated neuronal activation. Indeed, I speculate that monitoring may be essential for memory storage. Implicit and explicit memory should be distinguished here. On the energy model, implicit memory is mediated by changes in weights as energy flows through them, and does not require monitoring. Explicit memory only occurs when a local inhomogeneity becomes sufficiently great that energy flowing from it can reach and activate areas specialized for storage, such as the hippocampus (Knight and Nakada, 1998), before the activation of a motor output drains energy from the system. The model therefore predicts that prefrontal activation may be required for explicit but not implicit encoding.

### Consciousness is developmentally as well as evolutionarily graded

The hypothesis that the function of consciousness is to manipulate the energy field of awareness - to assist learning, fluid action and (for humans) description - has interesting consequences. For example, one implication is that monitoring is likely to be a learned ability. The PFC develops relatively late, reaching peak connectivity at about four months postnatally, and continuing to develop until adulthood (Rosenberg and Lewis, 1995; Woo *et al.*, 1997). Since the grain-of-consciousness - the discriminability of individual mental objects- is, on the energy model, correlated with the degree of development of prefrontal cortical function and connectivity, this provides a gradient of degree of consciousness- which fits the common-sense intuitions that babies and animals are less fully conscious than healthy adult humans.

## Consciousness is an entirely physical phenomenon

On the energy model, consciousness at a given time, that is, consciousness of a given mental object, is always due to a complex combination of the organism's current input state  $I(t)$ , the state of its energy field  $E(t)$ , and its past experience or history  $H(t)$ . Past experience, as stored in the cognitive networks of the brain, was itself generated by some combination of  $I(\tau)$ ,  $E(\tau)$  and  $H(\tau)$ , where  $\tau$  is the time at which the experience was generated. Although highly complex,  $I(\tau)$ ,  $E(\tau)$  and  $H(\tau)$  are in principle specifiable in purely physical terms for any given time  $\tau$  in the organism's lifetime. In practice, their complexity places any full specification of them far beyond the reach of current neuroscience.

## Unity of consciousness

Consciousness is generally thought of as unified. Indeed, self-consciousness is traditionally referred to as a "stream of consciousness" which encompasses only one mental object at a time. Indeed, when we introspect, "looking at ourselves to see what we are thinking about," we find that we appear to be able to focus on only one object at a time.

However, the "we" in the previous sentence is typically taken to refer to a mature human adult, equipped with full linguistic capabilities and years of experience in describing his (or her) inner life, to himself or for the benefit of others. If, as argued above, consciousness is a spectrum, a matter of degree, then taking this highly specialised case as typical is like assuming that all human beings have Richard Feynman's affinity for quantum mechanics. We view our consciousness as unified for two reasons. The first is that we are organisms who have a single physical body and a Darwinian imperative to preserve it which often dictates unified action - in that sense we are "one thing." The second is that we have been taught to think of ourselves as singularities. Human language and society teaches us about an attribute called the self from so early on, and so vigorously, that any challenge to the notion seems peculiar.

On the energy model, consciousness is singular, focussed on only one object at a time, because the act of introspection - focussing in on consciousness - makes it so. This focussing is something we learn to do with greater precision as our prefrontal lobes mature. One might expect, therefore, that consciousness feels less unified to a young child than to an adult; that is, that the sense of self develops with age. One might also expect that the ability to focus would be variable within the population, as are so many other features of brain function.

### Qualia

As noted at the start of this paper, it is written by a neuroscientist, not a philosopher. I have throughout made the basic physicalist assumption that consciousness is to do with the brain, while noting that it is also to do with social factors - in that it is a learned ability which may develop to different degrees in people dependent on their upbringing. The energy model gives a brusque and no doubt over-simplistic answer to the problem of qualia: if two individuals have the same pattern of energy flow through identical neuronal pathways in their brains, and if those energy flows are experienced as conscious, what the two individuals experience will be the same.

At one level this is so unlikely, given individual differences in both brains and environments, that consciousness must surely be unique at all times. However, the remarks earlier about the grain-or discriminability of awareness are worth reviewing here. This grain is refined by usage: the more the prefrontal monitor is used, the more adept it will be at differentiating patterns of energy flow. So the judgement of what is the same-and what different, in terms of mental objects, will differ between individuals. Thus two children could experience a shade of red as the same, while two trained artists experienced it differently from each other, according to the energy model.

### Plasticity

The energy model says that neuronal architecture is much more dependent on the pattern of inputs than has generally been thought. In principle, knowing the pattern of inputs over time to an area of cortex should allow the accurate prediction of its function, and changing those inputs should allow that function to be altered. (In practice, of course, the effects of noise will place limits on the accuracy of such predictions.) We already see this in plasticity following brain or body damage (Maldonado and Gerstein, 1996; Ramachandran *et al.*, 1992; Yang *et al.*, 1994). Another consequence is that cortical areas spatially further from the major sensory inputs (e.g. vision) should be less functionally committed, have a greater variety of inputs, and be more plastic in their neuronal properties. These properties are particularly characteristic of the parietal and frontal lobes, as one would expect from their locations relative to sensory inputs. This may help to explain why so many functions have traditionally been ascribed to these areas. Studies are already addressing this issue by looking at brain changes during learning (Zohary *et al.*, 1994), but determining the extent of brain plasticity remains a huge challenge.

### External interference

The energy model also predicts that adding energy to the field will distort the structure of awareness to a degree dependent on the amount of energy added. We can see this with electrical stimulation such as that used in electroconvulsive therapy, which at low levels causes conscious sensations, at higher levels seizures, unconsciousness and amnesia (Friedberg, 1977). Incidentally, similar interference occurs when energy is transferred to the skull and brain by a blow on the head. Light blows cause sensations (relatively small variations in the energy field's structure), heavy blows unconsciousness (catastrophic disruption of the energy field's structure). The analogy of waves passing through water is perhaps informative here. Adding a few droplets to the wave will not change its structure greatly. Pouring a bucket of water over the wave alters its structure in such highly complex ways that the original structure is effectively abolished.

### Conclusions

I have argued that viewing neurons as discrete information processors, while useful in some circumstances, is less useful when it comes to understanding and modelling large-scale phenomena such as consciousness. For such phenomena a field theory may be more useful. Central to field theory is the concept of energy flow between points in the field. I have given a simplistic example of such an energy model. The model generates numerous testable predictions, some of which have been described here. It has intriguing implications when applied to consciousness. Of course, the model described here is at a very preliminary stage. Much more needs to be said, in much more detail, about how it can be applied to clarifying what consciousness is and what it is for. However, the claim of this paper is that the model's usefulness makes it an important candidate for further theoretical and empirical research.

## References

1. Barbas, H. (1988). Anatomic organization of basoventral and mediodorsal visual recipient prefrontal regions in the rhesus monkey. *Journal of Comparative Neurology* **276**: 313-42.
2. Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences* **18**: 227-287.
3. Bottini, G., Paulesu, E., Sterzi, R., Warburton, E., Wise, R. J., Vallar, G., Frackowiak, R. S., and Frith, C. D. (1995). Modulation of conscious experience by peripheral sensory stimuli. *Nature* **376**: 778-81.
4. Buchel, C., and Friston, K. J. (1997). Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cerebral Cortex* **7**: 768-78.
5. Cai, Z. J. (1991). The functions of sleep: further analysis. *Physiology and Behavior* **50**: 53-60.
6. Carandini, M., Heeger, D. J., and Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience* **17**: 8621-44.
7. Carmichael, S. T., and Price, J. L. (1994). Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *Journal of Comparative Neurology* **346**: 366-402.
8. Cavada, C., Company, T., Hernandez-Gonzalez, A., and Reinoso-Suarez, F. (1995). Acetylcholinesterase histochemistry in the macaque thalamus reveals territories selectively connected to frontal, parietal and temporal association cortices. *Journal of Chemical Neuroanatomy* **8**: 245-57.
9. Elkington, P. T., Kerr, G. K., and Stein, J. S. (1992). The effect of electromagnetic stimulation of the posterior parietal cortex on eye movements. *Eye* **6**: 510-4.
10. Friedberg, J. M. (1977). Shock treatment, brain damage, and memory loss: a neurological perspective. *American Journal of Psychiatry* **134**: 1010-1013.
11. Gebhard, R., Zilles, K., Schleicher, A., Everitt, B. J., Robbins, T. W., and Divac, I. (1995). Parcellation of the frontal cortex of the New World

- monkey *Callithrix jacchus* by eight neurotransmitter-binding sites. *Anatomy and Embryology (Berlin)* **191**: 509-17.
12. Goldberg, M. E., and Bruce, C. J. (1990). Primate frontal eye fields. III. Maintenance of a spatially accurate saccade signal. *Journal of Neurophysiology* **64**: 489-508.
  8. Goldman-Rakic, P. S., Lidow, M. S., and Gallager, D. W. (1990). Overlap of dopaminergic, adrenergic, and serotonergic receptors and complementarity of their subtypes in primate prefrontal cortex. *Journal of Neuroscience* **10**: 2125-38.
  9. Guigon, E., Grandguillaume, P., Otto, I., Boutkhal, L., and Burnod, Y. (1994). Neural network models of cortical functions based on the computational properties of the cerebral cortex. *Journal of Physiology Paris* **88**: 291-308.
  10. Jouve, B., Rosenstiehl, P., and Imbert, M. (1998). A mathematical approach to the connectivity between the cortical visual areas of the macaque monkey. *Cerebral Cortex* **8**: 28-39.
  11. Kawamura, K., and Naito, J. (1984). Corticocortical projections to the prefrontal cortex in the rhesus monkey investigated with horseradish peroxidase techniques. *Neuroscience Research* **1**: 89-103.
  12. Kelso, J. A., Fuchs, A., Lancaster, R., Holroyd, T., Cheyne, D., and Weinberg, H. (1998). Dynamic cortical activity in the human brain reveals motor equivalence. *Nature* **392**: 814-8.
  13. Knight, R. T., and Nakada, T. (1998). Cortico-limbic circuits and novelty: a review of EEG and blood flow data. *Reviews in the Neurosciences* **9**: 57-70.
  14. Knight, R. T., Staines, W. R., Swick, D., and Chao, L. L. (1999). Prefrontal cortex regulates inhibition and excitation in distributed neural networks. *Acta Psychologica Amsterdam* **101**: 159-78.
  15. Lazarev, V. V. (1992). Informativeness of different approaches to EEG mapping in the study of human mental activity. *Human Physiology* **18**: 411-18.
  16. Leichnetz, G. R., and Gonzalo Ruiz, A. (1996). Prearcuate cortex in the *Cebus* monkey has cortical and subcortical connections like the macaque

frontal eye field and projects to fastigial-recipient oculomotor-related brainstem nuclei. *Brain Research Bulletin* **41**: 1-29.

17. Levine, D. N. (1990). Unawareness of visual and sensorimotor defects: a hypothesis. *Brain and Cognition* **13**: 233-81.
18. Liley, D. T. J., Cadusch, P. J., and Wright, J. J. (1998). A continuum theory of cortical activity. *CNS '98: Seventh Annual Computational Neuroscience Meeting*, Santa Barbara, CA.
19. Llinas, R. R., and Pare, D. (1991). Of dreaming and wakefulness. *Neuroscience* **44**: 521-35.
20. McCarthy, G., Luby, M., Gore, J., and Goldman-Rakic, P. (1997). Infrequent events transiently activate human prefrontal and parietal cortex as measured by functional MRI. *Journal of Neurophysiology* **77**: 1630-4.
21. Maldonado, P. E., and Gerstein, G. L. (1996). Reorganization in the auditory cortex of the rat induced by intracortical microstimulation: A multiple single-unit study. *Experimental Brain Research* **112**: 420-430.
22. Menon, V., Ford, J. M., Lim, K. O., Glover, G. H., and Pfefferbaum, A. (1997). Combined event-related fMRI and EEG evidence for temporal-parietal cortex activation during target detection. *Neuroreport* **8**: 3029-3037.
23. Mesulam, M. M. (1981). A cortical network for directed attention and unilateral neglect. *Annals of Neurology* **10**: 309-25.
24. Niedermeyer, E. (1997). Alpha rhythms as physiological and abnormal phenomena. *International Journal of Psychophysiology* **26**: 1-3.
25. Obermayer, K., Sejnowski, T., and Blasdel, G. G. (1995). Neural pattern formation via a competitive Hebbian mechanism. *Behavioral and Brain Research* **66**: 161-7.
26. Pennartz, C. M. (1997). Reinforcement learning by Hebbian synapses with adaptive thresholds. *Neuroscience* **81**: 303-19.
27. Platt, M. L., and Glimcher, P. W. (1997). Responses of intraparietal neurons to saccadic targets and visual distractors. *Journal of Neurophysiology* **78**: 1574-89.

28. Rajkowska, G., and Goldman-Rakic, P. S. (1995). Cytoarchitectonic definition of prefrontal areas in the normal human cortex: I. Remapping of areas 9 and 46 using quantitative criteria. *Cerebral Cortex* **5**: 307-22.
29. Rajkowska, G., and Goldman-Rakic, P. S. (1995). Cytoarchitectonic definition of prefrontal areas in the normal human cortex: II. Variability in locations of areas 9 and 46 and relationship to the Talairach Coordinate System. *Cerebral Cortex* **5**: 323-37.
30. Ramachandran, V. S., Stewart, M., and Rogers Ramachandran, D. C. (1992). Perceptual correlates of massive cortical reorganization. *Neuroreport* **3**: 583-6.
31. Rosenberg, D. R., and Lewis, D. A. (1995). Postnatal maturation of the dopaminergic innervation of monkey prefrontal and motor cortices: A tyrosine hydroxylase immunohistochemical analysis. *Journal of Comparative Neurology* **358**: 383-400.
32. Schellberg, D., Besthorn, C., Klos, T., and Gasser, T. (1990). EEG power and coherence while male adults watch emotional video films. *International Journal of Psychophysiology* **9**: 279-91.
33. Schlag Rey, M., Schlag, J., and Dassonville, P. (1992). How the frontal eye field can impose a saccade goal on superior colliculus neurons. *Journal of Neurophysiology* **67**: 1003-5.
34. Taylor, K. E. (2000). A continuum approach to modelling cognitive disorders. *Medical Hypotheses* **54**: 642-644.
35. Taylor, K. E., and Stein, J. F. (1999). Attention, intention and salience in the posterior parietal cortex. *Neurocomputing* **26-27**: 901-910.
36. Ungerleider, L. C., and Mishkin, M. (1982). Two cortical visual systems. Pages 549-586 in D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield, Eds., *Analysis of Visual Behaviour*. MIT Press, Cambridge, MA.
37. Vogt, B. A., and Gabriel, M. (1993). *Neurobiology of Cingulate Cortex and Limbic Thalamus*. Birkhauser, Boston.
38. Weinberger, D. R. (1993). A connectionist approach to the prefrontal cortex. *Journal of Neuropsychiatry and Clinical Neurosciences* **5**: 241-53.

39. Weiskrantz, L., Barbur, J. L., and Sahraie, A. (1995). Parameters affecting conscious versus unconscious visual discrimination with damage to the visual cortex (V1). *Proceedings of the National Academy of Sciences USA* **92**: 6122-6.
40. Woo, T. U., Pucak, M. L., Kye, C. H., Matus, C. V., and Lewis, D. A. (1997). Peripubertal refinement of the intrinsic and associational circuitry in monkey prefrontal cortex. *Neuroscience* **80**: 1149-58.
41. Wright, J. J. (1997). EEG simulation: variation of spectral envelope, pulse synchrony and approximately 40 Hz oscillation. *Biological Cybernetics* **76**: 181-94.
42. Yang, T. T., Gallen, C., Schwartz, B., Bloom, F. E., Ramachandran, V. S., and Cobb, S. (1994). Sensory maps in the human brain. *Nature* **368**: 592-3.
43. Zohary, E., Celebrini, S., Britten, K. H., and Newsome, W. T. (1994). Neuronal plasticity that underlies improvement in perceptual performance. *Science* **263**: 1289-92.