Ollscoil na Gaillimhe
University of Galway

Consciousness of brain

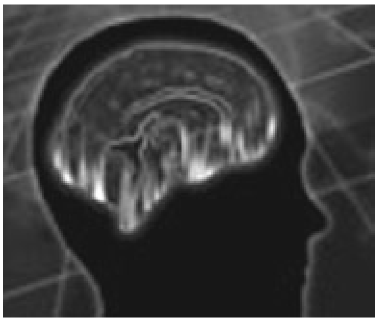
Title	Consciousness of brain
Author(s)	Hogan, Michael J.
Publication Date	2006
Publisher	Psychological Society of Ireland

ARTICLE

Consciousness of brain

Michael J. Hogan

Dr. Mike Hogan is a Lecturer in Psychology at the National University of Ireland, Galway. Correspondence regarding this article may be addressed to him by email to michael.hogan @nuigalway.ie.



> "(Feeling) *...takes the form of a rapturous amazement at the harmony of natural law, which reveals an intelligence of such superiority that, compared with it, all the systematic thinking and acting of human beings is an utterly insignificant reflection...This joy is the feeling from which true scientific research draws its spiritual sustenance, but which also seems to find expression in the song of birds.*"
>
> Einstein, *The world as I see it*

You plunge headlong into the cold Atlantic waves: beneath, surrounded, transformed. After 30 seconds you surface, at one with the ocean, swimming on with strength and gushing silence.

Evolved consciousness makes possible a rich variety of experience, transmissible from mind to mind via poetical and logical symbolism. Scientific experience involves mapping the energy of the system onto a correspondent mapping and manipulation of energy by the mind. The strength and gushing silence, the feeling of connection and oneness experienced by the swimmer is communicable via logical symbolism. But unlike the gestalt expressions of the poet, careful *analysis* precedes any *synthesis* in science, and in order to understand how the mind works in this context, it is necessary to study the brain. The source of both divine and earthly experience is the brain.

But is it really possible to achieve a scientific understanding of human behaviour and subjective experience in terms of neurons and other detailed physiological structures? If you have faith in science, the answer must be yes, but, as Andrew Coward argues, detailed understanding of cognitive and behavioural neuroscience will involve using hierarchies of description which make it possible to describe causal relationships within a phenomenon on many different levels of detail[1]. If God is unity, then the Devil is in the details. When considering the path from neurons to consciousness, we embrace the challenge of thinking about how the intricately designed, detailed parts of an evolved living system can produce wonderful experiences like the perception of controlled, integrated action, swimming in the sea.

The one and the many

Throughout evolution and lifespan development biology maintains a simple principle: the developing organism is a unified whole, a coming and going between the different parts. Even with evolution forcing greater differentiation and specialization, the dynamics of *interchange* keeps the organism together as an organised unit - a unified *doing* of the moment, undisturbed, until replaced by the next unified *doing* of the next moment. Here we can speak of a dynamic system, where layers and layers of detailed biology undergird the simplest of human actions as they emerge over time. In executing our standing, for example, thousands upon thousands of nerve-fibres and muscle-fibres co-act, each neurochemically and neuroelectrically controlled, and our awareness tells us little about how it is this happens.

And actions can be understood from any number of *different* time scales. The infant works here-and-now to maintain stability while holding on to daddy's leg; the young gymnast works for months to master a summersault on the balance beam; the active retiree battles for years to master and re-master their movement in the face of slow sensorimotor and musculoskeletal decline, placing more and more attention into the act of postural control while walking on wet autumn leaves.

As such, life may *be one*, and your life experiences may produce feelings of *being one*, but thinking about the unified action of a living system and the many time scales wherein action manifests can be exceedingly difficult.

Physical scientists experience the same practical constraints when describing phenomenon in the concrete world. Take the following general law of chemical engineering: any alkali plus any acid will produce a salt plus water. The categories "alkali" and "acid", and the law describing how they interact can be described using atomic theory, that is, by reference to less than 100 different types of atoms, and by means of the intermediate objects like hydroxyl ions and hydrogen ions that generate higher level chemical laws. Quantum mechanics, using a yet smaller set of objects (electrons, proton, and neutron) can generate the concepts of atoms, hydroxyl and hydrogen ions and the causal relations between them. String theory describes all quantum mechanical phenomena by reference to different states of vibration of a single string object, and even derives parameters which are ad hoc in all other theories, such as the number of observed dimensions to space.

Psychological scientists gravitate toward different levels of description, often called models, which span the realms of biological, behavioural, and social sciences. Ultimately, a deep understanding of the evolution of living systems – and the dynamic laws governing their structure, process, and function – implies that none of these realms are separable *per se* (Bertalanffy, 1968). Living systems are hierarchically ordered, self-organising systems, and the hierarchy of order observed is itself a product of evolution (Kauffman, 1993; Pettersson, 1996). Fortunately, in terms of our ability to manage conceptual and abstracted systems (Miller, 1978), evolution now grants human beings the opportunity to trace a meaningful path along the great web of interdependence from molecules to mind (De Duve, 2002).

In the hierarchy of order from non-living to living systems the objects and causal relationships at higher levels can be defined as combinations of more detailed objects and causal relationships. Naturally, being pragmatic, most scientists work to describe functional relations at one level or another, and the appropriate objects and causal relationships on any one level are strongly influenced by the parameters which define the phenomenological realm to be studied. Notably, for all models except perhaps the deepest it is clear that models are ways to organise phenomena rather than rules which the universe must follow.

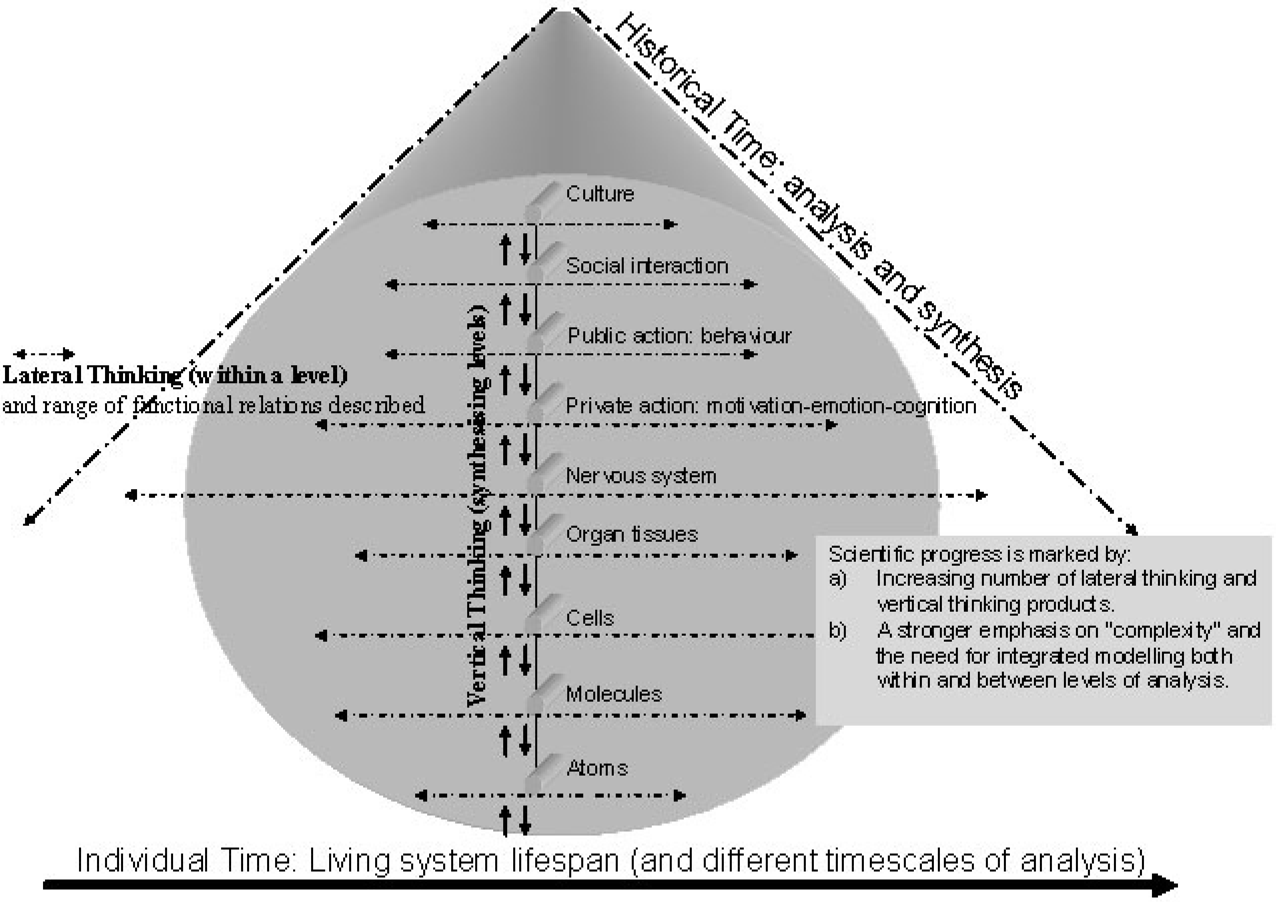
Nonetheless, in our example above, it makes sense for chemists to use higher-level descriptions, even when more detailed levels of description are available (or deemed *necessary)* for more complete/ comprehensive/complex understanding (Berry, Rice, & Ross, 2000). In reality, no theory is complete, and our criteria that good theory be both 'parsimonious' and 'comprehensive' is a self-designed paradox.

Complexity can always be increased by inclusion of more aspects of the system, but this isn't always useful. Many psychologists recognise this fact. For example, radical behaviourists (Chiesa, 1994), whose primary interest is the description, prediction, and control of behaviour, ignore detailed neural explanations, chiefly because they cannot directly control neural dynamics.

As such, detailed levels of description may not be necessary for reliable prediction and control. If you are to know how a car works, a description of Newton's laws and the laws of thermodynamics are less useful than a description of interactions between pistons, cylinders, and sparking plugs.

At the same time, it is impossible to quell humankind's joyous pursuit for more 'complete understanding', and it is inevitable that some people will attempt to cross the borders that make fuzzy and incomplete our map of functional relations. Psychologists and neuroscientists work together in this nebulous terrain: they must do if we are to understand *how the brain works the mind.*

Working to integrate levels of description relevant for a complete understanding of human psychology involves describing functional relations that cross levels of description in the hierarchy of living systems; it involves working *upward* and



Scientific progress is marked by:
a) Increasing number of lateral thinking and vertical thinking products.
b) A stronger emphasis on "complexity" and the need for integrated modelling both within and between levels of analysis.

Figure 1: Levels of analysis in human systems and two styles of thinking necessary for "comprehensive" understanding

downward in the hierarchy of ordered relations, describing ways in which observations at a lower level of description account for (or supplements) description of phenomena at a higher level, and how levels interact – bottom-up and top-down. In this sense - and converse to De Bono's meaning (De Bono, 1990) - *vertical* thinking and modelling works to *synthesise levels* of description, whereas lateral thinking works to *model relations within a level* of description. Inevitably, efforts to combine vertical and lateral thinking will produce more complex models (see figure 1). And notably, in describing how developing brains manage increasingly complex relational systems, Commons (Commons & Richards, 2002; Commons & White, 2003) has argued that the history of science provides clear evidence that many of the greatest innovations emerged after thinkers worked for years to synthesise different levels of description, thus constructing new systems of thought[2].

Describing brain-behaviour relations

> "*The key requirements for an adequate theory of cognition are firstly a consistent hierarchy of causal descriptions on many levels of detail, including both physiological and psychological, and a capability to bootstrap cognitive capabilities in such a way that learning a complex combination of cognitive capabilities can occur with limited resources and minimal interference between early and later learning. [No theory]…adequately addresses these requirements.*" (Coward, 2005)

Most folk psychological accounts of the way evolution has shaped brain structure, process, and function offer a misleading account of the causes of human behaviour. For example, although the essential micro- and macro-structures of the brain describe a 'modular' system, the hierarchy of modules that defines and detects conditions within the information available to the system - current raw sensory inputs, information indirectly activated by conditions detected within those inputs, and information about the current state of the system itself - does not imply that modules *correspond with* cognitive circumstances like categories (Pinker, 1997). Evolutionary psychology is still dominated by 'computational' models of mind that neglect the brain (Panksepp & Panksepp, 2000). It is a mistake to believe something is biologically real simply because one can computationally simulate the shadow of an end result. Importantly, because of the requirement to limit the use of resources, groups of modules that detect conditions in response to instances of different categories may overlap. 'User manual' models of the mind, which propose that evolution added new software modules to compute processes that *correspond* with different cognitive features, implies models that are very resource intensive and physiologically unrealistic. Andrew Coward's model of brain-behaviour relations makes explicit the practical considerations that constrain the design of the brain's hardware system. Some of these relations are captured in figure 2 and explained further below by reference to the numbering system used.

A system architecture approach to the brain

Human behaviour is flexible and varied. The need to *perform many behavioural features*[3] *with limited resources* (*1*) means that within the population of conditions relevant to performing the features, groups of similar conditions must be defined. For example, condition recording in the visual system involves detection of visual features, visual objects, groups of objects, and groups of groups. A module with customised physical resources is needed to detect conditions in each group. Groups of modules with some lesser degree of similarity between their conditions can share some resources and therefore form a higher level module, and so on. The optimal design solution is a *modular hierarchy of condition detection*, and a resource-friendly system is one that limits duplication of conditions between modules as much as possible.

Further, with the evolved drive to construct many copies of the system, it is necessary to *limit construction process specification* (*2*), so that many modules can be built from the same specification. Related is the need to *minimise construction complexity* (*3*) and to *optimise potential for recovery* (by limiting information exchange between modules), so that errors, failures or damage do not have catastrophic effects (*4*). Also, if requisite behavioural flexibility is to be achieved, the system must be designed such that a feature *change can occur without excessive undesirable side effects* on other features, and this implies protecting different meanings of the same information (*5*).

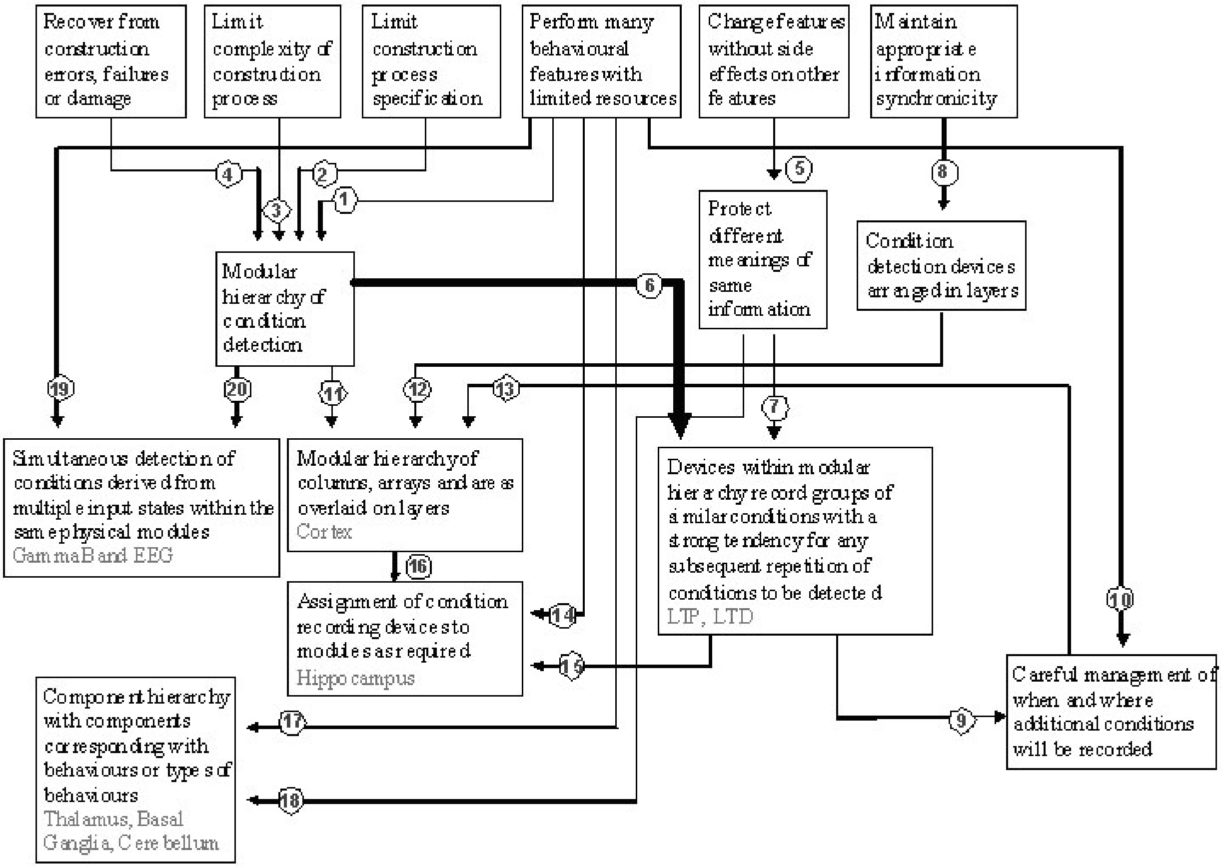


Figure 2: Practical considerations that constrain the design of the brain's hardware system

In physical terms, Coward argues that pyramidal neurons in the cortex correspond with devices that define and detect conditions. Pyramidal neurons are organised into layers, with columns of neurons penetrating several layers. These columns are organised into arrays in which the same input space is available to all the columns in the array. Such arrays of columns will discriminate between input states with different types of behavioural implications, although no one column will correspond with one type of input state. Successive arrays detect conditions of increasing complexity, where the complexity of a condition is the total number of raw sensory inputs that contribute to it, either directly or indirectly via intermediate conditions. Conditions with different levels of complexity will be effective for discriminating between different types of cognitive situations. Relatively simple conditions will be effective for discriminating between cognitive features, somewhat more complex conditions for discriminating between cognitive objects, yet more complex conditions for discriminating between different groups of objects, and so on.

The brain can only function if we assume that the same condition detection can have different behavioural meanings, depending on the context where information processing occurs. If all meanings are unambiguous (i.e. behavioural commands as in computers), resource usage is economical but feature changes are very difficult except with external intellectual control over design. The difficulty arises because without external intellectual control, the primary available guidance for changes to conditions is consequence feedback. At any point in time, consequence feedback is obtained in response to one behaviour, or in other words one meaning of the condition. If the detection of a particular condition has multiple unambiguous behavioural meanings, and the definition of the condition is changed by consequence feedback following just one of the behaviours, the side effects on other behaviours will be severe and uncontrollable. However, if condition detections have partially ambiguous behavioural meanings, module outputs indicating the presence of conditions can only be interpreted as behavioural *recommendations*. Multiple recommendations will be generated in response to an overall input condition in most cases and any changes to one condition definition will have limited, manageable behavioural side effects. Therefore, in Coward's architecture, a subsystem separate from the modular hierarchy is required which can *select* behaviour.

Coward distinguishes between *clustering* - the modular hierarchy within the cortex detecting conditions - and *competition* – the subcortical systems that receive behavioural recommendations and select between alternative behavioural options. To make behavioural selections the competition subsystem must have access to information in addition to the overall input state, such as the consequences of behaviours under similar conditions in the past. Learning can only occur either by gradual and limited changes to conditions within clustering and/or by consequence feedback changing recommendation weights in competition.

Again, the existence of the modular hierarchy means that condition detections will be shared between different features (6), and the need to protect information meanings means that changes to condition definitions must be very conservative, with careful management of when and where conditions can be recorded (*9, 10*). These considerations ultimately limit device learning algorithms (*7*). Long term depression (LTD) and long term potentiation (LTP) are appropriate mechanisms to support the required conservative, managed change algorithms. Detecting conditions in layers means that the memory-processing overhead required for computers can be avoided (*8*).

The combination of the need for a specific type of modular hierarchy (*11*), the layering of devices (*12*), and the need to manage condition recording (*13*) as well as externally directed behaviours shapes the detailed form of the modular hierarchy. A primary means for identifying these modules is greater internal connectivity relative to external connectivity (Tanaka, 1993; Van Essen & Anderson, 1995).

The need to limit resources (*14*) combined with the need for conservative device change algorithms (*15*) and the form of the modular hierarchy (*16*) requires some mechanism for resource assignment. The proposed mapping algorithm is a mechanism that appears consistent with the functions of the hippocampus. Damage to the system managing assignment of resources to clustering results in loss of ability to record new conditions, and therefore the ability to create new declarative memories, as is the case in Korsakov's Syndrome.

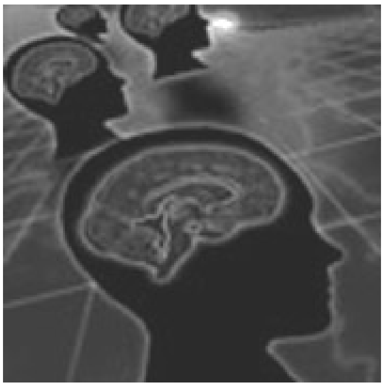
As noted, one module detects a group of conditions relevant to many different behaviours (17), and the information derived from many modules must be used to determine the currently appropriate behaviour. For learning, consequence feedback following behaviours must contribute. However, use of such feedback within the modular hierarchy would damage the multiple meanings of the same information (18). There must therefore be a separate subsystem (i.e., *competition*) that translates condition detections (from *clustering*) into behaviours but does not change condition definitions. Because consequence feedback is applied, competition cannot support exchanges of information with complex behavioural meanings. The components within the subsystem must therefore correspond with individual behaviours or types of behaviour and an information input to one component means either the component behaviour type is encouraged or it is discouraged. To minimize connectivity resources, separate subsystems manage a) selection of sensory information subsets to be released to the modular hierarchy (i.e., attention; thalamus), b) selection of the general type of behaviour (basal ganglia) and c) selection of the specific behaviour within a previously selected type (cerebellum). Because behaviour selection is a selection of an information subset, it is most resource economical for the behavioural selection subsystems to act through the attention subsystem.

The requirement that almost all sensory information must pass through the thalamus before reaching the cortex, combined with the strong projections from cortex to thalamus is consistent with the attention functions of the thalamus. The basal ganglia provides the crucial physiological link between the idea of movement and the motor expression of that idea, and the observation that cerebellar lesions prevent precise coordination of movement is consistent with its role in selecting detailed combinations of movements from alternatives provided by the cortex (Kingsley, 2000).

Some behavioural features will result from the presence of multiple different objects. The same resources must sometimes be used to detect conditions within multiple objects at the same time (19). The organisation into a modular hierarchy means that at least some modules must detect conditions within different objects without confusing information derived from different objects (20). A mechanism is therefore required to support this simultaneous, separate processing. The use of frequency modulation of spike outputs from devices is a relatively simple mechanism which can support this requirement (Shadlen & Movshon, 1999).

The conservative condition change algorithms mean that additional information is available in response to perception of an object over and above the conditions currently detected within the sensory experience of the object. This additional information could include other conditions often active in the past at the same time as the currently detected conditions, and conditions that were recorded in the past at the same time as the currently detected conditions. Coward's book describes how a cognitive model within the architectural constraints can use these additional sources of information to provide an account for semantic, episodic, and priming memory and the observed dissociations between them. For example, words chosen for connection with a specific experience will activate a population of conditions on the basis of simultaneous past activity. The resultant population can activate a secondary population on the basis of simultaneous past condition recording, and if the original words are well selected, the indirect population converges on this basis towards a population that approximates to the one active during the original experience. This population can then drive verbal behaviours to describe the content of the experience, in other words, an episodic memory.

Because indirect activations may interfere with activations recommending other immediately required behaviours, such indirect activations must be recommendations which often compete with alternative recommendations such as motor or speech behaviours for implementation. Generation of complex indirect activations will therefore depend upon consistency with other required behaviours at a given point in time.



Being One

A behavioural neuroscience account might well be able to explain the perceived complexity, uniqueness, and verbalization difficulty of some introspective experiences (or qualia), even though a detailed account of exact content might be impractical.

For example, in the case of the swimmer above, there are conditions which are activated in response to the direct sensory experience of the swimmer: the feel of the water on the body, the splashing in the face, the sounds of the waves, etc. These conditions have recommendation strengths in favour of indirect activation of other conditions including conditions recorded or often active in the past during swimming experiences, while watching other swimmers, while thinking about swimming, or during various moods while swimming. There may be enough total recommendation strength to indirectly activate some of these conditions. The indirectly activated conditions may themselves have recommendation strengths in favour of indirect activation of yet other conditions, such as conditions recorded when watching and listening to the ocean. Many of the directly or indirectly activated conditions have recommendation strengths in favour of different moods (because those moods have occurred shortly after the conditions were active in the past).

Hence the direct sensory experience of swimming results in activation of a population of conditions which includes fragments of many past sensory and internal experiences. None of the fragments may be large enough to constitute a memory of a specific past experience, and it will therefore be difficult to verbalise a description of the content of the population, but the population may generate an averaged mood across many such past experiences. Sometimes, one subset of the population may grow at the expense of the others and generate a memory of the corresponding event.

Condition stability plus the ability to activate conditions not just on the basis of presence in current sensory experience but also on the basis of past temporally correlated activity thus leads to complex mental states in response to sensory experience.

Not very poetical, but interesting nonetheless!

Footnotes

---

[1] Coward, A., *A System Architecture Approach to the Brain*. 2005, New York: Nova.

[2] Systems are constructed when a thinker coordinates more than one variable as input. Metasystem construction involves the synthesis of disparate systems. Some thinkers go further and fit metasystems together to form new paradigms. For example, Maxwell's construction of *electromagnetic fields* from two metasystems: electricity and magnetism; Darwin coordinated paleontology, geology, biology, and ecology to form the field of evolution which, in its turn, paved the way for chaos theory, evolutionary biology, and evolutionary psychology. None of this implies that new metasystems that are the offshoot of new paradigms – for example, evolutionary psychology as an offshoot of evolution – will produce correct explanations for specific phenomena. A paradigm can only frame a specific question. To answer it directly involves re-presentation of meta-systems, systems, and variables. When it comes to answering specific questions, evolutionary thinkers present to the observer a minefield of conflicting views (Laland & Brown, 2002).

[3] A feature is a consistent way in which a system responds to a set of similar environmental circumstances. The environmental circumstances and corresponding responses are similar from the point of view of an external observer, and are a useful way for such an observer to understand the system, but may not reflect similarities in the way the system detects conditions in the environment and generates behaviours on a more detailed level.

References

Berry, R. S., Rice, S. A., & Ross, J. (2000). *Physical chemistry* (2nd / ed.). New York: Oxford University Press.

Bertalanffy, L. v. (1968). *General system theory : foundations, development, applications*. New York: Braziller.

Chiesa, M. (1994). *Radical behaviorism: the philosophy and the science.*

Boston: Authors Cooperative.
Commons, M. L., & Richards, F. A. (2002). Organizing components into combinations: How stage transition works. *Journal of Adult Development, 9*(3), 159-177.
Commons, M. L., & White, M. S. (2003). A complete theory of tests for a theory of mind must consider hierarchical complexity and stage. *Behavioral and Brain Sciences, 26*(5), 606.
De Bono, E. (1990). *Lateral thinking : a textbook of creativity*. London ; New York: Penguin Books.
De Duve, C. (2002). *Life evolving : molecules, mind, and meaning*. Oxford: Oxford University Press.
Kauffman, S. A. (1993). *The origins of order : self-organization and selection in evolution*. New York ; Oxford: Oxford University Press.
Kingsley, R. E. (2000). *Concise text of neuroscience*. Baltimore, MA: Lippincott, Williams, and Wilkins.
Laland, K. N., & Brown, G. R. (2002). *Sense and nonsense : evolutionary perspectives on human behaviour*. Oxford ; New York: Oxford University Press.
Miller, J. G. (1978). *Living systems*. New York: McGraw-Hill.
Panksepp, J., & Panksepp, J. B. (2000). The seven sins of evolutionary psychology. *Evolution and Cognition, 6*(2), 108 - 131.
Pettersson, M. (1996). *Complexity and evolution*. New York: Cambridge University Press.
Pinker, S. (1997). *How the mind works*. New York: Norton.
Shadlen, M. N., & Movshon, J. A. (1999). Synchrony unbound: A critical evaluation of the temporal binding hypothesis. *Neuron, 24*(1), 67-77.
Tanaka, K. (1993). Neuronal mechanisms of object recognition. *Science, 262*, 685-688.
Van Essen, D. C., & Anderson, C. H. (1995). Information processing strategies and pathways in the primate visual system. In *An introduction to Neural and Electronic Networks*: Academic Press.