A model of the hierarchy of behaviour, cognition, and consciousness

Frederick Toates [*]

The Open University, Biological Sciences, Walton Hall, Milton Keynes MK7 6AA, UK

Abstract

Processes comparable in important respects to those underlying human conscious and non-conscious processing can be identified in a range of species and it is argued that these reflect evolutionary precursors of the human processes. A distinction is drawn between two types of processing: (1) stimulus-based and (2) higher-order. For 'higher-order,' in humans the operations of processing are themselves associated with conscious awareness. Conscious awareness sets the context for stimulus-based processing and its end-point is accessible to conscious awareness. However, the mechanics of the translation between stimulus and response proceeds without conscious control. The paper argues that higher-order processing is an evolutionary addition to stimulus-based processing. The model's value is shown for gaining insight into a range of phenomena and their link with consciousness. These include brain damage, learning, memory, development, vision, emotion, motor control, reasoning, the voluntary versus involuntary debate, and mental disorder.
© 2005 Elsevier Inc. All rights reserved.

Keywords: Consciousness; Automaticity; Behavioural hierarchy; Cognition; Determinism; Development; Goal; Habit; Modal action pattern; Motivation; Reflex; Will; Hippocampus; Prefrontal cortex

[*] Fax: +44 1908 654167.
E-mail address: F.Toates@Open.ac.uk.

1. Introduction

This paper: (a) examines some of the characteristics of processes, which, in humans, are described as either 'conscious' or 'unconscious,' (b) links this to their biological embodiment, and (c) speculates on their evolutionary history. To achieve this, it theorizes about the processes' evolutionary origins: it will show where certain 'precursor processes,' which are evident in a range of species, exhibit a similar distinction in performance to human non-conscious and conscious processes. That is, it will argue for evolutionary continuity. The paper will also look at interactions between conscious and non-conscious processing, in the context of the similar processes possessed by a range of species. As a framework, a model of behaviour and cognition (Toates, 1995, 1998, 2000, 2001, 2002a, 2002b, 2004, 2005) will be extended to the study of consciousness. It is suggested that researchers could benefit from a greater appreciation of design solutions employed more widely across species.

The paper will investigate how a system of multiple layers of interacting control can describe features of behaviour, cognition, and consciousness. As a start, it assumes that the simplest and evolutionarily oldest solutions to producing adaptive behaviour involve neither sophisticated cognition nor consciousness (cf. Reber, 1997). The paper then suggests what are the strengths and weaknesses of such relatively simple 'stimulus-based' solutions. It identifies the adaptive value of the evolutionary emergence of additional and *higher-order* processes. Such processes are commonly seen to involve *representations* of the environment, as implied by such terms as 'cognitive map' and 'expectation.' There is controversy on whether the term 'representation' really is needed to account for behaviour (Clark, 1998, 1999; Van Gelder, 1998). Therefore, the paper will largely avoid this issue by simply using the term 'higher order.' The higher-order processes are defined by their flexibility and also by exclusion, as those that are evolutionarily more recent and which do not mediate behaviour in a direct stimulus-based way. However, the fact that terms such as expectancy, error, model, disparity, and cognitive map will be used in various places implies a belief in something resembling a traditional representation, even though the strength of the model does not rest or fall with the validity of this notion.

Of course, the question of animal awareness ('phenomenal consciousness') is not a new one and we do not know whether non-humans experience phenomenal consciousness. Others have speculated about animal awareness on the basis of behaviour, brain mechanisms and evolutionary considerations (Dawkins, 1990; Griffin, 1981). However, the present paper adopts a somewhat different perspective from these authors by pursuing a contrast between processes that appear to map onto the human conscious/unconscious divide. We know that many species exhibit higher-order cognition of a kind that would be open to conscious introspection in humans, e.g., goal-directedness, flexible use of memory, extrapolation to beyond sensory stimulation, and resisting habitual behaviour. We can also identify the brain regions underlying these cognitive capacities. It is argued that a study of the bases of such information processing in non-humans is relevant to how, in humans, features of consciousness are associated with similar cognition and underlying brain regions.

These ideas represent a form of hierarchical control and this notion will provide a framework throughout. The hierarchical model conforms to evidence suggesting that new structures emerge in evolution by building onto existing structures, so-called 'tinkering' (Jacob, 1977; Rozin, 1976a). More recently evolved brain processes co-exist with, and exert modulation over, evolutionarily

older processes (Panksepp, 1998; Rinn, 1984). Reber (1997) argued that consciousness is an evolutionary stage that builds upon older unconscious cognition.

As a first approximation, the model will be based on the following distinction between processes. There are *stimulus-based* controls of behaviour, which are assumed to be evolutionarily old and appear to correspond to Reber's term 'unconscious cognition.' They provide rather direct prescriptions for behaviour triggered by raw stimulus input. Where memory is involved in their expression, such on-line processes exploit what is termed implicit or procedural memory (Squire, 1986).

There are also more evolutionarily recent *higher-order* controls that correspond in some ways to Reber's term 'conscious cognition.' They exploit what in humans is termed explicit or declarative memory. The higher-order controls appear to be an evolutionary development that builds upon, and co-exists with, older more mechanistic stimulus-based controls. Certain 'design requirements' associated with stimulus-based and higher-order control will be identified. It will be shown where higher-order information processing is intrinsically associated with human consciousness.

It will be argued that in the case of humans (and possibly some other species) conscious processing sets the context in which the stimulus-based process operates and the *product* of this processing can gain access to consciousness. For example, we set goals that facilitate stimulus-based processing that is compatible with the goal. However, we do not have conscious access to any of the stages of such processing. Under some conditions, conscious controls can also operate to direct behaviour even though this goes against the tendency arising from stimuli.

In discussions of consciousness, there is frequently a parallel consideration of the voluntary versus involuntary nature of different behavioural controls (Baars, 1988). In humans, this is sometimes framed in terms of 'intentional' versus 'automatic' controls and, in more philosophical discourse, as free-will versus determinism. Though this topic is somewhat beyond our brief, the ideas expressed here are relevant to it and so it will be addressed later. The paper will not concern itself with the 'hard' problem of how *phenomenal* consciousness arises (Chalmers, 1996). Rather, it will consider only a range of 'easy' problems concerning the kind of information processing and control of action that is described as either conscious or unconscious.

Of course, the reader might feel that it is almost tautological to claim that animals are moved by either stimuli or higher-order controls or both, for what else could there be? However, the article is designed to show *how* these controls interact, the functional significance of their interaction, and where this can provide a new source of insight and synthesis relevant to consciousness.

The paper discusses first the likely design considerations underlying the evolutionary emergence of stimulus-based and higher-order controls of behaviour.

2. Stimulus-based control: Behaviour as a reaction to stimuli

2.1. External stimuli

Some features of behaviour (e.g., behaviour described by the terms 'reflex' and 'modal action pattern') depend upon the fundamental design principle that organisms are constructed so as to *react to* stimuli (Barlow, 1977; Braitenberg, 1984; Gallistel, 1980; Goodale & Milner, 2003, p. 40; Rinn, 1984; Schneirla, 1959; Sherrington, 1948; Tulving, 1985). Survival requires that

certain key physical stimuli such as nutrients trigger approach, whereas escape is triggered from excesses of, amongst other things, heat, dryness, and acidity. Presumably, such stimulus-bound triggering characterizes much, if not all, of the behaviour of single-celled animals. With the emergence of relatively simple nervous systems came specialized sensory systems that detect a range of stimuli and specialized muscular systems that provide corresponding reactions (see Fig. 1A). One imagines that to 'design' a relatively simple animal, considerable (if not the entire) weight of control can be placed upon stimulus-based processes.

A range of such automatic reactions to particular biologically significant stimuli is evident throughout the phylogenetic scale, e.g., in humans as in coughing and sneezing, where clear trigger stimuli are evident and we are able to exert rather little in the way of voluntary control (Rinn, 1984).

2.2. Motor-related stimuli

For species with a nervous system, 'stimuli' can be taken to include not only those that arise in the outside world but also those that arise internally from motor effects. The most obvious of these is proprioceptive feedback. Lashley (1951) argued that skills, such as playing the piano, can be executed so rapidly that there is not time for proprioception from one response to act as a trigger for the subsequent one. However, as Willingham (1998) pointed out, this may well be true of skilled performance but need not apply during the skill acquisition stage when proprioceptive stimuli could be exploited. Neither, presumably, need it apply to slowly executed sequences of responses (Richman, 1989).
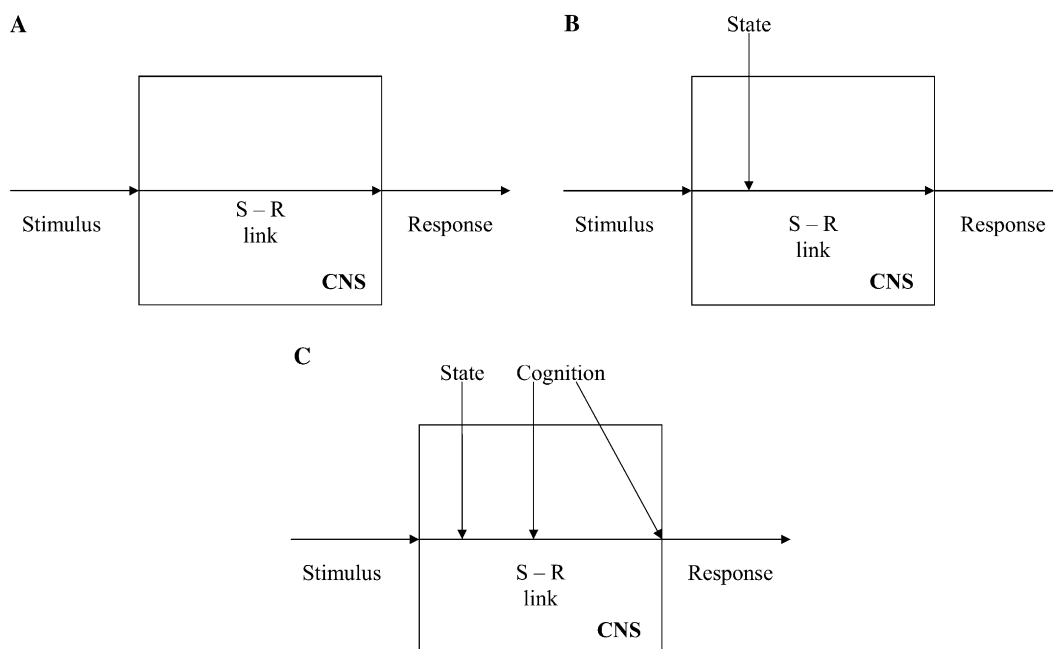

Fig. 1. Modes of control. (A) direct or stimulus response, (B) addition of modulation by 'state,' and (C) inclusion of higher-order control ('cognition').

Based as it is upon negative feedback, proprioception might act as a kind of scaffolding on which faster feedforward controls could build. Thus, in addition to proprioceptive stimuli, there is another source of input to motor control. For repeated sequences, e.g., R_1 followed by R_2 and then R_3, the efferent trigger for response R_1 comes to acquire some capacity to trigger response R_2 and so on. That is, within the CNS, response sequences can be formed, i.e., 'chunking' (Graybiel, 1998; Willingham, 1998).

2.3. Modulation of stimulus-based control

In evolutionary terms, the first appearance of such a *stimulus–response* (S–R) prescription for behaviour is perhaps inevitable. However, one can speculate that there was adaptive pressure to evolve the capacity to *modulate* the strength of such control, according to various circumstances. For example, a reflex does not trigger merely an *invariant* response to a given stimulus (Goldstein, 1995; Prochazka, Clarac, Loeb, Rothwell, & Wolpaw, 2000; Sherrington, 1948). Rather, the reaction depends also upon internal factors, i.e., the S–R link is modulated top-down. That is to say, an animal shows reaction R_1 to stimulus S_1 when condition C_1 is present but reacts with R_2 to the same stimulus S_1 when condition C_2 is present. For example, nutrients will be preferentially approached at times of energy depletion represented as 'state' modulating the strength of the S–R link (see Fig. 1B). For another example, in the startle reflex, there exists modulation by a descending pathway such that the magnitude of reaction depends upon central emotional processing (Davis, 1992; Lang, Bradley, & Cuthbert, 1990). In nociception (detection of tissue damage), although textbooks commonly show a simple pathway linking stimulus and response (e.g., Guyton, 1991), in reality it is not that simple (Glimcher, 2003, p. 146). The reaction of the arm to a nociceptive stimulus (e.g., pin-prick) on the top of the hand is diametrically opposite to that shown to the identical stimulus when the orientation of the hand is changed through 180°. Clearly, there is modulation in the reflex pathway such that different muscles can be activated by a given stimulus.

2.4. Advantages and disadvantages

There are adaptive advantages of stimulus-based control. There is simplicity of design: behaviour can be caused in a fairly straightforward way by physical stimuli. By reacting to stimuli, as in a reflex, a cost-effective process, a 'prescription,' arises as a 'pre-programmed' part of the design (Mayr, 1974). Since few synapses are between input and output, there is a high speed of reaction. Predictable features of the environment can be exploited, while modulation by context takes account of some fluctuating circumstances. This combines a high speed of response with a limited degree of flexible adjustment by controls outside the reflex itself.

There are, however, limitations on this mode of control. There is a limit to the number of viable trigger stimuli that represent biologically important regularities in the world and can be linked adaptively and predictably to particular responses. By definition, the solution of 'reaction to' is 'stimulus bound' and therefore rather inflexible. Even if a very wide range of appropriately adaptive trigger stimuli could be found, a design involving a proliferation of dedicated links would be cumbersome and problems of competition between triggers could become formidable.

Direct control does not permit spontaneity or autonomy in behaviour, i.e., action not tied to trigger stimuli. Similarly, under novel environmental circumstances, a novel behaviour would need to be synthesized but this would be beyond the system's capability. Although some flexibility is permitted by means of conditioning, this still needs stimuli to trigger behaviour and, except for some evolutionary specializations, conditioning that is attached to an S–R process cannot permit rapid change.

These advantages and disadvantages suggest the adaptive value of retaining a certain amount of straightforward stimulus-based control but evolving a higher-order process that is complementary to it and interactive with it. There would be an adaptive advantage for the higher-order processes to be able to: (a) provide solutions in situations where stimulus-based control is unable to do so, e.g., under novel circumstances or where stimulus-based control has failed, (b) allow extrapolation based upon fragmented stimulus input ('best guesses'), and (c) determine single goal-directed action by sensitizing certain S–R links and applying inhibition upon S–R links that run counter to the goal ('executive function'). Such properties are provided by higher-order control, as described next. However, the evolution of higher-order control is not at the expense of abandoning stimulus-based control. Rather, the final design reflects a combination of both processes. It will be argued that consideration of the properties of this composite design is fundamental to understanding certain features of consciousness.

3. Higher-order control

3.1. Bases

The term 'higher-order' refers to a system that is disengaged relative to the direct stimulus–response links. It permits such things as cognitive maps to be exploited. In either rudimentary or more sophisticated forms, such control is evident in both vertebrates and certain invertebrates, e.g., the digger wasp (Braitenberg, 1984; Gallistel, 1980). At least some version of higher-order control therefore appears to have been invented more than once in evolution. Since, by definition, within this mode of control, behaviour is not simply triggered by stimuli as such, two design considerations follow: (1) at a certain level of complexity, *goals* arise as part of the higher-order computation and (2) these goals need to have a means to link to action. Such exploitation of higher-order control and associated goals gives a degree of autonomy from direct stimulus-based control (Kaplan, 1987). By definition, higher-order controls do not tell the animal how to act. One is reminded of the cynic who suggested that the cognitive rat is left buried in thought in the maze! Theorists need to account for how higher-order controls link with muscles and, in the tradition of hierarchical explanation, this has been done (Gallistel, 1980; Wickelgren, 1979).

It is suggested that the higher-order control co-exists with the older stimulus-driven control and (a) exerts modulatory control over it and (b) can act even in the absence of stimuli or in competition with them for the control of behaviour. The evidence points to such joint control (see Fig. 1C). Although higher levels can modulate and even override lower levels, under some conditions a stimulus can take control of behaviour in a way that is at odds with the goal (e.g., following the sudden appearance of a powerful trigger stimulus).

3.2. Advantages and disadvantages

An advantage of higher-order control is that the animal can extrapolate beyond current sensory input, i.e., behaviour can be based on predictions. It allows some freedom from the 'tyranny of bondage to the present' (Edelman, 1989). With some autonomy from fixed links comes increased flexibility of behaviour, i.e., different means can be exploited to reach a given goal. Behaviour can be quickly modified in response to changed circumstances.

A disadvantage is that behaviour production via higher-order control can be relatively slow and costly in terms of its specialized serial processing requirements (Wagner, 1978). Under conditions in which the weight of control can reside with the more rapid parallel stimulus-based process, it is probably advantageous for this to happen. A comparable set of strengths and weaknesses is well recognized when the special case of unconscious versus conscious control is described (Baars, Fehling, LaPolla, & McGovern, 1997).

As has already been noted, reflexes mediate stimulus-based control. A condition under which weight shifts to stimulus-based control is described shortly: following extensive repetition of learned behaviour where a fixed solution has proved viable under constant conditions.

4. A dichotomy of processes but not of control

As a first approximation, two distinct types of process have just been described. The evidence (to be reviewed in this and later sections), from humans and non-humans, points to dual control of behaviour by stimulus-based and higher-order controls acting in concert. The model suggests that, in the competition for behavioural control, stimuli exert an influence in proportion to the modulated strength of the signal that they trigger in the CNS. In this way, a powerful stimulus can occasionally capture control in a way that is at odds with current intentions.

4.1. Non-humans

A principal source of insight has been provided by the study of learning (described in more detail shortly), where the notion of at least two layers of control in rats has long been recognized. Indeed, one of psychology's best known controversies concerned what rats learn: cognitions or stimulus–response associations (Hull, 1952; Tolman, 1932). A consensus has emerged that both types of learning can occur in parallel, reflecting adjustments to two different layers of control (Hirsh, 1974; Mishkin, Malamut, & Bachevalier, 1984; Packard & Knowlton, 2002; White & McDonald, 2002). The animal moves the weight of control between the layers according to various circumstances described later. Such modulation between layers of control suggests that behaviour is under joint control, as illustrated in Fig. 1C.

4.2. Humans

In words that can be interpreted as a special case of 'stimulus-based' and 'higher-order,' respectively, there is recognition that human behaviour can never be exclusively determined by either automatic or controlled processes (cf. Hommel, 2000; Monsell & Driver, 2000). Fig. 1C applies

here too. There is some kind of central decision making process that selects goals and biases the weight attached to particular higher-order controls and stimuli in terms of their candidature for controlling behaviour (Norman & Shallice, 1986).

Thus, the model fits theories that see intentions as modulating ('sculpting') S–R links (Rafal, Machado, Ro, & Ingle, 2000) and deny that S–R translation is invariably subordinate to intentions (Hommel, 2000; Monsell & Driver, 2000). Such a continuum, rather than absolute dichotomous responsibility, is reflected in theories of human information processing (Neumann, 1984).

5. The model

5.1. Background

The model fits a tradition of 'hierarchical control,' which describes different layers of control. Hierarchical theories and models of behaviour have a distinguished history in psychology (Freud, 1969; Gallistel, 1980; MacLean, 1990; Panksepp, 1998; Powers, 1973; Roitblat, 1991; Schore, 2001), ethology (Baerends, 1976; Tinbergen, 1969), and the neurosciences (Hughlings Jackson, see Rinn, 1984; Taylor, 1958). Some address the phenomenon of consciousness (Johnson-Laird, 1988; Norman & Shallice, 1986; Panksepp, 1998; Reber, 1997; Rozin, 1976a; Schneider & Pimm-Smith, 1997; Shallice, 1972). The notion that there exists a global workspace that broadcasts conscious information widely (Baars, 1988) appears to be in this tradition.

The paper will build on, integrate, and develop some of the ideas that these models express and do so in terms of the author's earlier model of hierarchical control, which only very briefly referred to consciousness. What is new in the paper is its use of hierarchical principles in order to present a framework for the integration of substantial areas of the scientific literature that normally remain distinct. This is based upon a comparative perspective involving both humans and non-humans.

5.2. Details of the model

Fig. 2 shows the details of the proposed model. Stimuli such as $S_1, S_2, S_3, \ldots$ impinge upon the sense organs. If sufficiently intense, novel, salient or modulated by attention, they enter conscious awareness as part of, in Edelman's terms, primary consciousness (Dehaene & Naccache, 2001; Edelman, 1989). See Box A. If not, their processing remains at a non-conscious level and they can contribute to the context for conscious processing (Baars, 1997). As stimulus-based control, a subset of stimuli such as $S_1, S_2, \ldots$ form links with responses $R_1, R_2, \ldots$ represented within Box C. The strength of one such linkage is represented as S_1–R_1. Responses provide proprioceptive feedback, labelled as PROP. There are 'external consequences of behaviour' that change the environment, indicated here by link 1 affecting S_3.

Consider the repetition in time of a sequence of stimuli $S_1, S_2, \ldots$ and performance of a corresponding sequence of responses (e.g., S_2–R_2 followed by S_3–R_3). Under these conditions, there can be chunking between responses (Graybiel, 1998), so that, for example, the trigger (S_2) to (R_2) comes to acquire some capacity to trigger R_3. This is labelled as MO, for motor output.

Box A, marked 'higher-order,' summarizes the role of expectations, goals, and high-level emotional processing in: (a) modulating the strength of S–R links (arrow COG_1) and (b) providing a
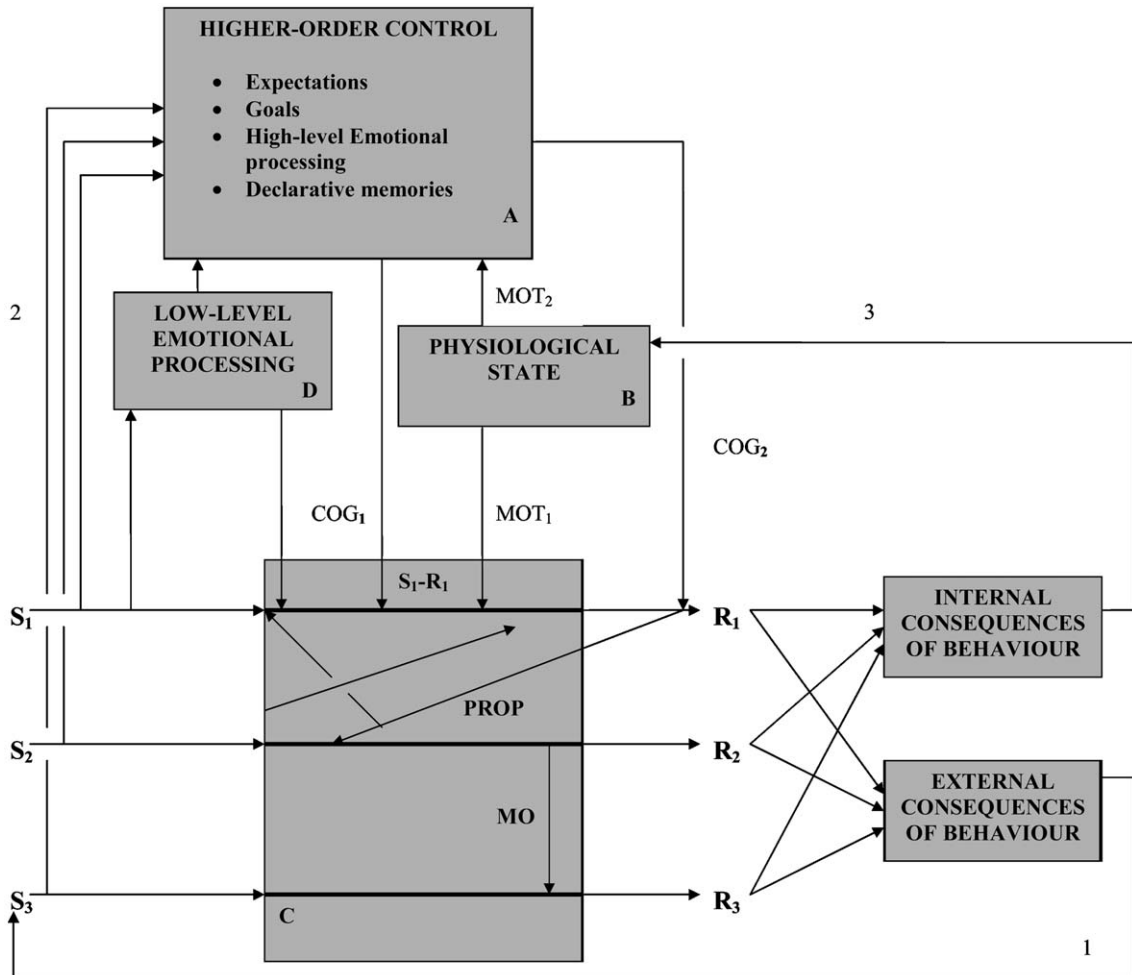
Fig. 2. Proposed system. See text for explanation.

direct control of motor output (arrow COG_2). Box A is informed of events in the world, indicated by the arrows (2) from stimuli. The model includes a goal at the top of a hierarchy of control. This accords with Shallice's (1972) assumption that consciousness is associated with the selector input to an action system.

In humans, Box A would be associated with conscious processing of the following kind. The entry of incoming stimuli into awareness can increase their candidacy for exerting control over behaviour. Memories (some theorists might like 'representations') of events can arise and occupy awareness even in the absence of corresponding stimulus inputs and they would also be part of control exerted from Box A. For example, memories of past events or plans for future events enter awareness and can gain control of behaviour, as described by the term 'higher-order consciousness' (Edelman, 1989).

Box B represents physiological state, in the form of nutrient, fluid or hormone levels. This modulates the strength of S–R links (MOT_1) (Pfaff, 1982) and also influences higher-order processing

(MOT$_2$) (Rolls, 2004). There are 'internal consequences of behaviour,' e.g., restoration of nutrient levels, changes in hormone secretion, summarized as link 3.

Box D represents low-level emotional processing as triggered by particular stimuli in the world such as loud sounds (LeDoux, 1989). This feeds back to modulate the strength of certain S–R links and also passes up to influence higher-order processing.

Though not yet simulated on a computer, the model is dynamic in the sense that cognitive processes mediate real action in real time (Clark, 1998, 1999; Van Gelder, 1998). Also the model conforms to van Gelder's assumption that the system is not passive, i.e., it does not just start when there is a stimulus input to it. Rather, the model is active and 'on-going' with continuous loops of information transfer. Stimulus processing and higher-order control will invariably both be present and the sensitivity of behaviour to either control will probably vary on a continuum.

Using the model as a framework, the next section considers the evidence that the relative sensitivity of behaviour to variations in stimulus-based control and higher-order control will change according to a number of identifiable circumstances.

6. Changes in weight of controls

Stimulus-based and higher-order controls are involved in any instance of behaviour. However, the evidence suggests that their relative weight ('efficacy as a control factor') can change with the following:

(a) *Time within the life-span*. Across various species, a similar shift of weight occurs with age. In dogs learning an instrumental task, there is a shift towards a more automatic ('stimulus–response') mode of control with increased age (Chan et al., 2002). In humans, there is a change in the power of explicit cognition (i.e., 'higher-order') with age but a more constant ability regarding implicit (i.e., 'stimulus-based') cognition (Howard, 1988; Reber, 1992). Evidence points to an inverse-U function, with developmental increases in the efficacy of 'higher-order' control, followed by a decline with old age (West, 1996). Howard (1988, p. 30) reports a particular "decline in the *conscious control* of retrieval" with increasing age. As a general principle, the stimulus-based controls mature earlier than the higher-order (cf. Perner & Clements, 2000). For example, the prefrontal cortex embodies certain higher-order controls (described later) and it matures slowly relative to other structures. As part of early development, the increasing capacity to exert higher-order control is exemplified by an increasing ability: (a) to resist reacting to powerful stimuli and (b) to utilize a memory in the control of behaviour even though a delay is imposed between stimulation and action (Bjorklund & Harnishfeger, 1995). There is an inverse-U relationship between development and the capacity to resist the capture of processing by task-extraneous stimuli (see Hazeltine, Poldrack, & Gabrieli, 2000; Jacoby, Yonelinas, & Jennings, 1997). Individual differences are particularly evident in what is here termed higher-order control (Rothbart & Posner, 2001).

(b) *Learning*. In humans, behaviour that starts out utilizing higher-order controls tends to become more automatic or 'habit-like' with repetition (Lieberman, 2000; Schneider & Shiffrin, 1977). This is also the case when rats learn instrumental tasks (Dickinson, 1985; Rich-

man, 1989). In humans under some circumstances, learning can proceed directly at an implicit level without an initial explicit phase (Van der Kamp, Oudejans, & Savelsbergh, 2003).

(c) *Phylogeny*. Within vertebrates, birds and mammals show evidence of forming expectations, whereas fish and reptiles do not (Bitterman, 1987; Wickelgren, 1979). It will later be argued that such cognitive skill is a higher-order factor. Humans appear to be the species with the most development of higher-order controls. This is evidenced by their flexibility (Rinn, 1984) and capacity to utilize goals in resisting the pull of stimuli (Diamond, 1996; Luria & Homskaya, 1964). Species differences are evident in the ability to bridge a gap in time and act on the basis of sensory stimuli no longer present, with, not surprisingly, humans (including young children) excelling over other species (Tolman, 1932, p.154).

(d) *Brain damage*. Implicit ('stimulus-based') controls are generally less susceptible to disruption than are explicit ('higher-order') controls (Reber, 1997). To some extent, this maps onto the assumption that more recently evolved structures are more vulnerable than older structures, i.e., 'Ribot's Law' (Ribot, 1885; Rozin, 1976a, 1976b). Damage to the prefrontal cortex is associated particularly with a failure of higher-order controls and a capture of behaviour by task-irrelevant stimuli (Dempster & Brainerd, 1995; Luria & Homskaya, 1964). However, there are some double-dissociations. Thus, certain brain lesions disrupt implicit controls, leaving the explicit intact (see Goodale & Milner, 2003; Lieberman, 2000).

(e) *Chemical influences*. For example, the level of prefrontal dopamine activity is a factor in the ability to utilize memories of events in making a response, where the sensory array does not itself provide the cue (Arnsten, 1998; Diamond, 1996), and in overcoming strong response tendencies (Cohen & Servan-Schreiber, 1992). Alcohol is notorious for inducing 'alcohol myopia,' a tendency to give increased weight to physically present cues relative to cognitions that represent temporally distant events (Steele & Joseph, 1990).

The next section will identify some examples of the control of behaviour that are observed across various species and then to relate these to the property of consciousness that accompanies such control in humans.

7. Stimulus/representation and conscious/unconscious control

7.1. Introduction

It is suggested that there is a broad generality of the stimulus/higher-order division of responsibility and, as a first approximation, in the case of humans, this maps onto non-conscious/conscious processing. Rather as a range of species show a shift of weight between stimulus-based and higher-order control as a function of such things as development and learning (Section 6), so humans shift between conscious and non-conscious control (Baars, 1988). As is implied by Fig. 2, any higher-order exertion of control inevitably occurs against a 'background' of stimulus-based controls. Similarly, Baars refers to unconscious processing as providing the *context* for conscious control.

7.2. General

The following are some situations that exemplify the use of higher-order control in various species:

 (i) Overcoming dominant, e.g., habitual, behaviour (Ach, see Hommel, 2000; Willingham, 1998), in the case of humans resisting temptation by inhibiting reactions to intrinsically attractive stimuli (Rothbart & Posner, 2001).
 (ii) Spontaneous generation of behaviour in the absence of obvious triggering by phasic sensory stimulation, i.e., using higher-order representations to guide behaviour as in cognitive maps (cf. O'Keefe & Nadel, 1978).
 (iii) Generating behaviour under novel conditions (Eichenbaum & Cohen, 2001).
 (iv) Altering behaviour rapidly and adaptively after 'tried-and-tested' solutions have failed (Norman, 2002; Ullsperger, Volz, & von Cramon, 2004).
 (v) Sampling or vicarious trial-and-error, i.e., testing 'simulations' of action prior to acting (Tolman, 1932).

From a reverse-engineering perspective, the emergence of higher-order control and thereby the capacity to perform behaviour under the circumstances (i)–(v) above would seem to necessitate the following properties of the nervous system:

1. A capacity for higher-order control to exert inhibition on behavioural tendencies arising from stimuli.
2. Closely related to (1) a capacity to modulate the strengths of candidates for behavioural control, such that, for example, a higher-order control can assume a greater strength than a physically present stimulus.
3. A capacity to set goals, compare actual outcomes of behaviour with desired and expected outcomes, and to alter behaviour according to any discrepancy. Whether this can be achieved in the absence of representations in higher-order control remains a moot point!
4. A feedback consequence in the form of positive and negative affect that arises when, respectively, adaptive and maladaptive courses of action are occurring.

These behavioural situations and design considerations link to the phenomenon of consciousness, discussed next.

7.3. Relation to consciousness

Features (i), (1), and (2) of Section 7.2 can be exemplified by the following situation. Participants wear distorting prisms and adopt the conscious strategy of resisting the 'pull' of the perceived location of the target by reaching at, say, 30° from this location (Willingham, 1998). Patients with damage to the frontal lobe are deficient at this task. Another example is that, where the stakes are particularly high, people tend to resist intuitive processing and go for a more analytical and consciously penetrable mode of reasoning (Lieberman, 2000).

Relating to (i) and (2), conscious processes coordinate the activity of lower-level modules (Baars et al., 1997; Schneider & Pimm-Smith, 1997; Willingham, 1998), as in resolving conflicts

arising from their outputs and resisting strong habits (Dehaene & Naccache, 2001) and temptation (Johnson-Laird, 1988).

In relation to (ii) above, a crucial feature of the processing of information in consciousness is the capacity to hold information in the absence of the corresponding direct stimulation (Dehaene & Naccache, 2001). This is usually seen as a feature of working memory and is mediated in important part by the prefrontal cortex. Relating to (ii) and (iii), Dehaene and Naccache suggest (p.10):

> The ability to combine several mental operations to perform a novel or unusual task is a second type of computation that seems to require consciousness.

In terms of the present model, this is comparable to tasks where an animal must extrapolate to beyond current sensory input in order to control behaviour, as in extrapolation based upon a cognitive map.

Relating to (ii), consciousness is associated with the ability to generate spontaneously intentional behaviour (Dehaene & Naccache, 2001). Blindsight patients exemplify the failure of non-conscious processes to be able to do this.

In relation to (2), links that exert top-down modulation over lower-level controls, map onto assumptions concerning the role of conscious processes in selective tuning of more modular controls (Baars, 1988; Dehaene & Naccache, 2001). As a general cross-species assumption, a central controller must have broad access to the range of S–R links for behaviour to show coordination. One might speculate that, as a development of this, consciousness offers the ultimate in autonomy from stimulus control. It does this by exploiting such a process in order to modulate and coordinate *memories* of stimuli even in their physical absence, i.e., a simulation of behaviour and its outcomes.

With regard to (3), intentionality is closely associated with conscious processes (Baars, 1988; Dehaene & Naccache, 2001). Concerning (iv) and (3), discrepancy monitoring and resolution is an essential feature of the conscious control of behaviour (Eisenberger & Lieberman, 2004). Baars (1988, 1997) emphasizes the importance of novelty and discrepancy for information to gain access to consciousness. A violation of an expectation can rapidly bring conscious control to bear on an otherwise unconsciously mediated behaviour production process. Concerning (iv) and (4), there is the assumption that the affect that colours consciousness arises in important part from a comparison of expected and actual outcomes and is based on the interaction of a number of brain regions including the amygdala, hypothalamus, prefrontal cortex, and anterior cingulate cortex (Dehaene, Kerszberg, & Changeux, 1998; Rolls, 2004). Relating to (3) and (4), the anterior cingulate cortex is associated with discrepancy monitoring and the affective colouring of pain arising from tissue damage or from social loss (Eisenberger & Lieberman, 2004).

8. Attention and layers of control

This section briefly relates the phenomena of attention to the proposed model.

8.1. Basics

Processes of attention are distinct from those underlying sensory and motor processing but interact with them (Posner & Petersen, 1990). The principles underlying both how attention arises

and the role of attentional processes might be better understood in terms of the present framework. That is, attention can be interpreted as a process of modulating the candidacy of particular stimuli and thoughts in their ability to capture the control of behaviour and consciousness.

8.2. Control of attention

Attentional processes are controlled by factors described by various authors as (1) 'active,' 'cognitive' or 'top-down' and (2) 'passive,' 'stimulus-driven' or 'bottom-up' (Dehaene & Naccache, 2001; Neumann, 1984; Rafal & Henik, 1994; Rizzolatti, 1983; Yantis, 1998), which can be partially segregated at the level of neural processes (Corbetta & Shulman, 2002). However, attention is invariably under both sets of control (Yantis, 2000). For example, certain powerful stimuli can gain the capture of attention at the expense of top-down factors.

8.3. Role of attention

As with the *determinants* of attention, there is not a sharp dichotomy in the *role* of attention. Rather, higher-order factors can modulate the sensitivity of the stimulus-based processes such that particular stimuli are more or less likely to engage attention ('perceptual set') (Corbetta & Shulman, 2002). That is, attention modulates the power of particular stimuli to control behaviour and to enter conscious awareness (Dehaene & Naccache, 2001; Posner & Petersen, 1990). This corresponds to increasing the amplification of a particular stimulus input. Thus, it could be argued that attention emerged in evolution precisely to adjust the weight between different types of candidate for behavioural control.

8.4. Brain processes

Posner and Rothbart (1991) describe three attentional networks: posterior, anterior, and vigilance networks.

The vigilance network exerts a bias in favour of the posterior network. In the terms developed here, the posterior network maps onto stimulus-based control. When one orients to a location in space, the location is privileged relative to other locations, a function of the posterior network.

The anterior network, associated with the midprefrontal cortex, is particularly active under conflict conditions as in the Stroop task. It maps onto higher-order control. Anatomical connections between 'anterior' and 'posterior' networks suggest interactions and indeed it is known that orienting can be commanded by means of high-level cognitive strategies.

Allport (1989) recognizes the design constraint that successful goal-directed action involves concerted effort over a period of time and resistance to interruption. In the production of motor output, Neumann (1987, p. 369) suggests that:

> . . .the prevention of physical interference may well be one of the tasks of attention.

However, the environment can suddenly change and therefore a facility for certain types of stimuli to be able to interrupt is important, i.e., to trigger attentional orientation. Allport (p. 654) writes:

This idea—of indirect priority assignment, implemented through the selective modulation (potentiation, tuning, output inhibition, and so on) of units in specific coding pathways—is a recurrent feature of many different theoretical approaches to attention.

8.5. Link with consciousness

8.5.1. General

To some extent, attention correlates positively with consciousness. In attending to something, it normally engages our conscious awareness (Shiffrin, 1997). Schneider and Pimm-Smith (1997) speculate that consciousness:

> ...may be an evolutionary extension of the attentional system...

This would appear to fit the line being advanced here. However, attention and consciousness are not simply synonyms. Thus, a subliminal visual stimulus presented at a particular location in visual space can increase the chances of detecting a subsequent stimulus at that location (Merikle, Smilek, & Eastwood, 2001). Such stimuli associated with a point in space appear to increase the local gain of the system corresponding to the stimulus.

The posterior network is concerned with orienting and, Posner and Rothbart argued, has rather little to do with conscious processing. For example, we experience little awareness of eye movements. By contrast, the anterior network is closely associated with conscious awareness. Although the posterior system is not associated with awareness as such, its activities facilitate the appearance of events in awareness, as mediated via the anterior system.

Baars et al. (1997, p. 428) argue that:

> ...selective attention can be viewed as a set of mechanisms designed to provide access to consciousness...

This suggests that conscious processes already exist and then attention evolves as something serving access to them or that they co-evolve. An alternative line would be to see the evolutionary basis of attention as a process of modulation of sensory information, which does not necessarily apply only to conscious processing but might represent an evolutionary precursor.

8.5.2. The orienting response

In a range of species, the orienting response (OR) consists of directing sense organs to a stimulus and changes in reactivity of the autonomic nervous system (Sokolov, 1976). The animal is described as 'attending to' a stimulus that triggers the OR. The response shows habituation (decreasing magnitude) with repetition of the given stimulus. A change in the stimulus is associated with a recovery of the response. The behaviour can best be understood in terms of the CNS's construction of an internal model of the stimulus (Sokolov, 1976). When incoming stimulation corresponds to the model orienting is not triggered, while orienting is triggered by disparity with the expectation. Baars (1988) considers habituation of the OR in humans to correspond to a shift of weight from conscious to unconscious processing.

8.6. Development

Evidence from attention points to increasing degrees of cognitive ('voluntary,' 'self-regulatory,' 'executive') control with development and thereby an increasing capacity to modulate the effects of stimulus ('reactive,' 'involuntary') control (Rothbart & Posner, 2001). Underlying this, there appears to be an increased weight of the anterior system (Posner & Rothbart, 1991). Early in development, attention is drawn particularly strongly to objects with rich contours such as a chess-board, so-called 'obligatory looking' (Rothbart & Posner, 2001). Posner and Rothbart subscribe to the view of Luria and Vygotsky that the development of a 'higher' form of attention occurs as part of socialization. More will be said on this in the section on vision.

8.7. Conclusion

The evidence reviewed in this section supports the model's assumptions, in that attention is under the control of both stimuli and 'higher-order' processes. Reciprocally, attention serves a role in the modulation of both stimulus-driven reactions and in the weight of higher-order control. There is a developmental shift of weight with age towards the 'higher-order' control. Attentional processes modulate access of information to consciousness and the higher-order aspect of attention is associated with consciousness.

9. Automaticity, habits, and capture

9.1. A move to stimulus-based control

With extensive repetition of flexible behaviour, a behavioural system can change its parameters to acquire some features of a reflex (Dickinson, 1985) or modal action pattern (Barlow, 1977). Behaviour becomes automatic, stereotyped, and under stimulus control, i.e., where an automatic process will suffice the system adjusts to become this. The adjustment points to the adaptive value of exploiting the benefits of speed and simplicity, captured by the term 'automaticity' (Baars, 1988; Schneider & Shiffrin, 1977). Neumann (1984, p. 258) notes:

> ...the course of an automatic process is determined by relatively permanent structural connections, either wired-in or acquired through practice.

For humans, Wilhelm Wundt described this process, the term 'automaticity' referring to (Neumann, 1984, p. 280):

> ...coordination between a sensory event and a motor action. Without automaticity, the two have to be linked by a conscious act... Automatization occurs during practice if there is an invariant relationship between the sensory event and the subsequent motor action. This results in the establishment of direct neural connections between them, so that the sensory event can lead to the proper action without any conscious mediation.

Reber (1992) argues that automatic processes lie "outside of consciousness and conscious control." This is true in part. However, in terms of Figs. 1C and 2 and considering conscious control to be a type of higher-order control, the link does not become entirely functionally severed following a move to automaticity. The probability of engaging automatic processes appears to be modulated by conscious intentions, pointing to joint responsibility and thereby integration between controls. Although the sequence of events is not mediated via conscious processing, it usually only occurs provided that some appropriate conscious intention is present.

9.2. Stimulus capture

The model suggests that a dynamic balance arises between stimulus-based and higher-order determinants. These can reinforce their effects or act antagonistically. Stimuli vary in their power to capture behaviour. If sufficiently powerful, they can do so even in the face of higher-order control to which the stimulus-triggered behaviour is at odds.

A form of capture by potent stimuli appears to be exhibited in classical 'misbehaviour' (Breland & Breland, 1961), termed a 'constraint on learning.' Animals were successfully trained to perform an instrumental task but then behaviour 'degenerated' and reverted to a more species-typical form. Typically, the task required the animal to insert tokens into a slot. After a while, the animal started to act towards the token as if it were food (e.g., by washing it) and thereby failed to perform the instrumental task.

Suppose that an operant contingency is arranged such that the cue of a light stimulus is paired with the presentation of reward. However, in order to earn reward, the response needs to be directed to a location other than that corresponding to the stimulus cue. Species and developmental differences are seen in the ability to train a subject an operant task under these conditions (Tomie, 1996). In humans, there is a developmental emergence of this capacity and humans are better at it than are non-human primates. This phenomenon exemplifies that the weight of control varies as a function of a number of circumstances: the capacity of an animal to overcome powerful stimulus–response links varies as a function of ontogeny and phylogeny. In the Piagetian 'A not B' task, the infant is required to direct behaviour at a location other than that rewarded in the past. Developmental differences are seen in the ability to solve this problem (Diamond, 1996).

Human examples of capture by stimuli are well known in terms of slips and capture errors (Reason, 1984; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997).

9.3. Conclusion

Stimuli vary in their capacity to control behaviour. A higher-order control (in humans, associated with conscious awareness) modulates the strength that they exert but at times stimuli can assume control even when to do so triggers action at odds with the goal set by the higher-order control.

10. Learning and memory

10.1. Basic principles

As indicated at various points already, whether taking the perspective of learning or memory, evidence points to stimulus-based and higher-order layers of control. Further development of this topic is presented in Section 11 where it is related to brain structure.

10.1.1. A memory perspective

A well-established tradition of research, based in part on brain damage, reveals different types of memory, with different susceptibilities to disruption (Claparède, 1911; Scoville & Milner, 1957) and serving different functions (Sherry & Schacter, 1987). Such distinctions have long been at the foundation of theorizing in memory (Ribot, 1885; Schacter & Tulving, 1982). Whether described in terms of 'implicit versus explicit' or 'procedural versus declarative,' it is clear that brain damage can disrupt the one whilst leaving the other intact.

According to Tulving's (1985) definition, procedural memory is a change in what is here termed 'stimulus-based processing.' For example, the memory underlying a motor skill involves changes in the links between sensory events and motor output. The memory tells you 'what to do.' Episodic memory maps onto a higher-order control. Episodic memory holds particular personal experiences, which need involve no link to the subsequent production of behaviour. If exploited in behaviour, the memory cannot tell you what to do but can be applied to a countless array of different reactions.

According to Tulving, semantic memory appears to be representation-based (p. 387):

> It permits the organism to construct mental models of the world...models that can be manipulated and operated on covertly, independently of any overt behaviour.

However, it could be that with repeated use in a given context a semantic memory becomes stimulus-based (cf. Baars, 1988). For example, on learning a foreign language, words are often at first attached to specific instances of experience but, following extensive repetition, their use becomes more automatic. Thus, semantic memory appears to have features of stimulus-based and higher-order control depending upon how it is exploited.

Exemplifying an evolutionary emergence of higher-order controls, Tulving suggests that semantic memory is an evolutionary development of procedural memory, whereas episodic memory is an evolutionary development of semantic memory.

10.1.2. A learning perspective

In the behaviourist era, debate raged on the question of what an animal learns when it is placed in a maze and earns food. Does it, as certain 'hard-nosed' behaviourists argued, learn a direct link (stimulus–response connection) (Hull, 1952)? Alternatively does it, as the 'softer' camp suggested, learn 'higher-order' cognitions that can be exploited flexibly in behaviour (Tolman, 1932)?

Evidence suggests that learning can be encoded at different levels, i.e., stimulus-response and higher-order (Amsel, 1994; Dickinson, 1985; Furedy, Shulhan, & Randall, 1989; Mackintosh, 1985; Wickelgren, 1979). Which level dominates depends upon such things as the context in which the maze is situated and the length of training, etc.

Learning based upon stimulus control would, by definition, be manifest as a direct link to behaviour. That is to say, suppose stimulus S_1 triggers response R_1 and neutral stimulus S_2 is paired with S_1. S_2 would be expected to come to evoke a reaction something like R_1. Such learning does occur.

Classical conditioning exemplifies the added value of higher-order control. By exploiting expectations, classical conditioning gives added flexibility to behaviour (cf. Eichenbaum & Cohen, 2001). This would not be possible if conditioning could only act by stimulus substitution. Thus, the conditional response (CR) to a CS can be very different from the unconditional response (UR) with which it has been associated (Mackintosh, 1974; Tolman, 1932). For example, in rats, the unconditional response to electric shock to the paws is jumping, flinching or attack, whereas the conditional response is freezing. Rather than simply repeating responses, behaviour that is based on expectations can be adaptively fine-tuned according to context.

Evidence points to S–R based and higher-order learning being found in various species but rats exhibit a greater weight of higher-order learning than do fish (Bitterman, 1975, 1987; cf. Schneirla, 1959; Wickelgren, 1979), though this point is controversial (Macphail, 1987). Fish and turtles conform more closely to the predictions of models based upon stimulus–response principles (Pert & Gonzalez, 1974). Extinction of a habit or its reversal in a discrimination task are relatively rapid for mammals but very slow for fish (Gleitman & Rozin, 1971), suggesting that the former but not the latter can exert a higher-order 'override' upon habit learning. In monkeys, there is a decline in such an override capacity with increasing age (Tsuchida, Kubo, & Kojima, 2002). For rats, early in development, stimulus–response processes dominate learning (Amsel, 1994; Schneirla, 1959). Some simple classical and instrumental tasks are more easily learned by decorticated (Goldstein & Oakley, 1985) or hippocampectomized (Hirsh, 1974) animals, suggesting that higher-order processes can intrude upon such learning or its expression.

A feature of the learning that relates to stimulus–response control is its gradual accumulation over numerous instances (Lieberman, 2000; O'Reilly & Norman, 2002; Tulving, 1985). Learning mediated by automatic processes is generally slow and 'unlearning' is also slow (Shiffrin, 1997).

10.2. Link to consciousness

The present line of argument maps well onto the distinction between memory types drawn by Squire, Hamann, and Knowlton (1994, p. 423):

> One system supports the acquisition, storage, and retrieval of conscious knowledge. Other systems support non-conscious learning and its various products: the dispositions, preferences, skills, and habits that are expressible through performance, but accessible to awareness only to the extent that subjects can (epiphenomenally) observe and monitor what they do.

Higher-order memories involve medial temporal lobe structures (e.g., the hippocampus) and prefrontal cortex, and, in humans, are conscious in their expression (Eichenbaum & Cohen, 2001; Overman & Bachevalier, 2001). Episodic memories are open to conscious introspection and have a 'self-knowing' frame of reference (Tulving, 1985). Striatum-based habit memories are expressed in the absence of conscious awareness of their content. Motor skill learning is gradually incremental and takes place largely outside of conscious awareness (Willingham, 1998), although of course the associated high-level goals and the outcome of the action are available to consciousness.

There is controversy concerning whether awareness is necessary for human conditioning and it involves hair-splitting criteria of when participants are really aware or how awareness can be indexed. Dawson and Furedy (1976) suggest that awareness can be necessary for the acquisition and extinction of an autonomic response but not for its maintenance under stable conditions. This would fit the interpretation advanced here, in that learning moves to an automatic mode with repetition, while extinction involves an inhibition on such a link. However, profoundly amnesic patients can still show evidence of classical conditioning even in the absence of being able to articulate any memory for the contingency (Claparède, 1911; Weiskrantz & Warrington, 1979).

For placebo effects, evidence favours parallel processes: conscious expectancies mediated via Pavlovian conditioning or verbal instruction and unconscious ('automatic') learning via a Pavlovian contingency (Stewart-Williams & Podd, 2004; Voudouris, Peck, & Coleman, 1990). The conscious expectancy can sometimes even act to oppose the effect of the unconscious Pavlovian association (Kirsch, 2004). This is evidence for parallel stimulus-based and higher-order controls, the effects of which might add or counter each other.

Interestingly, Tolman was prepared to speculate on the conditions under which a rat is conscious, and this is entirely compatible with the present argument (p. 205):

> ...primarily in moments of changing behaviour, the moments of learning, that consciousness will appear.

This is at times when the rat is "jammed" (anthropomorphically, 'ambivalent') in terms of behavioural options. Tolman continued in a way that is interesting given contemporary metaphors (p. 206):

"*We herewith define conscious awareness as consisting in the performance of a "sampling," or "running-back-and-forth," behaviour.*" The function and use of such sampling or running-back-and-forth behaviour, i.e., conscious awareness, will be to enhance, reinforce, throw a spot light upon, some section or area of an environmental field."

10.3. Conclusion

Learning and memory are embedded in processes that are organized at different levels. There is: (i) a habit-like process and (ii) a higher-order (cognitive) process that mediates learning not tied directly to behaviour. Learning can proceed at both levels in parallel. Learning exemplifies well cross-species generality of the principle of layers of control. In humans, there is conscious access to information processing within the higher-order learning/memory system.

11. Role of brain regions

11.1. Introduction

The role of brain regions in the stimulus/higher-order distinction has already been mentioned briefly in earlier sections. This section brings a focus to the topic. It is suggested that evolution confronted a fundamental design problem in the need to produce brain processes that can adaptively weight the control of behaviour between stimulus-based and higher-order determinants.

The present section looks at some brain regions underlying these roles, associated in humans with conscious processing as part of the higher-order aspect.

Evidence points to the first two regions to be discussed, the hippocampus and regions of the frontal cortex, as underlying higher-order control. Thus, in humans, a large percentage of brain tissue is dedicated to this task. This role can sometimes act to reinforce stimulus-based control but at other times it inhibits behavioural tendencies arising from stimuli. Historically, as a result of evidence from patients with brain damage, these regions have been held to be responsible for a type of cognition associated with consciousness (Stuss & Anderson, 2004) and so they form the bulk of the discussion. We then briefly consider the embodiment of stimulus-based control.

11.2. The hippocampus

11.2.1. Introduction
Controversy on the hippocampus is famous, with it being allocated roles as the basis of: (a) exerting contextual control (Hirsh, 1974), (b) forming a cognitive map (O'Keefe & Nadel, 1978), (c) providing a flexible memory (Eichenbaum & Cohen, 2001), (d) underlying working memory (Olton, Becker, & Handelmann, 1979), (e) providing a transient memory store (Rawlins, 1985), (f) mediating storage in, and retrieval from, long-term memory (Dehaene et al., 1998), (g) forming part of a complementary system acting with the cortex (McClelland, McNaughton, & O'Reilly, 1995; O'Reilly & Norman, 2002; Wickelgren, 1979), and (h) a source of behavioural inhibition (Douglas, 1967; Gray, 1982). It would be somewhat immodest and foolish to pretend that the present model can fully integrate or supersede this diversity of theories! However, it invites a new perspective that meshes with features of each of these accounts and shows areas of substantial overlap between them.

The paper will adopt the idea of *complementary* cortical and hippocampal learning systems (McClelland et al., 1995; O'Reilly & Norman, 2002). Cortical learning is slow, incremental, and associative and is modulated by hippocampal inputs. By contrast, hippocampal learning is rapid and can assimilate single instances of experience. Behaviour production, it is assumed, is influenced by a combination of the learning that has taken place within these two systems. Where there is a history of a clearly defined link between stimuli and consequences, in effect the hippocampus can be circumvented to some extent. Evidence suggests that the hippocampus comes into its own where the problem cannot be solved on the basis of associations formed with the stimulus array alone but requires complex computation and the contextualization of sensory information (Eichenbaum, 1999; Hirsh, 1974; Gluck & Myers, 2001). This can involve generating a novel course of action based upon extrapolation from earlier instances, or overriding the behavioural tendency arising from components of the stimulus array. The model suggests that the hippocampus has a role in higher-order control by biasing the weight of particular cognitions as determinants of behaviour.

11.2.2. Non-humans
Where a solution is possible based on a rather direct trigger from sensory input (e.g., lever → press; maze junction → turn left), rats with hippocampal damage perform perfectly well (Hirsh, 1974). They revert to a pattern of behaviour seen in early development (Altman, Brunner, & Bayer, 1973; Amsel, 1994; Spear & Miller, 1989). In simple conditioning, where a straightforward

S–R link is involved, hippocampally lesioned animals can sometimes even learn faster than intact controls (Douglas, 1967; Schmaltz & Theios, 1972). This evidence suggests that there can be antagonism between controls and hippocampal damage knocks out one control. This fits the argument concerning an inhibitory role for the hippocampus.

At times, expectations and cognitive maps of events in the world are given potency in the control of behaviour, e.g., 'what leads to what expectancies' (Wickelgren, 1979). Lesioned animals are deficient where extrapolation beyond the sensory array is needed.

In one laboratory simulation of natural foraging, the rat must avoid locations recently associated with gaining reward (Olton et al., 1979). This is disrupted by damage to the hippocampus. However, in the same apparatus, following hippocampal damage, there is a slight *enhancement* in a task in which a simple S–R link can provide the solution (McDonald and White, 1993). This indicates a role for the hippocampus both in using working memory and in exerting inhibition on another layer of control.

The behaviour of intact rats earning food reward in an instrumental task depends upon the size of reward *put into the context* of reward on the immediately previous trial(s). Thus, we see frustration and elation effects reflecting this comparison, i.e., the classical *Crespi effect* (Crespi, 1942). Hippocampal damage is associated with a loss of this feature of behaviour (Hirsh, 1974; Wickelgren, 1979).

Thus, rat-based data fits with the reasoning based on humans that the hippocampus encodes memories of particular *episodes of experience* whereas the cortical memory reflects a running tally over repeated trials (O'Reilly & Norman, 2002).

Simple delay conditioning (the CS starts before the UCS and they terminate together) remains intact in experimental animals after damage to the hippocampus. This indicates retention of a capacity to form S–R links with contiguity between CS and UCS. However, trace conditioning (i.e., there is a delay between CS offset and UCS onset) is disrupted by hippocampal damage (Clark & Squire, 1998), unless the damage occurs well after learning (Kim, Clark, & Thompson, 1995).

11.2.3. Human studies

Human amnesics, i.e., patients with damage to the medial temporal lobe involving the hippocampus, show a shift of weight towards stimulus control. They reveal progress on tasks that (Eichenbaum & Cohen, 2001):

> ...are highly structured, involve response strategies already in the subject's repertoire, and in which success can be achieved without reference to any particular event or episode.

Memory is intact provided that it can be cued by a trigger stimulus with an established associated response, e.g., word stem completion, reverse mirror-drawing or highly probable paired-associate learning of the kind 'man–woman' (Cutting, 1978; Warrington & Weiskrantz, 1982). Weiskrantz (1988) reports (p. 186):

> ...it appears that two general categories of tasks yield successful performance: (a) facilitation through priming and (b) acquisition of new stimulus–response and stimulus–stimulus relationships of the kind where the through-put from the stimulus can be made directly and with minimal interference to the associated stimulus or response, i.e., can ultimately be made routine and 'automatic.'

Such acquisition is of a kind described as 'incremental' over trials. Pair-associates can be learned if they are semantically or phonetically related but not if they are random. Memory is disrupted in situations where the patient is literally asked 'to remember something.'

Memory can also be impaired even in a stem completion task. This is seen when the experimenter requests a change in the completion response to the established cue (Warrington & Weiskrantz, 1982). This would involve overriding the existing association with a new one, in the present terms boosting the candidacy of an intrinsically weak association according to the experimenter's instruction. Warrington and Weiskrantz describe "cognitive mediation" to refer to the use of such things as imagery to solve the problem. They suggest that, in the intact brain, there is (p. 242):

> ...a dynamic cognitive mediational memory system, in which memoranda can be manipulated, inter-related, and stored in a continually changing record of events.

In this context, the amnesic patient is deficient when compared to controls.

Delay conditioning can occur in humans even in the absence of conscious insight into the existence of the contingency (Clark & Squire, 1998). However, learning of trace conditioning correlates with the extent to which the participants reveal conscious awareness of the CS–UCS contingency. Clark and Squire noted that, in trace conditioning, the CS–UCS link cannot be mediated in a direct way and they suggested that awareness is a prerequisite for it to occur. The system appears to model the CS and thereby uses such a representation to mediate the conditioned link. As indexed by the CR, amnesics can learn delay but not trace conditioning. They are unaware of the contingency in force. This again points to: (a) a role of conscious processes in exploiting representations in lieu of stimuli and (b) the role of the hippocampus in mediating this.

11.3. Frontal cortex

Regions of frontal cortex can be associated with higher-order control: (a) goal-setting, (b) monitoring outcomes and comparing with expectation, and (c) inhibiting competing tendencies.

A role of the prefrontal cortex (PFC) is to modulate the strength of candidacy of cognitions such that they can wrestle control from stimulus factors (Braver & Barch, 2002). Patients with PFC damage can sometimes identify a deficit in terms of their inappropriate responding but still be unable to correct their behaviour (Stuss & Anderson, 2004). This implies a functioning error detection system but an inadequacy in generating a signal to implement the right cognitive strategy in the face of competition. As part of goal-setting, PFC neurons represent the expected outcomes of actions, e.g., food reward (Rolls, 2004). The reduced capacity to exploit contextual information in the control of behaviour with increasing age is apparently embodied in changes in PFC and its dopaminergic innervation (Braver & Barch, 2002; Dempster & Brainerd, 1995).

PFC is particularly associated with the voluntary goal-setting aspect of human behaviour (Panksepp, 2003; Willingham, 1998). Damage to PFC is followed by inertia and apathy, loss of drive (Stuss & Benson, 1983), a loss of spontaneity in behaviour, loss of goal-directedness (Willingham, 1998), increased distractibility and stereotypy (Spence & Frith, 1999), and decreased inhibition on inappropriate social actions (Damasio & Van Hoesen, 1983). These all amount to decreased higher-order control and thereby a relative increase in weight of stimulus-based control.

Extinction and reversal are disrupted by damage to orbitofrontal division of the PFC. This suggests that normally the region plays a role in overriding the tendency to persist with formerly

reinforced behaviour. Again pointing to the use of a higher-order control, Schoenbaum, Setlow, and Ramus (2003) found evidence that (p. 27) "...many neurons in orbitofrontal cortex maintain representations of the value of the pending outcome across delays imposed between cue sampling and outcome delivery."

11.4. General points on the basis of higher-order control

The hippocampus and prefrontal cortex play a role in strengthening higher-order control in the face of competition from physically present stimuli. Both have a role in the inhibition of competing tendencies. Damage to either disrupts extinction (Douglas, 1967; Schoenbaum & Setlow, 2001). However, there are also differences in their roles. The prefrontal cortex is closely associated with goal-setting and the instigation of behaviour, roles not generally associated with the hippocampus. Patients suffering from hippocampal damage (Baddeley, 1982) do not exhibit the passivity and lack of initiative of frontal patients. From limited samples, it seems that patients with hippocampal damage are disrupted in so far as conscious access to classical conditioning contingencies to which they have recently been exposed, whereas only few frontal patients are so disrupted (Daum, Channon, Polkey, & Gray, 1991).

11.5. Basal ganglia

Damage to dorsal regions of the basal ganglia disrupts the capacity for response-based learning, leaving higher-order controls somewhat intact (Packard & Knowlton, 2002; White & McDonald, 2002). Lieberman (2000) argues that aspects of cognition are organized in a similar way to the control of behaviour. He presents evidence of disruption to implicit cognition with damage to the basal ganglia but preservation of explicit cognition. Lieberman suggests that the learning mediated via the basal ganglia occurs gradually over repeated trials even in the absence of a conscious intention to learn. He argues that skills associated with conscious supervision and which are embodied in the prefrontal cortex and hippocampus transfer with repetition to the striatal region of the basal ganglia.

11.6. The posterior parietal cortex

The posterior parietal cortex plays a role in mediating links between stimuli in the environment and the animal's responses to them. As such, it uses an egocentric representation of the body and exterior world (Willingham, 1998). One can assume that goals set higher in the hierarchy (i.e., allocentrically at prefrontal level) modulate parietal activity such as to vary the chances of functionally appropriate actions in the world.

11.7. The supplementary motor area

The supplementary motor area of the brain, acting together with the basal ganglia, is concerned with organizing response sequences (Willingham, 1998). In terms of the present model, the repeated occurrence of a sequence $R_1, R_2, \ldots$ means that the production of R_1 comes to lower the threshold for the initiation of R_2. Such sequencing operations are carried out outside conscious awareness.

11.8. Conclusion

In various species, different parts of the brain can be allocated responsibility for stimulus-based (e.g., regions of basal ganglia) and higher-order (prefrontal cortex and hippocampus) controls of behaviour. In humans, the processing associated with higher-order controls, within the prefrontal cortex and hippocampus, maps onto conscious processing whereas those regions associated with stimulus-based control perform their tasks at an unconscious level.

12. Motor control

Willingham (1998) presented a hierarchical model of motor control and skill learning. At the top, is a strategic process, involving goal-setting and biologically rooted in the dorsolateral frontal cortex. The performance of this layer is associated with conscious awareness. The high layer uses an allocentric representation of the world and the actor can articulate what (s)he is trying to do to change the world. At lower levels in the hierarchy, e.g., basal ganglia and spinal cord, representations are egocentric and motor control is by processes that are inaccessible to conscious awareness. Normally, the highest level control sets goals and delegates responsibility to the lower levels to put them into effect. However, the highest level can also assume control for moment-by-moment action, as represented by link COG$_2$ in Fig. 2. This occurs not only when the task is novel but also if the person is, for example, showing extreme care and caution in performing a task.

Much of Willingham's argument was concerned with explaining the role of vision in skill learning, and this forms the topic of the next section. The argument to be developed concerning vision is entirely compatible with that of Willingham on skill learning.

13. Vision

13.1. Basic principles

Vision and the control of eye movements show evidence of layers of control with different evolutionary histories but with shared responsibility. Paillard (1987) proposed a layer of control that (p. 43):

> …allows neural processing to step back from the immediate sensory input and to become progressively free from the environmental constraints under which sensorimotor analyzers have to work.

In humans, this maps onto older (non-conscious) stimulus-based controls and evolutionarily newer processing (Goodale & Milner, 2003). Adults are able to direct their eyes towards features of an image that are inherently less attention grabbing, thereby giving weight to higher-order control and inhibiting the pull of the stimulus-driven aspect (Morton & Johnson, 1991).

13.2. The division between cortical and subcortical layers

In their roles, a coarse distinction can be made between cortical and subcortical systems. The cortical system encodes the fine-grained features of the image, especially those that correspond to the fovea. The subcortical system detects peripheral stimuli of high salience and directs eye movements so that the fovea is brought into correspondence with them (Bridgeman, 1992; Bronson, 1974; Morton & Johnson, 1991; Paillard, 1987; Trevarthan, 1968). At first approximation, these parallel processing systems correspond to higher-order and stimulus-based control.

Neural systems of the superior colliculus controlling eye movements in a reflex mode are under descending control from the frontal cortex (Rafal et al., 2000). The anti-saccade task involves the dorsolateral prefrontal cortex inhibiting reflexive triggering via the superior colliculus (Roberts, Hager, & Heron, 1994). In human adults, damage to the frontal eye fields disrupts the capacities to inhibit automatic saccadic eye movements and to initiate eye movements based upon memory and anticipation (Guitton, Buchtel, & Douglas, 1985).

13.3. Division within the cortex

13.3.1. Basics

With the evolution of the cortex, something similar to the division *between* cortical and subcortical is reinvented *within* the cortex. The relatively fast dorsal stream has more responsibility for stimulus-based control of action (using an egocentric frame of reference) and the relatively slower ventral stream has more responsibility for representation-based control (using an allocentric frame) (Norman, 2002; Milner & Goodale, 1995; Van der Kamp et al., 2003).

In humans, activity of a certain level within regions of the ventral stream is associated with conscious awareness (Goodale & Milner, 2003). Changes in conscious perception (e.g., alternation between different interpretations of an ambiguous figure) correlate with changes in activity between parts of this stream. Patients with damage to the stream report only very limited conscious perception of objects but nonetheless interact almost normally with them in terms of motor responses. In general, it appears that conscious awareness is not associated with activity within the dorsal stream (Van der Kamp et al., 2003).

The dorsal stream is triggered by physically present stimuli and when they are extinguished it has very little memory that would enable a response to be made. By contrast, extinguishing target stimuli does not so quickly incapacitate the ventral stream as it can still control responding based upon a memory of stimuli (Norman, 2002).

Debate has raged concerning whether vision is direct or indirect ('constructivist'). The emerging evidence, interpreted in terms of multi-layered control means that both these positions have some validity (Norman, 2002). Like the debate on rat learning (see earlier), what started out as a dichotomy of perspectives in reality maps onto a duality of co-existing processes.

13.3.2. Interaction and allocation of weight

As represented in Figs. 1C and 2, there are interactions between the streams (Rossetti & Revonsuo, 2000). With conscious awareness, goals for action are set up by exploiting information processed within the ventral stream and these are then put into action via the dorsal stream. Norman (2002) argues:

> When the dorsal system is faced with difficulties in picking up the necessary information, due to, say, insufficient information or conflicting information, the ventral stream can be turned to for help…

In starting a task, ventral stream processing underlies control but, with skill acquisition, weight is given to the dorsal stream (Norman, 2002).

Norman notes the problem of "how a switch from dorsal to ventral system control occurs when conditions hamper dorsal stream functioning." The present model suggests that the system allocates weight to the dorsal stream provided that: (a) there are physical features of the world present sufficient to engage its processing and (b) over repeated experiences, the outcome of action is as intended. Thus, when the outcome deviates from the intention, full conscious processing would be brought to bear on the visual interaction with the world.

Similarly, Glover (2004) describes a distinction between *planning* versus *control* of action. The planning system is relatively slow and is engaged prior to an action, whereas the faster control system takes over once action is initiated. A gradual rather than abrupt change of responsibility is assumed. Glover suggests that the planning stage is associated with conscious awareness, whereas, being rapid, the control system operates outside conscious awareness (cf. Pisella et al., 2000). According to Glover, the planning system is responsible for adjusting the parameters of the control system.

13.4. Blindsight

Blindsight patients can locate targets by motor action such as pointing with hand or eyes even though they deny that, corresponding to the 'blind' region of visual field, they see anything (Weiskrantz, 1988). Similarly, given a straightforward mapping between stimulus (a question) and response, e.g., "if it is vertical say vertical" they show evidence of detection, whilst denying seeing anything. This is similar to the intact and damaged capacities of amnesics.

Although able to perform some simple stimulus-based tasks, patients fail to *instigate* behaviour based on the blind field and cannot exhibit flexibility in the use of visual information (Güzeldere, Flanagan, & Hardcastle, 1999; Van Gulick, 1994). As Kinsbourne (1997, p. 343) notes, the vision that does occur:

> …is not related to context. Therefore, it is not voluntarily initiated, but has to await the behest of the experimenter.

Blindsight appears to depend upon links of subcortical systems with the cortical dorsal stream. Evidence from blindsight has been used to argue that phenomenal awareness is necessary for the deliberate and flexible instigation of behaviour (Van Gulick, 1994).

13.5. Development

The visual system exemplifies developmental changes in weight of control. In new-born humans, vision is mainly mediated by an evolutionarily old system that involves the superior colliculi but not the cortex. The visual cortex is relatively undeveloped at birth and acquires control with increasing age (Bronson, 1974). The visual reaction of the new-born is determined largely by stimulus features (Johnson, 1995). Bronson (1974, p. 879) suggests:

...all neonatal reactions can satisfactorily be explained by reference to automatic mecha-
nisms, hence it is not necessary to assume volitional control in the new-born infant.

Development corresponds to the acquisition of a capacity to make eye movements in anticipa-
tion of changes in the stimulus (i.e., based upon memory) rather than simply in response to them.
By 4 months of age, humans can inhibit the automatic saccadic eye movements that otherwise
would have been triggered by current stimuli (Johnson, 1995). They can withhold a saccadic
eye movement in anticipation of a more interesting visual stimulus to follow (Johnson, 1995).

Atkinson (2000, p. 174), in effect, associates the emergence of consciousness with a developmen-
tal transition from stimulus-based to representation-based control:

> ...the changes we see in alertness and visual behaviour in the first 3–4 months look like the
> emergence of an active, self-controlled, information-seeking perceiver. If we can ever accept
> that consciousness is present in a creature who cannot describe it, we would have to accept it
> in an infant of this age.

13.6. Conclusion

Vision exemplifies well multiple but interacting layers of control. Subcortical and cortical dorsal
stream processing correspond to stimulus-driven control whereas ventral stream processing cor-
responds to higher-order control. In humans, the latter is associated with conscious awareness.
Development represents a shift of weight from stimulus-based to higher-order control.

14. Motivation, emotion, and affect

14.1. Motivation, rationality, stimuli, and higher-order controls

Philosophical discourse is concerned with the ability of humans (or lack of it) to make rational
('conscious') decisions. A feature of decision-making appears to be a capacity to weigh the prob-
able negative long-term future outcome of an action against the immediate positive gains by
employing somatic markers (Adolphs, 1999; Damasio, 1996), i.e., a capacity to weigh appropri-
ately stimuli and higher-order determinants is indicated. Evidence points to a disruption of this
capacity following damage to the ventromedial frontal lobe. According to McClure, Laibson,
Loewenstein, and Cohen (2004), a competition for control occurs between prefrontal cortical pro-
cesses that mediate delay of gratification and "lower level, automatic processes" that mediate
immediate gratification. The presence of an incentive can bias the competition in favour of
immediacy.

14.2. Emotion and affect

As represented in Fig. 2, the control of emotion reveals: (a) a hierarchical structure and (b)
modulation between levels. Emotions appear to be both generated within a hierarchical structure
and exert roles at different levels within this structure (Toates, 2002a). As noted earlier, the human
startle reflex is relatively simple and automatic with few synapses in the direct pathway. However,

its magnitude is under descending control from central emotional processing (Lang et al., 1990). In primates, extensive emotional processing occurs within the prefrontal cortex (Rolls, 2004). There exist reciprocal connections between the prefrontal cortex and subcortical structures.

A debate was over whether cognition is necessary for triggering emotion (Lazarus, 1984; Zajonc, 1980). A possible resolution is that emotion is triggered both directly (i.e., bypassing cognition) and indirectly (i.e., involving cognition). There are both direct triggers to activity of the amygdala, a brain region important in the control of emotion (i.e., sensory input → thalamus → amygdala), as well as indirect pathways (i.e., sensory input → thalamus → cortex → amygdala) (LeDoux, 1989). These are sometimes termed respectively the 'quick and dirty' and 'slow and clean' routes.

Evidence points to universal emotions that are shared between humans and non-humans (Panksepp, 1998). However, in humans, there is cultural relativity in the triggers to them and how they are expressed (Ekman, 1992). This suggests a layer of culture-sensitive control that modulates the basic emotional processing.

Facial expressions corresponding to the basic emotions are organized at a subcortical level, are relatively stereotyped, appear early in development and are similar across cultures (Rinn, 1984). They are difficult to produce by conscious intention. However, their expression can be modulated by social learning that takes the form of cortically organized 'display rules.' This social modulation appears developmentally with the maturation of the frontal cortex. This input is available to conscious awareness. Baars (1997) discusses the case of spontaneous versus voluntary ('willed') smiling and their distinct neural bases. The willed kind usually look faked, except in the good movies.

14.3. Pain

In so far as humans are able to muster opposition to the pain arising from nociceptive stimuli, this is reflected in activity within the dorsolateral prefrontal cortex (Lorenz, Minoshima, & Casey, 2003). Participants able to counter the 'pain' of social rejection exhibit elevated activity in the right ventral prefrontal cortex, which correlates negatively with self reports of distress (Eisenberger, Lieberman, & Williams, 2003). Eisenberger et al. suggested that prefrontal cortical activity inhibits that of lower brain regions embodying the pain of social exclusion, i.e., conscious strategies organized at a high level are able to inhibit lower-level controls.

14.4. Conclusion

Motivation, emotion and pain provide evidence of different layers of control. Higher-order control associated with full conscious awareness is to some extent able to modulate, refine and inhibit the activity organized at a lower level.

15. Ontogeny and phylogeny

15.1. The basic idea

A notion passionately advanced and criticized at one time was that ontogeny recapitulates ('recaptures') phylogeny. That is, changes over evolutionary time can be observed in speeded-up form

over development. This idea fell into disfavour, though it then saw a qualified revival (Gould, 1977; cf. Bjorklund & Pellegrini, 2002; Rozin, 1976b). Evidence points to one respect in which the principle is valid: in both ontogeny and phylogeny, there emerges an increasing degree of higher-order control and thereby autonomy from stimulus control. The present model could provide a useful framework for reconsidering the general principle. For example, it has been argued that, in both phylogeny and ontogeny, episodic memory emerges last (Schacter & Tulving, 1982). Also, earliest memories in terms of ontogeny and phylogeny are usually the last to be lost in disease (Rozin, 1976b).

Flexibility appears to reach its peak in humans with the evolution of a large cerebral cortex, particularly its frontal lobe. This gives a greater weight to voluntary controls that involve learning, relative to the more involuntary controls organized at a lower level (Rinn, 1984). Relevant to a consideration of ontogeny and phylogeny and their link with consciousness, Rinn argues (p. 67):

> In general, it may be said that cortically mediated behaviours (e.g., language) are not present in infancy and must be learned. They are generally highly flexible and readily changeable, thus increasing the organism's adaptability to novel situations. They may also show considerable cultural variability. Typically, we have good conscious awareness of these behaviours, and they can easily be produced or inhibited on command. By contrast, behaviours mediated primarily by other motor systems (e.g., sneezing, heart beat, etc.) are generally present very early in development, not substantially influenced by learning, rigidly stereotyped in topography, inflexible, and show little cultural variability. In many cases, we have poor conscious awareness of the behaviour. Generally, they can only be approximated on command, and are difficult to inhibit when they occur spontaneously.

15.2. Conclusion

Both phylogeny and ontogeny reveal the emergence of higher-order controls. They are characterized by flexibility and are associated in humans with conscious processing.

16. Complex human cognition

16.1. Introduction

With the evolution of complex cognition in humans, the solution of joint control by stimulus-based and higher-order controls appears to be conserved. Features of cognition, e.g., reasoning, moral judgements, and attitudes, reveal parallel layers of control (for an application to evolutionary psychology, see Toates, 2005). That cognition reinvents some principles of behavioural control fits with the notion that selection by the basal ganglia applies to both actions and thoughts (Graybiel, 1998; Lieberman, 2000). A theorist closely associated with a dual processing model within cognition is Fodor (1985). Dedicated encapsulated modules are said to provide rapid information to a central cognitive system.

16.2. Reasoning

Reasoning involves dual control (Epstein, 1994; Evans, 2003; Lieberman, 2000; Over, 2003; Sloman, 1996; Stanovich & West, 2000), with different neural processes at the base of each system (Goel & Dolan, 2003; Lieberman, 2000). Systems are dichotomized into an evolutionarily old and intuitive System 1 ('implicit, rapid, parallel, and automatic') and an evolutionarily new System 2 ('explicit, slow, and sequential in nature'; 'low processing capacity'; 'volitional') (Evans, 2003; Stanovich & West, 2000). For System 1, only the product of reasoning becomes available to consciousness whereas, within System 2, we have conscious insight into both the outcome and the means by which we reason (Evans, 2003; Sloman, 1996). Thus, conclusions arrived at by intuition are inaccessible to rational justification on the basis of cognitive insight into underlying processing (Lieberman, 2000). There appear to be wider individual differences in System 2 reasoning than with System 1, e.g., variation with intelligence and development (Evans, 2003). Reasoning is normally the outcome of interactions between the two systems. There can be competition for expression between the systems (Goel & Dolan, 2003) and what starts out being computed in System 2 can, with repetition, move to System 1 (Sloman, 1996).

16.3. Conclusion

Advanced human cognition shows evidence of a joint control that is similar to that underlying behaviour. Something like the higher-order control described for behaviour acts in parallel with a rapid and more intuitive system. The processing within the higher-order control is accessible to consciousness whereas that of the rapid system is not.

17. Pathology of thought and behaviour

Insight can sometimes be gained by observing how brain and behavioural systems can 'go wrong.' Space precludes a detailed discussion but this section attempts to give a flavour of the argument.

17.1. Neurotic disorders

Brewin (1996) suggests that mood depends upon both consciously accessible memories which (p. 38) "...can in principle be deliberately interrogated and retrieved," and implicit memories that are situationally dependent, i.e., triggered only by particular stimuli. Phobics commonly hold simultaneously two contradictory beliefs on danger and safety (Beck, 1976). The respective dominance of the beliefs is situation-dependent, i.e., whether or not the patient is in the presence of the object concerned.

A depressed mood is associated with a greater difficulty in the use of active strategies in memory search but deficiencies of memory are not seen when there is stimulus triggering (Hertel & Hardin, 1990). Beck (1976, p. 237) refers to the experience of depressed subjects in entertaining repetitive negative thoughts:

> The person perceives these thoughts as though they arise by reflex—without any prior reflection or reasoning...

Beck's therapy challenges the validity of negative automatic thoughts, thereby allowing the patient to substitute more rational responses, i.e., there is a change of weight across levels. For obsessive-compulsive disorder, Wegner (1997) distinguishes between (p. 148):

> ...an *intentional operating process* that is conscious, effortful, and interruptible and an *ironic monitoring process* that is unconscious, less effortful, and uninterruptible.

17.2. Psychotic disorders

Schizophrenia can be characterized in part by a chronic shift of weight from higher-order to stimulus-based controls (cf. Cohen & Servan-Schrieber, 1992).

Positive symptoms include being bombarded with excessive stimulation from the outside world and having difficulty organizing cognition (Andreasen, 1997, 1999). Negative symptoms such as apathy and avolition are associated with a dysfunction of the frontal lobes (see Schmajuk, 2001). There is a reduced capacity to exploit higher-order control in the form of working memory (Andreasen, 1997). As its biological basis, there is also evidence for reduced size of the hippocampus, a region interacting with the prefrontal cortex (see Schmajuk, 2001).

17.3. Conclusion

Evidence from psychiatric disorder lends support to multiple layers of control with some competition between them for access to consciousness.

18. Intentions, free-will, and determinism

18.1. Introduction

This section points to the relevance of the model to the philosophical discourse. This includes the link between conscious and non-conscious processing, the role of stimuli in capturing behaviour and a framework for understanding resistance to temptation. It would seem a necessary condition for suggesting free-will that a system is able to anticipate the consequences of its actions, a feature of higher-order control.

Traditionally, the notion of free-will is closely associated with that of consciousness (Bargh & Ferguson, 2000), the usual assumption being that we consciously and freely decide a course of action (Baars, 1988; Claxton, 1999; Willingham, 1998). Dehaene and Naccache (2001, p. 29) suggest:

> One may hypothesize that subjects label an action or a decision as 'voluntary' whenever its onset and realization are controlled by higher-level circuitry and are therefore easily modified or withheld, and as 'automatic' or 'involuntary' if it involves a more direct or hardwired pathway.

The topic has usually been discussed in dichotomies and it lies at the heart of philosophical, ethical, legal, and theological belief and doctrine systems (see Prochazka et al., 2000). Thus, both external causation and unconscious processes are seen to undermine free-will. The *external* triggering of behaviour via reflexes is seen as in opposition to free-will. Descartes dichotomized between external (reflex controls) and the actions of the soul, and, outside behavioural science and philosophy, such a dichotomy has been popular since.

In proposing modulation of a stimulus-based pathway by higher-order processing, including that associated with consciousness, the model heavily qualifies the dichotomy 'will versus reflex' (cf. Hommel, 2000). It suggests continuity rather than a clear dichotomy. As Umiltà (1988) points out, engaging in one activity while 'keeping an eye and ear open' to react to something else involves setting a bias to attention. He continues (p. 338):

> Here we have a stimulus-driven response that is rendered more probable because attentional resources have been intentionally allocated to the relevant cognitive process. In some sense one might say that here the behaviour is not volitional, but it is not unintentional either because it is guided by a conscious goal.

Even something resembling a reflex can have a voluntary element if it accords with the setting of a goal. Similarly, Zhu (2003) suggests that movements that are externally triggered do not necessarily represent pure unwilled reactions. Zhu gives the example of an experiment measuring choice reaction time. The participants are trained to respond to the external stimulus as quickly as possible, i.e., 'reflex like.' However, there is a clear prior intention established in liaison with the researcher at the start of the experiment.

The failures of free-will to link one-to-one to an appropriate action cause some to doubt the validity of the notion. According to the model, any free-will acts together with environmental trigger stimuli and hence there will be a less than perfect association between free-will and action. There is a moral dimension to this. As Rothbart and Posner (2001, p. 361) point out:

> …the programming of effortful control is critical to socialization, which often demands that behaviour proceed in a manner that is counter to reactive tendencies. Indeed, the ability to rise above prepotent emotional drives may underlie human conscientiousness and empathy.

18.2. Wegner's critique

Wegner (2002) argued that our feeling of being in control of our actions is largely illusory. Rather, conscious awareness of an act appears too late to form part of the causal sequence.

Consider that processes embodying conscious will modulate the causal pathway from input to output. This would fit the evidence of Section 13 on vision, where planning is described as higher-order and then action is triggered by stimuli. Much organization occurs prior to the action, as in formulating the intention (Zhu, 2004). The notion of conscious will is applicable here. Wegner suggests that (p. 20) "…the beliefs, desires, and plans that may serve as the scaffolding for the intention need not be in consciousness." Presumably these were once conscious and thereby, it might be argued, they form part of the basis of free-will. A crucial experiment, by Libet is discussed shortly.

18.3. Types of intention

Spence and Frith (1999) associate the dorsolateral prefrontal cortex with the "subjective experience of deciding when to act and which action to perform." They refer to Searle's distinction between:

> …an 'intention in action' (the implicit step that precedes an overtly executed act), and a 'prior intention' to act (the conscious desire to do something).

and suggest that with regard to the posterior parietal cortex (PPC):

> It may be most parsimonious at this stage to regard the 'intentions' localized to some PPC neurons as being components of the former implicit variety.

The 'intention in action' is triggered by an environmental cue such as door handle. Without there being a prior intention, then the cue would not normally trigger this action.

18.4. Does consciousness come too late?

Libet (1985) suggested that a person becomes consciously aware of the urge to perform an action only *after* the brain has already decided to act. This appears to cast doubt on the common-sense notion that we consciously form an intention and then later act as a result of it. However, higher-order processes can *modulate* the probability of responding rather than necessarily instigating it.

In agreeing to participate in Libet's study and perform a certain movement, the participants had already implemented a consciously accessible goal (Näätänen, 1985; Zhu, 2003). The element of uncertainty concerned exactly *when* to act. Presumably, given the intention, something like the position of the arm on the clock that they were asked to monitor or the passage of time per se triggered the response. The notion of agency might require revision but the study hardly reveals that conscious agency is not causally effective. As argued by Glover (2004) (see earlier), consciousness can be associated with the planning stage of action, involving setting parameters of action control. Once set, the system comes under the control of the physical properties of currently present stimuli. In discussing the result of Libet, Wegner does not consider the possibility of awareness serving a modulatory, and thereby causally efficacious (albeit indirect), role.

So do I 'only know what I am going to say after I have heard myself saying it?' No. I can plan in advance my exact words but, even for high-speed speech, usually I consciously know roughly what *I intend to say* even though the exact choice of words is delegated to lower controls. By analogy, I know that I am getting up right now to check an absent neighbour's house security but have no idea as to the exact leg movements that I shall employ.

18.5. The alien hand syndrome

The alien hand syndrome (or 'anarchic hand syndrome') is exhibited by patients with brain damage (e.g., to prefrontal cortex). It consists of one hand 'taking on a life of its own' such as suddenly groping and grasping, at odds with the conscious intention, and perceived as alien by the patient (Bogen, 1993; Leiguarda, Starkstein, Nogués, Berthier, & Arbelaiz, 1993). It appears

that the hand acts according to the *affordances* provided by objects, reflecting their nature and the normal reaction involved in interaction with them (Riddoch, Edwards, Humphreys, West, & Heafield, 1998) but in the absence of high-level inhibition.

Wegner considers the 'alien hand syndrome' to be evidence against the causal efficacy of the will. However, it points to considerable organization of movement at a level below the top of the hierarchy, a level that can gain excitation as a result of stimulus factors. It does not demonstrate the impotence of the will in modulating such processes under normal conditions.

18.6. Utilization behaviour

What is termed 'utilization behaviour' can follow damage to the frontal lobe (Lhermitte, 1983; Shallice, Burgess, Schon, & Baxter, 1989). In inappropriate contexts, patients suddenly grasp objects put before them. This demonstrates the ability of objects that offer 'affordance' to capture the behaviour and points to a pathological disruption of the will rather than an undermining of the notion (Blakemore, Wolpert, & Frith, 2002). Spence and Frith (1999, p. 16) note:

> Such patients might be said to have a major impairment of free-will since they have become slaves to their environment.

18.7. Conclusion

Insight into philosophical issues can be gained by considering the feature of the model that behaviour is determined jointly by stimuli and higher-order controls. Normally, there will be compatibility between the tendencies arising from each control but, under some conditions, the stimulus control can trigger behaviour that is at odds with intentions.

19. Discussion

19.1. General

A range of evidence, human and non-human, points to a dichotomy in *types of control* that is manifest in their *shared responsibility* for behaviour and cognition. There can be a shift of weight between types of control and competition between them. What appears to be an adaptive ('dual') solution for the control of behaviour can sometimes be reinvented by evolution. This is revealed in: (i) vision in the dichotomy *between* subcortical and cortical being reinvented *within* cortical processing and (ii) the cognition underlying human reasoning, which employs a similar design to that underlying behavioural control. In a possibly non-adaptive sense, dual control is revealed in certain mental disorders.

For some reason, not at all clear, higher-order information processing in humans is associated with conscious awareness of the *processing itself*. Stimulus-based control is not necessarily associated with conscious awareness, though the probability of its expression can be biased by conscious factors and its outcome can become accessible to awareness. This is particularly the case, where the outcome departs from an expectation. This general principle is applicable equally

to the control of action and to certain human cognitive skills such as reasoning, where conscious processing sets the context and the *outcome* of stimulus-based processing is made available to conscious awareness. Attributing regions such as the hippocampus, prefrontal cortex, and ventral stream of visual processing with conscious processing does not imply that consciousness is exclusively rooted in higher structures. Rather, it appears that dynamic hierarchical interactions between old and new structures underlie the experience of consciousness (Panksepp, 1998). However, we can identify particular features of conscious awareness (such as that accompanying resisting temptation) with the more recently evolved structures.

Some evidence reviewed, i.e., on trace conditioning (Section 10.1.2) and use of the ventral stream in directing eye movements (Section 13.3.1), points to conscious processing being necessary for extrapolation to beyond the sensory information currently impinging on the sense organs. Similar tasks can be set to non-humans such as to require extrapolation and bridging a gap between stimulus and action. The hippocampus and prefrontal cortex are implicated in their solution.

The architecture that has been described in the paper is rather different from that which emerges by adoption of the computer metaphor beloved of both traditional cognitive psychology and evolutionary psychology. Yet the evidence points to the conclusion that a hierarchical architecture best fits biological reality. The kind of architecture that evolved to extract information can be best viewed in the context of how it is to be exploited in behaviour. Behaviour is inextricably a function of stimulus input, cognition, and available modes of motor control. To study one of these to the exclusion of the others is to risk devising an architecture that is unrealized in the real animal.

Although a dichotomy of types of control has been constructed there exists joint responsibility between them. Weight shifts according to circumstances. By means of the creative aspect of consciousness, we are able to incorporate certain stimulus–response links into our engineered environment (Clark, 1998). From knots tied in handkerchiefs to remind us to do something, to stop signs on the highway, we introduce elements of stimulus–response into our culture. This gives an increased degree of freedom to conscious processes.

19.2. Evolutionary considerations

Taking the perspective of reverse-engineering and speaking metaphorically, evolution confronts a dilemma. How much trust can be put in the regularities of the world? Given certain regularities, how much trust can be put in constructing a nervous system that triggers a prescribed response as the optimal solution for dealing with them? Where both the world and the optimal solution are fairly constant, there can be a built-in prescription. For example, tissue damage is a ubiquitous threat from the environment and escaping from it by exploiting reflexes is usually the adaptive response exhibited, though top-down modulation is also incorporated. As the complementary process and in a more open-ended way, anticipation of possible future damage allows flexibility in taking avoidance action. When, however, something with a capacity to show flexible action is called upon under repeated constant conditions, parameters sometimes shift making this more like a reflex.

From the range of evidence reviewed in this paper, evolution appears to 'hedge its bets' by putting some weight on stimulus-based processes but takes care to monitor their efficacy, modulate their strength, and have fail-safe processes to call on when they fail.

19.3. A question of animal awareness

The paper has not explicitly set out to address the issue of whether non-humans possess conscious awareness comparable to our own. However, indirect evidence on this issue, which was presented throughout the paper, can be cited, such as that the very same brain mechanisms crucial for our state of awareness are broadly represented, especially in mammalian species (Baars, 1997). The evidence presented here might be used to reinforce such a case: the *kind of information processing* that is associated with human consciousness is also carried out by non-human species and damage to particular brain regions (e.g., hippocampus and prefrontal cortex) disrupts the capacity to perform such processing across species.

Acknowledgments

References

Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences, 3*, 469–479.

Allport, A. (1989). Visual attention. In M. Posner (Ed.), *Foundations of cognitive science* (pp. 631–682). Cambridge: The MIT Press.

Altman, J., Brunner, R. L., & Bayer, S. A. (1973). The hippocampus and behavioural maturation. *Behavioural Biology, 8*, 557–596.

Amsel, A. (1994). Précis of ''Frustration Theory: An analysis of dispositional learning and memory''. *Psychonomic Bulletin and Review, 1*, 280–296.

Andreasen, N. C. (1997). Linking mind and brain in the study of mental illnesses: A project for a scientific psychopathology. *Science, 275*, 1586–1592.

Andreasen, N. C. (1999). A unitary model of schizophrenia. *Archives of General Psychiatry, 56*, 781–787.

Arnsten, A. F. T. (1998). Catecholamine modulation of prefrontal cortical cognitive function. *Trends in Cognitive Sciences, 2*, 436–447.

Atkinson, J. (2000). *The developing visual brain*. Oxford: Oxford University Press.

Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.

Baars, B. J. (1997). *In the theatre of consciousness*. New York: Oxford University Press.

Baars, B. J., Fehling, M. R., LaPolla, M., & McGovern, K. (1997). Consciousness creates access: Conscious goal images recruit unconscious action routines, but goal competition serves to ''liberate'' such routines, causing predictable slips. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 423–444). Mahwah, NJ: Erlbaum.

Baddeley, A. (1982). Amnesia: A minimal model and an interpretation. In L. S. Cermak (Ed.), *Human memory and amnesia* (pp. 305–336). Hillsdale: Lawrence Erlbaum.

Baerends, G. P. (1976). The functional organization of behaviour. *Animal Behaviour, 24*, 726–738.

Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviourism: On the automaticity of higher mental processes. *Psychological Bulletin, 126*, 925–945.

Barlow, G. W. (1977). Modal action patterns. In T. A. Sebeok (Ed.), *How animals communicate* (pp. 98–134). Bloomington: Indiana University Press.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.

Bitterman, M. E. (1975). The comparative analysis of learning. *Science, 188*, 699–709.

Bitterman, M. E. (1987). Evidence of divergence in vertebrate learning. *Behavioral and Brain Sciences, 10*, 659–660.

Bjorklund, D. F., & Harnishfeger, K. K. (1995). The evolution of inhibition mechanisms and their role in human cognition and behaviour. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 141–173). San Diego: Academic.

Bjorklund, D. F., & Pellegrini, A. D. (2002). *The origins of human nature*. Washington: American Psychological Association.

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences, 6*, 237–242.

Bogen, J. E. (1993). The callosal syndromes. In K. M. Heilman & E. Valenstein (Eds.) (3rd ed.. *Clinical neuropsychology*, pp. 337–407). New York: Oxford University Press.

Braitenberg, V. (1984). *Vehicles: Experiments in synthetic psychology*. Cambridge: The MIT Press.

Braver, T. S., & Barch, D. M. (2002). A theory of cognitive control, aging cognition, and neuromodulation. *Neuroscience and Biobehavioural Reviews, 26*, 809–817.

Breland, K., & Breland, M. (1961). The misbehaviour of organisms. *American Psychologist, 16*, 681–684.

Brewin, C. R. (1996). Theoretical foundations of cognitive-behaviour therapy for anxiety and depression. *Annual Review of Psychology, 47*, 33–57.

Bridgeman, B. (1992). Conscious vs unconscious processes. *Theory and Psychology, 2*, 73–88.

Bronson, G. (1974). The postnatal growth of visual capacity. *Child Development, 45*, 873–890.

Chan, A. D. F., Nippak, P. M. D., Ikeda-Douglas, C. J., Head, E., Cotman, C. W., Murphey, H., Muggenburg, B., & Milgram, N. W. (2002). Visuospatial impairments in aged canines (Canis familiaris): The role of cognitive-behavioural flexibility. *Behavioural Neuroscience, 116*, 443–454.

Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.

Claparède, M. E. (1911). Récognition et Moité. *Archives de Psychologie, 11*, 79–80.

Clark, A. (1998). Where brain, body and world collide. *Daedalus, 127*, 257–280.

Clark, A. (1999). Visual awareness and visuomotor action. *Journal of Consciousness Studies, 6*, 1–18.

Clark, R. E., & Squire, L. R. (1998). Classical conditioning and brain systems: The role of awareness. *Science, 280*, 77–81.

Claxton, G. (1999). Whodunnit? Unpicking the 'seems' of free will. *Journal of Consciousness Studies, 6*, 99–113.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behaviour and biology in schizophrenia. *Psychological Review, 99*, 45–77.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*, 201–215.

Crespi, L. P. (1942). Quantitative variation of incentive and performance in the white rat. *The American Journal of Psychology, 55*, 467–517.

Cutting, J. (1978). A cognitive approach to Korsakoff's syndrome. *Cortex, 14*, 485–495.

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London B, 351*, 1413–1420.

Damasio, A. R., & Van Hoesen, G. W. (1983). Emotional disturbances associated with focal lesions of the limbic frontal lobe. In K. H. Heilman & P. Satz (Eds.), *Neuropsychology of human emotion* (pp. 85–110). New York: The Guilford Press.

Daum, I., Channon, S., Polkey, C. E., & Gray, J. A. (1991). Classical conditioning after temporal lobe lesions in man: Impairment in conditional discrimination. *Behavioural Neuroscience, 105*, 396–408.

Davis, M. (1992). Analysis of aversive memories using the fear-potentiated startle paradigm. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (pp. 470–484). New York: The Guilford Press.

Dawkins, M. (1990). From an animal's point of view: Motivation, fitness, and animal welfare. *Behavioural and Brain Sciences, 13*, 1–61.

Dawson, M. E., & Furedy, J. J. (1976). The role of awareness in human differential conditioning: The necessary gate hypothesis. *Psychophysiology, 13*, 50–57.

Dehaene, S., Kerszberg, M., & Changeux, J. P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences of the United States of America, 95*, 14529–14534.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition, 79*, 1–37.

Dempster, F. N., & Brainerd, C. J. (1995). *Interference and inhibition in cognition*. New York: Academic Press.

Diamond, A. (1996). Evidence for the importance of dopamine for prefrontal cortex functions early in life. *Philosophical Transactions of the Royal Society of London B, 351*, 1483–1494.

Dickinson, A. (1985). Actions and habits: The development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London B, 308*, 67–78.

Douglas, R. J. (1967). The hippocampus and behaviour. *Psychological Bulletin, 67*, 416–442.

Edelman, G. M. (1989). *The remembered present: A biological theory of consciousness*. New York: Basic Books.

Eichenbaum, H. (1999). The hippocampus and mechanisms of declarative memory. *Behavioural Brain Research, 103*, 123–133.

Eichenbaum, H., & Cohen, N. J. (2001). *From conditioning to conscious recollection: Memory systems of the brain*. Oxford: Oxford University Press.

Eisenberger, N. I., & Lieberman, M. D. (2004). Why rejection hurts: A common neural alarm system for physical and social pain. *Trends in Cognitive Sciences, 8*, 294–300.

Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science, 302*, 290–292.

Ekman, P. (1992). Facial expressions of emotion: New findings, new questions. *Psychological Science, 3*, 34–38.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49*, 709–724.

Evans, J. S. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7*, 454–459.

Fodor, J. A. (1985). Precis of the modularity of mind. *The Behavioural and Brain Sciences, 8*, 1–42.

Freud, S. (1969). *An outline of psycho-analysis*. London: The Hogarth Press.

Furedy, J. J., Shulhan, D., & Randall, D. C. (1989). Human Pavlovian HR decelerative conditioning with negative tilt as US: A review of some S–R, stimulus-substitution evidence. *International Journal of Psychophysiology, 7*, 19–23.

Gallistel, C. R. (1980). *The organization of action—A new synthesis*. Hillsdale: Lawrence Erlbaum.

Gleitman, H., & Rozin, P. (1971). Learning and memory. In W. S. Hoar & D. J. Randall (Eds.). *Fish physiology* (Vol. VI, pp. 191–278). Academic Press: New York.

Glimcher, P. W. (2003). *Decisions, uncertainty, and the brain: The science of neuroeconomics*. Cambridge, MA: The MIT Press.

Glover, S. (2004). Separate visual representations in the planning and control of action. *Behavioural and Brain Sciences, 27*, 3–78.

Gluck, M. A., & Myers, C. E. (2001). *Gateway to memory*. Cambridge: The MIT Press.

Goel, V., & Dolan, R. J. (2003). Explaining modulation of reasoning by belief. *Cognition, 87*, B11–B22.

Goldstein, K. (1995). *The organism: A holistic approach to biology derived from pathological data in man*. New York: Zone Books.

Goldstein, L. H., & Oakley, D. A. (1985). Expected and actual behavioural capacity after diffuse reduction in cerebral cortex: A review and suggestions for rehabilitative techniques with the mentally handicapped and head injured. *British Journal of Clinical Psychology, 24*, 13–24.

Goodale, M. A., & Milner, A. D. (2003). *Sight unseen: An exploration of conscious and unconscious vision*. Oxford: Oxford University Press.

Gould, S. J. (1977). *Ontogeny and phylogeny*. Cambridge: Harvard University Press.

Gray, J. A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. Oxford: Oxford University Press.

Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory, 70*, 119–136.

Griffin, D. R. (1981). *The question of animal awareness*. New York: Rockefeller University Press.

Guitton, D., Buchtel, H. A., & Douglas, R. M. (1985). Frontal lobe lesions in man cause difficulties in suppressing reflexive glances and in generating goal-directed saccades. *Experimental Brain Research, 58*, 455–472.

Guyton, A. C. (1991). *Textbook of medical physiology*. Philadelphia: W.B. Saunders.

Güzeldere, G., Flanagan, O., & Hardcastle, V. G. (1999). The nature and function of consciousness: Lessons from blindsight. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 1277–1284). Cambridge: The MIT Press.

Hazeltine, E., Poldrack, R., & Gabrieli, J. D. E. (2000). Neural activation during response competition. *Journal of Cognitive Neuroscience, 12*(Suppl. 2), 118–129.

Hertel, P. T., & Hardin, T. S. (1990). Remembering with and without awareness in a depressed mood: Evidence of deficits in initiative. *Journal of Experimental Psychology: General, 119*, 45–59.

Hirsh, R. (1974). The hippocampus and contextual retrieval of information from memory: A theory. *Behavioural Biology, 12*, 421–444.

Hommel, B. (2000). The prepared reflex: Automaticity and control in stimulus–response translation. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 247–273). Cambridge, MA: The MIT Press.

Howard, D. V. (1988). Implicit and explicit assessment of cognitive aging. In M. L. Howe & C. J. Brainerd (Eds.), *Cognitive development in adulthood* (pp. 3–37). New York: Springer-Verlag.

Hull, C. L. (1952). *A behaviour system.* New Haven: Yale University Press.

Jacob, F. (1977). Evolution and tinkering. *Science, 196*, 1161–1166.

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. M. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 13–47). Mahwah, NJ: Erlbaum.

Johnson, M. H. (1995). The inhibition of automatic saccades in early infancy. *Developmental Psychobiology, 28*, 281–291.

Johnson-Laird, P. N. (1988). A computational analysis of consciousness. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 357–368). Oxford: Clarendon Press.

Kaplan, S. (1987). Associative learning and the cognitive map: Differences in intelligence as expressions of a common learning mechanism. *Behavioral and Brain Sciences, 10*, 672–673.

Kim, J. J., Clark, R. E., & Thompson, R. F. (1995). Hippocampectomy impairs the memory of recently, but not remotely, acquired trace eyeblink conditioned responses. *Behavioural Neuroscience, 109*, 195–203.

Kinsbourne, M. (1997). What qualifies a representation for a role in consciousness. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 335–355). Mahwah, NJ: Erlbaum.

Kirsch, I. (2004). Conditioning, expectancy, and the placebo effect: Comment on Stewart-Williams and Podd (2004). *Psychological Bulletin, 130*, 341–343.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention and the startle reflex. *Psychological Review, 97*, 377–395.

Lashley, K. S. (1951). The problem of serial order in behaviour. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behaviour* (pp. 112–146). New York: Wiley.

Lazarus, R. S. (1984). On the primacy of cognition. *American Psychologist, 39*, 124–129.

LeDoux, J. E. (1989). Cognitive–emotional interactions in the brain. *Cognition and Emotion, 3*, 267–289.

Leiguarda, R., Starkstein, S., Nogués, M., Berthier, M., & Arbelaiz, R. (1993). Paroxysmal alien hand syndrome. *Journal of Neurology, Neurosurgery, and Psychiatry, 56*, 788–792.

Lhermitte, F. (1983). "Utilization behaviour" and its relation to lesions of the frontal lobes. *Brain, 106*, 237–255.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *The Behavioural and Brain Sciences, 8*, 529–566.

Lieberman, M. D. (2000). Intuition: A social cognitive neuroscience approach. *Psychological Bulletin, 126*, 109–137.

Lorenz, J., Minoshima, S., & Casey, K. L. (2003). Keeping pain out of mind: The role of the dorsolateral prefrontal cortex in pain modulation. *Brain, 126*, 1079–1091.

Luria, A. R., & Homskaya, E. D. (1964). Disturbance in the regulative role of speech with frontal lobe lesions. In J. M. Warren & K. Akert (Eds.), *The frontal granular cortex and behaviour* (pp. 353–371). New York: McGraw-Hill.

Mackintosh, N. (1974). *The psychology of animal learning.* London: Academic Press.

Mackintosh, N. J. (1985). Varieties of conditioning. In N. M. Weinberger, J. L. McGaugh, & G. Lynch (Eds.), *Memory systems of the brain: Animal and human cognitive processes* (pp. 335–350). New York: The Guilford Press.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*, 419–457.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306*, 503–507.

McDonald, R. J., & White, N. M. (1993). A triple dissociation of memory systems: Hippocampus, amygdala, and dorsal striatum. *Behavioral Neuroscience, 107*, 3–22.

MacLean, P. D. (1990). *The triune brain in evolution*. New York: Plenum Press.

Macphail, E. M. (1987). The comparative psychology of intelligence. *Behavioral and Brain Sciences, 10*, 645–695.

Mayr, E. (1974). Behavioural programs and evolutionary strategies. *American Scientist, 62*, 650–659.

Merikle, P. M., Smilek, D., & Eastwood, J. D. (2001). Perception without awareness: Perspectives from cognitive psychology. *Cognition, 79*, 115–134.

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

Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. In G. Lynch, J. L. McGaugh, & N. M. Weinberger (Eds.), *Neurobiology of learning and memory* (pp. 6577). New York: The Guilford Press.

Monsell, S., & Driver, J. (2000). Banishing the control homunculus. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 3–32). Cambridge, MA: The MIT Press.

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review, 98*, 164–181.

Näätänen, R. (1985). Brain physiology and the unconscious initiation of movements. *Behavioural and Brain Sciences, 8*, 549.

Neumann, O. (1984). Automatic processing: A review of recent findings and a plea for an old theory. In W. Prinz & A. F. Sanders (Eds.), *Cognition and motor processes* (pp. 255–293). Berlin: Springer-Verlag.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 361–394). Hillsdale: Lawrence Erlbaum.

Norman, D. A., & Shallice, T. (1986). Attention to action—willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.). *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 1–18). New York: Plenum Press.

Norman, J. (2002). Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological perspectives. *Behavioural and Brain Sciences, 25*, 73–144.

Olton, D. S., Becker, J. T., & Handelmann, G. E. (1979). Hippocampus, space and memory. *The Behavioral and Brain Sciences, 2*, 313–365.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: The Clarendon Press.

O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Sciences, 6*, 505–510.

Over, D. E. (2003). From massive modularity to metarepresentation: The evolution of higher cognition. In D. E. Over (Ed.), *Evolution and the psychology of thinking* (pp. 121–144). Hove: Psychology Press.

Overman, W. H., & Bachevalier, J. (2001). Inferences about the functional development of neural systems in children via the application of animal tests of cognition. In C. A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 109–124). Cambridge: The MIT Press.

Packard, M. G., & Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. In W. M. Cowan, et al. (Eds.), *Annual review of neuroscience, Vol. 25, Annual Reviews*, Palo Alto.

Paillard, J. (1987). Cognitive versus sensorimotor encoding of spatial information. In P. Ellen & C. Thinus-Blanc (Eds.). *Cognitive processes and spatial orientation in animal and man. Vol. II Neurophysiology and developmental aspects* (pp. 43–77). Dordrecht: Martinus Nijhoff.

Panksepp, J. (1998). *Affective neuroscience*. New York: Oxford University Press.

Panksepp, J. (2003). Damasio's error? *Consciousness and Emotion, 4*, 111–134.

Perner, J., & Clements, W. A. (2000). From an implicit to an explicit "theory of mind". In Y. Rossetti & A. Revonsuo (Eds.), *Beyond dissociation* (pp. 273–293). Amsterdam: John Benjamins.

Pert, A., & Gonzalez, R. C. (1974). Behaviour of the turtle (*Chrysemys picta picta*) in simultaneous, successive and behavioural contrast situations. *Journal of Comparative and Physiological Psychology, 87*, 526–538.

Pfaff, D. W. (1982). Neurobiological mechanisms of sexual motivation. In D. W. Pfaff (Ed.), *The physiological mechanisms of motivation* (pp. 287–317). New York: Springer-Verlag.

Pisella, L., Gréa, H., Tilikete, C., Vighetto, A., Desmurget, M., Rode, G., Boisson, D., & Rossetti, Y. (2000). An 'automatic pilot' for the hand in human posterior parietal cortex: Toward reinterpreting optic ataxia. *Nature Neuroscience, 3*, 729–736.

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

Posner, M. I., & Rothbart, M. K. (1991). Attentional mechanisms and conscious experience. In A. D. Milner & M. D. Rugg (Eds.), *The neuropsychology of consciousness* (pp. 91–111). London: Academic Press.

Powers, W. (1973). *Behaviour: The control of perception*. London: Wildwood House.

Prochazka, A., Clarac, F., Loeb, G. E., Rothwell, J. C., & Wolpaw, J. R. (2000). What do reflex and voluntary mean? Modern views on an ancient debate. *Experimental Brain Research, 130*, 417–432.

Rafal, R., & Henik, A. (1994). The neurology of inhibition: Integrating controlled and automatic processes. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 1–51). San Diego: Academic Press.

Rafal, R., Machado, L., Ro, T., & Ingle, H. (2000). Looking forward to looking: Saccade preparation and control of the visual grasp reflex. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 155–174). Cambridge, MA: The MIT Press.

Rawlins, J. N. P. (1985). Associations across time: The hippocampus as a temporary memory store. *The Behavioral and Brain Sciences, 8*, 479–496.

Reason, J. (1984). Lapses of attention in everyday life. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 515–549). Orlando: Academic Press.

Reber, A. S. (1992). The cognitive unconscious: An evolutionary perspective. *Consciousness and Cognition, 1*, 93–133.

Reber, A. S. (1997). How to differentiate implicit and explicit modes of acquisition. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 137–159). Mahwah, NJ: Erlbaum.

Ribot, T. H. (1885). *Diseases of memory*. London: Kegan Paul Trench and Co..

Richman, C. L. (1989). SAB, reward and learning. In W. N. Dember & C. L. Richman (Eds.), *Spontaneous alternation behaviour* (pp. 59–71). New York: Springer-Verlag.

Riddoch, M. J., Edwards, M. G., Humphreys, G. W., West, R., & Heafield, T. (1998). Visual affordances direct action: Neuropsychological evidence from manual interference. *Cognitive Neuropsychology, 15*, 645–683.

Rinn, W. E. (1984). The neuropsychology of facial expression: A review of neurological and psychological mechanisms for producing facial expressions. *Psychological Bulletin, 95*, 52–77.

Rizzolatti, G. (1983). Mechanisms of selective attention in mammals. In J.-P. Ewert, R. R. Capranica, & D. J. Ingle (Eds.), *Advances in Vertebrate Neuroethology* (pp. 261–297). New York: Plenum Press.

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General, 123*, 374–393.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). 'Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia, 35*, 747–758.

Roitblat, H. L. (1991). Cognitive action theory as a control architecture. In J.-A. Meyer & S. W. Wilson (Eds.), *From animals to animats* (pp. 444–450). Cambridge: The MIT Press.

Rolls, E. T. (2004). The functions of the orbitofrontal cortex. *Brain and Cognition, 55*, 11–29.

Rothbart, M. K., & Posner, M. I. (2001). Mechanism and variation in the development of attentional networks. In C. A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 353–363). Cambridge: The MIT Press.

Rossetti, Y., & Revonsuo, A. (2000). *Beyond dissociation*. Amsterdam: John Benjamins.

Rozin, P. (1976a). The evolution of intelligence and access to the cognitive unconscious. In J. M. Sprague & A. N. Epstein (Eds.), *Progress in psychobiology and physiological psychology* (pp. 245–280). New York: Academic Press.

Rozin, P. (1976b). The psychobiological approach to human memory. In M. R. Rosenzweig & E. L. Bennett (Eds.), *Neural mechanisms of learning and memory* (pp. 3–48). Cambridge: The MIT Press.

Schacter, D. L., & Tulving, E. (1982). Amnesia and memory research. In L. S. Cermak (Ed.), *Human memory and amnesia* (pp. 1–32). Hillsdale: Lawrence Erlbaum.

Schmajuk, N. A. (2001). Hippocampal dysfunction in schizophrenia. *Hippocampus, 11*, 599–613.

Schmaltz, L. W., & Theios, J. (1972). Acquisition and extinction of a classically conditioned response in hippocampectomized rabbits (*Oryctolagus cuniculus*). *Journal of Comparative and Physiological Psychology, 79*, 328–333.

Schneider, W., & Pimm-Smith, M. (1997). Consciousness as a message aware control mechanism to modulate cognitive processing. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 65–80). Mahwah, NJ: Erlbaum.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review, 84*, 1–66.

Schneirla, T. C. (1959). An evolutionary and developmental theory of biphasic processes underlying approach and withdrawal. In M. R. Jones (Ed.), *Nebraska symposium on motivation* (pp. 1–42). Lincoln: University of Nebraska Press.

Schoenbaum, G., & Setlow, B. (2001). Integrating orbitofrontal cortex into prefrontal theory: Common processing themes across species and subdivisions. *Learning and Memory, 8*, 134–147.

Schoenbaum, G., Setlow, B., & Ramus, S. J. (2003). A systems approach to orbitofrontal cortex function: Recordings in rat orbitofrontal cortex reveal interactions with different learning systems. *Behavioural Brain Research, 146*, 19–29.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery and Psychiatry, 20*, 11–21.

Shallice, T. (1972). Dual functions of consciousness. *Psychological Review, 79*, 383–393.

Shallice, T., Burgess, P. W., Schon, F., & Baxter, D. M. (1989). The origins of utilization behaviour. *Brain, 112*, 1587–1598.

Sherrington, C. (1948). *Integrative action of the nervous system*. Liverpool: Liverpool University Press.

Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems. *Psychological Review, 94*, 439–454.

Shiffrin, R. M. (1997). Attention, automatism, and consciousness. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 49–64). Mahwah, NJ: Erlbaum.

Schore, A. N. (2001). Effects of a secure attachment relationship on right brain development, affect regulation, and infant mental health. *Infant Mental Health Journal, 22*, 7–66.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3–22.

Sokolov, E. N. (1976). Learning and memory: Habituation as negative learning. In M. R. Rosenzweig & E. L. Bennett (Eds.), *Neural mechanisms of learning and memory* (pp. 475–479). Cambridge: The MIT Press.

Spear, N. E., & Miller, J. S. (1989). Ontogeny of spontaneous alternation behaviour. In W. N. Dember & C. L. Richman (Eds.), *Spontaneous alternation behaviour* (pp. 131–144). New York: Springer-Verlag.

Spence, S. A., & Frith, C. D. (1999). Towards a functional anatomy of volition. *Journal of Consciousness Studies, 6*, 11–29.

Squire, L. R. (1986). Mechanisms of memory. *Science, 232*, 1612–1619.

Squire, L. R., Hamann, S., & Knowlton, B. (1994). Dissociable learning and memory systems of the brain. *Behavioral and Brain Sciences, 17*, 422–423.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioural and Brain Sciences, 23*, 645–726.

Steele, C. M., & Joseph, R. A. (1990). Alcohol myopia: Its prized and dangerous effects. *American Psychologist, 45*, 921–933.

Stewart-Williams, S., & Podd, J. (2004). The placebo effect: Dissolving the expectancy versus conditioning debate. *Psychological Bulletin, 130*, 324–340.

Stuss, D. T., & Anderson, V. (2004). The frontal lobes and theory of mind: Developmental concepts from adult focal lesion research. *Brain and Cognition, 55*, 69–83.

Stuss, D. T., & Benson, D. F. (1983). Emotional concomitants of psychosurgery. In K. H. Heilman & P. Satz (Eds.), *Neuropsychology of human emotion* (pp. 111–140). New York: The Guilford Press.

Taylor, J. (1958). *Selected writings of John Hughlings Jackson*. London: Staples Press.

Tinbergen, N. (1969). *The study of instinct*. Oxford: Clarendon Press.

Toates, F. (1995). Animal motivation and cognition. In H. Roitblat & J.-A. Meyer (Eds.), *Comparative approaches to cognitive science* (pp. 435–464). Cambridge: The MIT Press.

Toates, F. (1998). The interaction of cognitive and stimulus-response processes in the control of behaviour. *Neuroscience and Biobehavioural Reviews, 22*, 59–83.

Toates, F. (2000). Multiple factors controlling behaviour: Implications for stress and welfare. In G. P. Moberg & J. A. Mench (Eds.), *The biology of animal stress* (pp. 199–226). Wallingford: CAB International.

Toates, F. (2001). *Biological psychology: An integrative approach*. Harlow: Pearson Education.

Toates, F. (2002a). Application of a multilevel model of behavioural control to understanding emotion. *Behavioural Processes, 60*, 99–114.

Toates, F. (2002b). Setting Norman's dual model into a broader context. *Behavioural and Brain Sciences, 25*, 119–120.

Toates, F. (2004). Cognition, motivation, emotion and action: A dynamic and vulnerable interdependence. *Applied Animal Behavioural Science, 86*, 173–204.

Toates, F. (2005). Evolutionary psychology—towards a more integrative model. *Biology and Philosophy*, in press.

Tolman, E. C. (1932). *Purposive behaviour in animals and men*. New York: The Century, Co.

Tomie, A. (1996). Locating reward cue at response manipulandum (CAM) induces symptoms of drug abuse. *Neuroscience and Biobehavioural Reviews, 20*, 505–535.

Trevarthan, C. B. (1968). Two mechanisms of vision in primates. *Psychologische Forschung, 31*, 299–337.

Tsuchida, J., Kubo, N., & Kojima, S. (2002). Position reversal learning in aged Japanese macaques. *Behavioural Brain Research, 129*, 107–112.

Tulving, E. (1985). How many memory systems are there? *American Psychologist, 40*, 385–398.

Ullsperger, M., Volz, K. G., & von Cramon, D. Y. (2004). A common neural system signalling the need for behavioural changes. *Trends in Cognitive Sciences, 8*, 445–446.

Umiltà, C. (1988). The control operations of consciousness. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 334–356). Oxford: Clarendon Press.

Van der Kamp, J., Oudejans, R., & Savelsbergh, G. (2003). The development and learning of the visual control of movement: An ecological perspective. *Infant Behaviour and Development, 26*, 495–515.

Van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences, 21*, 615–665.

Van Gulick, R. (1994). Deficit studies and the function of phenomenal consciousness. In G. Graham & G. L. Stephens (Eds.), *Philosophical psychology* (pp. 25–49). Cambridge: The MIT Press.

Voudouris, N. J., Peck, C. L., & Coleman, G. (1990). The role of conditioning and verbal expectancy in the placebo response. *Pain, 43*, 121–128.

Wagner, A. R. (1978). Expectancies and the priming of STM. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behaviour* (pp. 177–209). Hillsdale: Lawrence Erlbaum.

Warrington, E. K., & Weiskrantz, L. (1982). Amnesia: A disconnection syndrome? *Neuropsychologia, 20*, 233–248.

Wegner, D. M. (1997). When the antidote is the poison: Ironic mental control processes. *Psychological Science, 8*, 148–150.

Wegner, D. (2002). *The illusion of conscious will*. Cambridge: The MIT Press.

Weiskrantz, L. (1988). Some contributions of neuropsychology of vision and memory to the problem of consciousness. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 183–199). Oxford: Clarendon Press.

Weiskrantz, L., & Warrington, E. K. (1979). Conditioning in amnesic patients. *Neuropsychologia, 17*, 187–194.

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin, 120*, 272–292.

White, N. M., & McDonald, R. J. (2002). Multiple parallel memory systems in the brain of the rat. *Neurobiology of Learning and Memory, 77*, 125–184.

Wickelgren, W. A. (1979). Chunking and consolidation: A theoretical synthesis of semantic networks, configuring in conditioning, S–R versus cognitive learning, normal forgetting, the amnesic syndrome, and the hippocampal arousal system. *Psychological Review, 86*, 44–60.

Willingham, D. B. (1998). A neurophysiological theory of motor skill learning. *Psychological Review, 105*, 558–584.

Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), *Attention* (pp. 223–256). Hove: Psychology Press.

Yantis, S. (2000). Goal-directed and stimulus-driven determinants of attentional control. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 73–103). Cambridge, MA: The MIT Press.

Zajonc, R. B. (1980). Feeling and thinking—Preferences need no inferences. *American Psychologist, 35*, 151–175.

Zhu, J. (2003). Reclaiming volition. *Journal of Consciousness Studies, 10*, 61–77.

Zhu, J. (2004). Is conscious will an illusion? *Disputatio, 16*, 58–68.