

Mechanisms of madness: evolutionary psychiatry without evolutionary psychology

P. GERRANS

Department of Philosophy, Adelaide University, Level 1, Hughes Building, North Terrace, Adelaide, Australia; (e-mail: philip.gerrans@adelaide.edu.au)

Received 7 October 2005; accepted in revised form 27 January 2006

Key words: Delusions, Dopamine, Evolutionary psychiatry, Mental time travel, Prefrontal cortex, Neural networks

Abstract. Delusions are currently characterised as false beliefs produced by incorrect inference about external reality (DSM IV). This inferential conception has proved hard to link to explanations pitched at the level of neurobiology and neuroanatomy. This paper provides that link via a neurocomputational theory, based on evolutionary considerations, of the role of the prefrontal cortex in regulating offline cognition. When pathologically neuromodulated the prefrontal cortex produces hypersalient experiences which monopolise offline cognition. The result is characteristic psychotic experiences and patterns of thought. This bottom-up account uses neural network theory to integrate recent theories of the role of dopamine in delusion with the insights of inferential accounts. It also provides a general model for evolutionary psychiatry which avoids theoretical problems imported from evolutionary psychology.

Introduction

The cognitive and behavioural flexibility of humans depends on our abilities to trawl our memories, imagine alternative histories and futures for ourselves, consult with others using linguistic representations and, where necessary, resort to abstract reasoning using symbolic representation. These capacities all depend on the prefrontal cortex. The prefrontal cortex evolved to allow us to escape the stimulus-bound present and manipulate representations in *offline mode*. Offline cognition, under prefrontal control, inhibits automatic cognitive routines activated by a current stimulus and engages cognitive processes such as attention, working memory, executive functions, metarepresentation, language and inference (Knight 1999; Wood 2003).

In this paper I give an evolutionary account of offline cognition and provide a neurocomputational framework which links offline cognition to its biological substrate in the prefrontal cortex. The account underpins a novel account of psychiatric disorder, especially delusion, and provides the basis for a principled incorporation of evolutionary theory into psychiatry. I shall argue that the best way to explain delusion is as

The monopoly of offline cognition by hypersalient experiential representations caused by dysregulation of mechanisms which manage the *transition* between on and offline cognition.

A proper understanding of these mechanisms explains delusions without invoking psychological concepts like belief, desire and inference which are part of the everyday and clinical account of delusion as a

False belief about external reality based on incorrect inference (DSM IV, 1994)

The approach I recommend is consistent with biological and pharmacological approaches to psychiatric disorder because I locate the source of delusion in the regulation of prefrontal cortical function by basic neurotransmitter systems. However treating psychiatric disorders simply as a phenomenon of neurochemical imbalance does not explain their cognitive properties. I provide an integrative cognitive framework which links biology to the phenomenology of psychiatric disorder (Halligan and Marshall 1996).

The evolutionary account I propose has consequences for evolutionary psychiatry. One strand of evolutionary psychiatry explains apparently irrational beliefs by replacing the unconscious psychological mechanisms of Freudian theory with innate psychological mechanisms postulated by evolutionary psychology (Stevens 1996; Charlton 2000). The account I provide, however, suggests that much psychiatric disorder, indeed much irrationality, is not essentially psychological at all. However the account also suggests that evolutionary psychiatrists are correct to trace the origin of psychiatric disorder to the influence on cognition of phylogenetically ancient systems which manage basic cognitive and behavioural routines for foraging, caregiving, affiliation, cooperation and competition. Thus there is an evolutionary foundation for psychiatry, but it is best thought of as ethological rather than psychological.

The paper is organised in six sections. Section 1 gives an example of the replacement of a psychological explanation by an evolutionary account of a neurobiological mechanism. The example is the phenomenon of selective amnesia following trauma, which can be potentially be explained in terms of the effects of glucocorticosteroids on the mammalian hippocampus-amygdala system. This replaces the need to postulate psychological mechanisms such as repression which postulate unconsciously represented goals and inferences. In Section 2, I begin to address the worry that while this approach may work for selective amnesia it might not be adequate for forms of psychiatric disorder, such as delusion, which present clinically as beliefs generated by faulty inference. In this section, I explain the distinction between offline and online cognition and show how offline cognition depends on prefrontal cortical function regulated by neurotransmitter systems projecting from phylogenetically ancient structures in the brainstem. In Section 3, I provide a neurocomputational framework which links this evolved neurobiology to the cognitive capacities

provided by the prefrontal cortex. The interim conclusion of these two sections is that cognitive functions which depend on the prefrontal cortex are influenced by the way the cortex is layered over more ancient mechanisms which regulate online cognition. In Section 4, I show that the nature of this influence can explain delusion. In Section 5, I discuss Shitij Kapur's theory of delusion as an instance of dopaminergic dysregulation and show it fits with my account. Finally I consider the consequences of the account developed here for evolutionary psychiatry.

Section 1. The mechanisms of memory. Selective amnesia following trauma

One of the central ideas of psychoanalysis is that memories of traumatic experiences can be repressed (Freud 1932). The notion of repression postulates a psychological mechanism whose function is to prevent intact memories of distressing episodes from surfacing into conscious experience. The mechanism is psychological because it is characterised in terms of the goals, beliefs and inferences (albeit unconscious) of an agent (Freud 1963).

To give this theory a biological foundation requires an understanding of memory which integrates its cognitive and neural properties. The basic insight of the neural study of memory is that it depends on synaptic plasticity (Gaiarsa et al. 2002). The neural representation of a stimulus is the result of long term changes in the synaptic properties of interconnected neurons. Learning, whether of a conditioned reflex, a skill, a language or a theory is a matter of constructing a neural circuit which exhibits a particular pattern of excitation for each mnemonic task.

The molecular mechanisms of long-term memory are the same for all types of memory in mammalian species. Long term memory involves long term potentiation (LTP) of synapses which depend on the synthesis of proteins within the cell driven by gene expression consequent on chemical signalling between the membrane surface and the nucleus (Frey 1997; Bennett 2000). Every neuron which undergoes LTP, does so via the same mechanisms of protein synthesis.

Unlike very simple organisms, mammals can be conditioned by attending to positively or negatively reinforcing stimuli. In mammals the basic mechanism of reinforcement learning is the distributed circuitry connecting the amygdala, hippocampus and medial temporal systems (Clark 1998; Eichenbaum 1998). The role of the amygdala is to tag relevant memories with affective valence via circuitry which initiates aversive or attractive behaviour (LeDoux 1999).

The relevance of these facts to repression is seen in experiments with juvenile rodents. Mammalian juveniles, unlike reptiles and fish which are born mini-adults, use the adult as a developmental resource, both metabolically and cognitively. This feature of warm-blooded life means that infants need to be

attached to caregivers while they learn how to cope with their environment. Grooming and feeding reinforce attachment and separation reduces it. Prolonged separation produces distress whose molecular basis is corticotropin releasing factor (CRF) and the resultant production of glucocorticoids (Plotsky 1993). In insecurely attached rats the hypothalamus and limbic regions, show higher levels of corticotropin releasing factor (CRF) and resultant glucocorticoids. Heightened CRF and glucocorticoids resulting from prolonged separation, lasts for life and causes cellular atrophy and loss of long term potentiation (Nemeroff 1996). In such animals hippocampal volume and density of synaptic connections are reduced. The neurobiological basis for memory has been eroded by the hormonal basis of separation distress (Sapolsky 1996; Kandel 1999).

Of course the fact that distressed rats have damaged memory systems does not by itself refute a theory which invokes psychological mechanisms of repression (not least since rats do not have the neural basis for episodic memory and its employment in higher cognition) but the temporal lobe-amygdala circuitry in all mammals is essentially the same, which does suggest that high levels of stress can produce amnesia. Thus Eric Kandel concluded that “what initially appears as repression may actually prove to be true amnesia: damage to the temporal lobe system of the brain” (Kandel p. 514). The onus is on the theorist of repression to show that this type of mechanistic explanation is incomplete.

The purpose of this example is to show that, in principle, accounts of neurobiological mechanisms can replace psychological theories by focusing on a simple case of amnesia for procedural memory in rats. If it is to be plausibly extended to more sophisticated types of memory we need to show the mechanisms which damage procedural memories in rats have similar effects on episodic memory in humans. This is in fact the case because more sophisticated forms of memory (episodic, declarative) are layered over phylogenetically older ones, use the same molecular mechanisms of LTP and inherit the same reinforcement circuitry.

However higher cognition does not consist in memory alone, but in the integration of relevant memories with novel representations and their manipulation in order to solve problems (Waltz 1999; Wood 2003). Thus the main focus of the rest of the paper is on these processes of integration and manipulation. However I will claim that the very same learning and reinforcement mechanisms which explain the adaptive behaviour of rats influence the intelligent behaviour of humans. The reason is that the cortical structures which manage the integration and manipulation of representations in offline cognition are layered over phylogenetically ancient systems for the control of behaviour by procedural memory systems.

The progressively more abstract forms of cognition enabled by the slowly developing human prefrontal cortex are Mental Time Travel, Decontextualisation and Procedural Rationality. These abstract forms of cognition arrive in this order, both phylogenetically and ontogenetically because of the increasing

demands they place on the prefrontal cortex, firstly in the degree of inhibition of online cognition and, secondly, in the complexity of the interactions which enable representations to be manipulated offline.

Section 2. Higher cognition and the prefrontal cortex

The functioning of the prefrontal cortex can be explained at different levels of abstraction. Most neuroscientists have given *processing accounts*, concentrating on the nature of basic properties like inhibition and activation, firing rates, sequencing, connectivity and neural architecture. Cognitive neuropsychologists have given *representational accounts* which concentrate on the higher cognitive functions which compute responses to experience. Examples are studies which show that different areas of the prefrontal cortex are engaged by different reasoning tasks. Left and right hemispheres are engaged by deductive and inductive reasoning tasks for example. The representations here are characterised at a high level of abstraction: logical rules or rules of probability theory (Harrington et al. 1998; Knight 1999; Wood 2003).

The account I provide is a *neurocomputational account* which focuses on the interface between basic processing and representation. My aim is to explain how the processing mechanisms influence the representational ones, since this is where psychiatric disorder originates. As a device of exposition we can organise the neurocomputational account around the role of progressively more abstract forms of memory and the uses to which they can be put in progressively more abstract forms of cognition. More complex environments pose greater cognitive demands which require sophisticated procedures for learning, retrieving and manipulating information to cope with complexity. Animals with these capacities require a longer learning history and period of dependence supported by neurobiological mechanisms of attachment.

Phylogenetically a larger and more densely connected prefrontal cortex in mammals, primates, apes, hominids and humans, enables progressively greater degrees of inhibition of online cognition which in turn allows progressively more behavioural and cognitive flexibility (Prabakaran 1997). The price of human intelligence is helplessness (Allman 1999) while the large and relatively plastic prefrontal cortex is wired up in development.

Ontogenetically humans follow the same pattern. The progression from stimulus bound infant governed by procedural memories acquired by conditioning within a narrow range of parameters set by attachment mechanisms, to adults capable of mental time travel, declarative cognition and ultimately procedural rationality depends on the development of the prefrontal cortex into adolescence.

The diagram below shows the relationships between different types of long term memory and their neural correlates (Figure 1). Notably only one of the branches, declarative memory is unique to humans.

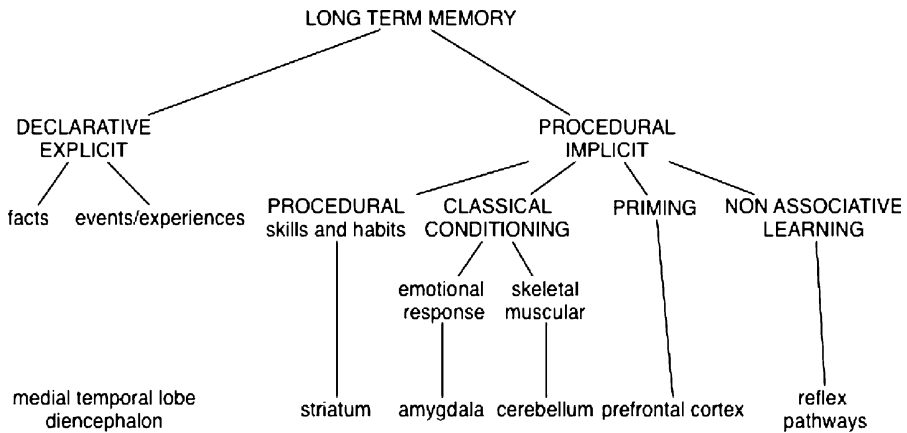


Fig. 1.

Procedural memory supports habitual behavioural routines, like foraging or navigation. An animal with a vast procedural repertoire can negotiate its environment provided that each situation it encounters activates an appropriate procedurally regulated memory. Such an animal will appear highly intelligent unless it encounters situations which activate inappropriate, conflicting procedural memories or no relevant memories at all.

In such a situation an organism needs to be able to determine which of its previous behavioural strategies are appropriate or, if they are inappropriate, to generate a new one. Procedural memory alone is inadequate for this task. At the very least an organism needs to inhibit automatic behaviour, generate alternate behaviours, test them and remember successful strategies.

Learning in rats with their proto-prefrontal cortices is modulated by attention. It creates a window for conditioning to take place as a result of reinforced exploratory behaviour. In order for this to be the case competing behavioural routines must be *inhibited*. Attention and inhibition are correlated: a stimulus can be made the focus of attention by inhibiting attention to competitors or by enhancing attention to a salient stimulus.

The ability to make a representation salient using mechanisms of selective inhibition and activation is the basic cognitive ability provided by the prefrontal cortex. It is a precondition for all the other sophisticated processes in which representations are manipulated. The prefrontal cortex evolved in mammals, firstly, to inhibit the online systems while maintaining activation in assemblies which produce represent a salient stimulus and, secondly, to enable the increasingly complex computation of appropriate responses to that stimulus (Knight 1999; Wood 2003).

Fortunately humans can not only learn from experience but encode those experiences in forms which enable them to be retrieved. This is episodic memory. One of the features of episodic memory is that it is intimately connected with imagination. Both involve the construction of perceptual or

sensory imagery *in the absence of a stimulus*. This ability confers enormous behavioural flexibility. Instead of performing a potentially dangerous action and hoping for conditioning by trial and error one can recall the consequences of previous efforts or rehearse possible outcomes in imagination. Thus episodic memory is intimately involved in planning and deliberation and *exploits the same mechanisms*. Subjects who selectively lose episodic memory as a result of neural damage are also impaired in planning (Klein 2002). As we shall see below, this combination of episodic memory and imagination recruited in the service of intelligent behaviour is unique to humans.

Also unique to language using humans is declarative memory: memory for facts or propositions. As with episodic memory, the role of declarative memory co-evolved with the use of declarative representation in planning and deliberation. The mind of an individual with declarative representation is not restricted to manipulating information encoded in episodic memory or imagination. In principle she has access to any fact which can be represented linguistically.

Of course having a vastly expanded database of experiential and declarative representations does not in itself confer any advantages over organisms with purely procedural memory systems. It is the ability to retrieve and manipulate information relevant to problems which exceed procedural abilities which gives humans their cognitive advantage. The evolution of more sophisticated and abstract forms of memory co-evolves with the capacities to generate and manipulate novel representations. These cognitive capacities depend on the prefrontal cortex.

2.1 Three modes of offline cognition

Thomas Suddendorf has pointed out that a crucial difference between offline cognition in humans and other primates, conferred by the size and connectivity of the human prefrontal cortex, is that offline cognition in humans involves self-representation. In order to reflect plan and deliberate one needs to be able to represent oneself as the owner of the relevant memories or the executor of plans and intentions. In offline cognition humans retrieve information from the past and project themselves imaginatively into the future to gather, compare and evaluate information which enables us to deal with present experience. As he puts it we can engage in Mental Time Travel (Suddendorf 2000).

A crucial aspect of mental time travel is that the episodic representations it manipulates are associated with the basic mechanisms of reinforcement shared with other mammals. This is essential to its success, since unless those affective associations were present, the episodic representations characteristic of mental time travel could not play their essential role in deliberation. Imagining or remembering oneself in a dangerous situation is no guide to action in that situation unless the representation of the experience is characteristically affect-laden.

The representations manipulated in mental time travel are *essentially indexical* because they are actual or imagined autobiographical episodes. The autobiographical representations involved retain affective and hence motivational traces of the subject's biographical history. This indexicality is good when it gives the subject information about the relevance of representations. If you remember or imagine yourself going on holiday or getting married the affective tone of the representations is essential information for decision-making.

However, mental time travel can also be a poor source of information where the subject's biography provides incomplete or distorted information with inappropriate affective associations. The classic cases here are mood disorders in which self-representations and their affective associations make it impossible for subjects to generate and use alternative representations (Seligman et al. 1979; Beck 1989; Gerrans 2000).

In order to access information which is not essentially autobiographical we need representations which are not indexical. For this we need non-episodic information, communicated by third parties and here language is the obvious medium. This gives us a declarative medium of representation which allows us to access a wider range of information than that to which we have been exposed in experience. Everyone's declarative database is larger than their experiential database.

Declarative representation also allows us to track our episodic representations, combining mental time travel with linguistic representation in offline mode. Or it might allow us to dispense with episodic representation by allowing linguistic representations to stand in for episodic ones. I can remember what Jon looks like episodically, I can simply say to myself "Jon is tall" or I can solve problems "Jon is too tall to borrow my trousers" in either mode. It is a telling feature of cognitive development that children's improvement on tasks which require them to detach from the current stimulus is language-dependent (Deák 2003).

In the normal case our declarative representations are interpreted for us by episodic ones for which they stand in. The declarative sentence "there are sharks at Cactus beach" is by itself a behaviourally and affectively inert symbol string. However it is interpreted for a subject by the episodic representations of sharks and their affective associations.

In some cases a subject needs to be able to respond to experience by manipulating information without implicating herself in its representation. She can do this by thinking entirely in declarative mode, inhibiting the episodic interpretation of those sentences by posterior systems and their links with reinforcement systems.

The ability to evaluate and act on declarative information *without activating episodic interpretations and their affective associations* is called *decontextualisation* by cognitive psychologists and its incidence varies widely within and across human populations (Stanovich 1999). Full decontextualisation requires that offline cognition be *uninterpreted* by not being linked to episodic representations (Stanovich 1999) and requires a greater degree of inhibition than mental time travel.

The most abstract form of decontextualisation the manipulation of uninterpreted symbols to test beliefs for consistency. It is formalised by theories of logic or decision theory and competence is tested by tasks whose solutions require this type of symbol manipulation. The relationship between these formalisations and actual human cognition is unclear (Stein 1996). Nonetheless it does seem that there are some circumstances in which humans do use inferential rules to govern behaviour. I shall reserve the term procedural rationality for the application of inferential rules to uninterpreted symbols in order to solve a problem.

The application of inferential rules to uninterpreted symbols is the most extreme form of form of decontextualisation, and as Dan Sperber points out, it is the last communicative resort. An example might be two persons arguing about the shortest route to a destination. Both rely on their own resources for mental time travel and declarative cognition to evaluate alternatives but cannot agree. As a last resort I might demonstrate that one route is shorter than another by inviting my feisty interlocutor multiply average speed by time.

Offline processing is not monolithic but involves a hierarchy of progressively more abstract operations which enable the subject to detach from the stimulus and mobilise representations relevant to the explanation of experience. Mental time travel requires less inhibition than declarative cognition and declarative cognition requires less inhibition than procedural rationality. The progressively greater inhibitory demands of each type of cognition go hand in hand with their cognitive complexity. Even if we could equip a chimpanzee with the ability to compute an algorithm to determine how to maximise expected utility the first sight of a ripe banana or anogenital swellings would disrupt his contemplations. And behaviour would revert to regulation by online processes.

Psychological accounts of delusion assimilate the delusional response to procedural rationality Cognitive psychologists also do this when they probe the cognitive structure of disorders by testing the ability to use inferential rules, or interpret the results of tests as faulty application of an inferential rule (Huq et al. 1988; Garety et al. 1991; Maher 1992; Mujica-Parodi 2000; Gerrans 2001).

In fact delusions result from the effect on mental time travel of mechanisms which make some representations hypersalient. That is to say they attract frontal resources, dominating attention, executive function and working memory. Rather than being selectively inhibited, some representations are inappropriately enhanced. Furthermore the very same failures of inhibitory mechanisms which make those representations hypersalient *prevent* the engagement of more abstract forms of cognition such as decontextualisation or procedural rationality, for these depend on inhibition.

In order to substantiate this claim and apply it to psychiatric disorder we need to show how neurobiological processes influence the construction and manipulation of representations. Thus I first describe relevant anatomical features of the prefrontal cortex and show how they influence its computational properties. We then apply this neurocomputational account to the explanation of delusion.

2.2 The anatomical substrate of offline cognition

Neurally, offline cognition depends on structures in the prefrontal cortex which evolved to inhibit processing in the posterior structures which implement on-line cognition. While there is a controversy about the extent of differences in prefrontal cytoarchitecture between humans and great apes (McKinney 2002) there is consensus that normal prefrontal function is essential for the higher cognitive processes of deliberation, reflection, planning and inference which are unique to our cognitive phenotype.

It is not just the size of the prefrontal cortex but its dense interconnectivity with posterior, limbic and brainstem areas which enables offline cognition. These connections are both afferent and efferent which enables bi-directional signalling between the prefrontal cortex and posterior areas. Furthermore while most connections from posterior networks to the prefrontal cortex are excitatory the prefrontal cortex has extensive *inhibitory connections* (via GABA interneurons) to posterior areas. This interconnectivity enables construction of transient recurrent circuits distributed across the prefrontal cortex and posterior assemblies (Friston 2002). The prefrontal cortex maintains salient representations by enhancing the level of activation in their implementation circuitry and inhibiting activation levels in other circuits competing for prefrontal resources (Fuster 1997).

Moment to moment experience is a product of competition between more or less transient coalitions of active neurones, distributed across prefrontal-posterior circuitry, to monopolise offline cognition by attracting prefrontal processing resources. Success in the competition can be produced volitionally, from above, or from below by the novelty or salience of a posterior representation.

The ability of a representation to attract and maintain the frontal resources necessary to monopolise offline cognition depends on a combination of factors. Most obviously the level of activity within a coalition must exceed a threshold sufficient to defeat competitors. Thus raising or lowering activation in a neural assembly affects the cognitive salience of the representations it implements, according to the level of activation in competing coalitions. Normal experience is a constant dynamic interaction between these processing modes: a train of reflective thought is interrupted by a voice or movement, attention shifts, directing frontal resources to the neural activation which represents the new stimulus, which then recedes from attention as we habituate to it, or monopolises attention further as we mobilise more frontal resources to determine how to respond.

Once again the basic neurotransmitter mechanisms which regulate activation in these distributed networks are phylogenetically conserved.

In organisms without a prefrontal cortex these systems regulate basic behavioural systems and modes of online cognition. The Serotonergic (5Hydroxytryptamine, 5HT) system has been present in the brainstem for 500 million years. Like all neurotransmitters it operates by binding to receptors at different sites and changing levels of activity in postsynaptic neurons. It does

not do this directly but by modulating the responses of neurones to other neurotransmitters. Thus, depending on the actions of receptors (14 in humans) to which it binds, it can function as an agonist or antagonist to other neurotransmitters. Its most basic effect however is to modulate levels of arousal by modulating activity in neural circuits which control wakeful exploratory activity. The *raphe nuclei* from which 5HT systems project are inactive during sleep, during which time the organism is essentially cholinergically regulated, and active when the organism is awake (Hobson 1999). Even in animals with almost no brain, serotonin neurons are involved in orienting animal towards nutrient sources and controlling digestion. In mammals with brains 5HT modulates online cognition at various levels of complexity. For example if 5HT levels are reduced in rats their foraging, sexual and social activity is reduced (Allman 1999).

The diffuse projections of the serotonin system in the human prefrontal cortex perform the same type of function for offline cognition. It is helpful to think of offline cognition as *cognitive foraging*, modulated by the same systems which enable search through behavioural space in phylogenetically older systems. Once a system is able to attend to a salient experience it can then *forage among its representations* for a solution to the problem set by experience, or if the experience is a rewarding one, hold it in working memory while strengthening associations with relevant representations. Planning, deliberation, reasoning are ways of exploring representational space.

Just as offline cognition inherits its foraging and reinforcement mechanisms from the online mechanisms of behavioural foraging it also inherits the mechanisms on which salience depends. For example a foraging rodent who successfully finds a food source will have the successful pattern of neural activity marked by dopaminergic innervation. Thus the dopamine system interacts with the limbic system to reinforce procedural memory for rewarding behaviour. Similarly in animals with a larger prefrontal cortex, capable of offline cognition, dopamine makes representations salient by increasing activation in the relevant neural circuitry. In effect when an organism searches its store of representations those representations which are enhanced by dopamine dominate the window of attention and working memory.

So far we have only described these phylogenetically ancient systems, not the computational system they regulate. Before doing that I will give two examples which show just how important these regulatory systems are.

2.3 Dreaming and depression

The most radical example of the role of 5HT is dreaming in which the 5HT system is down regulated and the cholinergic system (which enhances activation in local neural assemblies and is antagonistic to the formation of large distributed networks such as those between the prefrontal and posterior systems) takes over. Motor systems are deactivated and perceptual systems and

episodic memory run offline since there is no perceptual stimulus. The dopamine system is also active in dreaming, but interestingly, there is no online or offline stimulus or problem to organise cognition. The “theme” of rapid eye movement (REM) dreams is just the contingent association of cholinergically regulated representations and affective states ungoverned by processes like executive inhibition, selective attention or inference. The subject drifts from scene to scene through an eerie dreamscape, spectator to representations experienced through a dopamine haze (Hobson 1999).

Interestingly in non rapid eye movement (NREM) sleep the 5HT system is more active with the consequence that the pattern of representational association is more coherent, although repetitive and perseverative, because unbroken by the intrusion of representations of environmental stimuli. But, even in NREM sleep, offline cognition, understood as the systematic exploration of solutions to problems set by experience, does not occur. The prefrontal cortex does not have the neurochemical resources to perform its executive tasks.

The example of dreaming shows how cognitive subsystems interact in different mode when their interactions are differentially regulated by neurotransmitter systems. In REM dreaming neuromodulators deactivate offline cognition while online systems continue to produce their characteristic representations in the absence of inputs from the perceptual periphery.

This endogenous production of representations experienced as having special salience and affective tone has led many to compare schizophrenic delusions with dreaming.

While it is not correct to say that delusional experience is simply a waking dream, because the prefrontal-posterior network is configured differently during delusions, dreams are an important source of evidence about the connection between fundamental neuropharmacology and cognition. It is not coincidental that John Nash, describing the auditory hallucinations he experienced in schizophrenia, said: “You’re really talking to yourself is what the voices are, but it’s also parallel to a dream. In a dream it’s typical not to be rational.” It’s typical because it’s impossible to be rational when the frontal systems required for higher order cognition are deactivated.

Another example of the way the neuropharmacology affects waking cognition by restructuring the prefrontal-posterior network in a more permanent way is the role of norepinephrine in depression. Via projections to the amygdala-hippocampus loop and prefrontal areas, the over active norepinephrine system enhances amygdala activation with the consequence that representations with negative valence dominate mental time travel. When the depressive subject remembers or imagines episodes they are automatically associated with negative affect, due to connections between the amygdala and hippocampus whose role is to “tag” representations stored in memory with affective valence. If the imbalance of neurochemicals is maintained the actual circuitry system may change as neurochemical imbalance triggers gene expression which leads to the reconfiguration of pathways between the amygdala and hippocampus (Ressler and Nemeroff 1999).

So far we have described neuromodulatory effects on the distributed neural system comprised of prefrontal and posterior circuitry in order to show how representations acquire salience and how that network can be transiently or permanently reconfigured to deal with the resultant experience. It is now time to provide the computational framework which links this processing account to representational accounts and allows us to explain delusion in terms of the way the prefrontal-posterior network handles particular representations.

Section 3. A computational architecture for offline cognition

Computationally, we can treat the neural systems which manage online cognition as interconnected neural networks whose automatic input-output relationships are determined by inflexible weights in their hidden layers (Quartz and Sejnowski 1997; Arbib 2003). A weight is the propensity, described as a probability, of a unit in a neural network to become active when another to which it is connected becomes active. Automaticity of processes implemented in neural networks is enabled by fixing weights. This inflexibility is an advantage when processing stereotypical stimuli (a visual system producing different representations of the same stimulus would not be fitness enhancing, nor would a sensori-motor loop which sometimes produced aversive and sometimes exploratory behaviour towards a predator).

The advantage of cognition governed by this automatic, weight-based processing of tacit representations has been well described by Robert Zajonc (Zajonc 1980).

If the rabbit is to escape, action must be undertaken long before the completion of even a simple cognitive process... before, in fact, the rabbit has fully established and verified that a nearby movement might reveal a snake in all its coiled glory. The decision to run must be made on the basis of minimal cognitive engagement (Zajonc 1980, p. 156).

Zajonc's example allows us to make a further important point about online weight-based processing. A familiar stimulus is encoded as a pattern of activation across an input layer of a network. This pattern is fed forward via intermediate "hidden" layers to an output layer whose activation represents the response, which may vary from an motor instruction to a perceptual image (This is an oversimplification. Recurrence and feedback between layers are the norm as is influence from other networks). Thus the *hidden layers compute the response to the input*.

Most representational phenomenology is awareness of activation in *output layers* of these systems. We have no introspective access to representations computed by hidden layers, which are thus tacit. Most automatic, online, cognition does not require the complete computation of a response expressed as activation in an output layer. Activation in hidden layers is sufficient to

trigger responses managed by other networks to which the hidden layers are connected. Zajonc's rabbit ran because activation in hidden layers reached a threshold which initiated a flight response via connections between hidden layers and the amygdala.

However maintaining activation in output layers enables us to explicitly represent the causes of experience and experience the link between a representation and its affective valence. This has obvious advantages in conditioning or more elaborate forms of learning. Maintaining activation in output layers allows us to become aware of the motivational character of explicitly represented objects and allows for weights to be appropriately changed by Hebbian or other learning mechanisms (Clark 1998).

In offline cognition the outputs of different weight-based systems are reconstructed by memory or imagination, simultaneously held in working memory, metarepresented, compared and manipulated into novel combinations. Hence it can be described as *activation-based* processing because it maintains activation in output layers of online networks which would otherwise decay as online cognition copes with successive ecological challenges (O'Reilly and Munakata 2000; O'Brien 2004). Computationally the most basic role of the prefrontal cortex is to selectively enhance or inhibit activation, allowing us to manipulate representations in mental time travel.

Section 4. Offline cognition as activation-based processing and the role of dopamine

The concepts of static and dynamic coalitions between prefrontal-posterior systems and the contrast between weight and activation-based processes help to clarify the influence of neurochemical delivery systems on cognition. For the rest of the paper we will use this framework to explain how dopamine makes some representations hypersalient in offline cognition. Results from neural network models of the interaction between dopamine delivery systems and the PFC (Braver 1999) are the basis of the arguments in this section.

While posterior systems themselves have relatively inflexible weights, since they perform specialised routines the larger offline systems comprised of the interconnected prefrontal and posterior systems are more plastic. This plasticity enables transition between on and offline cognition.

The weights within the networks which implement transient prefrontal-posterior coalitions can be reset by neurotransmitters which operate on different time scales. When we describe some of these processes below it is worth recalling that the concept of a weight in a neural network abstracts from the mechanisms by which regulate the threshold of activation of a neurone: as well as neurotransmitter blockade, the density, sensitivity and type of receptors, dendritic growth or death, cell metabolism and internal protein structure and currents all contribute to the likelihood that an axon potential will be generated

by depolarisation at a synapse. The concept of a weight expresses the sum of these changes as a probability *averaged across a neural assembly*.

Dopamine alters the explicit (activation-based) and tacit (weight based) computational properties of the prefrontal-posterior networks (Braver 1999) by resetting the weights in a network or by simply increasing or lowering activation in a particular neural assembly, leaving the weights system intact.

Dopamine works by enhancing the signal to noise ratio between communicating neural assemblies (because the SNR enhancement effect operates at the level of neural assemblies rather than single neurones, it is amenable to modelling in neural networks which abstract from the biochemistry of inter-neural signalling). It does so via the interaction of at least two types of dopamine action. Phasic dopamine, delivered in short bursts, binds to D2 receptors on the postsynaptic membrane. It is rapidly removed by reuptake from the synaptic cleft and acts quickly. Tonic dopamine which acts over longer time scales, accumulates in the synaptic cleft and binds to presynaptic DI autoreceptors triggering reuptake. Thus short bursts of phasic dopamine make a target assembly more receptive to incoming signals, sustained steady levels of tonic dopamine make a target assembly less receptive.

Phasic and tonic dopamine are thus antagonists and have different effects on the assemblies they afferent: phasic dopamine, acting on prefrontally regulated assemblies, produces a gating effect. It allows new activation patterns in the prefrontal cortex to be formed, producing representations of new stimuli. Tonic dopamine maintains an occurrent activation pattern, allowing a representation to be sustained against interference or competition. The hypothesis follows, and is confirmed by neural network models, that the balance of tonic and phasic dopamine is responsible for the rate of turnover of representations in the PFC-posterior networks (Grace 1991).

Tonic dopamine effects may increase the stability of maintained representations through an increase in the SNR of background versus evoked activity patterns. In contrast, phasic dopamine effects may serve as a gating signal indicating when new inputs should be encoded and maintained. (Braver 1999) 226.

Dopaminergic modulation of target neural systems in the prefrontal cortex produces both gating and learning effects. Gating effects ensure that representations of stimuli, implemented in transient neural assemblies, monopolise cognition. Learning effects are longer term consequences of dopamine produced by Hebbian modification of synaptic connections: when a representation is repeatedly sustained the relevant synaptic pathways are strengthened via a cascade of chemical changes which ultimately change the receptive properties of a target neurons (Braver 1999) p. 317). Computationally we can say that gating effects increase and maintain local activation and learning effects change the weights in a neural network.

To summarise, offline cognition requires the construction of a transient representation in a coalition of prefrontal-posterior neurones. Whether that representation survives depends on dopamine modulated gating and maintenance effects. Pathological versions of these effects can lead to offline cognition computing over inappropriately maintained representations. Dopamine delivery is thus a mechanism by which representations become *hypersalient*.

The best example of how this neurotransmitter-induced hypersalience produces psychiatric disorder is Shitij Kapur's account of the role dopamine plays in the psychotic symptoms of schizophrenia (Kapur 2003).

Section 5. Delusions dopamine and hypersalience

Kapur's starting point is the "dopamine hypothesis" of psychosis, the foundation of biological approaches. Evidence for the role of dopamine comes from two sources. The first is the fact that antipsychotic drugs are dopamine antagonists, whose effects are achieved by targeting dopamine receptors, the second is the role of dopamine in triggering psychosis (evidenced by hallucinogenic drugs and heightened dopamine synthesis during psychosis).

Kapur's account builds on recent theories of the role of dopamine which supersede an older idea that its role is essentially to encode representations of particular events or behaviours as rewarding by reinforcing them with pleasurable associations. Although it has a reward function, dopamine is not a simple hedonic reward system. In fact suppressing dopamine with antipsychotics does not change the hedonic associations of stimuli. Rather lack of dopamine means that people no longer pursue objects with positive affective valence. "i.e. they change the wanting without necessarily changing the liking" (Kapur 2003, p. 14). Furthermore dopamine is also involved in the representation of events with negative affective valence.

In the light of these facts Kapur endorses the hypothesis that

the mesolimbic dopamine system is seen as a critical component in the 'attribution of salience,' a process whereby events and thoughts come to grab attention, drive action and influence goal directed behaviour because of their association with reward or punishment... it provides an interface whereby the hedonic subjective pleasure, the ability to predict reward and the learning mechanisms allow the organism to focus on what it deems valuable and allows a seamless conversion of motivation into action (Kapur 2003, p. 14)

Kapur's theory integrates several features of schizophrenia, including the characteristic phenomenology of the prodromal period in which subjects feel that events or objects are extremely significant and/or that their senses are hypersensitive. As Kapur points out transient episodes of this nature are not abnormal (and an uncanny feeling is characteristic of dreams as we pointed out

earlier) but in delusional subjects dopamine dysregulation ensures that their hypersalience gives representations of objects or scenes a halo of significance.

Before we turn to Kapur's explanation of delusion note that his explanation has the potential to illuminate some puzzling phenomena which are obscured by theories which draw an artificially sharp distinction between the processes which produce an experience and the processes the subject uses to respond to experience.

Many inferential theorists are aware that the distinction between experience and response is hard to draw for cognitive theories but their solutions suffer from identifying offline cognition with an implementation of procedural rationality. For example "attribution" theories correctly note the role of mechanisms of paranoia or depression in some delusions (Beck 1989; Bentall 1994; Bentall 2004). These theories note that paranoid subjects focus on threatening stimuli and interpret neutral stimuli as threatening and depressive subjects focus on negative stimuli and interpret neutral stimuli as negative.

However when theorists come to model the processes by which these attentional biases influence cognition they reconstruct the delusional response as an instance of the way background hypotheses exert a confirmation bias in scientific theorising.

The delusional subject is equated to a paranoid scientific community interpreting evidence to support an existing theory.

As a device of exposition this is a very handy metaphor but it is not a description of a cognitive process, *unless delusional subjects actually implement the logic of scientific discovery in coping with their experience.*

It seems clear that these cases of attributional bias are the result of low-level neuromodulatory processes which affect the interaction between posterior and prefrontal systems. In depressive reasoning involving negative self attribution, for example, there is a cascade of changes consequent on norepinephrinergic dysregulation which modify amygdala-hippocampus loops (Ressler and Nemeroff 1999; Davidson et al. 2002). Similarly in the case of paranoia, the experience of stimuli as threatening is not the result of a theoretically constrained judgement but the influence on offline cognition of processes which put the organism into a vigilant aversive mode. In both cases the effect is to make particular types of representation hypersalient.

Kapur initially phrases his characterisation of delusional responses to experience neutrally as a "top down cognitive phenomenon that the individual imposes on these experiences of aberrant salience in order to make sense of them" (Kapur 2003, p. 15). Or as I have put it hallucinations are experiences produced by unusual activation of posterior systems which then monopolise frontal resources to become hypersalient, and delusions are responses to that experience produced by a mind in offline mode which cannot inhibit those experiences.

Kapur then follows the standard inferential approach in reconstructing the delusional response to hypersalient experiences as the adoption and maintenance of an explanatory hypothesis. The irrationality of delusion measured

against norms of inference is explained by the fact that, once adopted, the delusional hypothesis “serves as a cognitive scheme for further thoughts and actions. It drives the patients to find further confirmatory evidence – in the glances of strangers, the headlines of newspapers and the tiepins of news-readers”(Kapur 2003, p. 16). The experience constrains the initial abductive hypothesis, but does not determine it. Someone preoccupied by the experiences of alienation might attribute them to the occult ministrations of a shaman or to the CIA depending on their background: “the same neurochemical dysregulation leads to different phenomenological expression”(Kapur 2003, p. 15).

A natural way for an inferential theorist to accommodate Kapur’s account, adopted by Kapur himself would be to say that the delusional hypothesis acquires a level of credence which makes it effectively disconfirmable by processes of probabilistic reasoning. The effect of hypersalience is to give the delusional belief a probability approaching 100% with the result that other beliefs are revised to accommodate the delusional belief rather than the delusional belief being revised or rejected in the face of the web of background beliefs (Gerrans 2001).

One difficulty with this approach is that that in order to implement this type of procedural rationality a subject would have to reconfigure her offline pre-frontal posterior-posterior network in quite a radical way. She would have to treat hypersalient experiences as uninterpreted propositions and establish their degree of consistency with other background propositions using rules of inference.

However, delusional subjects are not engaged in persuading anyone, even themselves, of the truth or probability of their delusion by demonstrating its consistency with a set of other beliefs. Richard Bentall observed, after a lifetime’s experience with delusional subjects, that that “the unusual beliefs and experiences of psychiatric patients all seem to reflect preoccupations about the position of the self in the social universe”. Bentall is surely right that delusions are personal responses to experience. That is, the subject is using her own cognitive resources to respond to her experience rather than the impersonal ones of procedural rationality to demonstrate the truth of a theoretical hypothesis to others (Kaney and Bentall 1992; Lyon et al. 1994; Kinderman and Bentall 1996). And as we have seen those resources consist, in the first case in the offline manipulation of indexical representations by mental time travel.

Thus the cognitive culprit in delusion is the changes wrought in the representational architecture of offline cognition by dopaminergic or other influences. The effect as we saw of dopaminergic dysregulation is to make some representations monopolise offline cognition.

The delusional subject is confronted with an unusual experience typically as a result of sensory malfunction in a posterior subsystem (e.g. auditory hallucination in schizophrenia or face recognition in delusions of misidentification). That experience commands offline resources as part of the shift to mental time travel. But the delusional subject cannot inhibit the problem representation and allow others to assume salience in the same way as the normal person

would. For neurotransmitter delivery reconfigures her prefrontal-posterior network. It makes no sense to say that someone who is dreaming has a deficit in procedural rationality because dreaming is not a mode of cognition for which the theories of procedural rationality are appropriate explanations. Similarly, although delusions are not dreams, it is very likely that the mind of the delusional subject is not configured for procedural rationality. The delusional subject cannot make full use of decontextualised offline cognition because the declarative representations involved are interpreted by posterior representations some of which have acquired a dopaminergic halo.

In the delusional case the process which make an experience hypersalient also makes it harder to decontextualise sufficiently to apply procedural rationality to that experience because hypersalience is caused by mechanisms which affect inhibition.

Section 6. Evolutionary psychiatry

The evolutionary psychiatry movement (exemplified in the work of Price, Stevens and Charlton) aims to use evolutionary theory as the basis for a cognitive explanation of psychiatric disorder which makes no appeal to the unconscious psychological mechanisms postulated by psychoanalysis.

In their search for a biologically plausible foundation for cognitive theories of psychiatric disorder, evolutionary psychiatrists are drawn in two directions which we might call evolutionary ethology and evolutionary psychology. Ethology emphasises the role in development of affectively scaffolded cognitive and behavioural routines which regulate attachment, affiliation, cooperation and competition for resources and mates. The ethological approach invites us to explain human behaviour in terms of basic phylogenetically conserved mechanisms which influence higher level cognition “from below”.

A prime example of evolutionary ethology is John Bowlby’s attachment theory, which dovetails perfectly with the account of amnesia given in Section 1. For Bowlby insecure attachment produced a range of adverse effects on maturation and learning which subsequently influence mature patterns of behaviour. This type of theory lends itself to explanation in terms of mechanisms which underlie the behavioural responses to distress arising from insecure attachment. For example there is a correlation between insecure early attachment and the subsequent development of borderline personality disorder which invites investigation of the effect of sustained negative affect on cognitive development (Hobson 2002). It also invites investigation of the contribution of genetic factors which seem to buffer some infants against adverse environments.

Evolutionary psychology is a much more problematic theoretical ally for psychiatry. Evolutionary psychologists posit the presence of tacit inference systems evolved by hominids to cognise Pleistocene environments. (Cosmides and Tooby 1994). Such systems compute things like status, genetic distance

from conspecifics, reciprocal social obligations and solutions to problems requiring decision under uncertainty (Stanovich 1999; Stanovich 2000).

The evolutionary psychiatrists of the mid 1990s endorsed this approach to the mind in an attempt to root their psychological theories in biological reality. The idea that our conscious rationalisations of experience are influenced by more primitive inferential systems whose structure is opaque to introspection retains some of the apparent insights of psychoanalysis but tethers them to an apparently more biologically plausible psychological theory

However the unconscious mechanisms which play the most important role in psychiatric disorder are not repressed sexual drives or frustrations, or even inferential systems evolved to cope with Pleistocene environments. Rather they are the mechanisms which make experiences salient via an interplay between phylogenetically ancient systems and the recently evolved prefrontal cortex.

References

- Allman 1999. *Evolving Brains*. Scientific American Library, New York.
- Arbib M.E. 2003. *The Handbook of Brain Theory and Neural Networks*. London, MIT Press, Cambridge, Mass.
- Beck A.T. 1989. *Cognitive Therapy and the Emotional Disorders*. Peguin, Harmondsworth, UK.
- Bentall R.P. 1994. Cognitive biases and abnormal beliefs: towards a model of persecutory delusions. In: David and A., Cutting, J. (eds), *The Neuropsychology of Schizophrenia*. Hove, E. Sussex, Psychology Press, pp. 337–360.
- Bentall R.P. 2004. *Madness Explained*.
- Bennett M.R. 2000. The concept of long term potentiation of transmission at synapses 60(2):109–37. *Progress in Neurobiology* 60: 108–137.
- Braver T.S., Barch D.M. and Cohen J.D. 1999. Cognition and Control in Schizophrenia: a Computational Model of Dopamine and Prefrontal Function. *Biol. Psychiat.* 46: 312–328.
- Braver T.C.J. 1999. On the control of control: the role of dopamine in regulating prefrontal function and working memory. In: Monsell, S.D. (ed.), *Attention and Performance XVIII*, J. Cambridge, MA, MIT Press.
- Charlton B. 2000. *Psychiatry and the Human Condition*. Radcliffe Medical Press, Oxford, UK.
- Clark R. and Squire L.R. 1998. Classical conditioning and brain systems: the role of awareness. *Science* 280: 77–81.
- Cosmides L. and Tooby J. 1994. Beyond intuition and instinct blindness: Towards an evolutionarily rigorous cognitive science. *Cognition* 58: 1–73.
- Davidson R.J. and Pizzagalli D. et al. 2002. Depression: perspectives from affective neuroscience. *Annual Review of Psychology* 53: 545–574.
- Deák G.O., Ray S.D. and Brenneman K. 2003. Children's perseverative appearance-reality errors are related to emerging language skills. *Child Development* 74: 944–964.
- Eichenbaum H. 1998. Amnesia the hippocampus and episodic memory (editorial). *Hippocampus* 8: 197.
- Freud A. 1932. *The Ego and the Mechanisms of Defence*. Hogarth Press, London.
- Freud S. 1963. *Three Case Histories*. Collier Books, New York.
- Frey U. and Morris R.G.M. 1997. Synaptic tagging and long-term potentiation. *Nature* 385(6616): 533–5366.

- Friston K. 2002. Beyond Phrenology: What can neuroimaging tell us about distributed circuitry?. *Ann. Rev. Neurosci.* 25: 221–250.
- Fuster J.M. 1997. *The Prefrontal Cortex: Anatomy, Physiology and Neuropsychology of the Frontal Lobe*. Raven, New York.
- Gaiarsa J.L., Caillard O. and Ben-Ari Y. 2002. Long-term plasticity at GABAergic and glycinergic synapses: mechanisms and functional significance. *Trends Neurosci.* 25: 564–570.
- Garety P.A. and Hemsley D.R. et al. 1991. Reasoning in deluded schizophrenic and paranoid patients: Biases in performance on a probabilistic inference task. *J. Nervous Mental Disease* 179: 194–201.
- Gerrans P. 2000. Refining the explanation of Cotard's delusion. *Mind Language* 15(1): 111–122.
- Gerrans P. 2001. Delusions as performance failures. *Cognitive Neuropsychiatry*. 6(3): 161–173.
- Grace A. 1991. Phasic versus tonic dopamine release and the modulation of dopamine responsiveness: a hypothesis of the etiology of schizophrenia. *Neuroscience* 41: 1–24.
- Halligan P.W. and Marshall J.C. (eds), 1996. *Method in Madness: Case Studies in Cognitive Neuropsychiatry*. Psychology Press, Hove, East Sussex.
- Harrington D.L. and Haaland K.Y. et al. 1998. Cortical networks underlying mechanisms of time perception. *J. Neurosci.* 18: 1085–1095.
- Hobson A. 1999. *Dreaming as Delirium*. MIT Press, Cambridge, Mass.
- Hobson P. 2002. *The Cradle of Thought*. Pan Macmillan, London.
- Huq S.F. and Garety P.A. et al. 1988. Probabilistic judgements in deluded and non-deluded subjects. *Quart. J. Exp. Psychol.* 40(A): 801–813.
- Kandel E. 1999. Biology and the Future of Psychoanalysis: A New Intellectual Framework for Psychiatry Revisited. *Am. J. Psychiat.* 156: 505–523.
- Kaney S. and Bentall R.P. 1992. Persecutory delusions and the self-serving bias: Evidence from a contingency judgement task. *J. Nervous Mental Disease* 180: 773–780.
- Kapur J. 2003. Psychosis as a State of Aberrant Salience: A Framework for Linking Biology, Phenomenology and Pharmacology in Schizophrenia. *Am. J. Psychiat.* 160: 13–23.
- Kinderman P. and Bentall R.P. 1996. Self-discrepancies and persecutory delusions: Evidence for a model of paranoid ideation. *J. Abnormal Psychol.* 105: 106–113.
- Klein S.e.a. 2002. Memory and temporal experience: the effects of episodic memory loss on an amnesic patient's ability to remember the past and imagine the future. *Social Cognition* 20: 353–379.
- Knight R.T. 1999. Prefrontal cortex regulates inhibition and excitation in distributed neural networks. *Acta Psychiatrica Scandinavia* 101: 159–178.
- LeDoux J. 1999. *The Emotional Brain*. Phoenix, New York.
- Lyon H.M. and Kaney S. et al. 1994. The defensive function of persecutory delusions: Evidence from attribution tasks. *Br. J. Psychiat.* 164: 637–646.
- Maher B.A. 1992. Models and methods for the study of reasoning in delusions. *Eur. Rev. Appl. Psychol.* 42: 97–102.
- McKinney M. 2002. Brain Evolution by Stretching the Global Mitotic Clock of Development. In: McNamara N.M.-P.a.K. (ed.), *Human Evolution through Developmental Change*. Baltimore and London, Johns Hopkins Press.
- Mujica-Parodi L.R., Malapisina D. and Sackheim H.A. 2000. Logical processing, affect and delusion in schizophrenia. *Harvard Rev. Psychiat.* 8: 73–83.
- Nemeroff C. 1996. The corticotrophin releasing factor (CRF) hypothesis of depression: new findings and new directions. *Mol. Psychiat.* 1: 326–342.
- O'Brien G. O. and J., Elsevier (2004). Notes towards a structuralist theory of mental representation. In: Clapin P.S.P.S.H. (ed.), *In Representation in Mind: New Approaches to Mental Representation*. Elsevier.
- O'Reilly R.C. and Munakata Y. 2000. *Computational Explorations in Cognitive Neuroscience*. MIT Press, Cambridge, Mass.
- Plotsky P. and Meaney M.J. 1993. Early postnatal experience alters hypothalamic corticotrophin releasing factor (CRF) mRNA, median eminence CRF content and stress induced release in rats. *Brain Research Molecular Brain Research* 18: 195–200.

- Prabakaran V., Smith J., Desmond J., Glover G. and Gabrieli J. 1997. Neural Substrates of Fluid Reasoning: An fMRI Study of Neocortical Activation during Performance of the Raven's Progressive Matrices Test. *Cognitive Psychol.* 33: 43–63.
- Quartz S.R. and Sejnowski T.J. 1997. A Neural Basis of Cognitive Development: a neuroconstructivist manifesto. *Behav. Brain Sci.* 20: 537–556.
- Ressler K.J. and Nemeroff C.B. 1999. Role of norepinephrine in the pathophysiology and treatment of mood disorders. *Biol. Psychiat.* 46(9): 1219–1233.
- Sapolsky R. 1996. Why stress is bad for your brain. *Science* 273: 749–750.
- Seligman M.E.P. and Abramson L.Y. et al. 1979. Depressive attributional style. *J. Abnormal Psychol.* 88: 242–247.
- Stanovich 1999. *who is rational? studies of individual differences in reasoning*. Mahwah, New Jersey, Lawrence Erlbaum and Associates.
- Stanovich K. and West R. 2000. Individual differences in reasoning: implications for the rationality debate?. *Behav. Brain Sci.* 23: 643–726.
- Stein E. 1996. *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*. Oxford University Press, Oxford.
- Stevens A.P.J. 1996. *Evolutionary Psychiatry. A new beginning*. Routledge, London.
- Suddendorf T. and Busby J. 2003. Mental time travel in animals? *Trends Cognitive Sci.* 7: 391–396.
- Waltz J., Knowlton B., Holyoak K., Boone K., Mishkin F., De Menezes Santos M., Thomas C. and Miller B. 1999. A System for Relational Reasoning in the Prefrontal Cortex. *Psychol. Sci.* 10: 119–125.
- Wood J.N. and Grafman J. 2003. Human Prefrontal Cortex. *Nature Rev. Neurosci.* 4: 139–147.
- Zajonc R. 1980. Feeling and thinking: Preferences need no inferences. *American Psychologist* 35: 151–175.