EJN
European Journal of Neuroscience

European Journal of Neuroscience, Vol. 28, pp. 782–793, 2008 doi:10.1111/j.1460-9568.2008.06379.x

An attractor hypothesis of obsessive–compulsive disorder

Edmund T. Rolls,[1] Marco Loh[2] and Gustavo Deco[2,3]
[1]Oxford Centre for Computational Neuroscience, Oxford, UK
[2]Department of Technology, Computational Neuroscience, Universitas Pompeu Fabra, Barcelona, Spain
[3]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Keywords: emotion, glutamate, neural network, obsession, OCD, stability

Abstract

We propose a top-down approach to the symptoms of obsessive–compulsive disorder (OCD) based on a statistical dynamical framework. An increased depth in the basins of attraction of attractor network states in the brain makes each state too stable, so that it tends to remain locked in that state and can not easily be moved on to another state. We suggest that the different symptoms that may be present in OCD could be related to changes of this type in different brain regions. For example, the difficulty in attentional and cognitive set switching could be related to networks operating in this way in the prefrontal cortex. Repetitive actions and a difficulty in moving to new actions could be related to overstability in networks in the higher order motor, including cingulate areas. In integrate-and-fire network simulations, an increase in the *N*-methyl-D-aspartate (NMDA) and/or α-amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid (AMPA) receptor conductances, which increases the depth of the attractor basins, increases the stability of attractor networks, and makes them less easily moved on to another state by a new stimulus. Increasing γ-aminobutyric acid (GABA)-receptor activated currents can partly reverse this overstability. There is now some evidence for overactivity in glutamate transmitter systems in OCD, and the hypothesis presented here shows how some of the symptoms of OCD could be produced by the increase in the stability of attractor networks that is produced by increased glutamatergic activity.

Introduction

Obsessive–compulsive disorder (OCD) is a chronically debilitating disorder with a lifetime prevalence of 2–3% (Robins *et al.*, 1984; Karno *et al.*, 1988; Weissman *et al.*, 1994). It is characterized by two sets of symptoms, obsessive and compulsive. Obsessions are unwanted, intrusive, recurrent thoughts or impulses that are often concerned with themes of contamination and 'germs', checking household items in case of fire or burglary, order and symmetry of objects, or fears of harming oneself or others. Compulsions are ritualistic, repetitive behaviours or mental acts carried out in relation to these obsessions, e.g. washing, household safety checks, counting, rearrangement of objects in symmetrical array or constant checking of oneself and others to ensure no harm has occurred (Menzies *et al.*, 2008). Patients with OCD experience the persistent intrusion of thoughts that they generally perceive as foreign and irrational but which cannot be dismissed. The anxiety associated with these unwanted and disturbing thoughts can be extremely intense; it is often described as a feeling that something is incomplete or wrong, or that terrible consequences will ensue if specific actions are not taken. Many patients engage in repetitive, compulsive behaviours that aim to discharge the anxieties associated with these obsessional thoughts. Severely affected patients can spend many hours each day in their obsessional thinking and resultant compulsive behaviours, leading to marked disability (Pittenger *et al.*, 2006).

Correspondence: Professor E. T. Rolls, as above.
E-mail: Edmund.Rolls@oxcns.org; http://www.oxcns.org

Received 8 May 2008, revised 6 June 2008, accepted 23 June 2008

While patients with OCD exhibit a wide variety of obsessions and compulsions, the symptoms tend to fall into specific clusters. Common patterns include obsessions of contamination, with accompanying cleaning compulsions; obsessions with symmetry or order, with accompanying ordering behaviours; obsessions of saving, with accompanying hoarding; somatic obsessions; aggressive obsessions with checking compulsions; and sexual and religious obsessions (Pittenger *et al.*, 2006).

In this paper we describe a theory of how OCDs arise, and of the different symptoms. The theory is based on the top-down proposal that there is overstability of attractor neuronal networks in cortical and related areas in OCDs. The approach is top-down in that it starts with the set of symptoms and maps them onto the dynamical systems framework, and only after this considers detailed underlying biological mechanisms, of which there could be many, that might produce the effects. (In contrast, a complementary bottom-up approach starts from detailed neurobiological mechanisms, and aims to interpret their implications with a brain-like model for higher level phenomena.) We show by integrate-and-fire neuronal network simulations that the overstability could arise by, for example, overactivity in glutamatergic excitatory neurotransmitter synapses, which produces an increased depth of the basins of attraction, in the presence of which neuronal spiking-related and potentially other noise is insufficient to help the system move out of an attractor basin. We relate this top-down proposal, related to the stochastic dynamics of neuronal networks, to new evidence that there may be overactivity in glutamatergic systems in OCDs, and consider the implications for treatment.

There has been interest for some time in the application of complex systems theory to understanding brain function and behaviour

© The Authors (2008). Journal Compilation © Federation of European Neuroscience Societies and Blackwell Publishing Ltd

(Blackerby, 1993; Riley & Turvey, 2001; Peled, 2004; Heinrichs, 2005; Lewis, 2005). Dynamically realistic neuronal network models of working memory (Wang, 1999, 2001; Compte *et al.*, 2000; Durstewitz *et al.*, 2000a, b; Brunel & Wang, 2001) and decision-making (Wang, 2002; Deco & Rolls, 2006), and attention (Rolls & Deco, 2002; Deco & Rolls, 2005), and how they are influenced by neuromodulators such as dopamine (Durstewitz *et al.*, 1999, 2000a; Brunel & Wang, 2001; Durstewitz & Seamans, 2002; Deco & Rolls, 2003) have been produced, and it has been suggested that alterations in the attractor landscape may contribute to schizophrenia (Seamans & Yang, 2004; Winterer & Weinberger, 2004; Rolls, 2005). Indeed we have recently considered how a decreased depth in the basins of attraction of prefrontal cortical networks might contribute to, for example, the cognitive symptoms of schizophrenia, by making short-term memory systems unstable, and thereby contributing to the failure to maintain attention and the disorganization that characterizes the dysexecutive symptoms that are among the cognitive symptoms of schizophrenia (Loh *et al.*, 2007). This decreased depth in the basins of attraction in schizophrenia is attributed inter alia to decreased currents in *N*-methyl-D-aspartate (NMDA) receptor-activated channels (Coyle, 2006), and the associated decrease in firing rates when present in the orbitofrontal and anterior cingulate cortices is related to the negative symptoms, such as anhedonia, and flattening of affect, motivation and volition, which contribute to the psychomotor poverty in schizophrenia according to this computational approach (Rolls, 2005; Loh *et al.*, 2007).

Attractor networks, and their stability

The attractor framework is based on dynamical systems theory. In a network of interconnected neurons, a memory pattern (represented by a set of active neurons) can be stored by synaptic modification, and later recalled by external inputs. Furthermore, a pattern activated by an input is then stably maintained by the system even after input offset. These patterns could correspond to memories, perceptual representations or thoughts.

The architecture of an attractor or autoassociation network is as follows (Fig. 1A). External inputs e_i activate the neurons in the network, and produce firing y_i, where i refers to the ith neuron. The neurons are connected to each other by recurrent collateral synapses w_{ij}, where j refers to the jth synapse on a neuron. By these synapses an input pattern on e_i is associated with itself, and thus the network is referred to as an autoassociation network. Because there is positive feedback implemented via the recurrent collateral connections, the network can sustain persistent firing. These synaptic connections are assumed to build up by an associative (Hebbian) learning mechanism (Hebb, 1949). The inhibitory interneurons are not shown. They receive inputs from the pyramidal cells, and make inhibitory negative feedback connections onto the pyramidal cells to keep their activity under control. Hopfield (1982) showed that the recall state in a simple attractor network can be thought of as the local minimum in an energy landscape, where the energy would be defined as

$$E = -\frac{1}{2}\sum_{i,j} w_{ij}(y_i - <y>)(y_j - <y>). \quad (1)$$

Autoassociation attractor systems have two types of stable fixed points: a spontaneous state with a low firing rate, and one or more attractor states with high firing rates in which the positive feedback implemented by the recurrent collateral connections maintains a high firing rate. We sometimes refer to this latter state as the persistent state. The area in the energy landscape within which the system will move to a stable attractor state is called its basin of attraction.

The attractor dynamics can be pictured by energy landscapes, which indicate the basin of attraction by valleys, and the attractor states or fixed points by the bottom of the valleys. [Although energy functions apply to recurrent networks with symmetric connections between the neurons (Hopfield, 1982), as would be the case in a fully connected network with associative synaptic modification, and do not necessarily apply to more complicated networks with, for example, incomplete connectivity, nevertheless the properties of these other recurrent networks are similar (Treves, 1991; Treves & Rolls, 1991; Rolls & Treves, 1998), and the concept of an energy function and landscape is useful for discussion purposes. In practice, a Lyapunov function can be used to prove analytically that there is a stable fixed point such as an attractor basin (Khalil, 1996), and even in systems where this can not be proved analytically, it may still be possible to show numerically that there are stable fixed points, to measure the flow towards those fixed points that describes the depth of the attractor basin as we have done for this type of network (Loh *et al.*, 2007), and to use the concept of energy or potential landscapes to help visualize the properties of the system.] The stability of an attractor is characterized by the average time in which the system stays in the basin of attraction under the influence of noise. The noise provokes transitions to other attractor states. One source of noise results from the interplay between the Poissonian character of the spikes and the finite-size effect due to the limited number of neurons in the network. Two factors determine the stability. First, if the depths of the attractors are shallow (as in the left compared with the right valley in Fig. 1B), then less force is needed to move a ball from one valley to the next. Second, high noise will make it more likely that the system will jump over an energy boundary from one state to another. We envision that the brain as a dynamical system has characteristics of such an attractor system, including statistical fluctuations. The noise could arise not only from the probabilistic spiking of the neurons, which has significant effects in finite size integrate-and-fire networks (Deco & Rolls, 2006), but also from any other source of noise in the brain or the environment (Faisal *et al.*, 2008), including the effects of distracting stimuli.

A hypothesis about the increased stability of attractor networks and the symptoms of OCD

We hypothesize that cortical and related attractor networks become too stable in OCD, so that once in an attractor state, the networks tend to remain there too long. The hypothesis is that the depths of the basins of attraction become deeper, and that this is what makes the attractor networks more stable. We further hypothesize that part of the mechanism for the increased depth of the basins of attraction is increased glutamatergic transmission, which increases the depth of the basins of attraction by increasing the firing rates of the neurons, and by increasing the effective value of the synaptic weights between the associatively modified synapses that define the attractor, as is made evident in Eq. (1) above. The synaptic strength is effectively increased if more glutamate is released per action potential at the synapse, or if in other ways the currents injected into the neurons through the NMDA and/or α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) synapses are larger. In addition, if NMDA receptor function is increased, this could also increase the stability of the system because of the temporal smoothing effect of the long time constant of the NMDA receptors (Wang, 1999).

This increased stability of cortical and related attractor networks, and the associated higher neuronal firing rates, could occur in different brain regions, and thereby produce different symptoms, as follows.

© The Authors (2008). Journal Compilation © Federation of European Neuroscience Societies and Blackwell Publishing Ltd

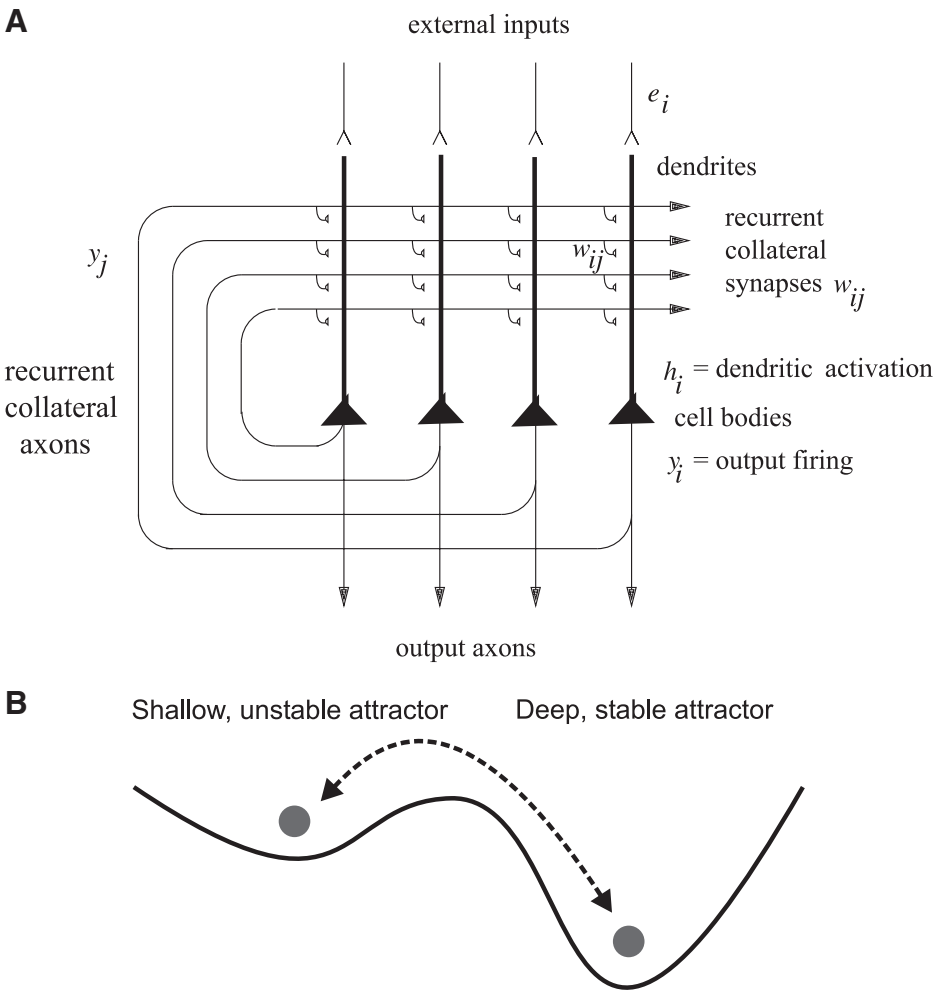

FIG. 1. (A) Architecture of an attractor network. External inputs e_i activate the neurons in the network, and produce firing y_i, where i refers to the ith neuron. The neurons are connected by recurrent collateral synapses w_{ij}, where j refers to the jth synapse on a neuron. By these synapses an input pattern on e_i is associated with itself, and thus the network is referred to as an autoassociation network. Because there is positive feedback via the recurrent collateral connections, the network can sustain persistent firing. These synaptic connections are assumed to be formed by an associative (Hebbian) learning mechanism. The inhibitory interneurons are not shown. They receive inputs from the pyramidal cells, and make negative feedback connections onto the pyramidal cells to control their activity. The recall state (which could be used to implement short-term memory or memory recall) in an attractor network can be thought of as the local minimum in an energy landscape. (B) Energy landscape. The first basin (from the left) in the energy landscape is the spontaneous state, and the second basin is the high firing rate attractor state, which is 'persistent' in that the neurons that implement it continue firing. The vertical axis of each landscape is the energy potential. The horizontal axis is the firing rate, with high to the right. In the normal condition, the valleys for both the spontaneous and the high firing attractor state are equally deep, making both states stable. In the situation that is hypothesized to be related to some of the symptoms of OCD, the basin for the high firing attractor state is deep, making the high firing rate attractor state that implements for example short-term memory too stable and very resistant to distraction. This increased depth of the basin of attraction of the persistent state may be associated with higher firing rates of the neurons, if for example the state is produced by increased currents in NMDA receptors. In general, there will be many different high firing rate attractor basins, each corresponding to a different memory.

If these effects occurred in high-order motor areas, the symptoms could include inability to move out of one motor pattern, resulting for example in repeated movements or actions. In parts of the cingulate cortex and dorsal medial prefrontal cortex, this could result in difficulty in switching between actions or strategies (Rushworth *et al.*, 2007a, b), as the system would be locked into one action or strategy. If an action was locked into a high-order motor area due to increased stability of an attractor network, then lower order motor areas might thereby not be able to escape easily what they implement, such as a sequence of movements, so that the sequence would be repeated.

A similar account, of becoming locked in one action and having difficulty in switching to another action, can be provided for response inhibition deficits, which have been found in OCD. The response inhibition deficit has been found in tasks such as go/no-go and stop-signal reaction time (SSRT), which examine motor inhibitory processes, and also the Stroop task, a putative test of cognitive inhibition (Hartston & Swerdlow, 1999; Bannon *et al.*, 2002, 2006; Penades *et al.*, 2005, 2007; Chamberlain *et al.*, 2006, 2007). For example, response inhibition deficits have been reported in patients with OCD when performing the SSRT, which measures the time taken to internally suppress pre-potent motor responses (Chamberlain *et al.*, 2006). Unaffected first-degree relatives of patients with OCD are also impaired on this task compared with unrelated healthy controls, suggesting that response inhibition may be an endophenotype (or intermediate phenotype) for OCD (Chamberlain *et al.*, 2007; Menzies *et al.*, 2008).

If occurring in the lateral prefrontal cortex (including the dorsolateral and ventrolateral parts), the increased stability of attractor

© The Authors (2008). Journal Compilation © Federation of European Neuroscience Societies and Blackwell Publishing Ltd

networks could produce symptoms that include a difficulty in shifting attention and in cognitive set shifting. These are in fact important symptoms that can be found in OCD (Menzies *et al.*, 2008). These have been concerned with two quite different forms of shift: affective set shifting, where the affective or reward value of a stimulus changes over time (e.g. a rewarded stimulus is no longer rewarded; intradimensional or ID set shifting); and attentional set shifting, where the stimulus dimension (e.g. shapes or colours) to which the subject must attend is changed (extradimensional or ED set shifting). Deficits of attentional set shifting in OCD have been found in several neurocognitive studies using the CANTAB ID/ED set shifting task (Veale *et al.*, 1996; Watkins *et al.*, 2005; Chamberlain *et al.*, 2006, 2007). This deficit is most consistently reported at the ED stage (in which the stimulus dimension, e.g. shape, colour or number, alters and subjects have to inhibit their attention to this dimension and attend to a new, previously irrelevant dimension). The ED stage is analogous to the stage in the Wisconsin Card Sorting Task where a previously correct rule for card sorting is changed and the subject has to respond to the new rule (Berg, 1948). This ED shift impairment in patients with OCD is considered to reflect a lack of cognitive or attentional flexibility and may be related to the repetitive nature of OCD symptoms and behaviours. Deficits in attentional set shifting are considered to be more dependent upon dorsolateral and ventrolateral prefrontal regions than the orbital prefrontal regions included in the orbitofronto-striatal model of OCD (Pantelis *et al.*, 1999; Rogers *et al.*, 2000; Nagahama *et al.*, 2001; Hampshire & Owen, 2006), suggesting that cognitive deficits in OCD may not be underpinned exclusively by orbitofrontal cortex pathology. Indeed, ID or affective set shifting may not be consistently impaired in OCD (Menzies *et al.*, 2008).

Planning may also be impaired in patients with OCD (Menzies *et al.*, 2008), and this could arise because there is too much stability of attractor networks in the dorsolateral prefrontal cortex concerned with holding in mind the different short-term memory representations that encode the different steps of a plan (Rolls, 2008b). Indeed, there is evidence for dorsolateral prefrontal cortex dysfunction in patients with OCD, in conjunction with impairment on a version of the Tower of London, a task often used to probe planning aspects of executive function (van den Heuvel *et al.*, 2005). Impairment on the Tower of London task has also been demonstrated in healthy first-degree relatives of patients with OCD (Delorme *et al.*, 2007).

An increased firing rate of neurons in the orbitofrontal cortex and anterior cingulate cortex, produced by hyperactivity of glutamatergic transmitter systems, would increase emotionality, which is frequently found in OCD. Part of the increased anxiety found in OCD could be related to an inability to complete tasks or actions in which one is locked. But part of our unifying proposal is that part of the increased emotionality in OCD may be directly related to increased firing produced by the increased glutamatergic activity in brain areas such as the orbitofrontal and anterior cingulate cortex. The orbitofrontal cortex and anterior cingulate cortex are involved in emotion, in that they are activated by primary and secondary reinforcers that produce affective states (Rolls, 2004, 2005, 2008a), and in that damage to these regions alters emotional behaviour and emotional experience (Rolls *et al.*, 1994; Hornak *et al.*, 1996, 2003). Indeed, negative emotions as well as positive emotions activate the orbitofrontal cortex, with the emotional states produced by negative events tending to be represented in the lateral orbitofrontal cortex and dorsal part of the anterior cingulate cortex (Kringelbach & Rolls, 2004; Rolls, 2005, 2008a). We may note that stimulus-reinforcer reversal tasks (also known as ID shifts or affective reversal) are not generally impaired in patients with OCD (Menzies *et al.*, 2008), and this is as predicted, for the machinery for the reversal including the detection of non-reward (Rolls, 2005) is present even if the system is hyperglutamatergic.

If the increased stability of attractor networks occurred in temporal lobe semantic memory networks, then this would result in a difficulty in moving from one thought to another, and possibly in stereotyped thoughts, which again may be a symptom of OCD (Menzies *et al.*, 2008).

The obsessional states are thus proposed to arise because cortical areas concerned with cognitive functions have states that become too stable. The compulsive states are proposed to arise partly in response to the obsessional states, but also partly because cortical areas concerned with actions have states that become too stable. The theory provides a unifying computational account of both the obsessional and compulsive symptoms, in that both arise due to increased stability of cortical attractor networks, with the different symptoms related to overstability in different cortical areas. The theory is also unifying in that a similar increase in glutamatergic activity in the orbitofrontal and anterior cingulate cortex could increase emotionality, as described above.

Alterations to glutamatergic transmitter systems that may increase the depth of the basins of attraction of cortical and related attractor networks

To demonstrate how alterations of glutamate as a transmitter for the connections between the neurons may influence the stability of attractor networks, we performed integrate-and-fire simulations. A feature of these simulations is that we simulated the currents produced by activation of NMDA and AMPA receptors in the recurrent collateral synapses, and took into account the effects of the spike-related noise, which is an important factor in determining whether the attractor stays in a basin of attraction or jumps over an energy barrier into another basin (Loh *et al.*, 2007). These attractors are likely to be implemented in many parts of the cerebral cortex by the recurrent collateral connections between pyramidal cells, and have short-term memory properties with basins of attraction that allow systematic investigation of stability and distractibility. The particular neural network implementation we adopt includes channels activated by AMPA, NMDA and γ-aminobutyric acid $(GABA)_A$ receptors, and allows not only the spiking activity to be simulated but also a consistent mean-field approach to be used (Brunel & Wang, 2001).

Methods

Our aim is to investigate stability and distractibility in a biophysically realistic attractor framework, so that the properties of receptors, synaptic currents and the statistical effects related to the noisy probabilistic spiking of the neurons can be part of the model. We use a minimal architecture, a single attractor or autoassociation network (Hopfield, 1982; Amit, 1989; Hertz *et al.*, 1991; Rolls & Treves, 1998; Rolls & Deco, 2002). We chose a recurrent (attractor) integrate-and-fire network model, which includes synaptic channels for AMPA, NMDA and $GABA_A$ receptors (Brunel & Wang, 2001). The integrate-and-fire model is necessary to characterize and exploit the effects of the spiking noise produced by the neurons in a finite-sized network. However, to initialize the parameters of the integrate-and-fire model such as the synaptic connection strengths to produce stable attractors, and to ensure that the spontaneous activity is in the correct range, we used a mean-field approximation consistent with the integrate-and-fire network, as described in the Supplementary material, Appendix S1.

© The Authors (2008). Journal Compilation © Federation of European Neuroscience Societies and Blackwell Publishing Ltd

Both excitatory and inhibitory neurons are represented by a leaky integrate-and-fire model (Tuckwell, 1988). The basic state variable of a single model neuron is the membrane potential. It decays in time when the neurons receive no synaptic input down to a resting potential. When synaptic input causes the membrane potential to reach a threshold, a spike is emitted and the neuron is set to the reset potential at which it is kept for the refractory period. The emitted action potential is propagated to the other neurons in the network. The excitatory neurons transmit their action potentials via the AMPA and NMDA glutamatergic receptors, which are both modelled by their effect in producing exponentially decaying currents in the postsynaptic neuron. The rise time of the AMPA current is neglected, because it is typically very short. The NMDA channel is modeled with an alpha function including both a rise and a long decay term. In addition, the synaptic function of the NMDA current includes a voltage dependence controlled by the extracellular magnesium concentration (Jahr & Stevens, 1990). The inhibitory postsynaptic potential is mediated by a $GABA_A$ receptor model and is described by a decay term.

The single attractor network contains 400 excitatory and 100 inhibitory neurons, which is consistent with the observed proportions of pyramidal cells and interneurons in the cerebral cortex (Abeles, 1991; Braitenberg & Schütz, 1991). The connection strengths are adjusted using mean-field analysis (Brunel & Wang, 2001), so that the excitatory and inhibitory neurons exhibit a spontaneous activity of 3 Hz and 9 Hz, respectively (Koch & Fuster, 1989; Wilson *et al.*, 1994). The recurrent excitation mediated by the AMPA and NMDA receptors is dominated by the long time constant NMDA currents to avoid instabilities during the delay periods (Wang, 1999, 2002).

Our cortical network model features a minimal architecture to investigate stability and distractibility, and consists of two selective pools S1 and S2 (Fig. 2A). We use just two selective pools to eliminate possible disturbing factors. Pool S1 is used for the short-term memory item to be remembered, sometimes called the target; and pool S2 is used for the distractor. The non-selective pool NS models the spiking of cortical neurons and serves to generate an approximately Poisson spiking dynamics in the model (Brunel & Wang, 2001), which is what is observed in the cortex. The inhibitory pool IH contains the 100 inhibitory neurons. There are thus four populations or pools of neurons in the network, and the connection weights are set up as described next using a mean-field analysis to make S1 and S2 have stable attractor properties. The connection weights between the neurons of each selective pool or population are called the intra-pool connection strengths w_+. The increased strength of the intra-pool connections is counterbalanced by the other excitatory connections (w_-) to keep the average input to a neuron constant.

The network receives Poisson input spikes via AMPA receptors, which are envisioned to originate from 800 external neurons at an average spontaneous firing rate of 3 Hz from each external neuron, consistent with the spontaneous activity observed in the cerebral cortex (Wilson *et al.*, 1994; Rolls & Treves, 1998). A detailed mathematical description is provided in the supplementary Appendix S1.

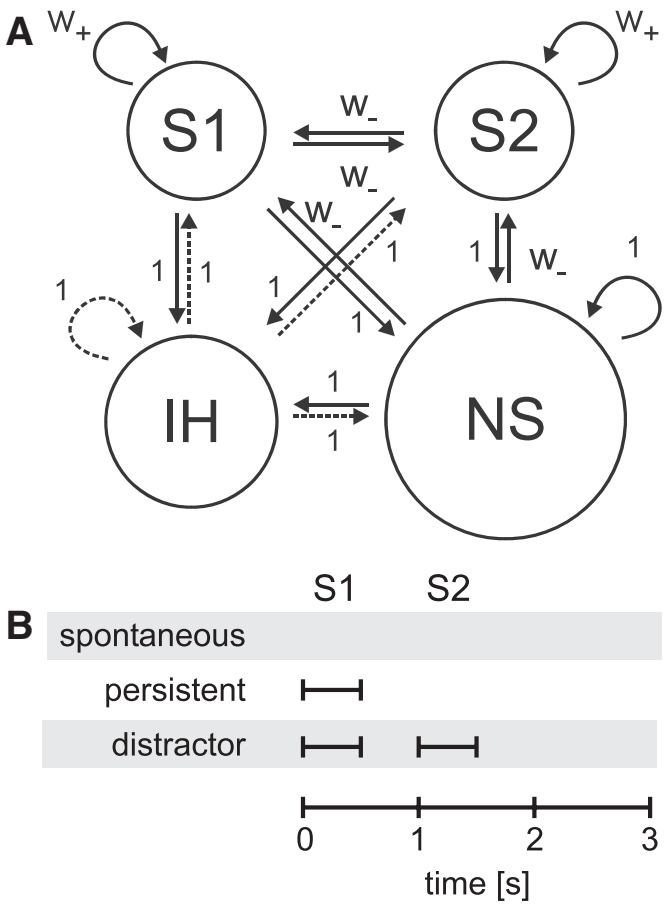

FIG. 2. (A) The attractor network model. The excitatory neurons are divided into two selective pools S1 and S2 (with 40 neurons each) with strong intra-pool connection strengths w_+ and one non-selective pool (NS) (with 320 neurons). The other connection strengths are 1 or weak w_-. The network contains 500 neurons, of which 400 are in the excitatory pools and 100 are in the inhibitory pool IH. The network also receives inputs from 800 external neurons, and these neurons increase their firing rates to apply a stimulus or distractor to one of the pools S1 or S2. The supplementary Appendix S1 contains the synaptic connection matrices. (B) The simulation protocols. Stimuli to either the S1 or S2 population of neurons are applied at different times depending on the type of simulations. The spontaneous simulations include no input. The persistent simulations assess how stably a stimulus is retained by the network. The distractor simulations add a distractor stimulus to further address the stability of the network activity, when it has been started by S1.

Analysis

Our analysis is targeted to investigate the stability and distractibility with respect to NMDA and GABA receptor modulations. Multiple trial integrate-and-fire spiking simulations were used to integrate the complete neural and synaptic dynamics over time, including statistical components of the network model. Therefore the spiking simulations are needed to assess the stability and distractibility of the dynamical system, for this depends in part on the statistical fluctuations that occur in a network of spiking neurons (Deco & Rolls, 2006). This is done by simulating a network configuration for several trials each run with different random seeds and running a statistical analysis on the data.

We simulate three different conditions: the spontaneous, persistent and distractor conditions (Fig. 2B).

In spontaneous simulations, we run spiking simulations for 3 s without any extra external input. The aim of this condition is to test whether the network is stable in maintaining a low average firing rate in the absence of any inputs, or whether it falls into one of its attractor states without any external input.

In persistent simulations, an external cue of 120 Hz above the background firing rate of 2400 Hz is applied to each neuron in pool S1

© The Authors (2008). Journal Compilation © Federation of European Neuroscience Societies and Blackwell Publishing Ltd

during the first 500 ms to induce a high activity state and then the system is run for another 2.5 s. The 2400 Hz is distributed across the 800 synapses of each S1 neuron for the external inputs, with the spontaneous Poisson spike trains received by each synapse thus having a mean rate of 3 Hz. The aim of this condition is to investigate whether once in an attractor short-term memory state the network can maintain its activity stably, or whether it falls out of its attractor, which might correspond to an inability to maintain attention.

The distractor simulations start off like the persistent simulations with a 500 ms input to pool S1 to start the S1 short-term memory attractor state, but between 1 s and 1.5 s we apply a distracting input to pool S2 with varying strengths. The aim of this condition is to measure how distractible the network is. The degree of distractibility is measured parametrically by the strength of the input to S2 required to remove the high activity state of the S1 population. These simulation protocols serve to assess the generic properties of the dynamical attractor system rather than to model specific experimental data obtained in particular paradigms.

The mean-field approach (described in the supplementary Appendix S1) was used to calculate the synaptic weights to set the normal conditions for the operation of the network to be as follows. For the spontaneous state, the conditions for the numerical simulations of the mean-field method were set to 3 Hz for all excitatory pools and 9 Hz for the inhibitory pool. These values correspond to the approximate values of the spontaneous attractors when the network is not driven by stimulus-specific inputs. For the persistent state, a selective pool was set to a higher initial value (30 Hz) to account for the excitation of these neurons during the preceding cue period.

Results

To clarify the concept of stability, we show examples of trials of spontaneous and persistent simulations in which the statistical fluctuations have different impacts on the temporal dynamics. Figure 3 shows the possibilities, as follows.

In the spontaneous state simulations, no cue is applied, and we are interested in whether the network remains stably in the spontaneous firing state, or whether it is unstable and on some trials due to statistical fluctuations enters one of the attractors, thus falsely retrieving a memory. Figure 3 (top) shows an example of a trial on which the network correctly stays in the low spontaneous firing rate regime, and (bottom) another trial (labelled spontaneous unstable) in which statistical spiking-related fluctuations in the network cause it to enter a high activity state, moving into one of the attractors even without a stimulus.

In the persistent state simulations, a strong excitatory input is given to the S1 neuronal population between 0 and 500 ms (see 'Analysis' section). Two such trials are shown in Fig. 3. In Fig. 3 (top), the S1 neurons (correctly) keep firing at approximately 30 Hz after the retrieval cue is removed at 500 ms. However, due to statistical fluctuations in the network related to the spiking activity, on the trial labelled persistent unstable the high firing rate in the attractor for S1 was not stable, and the firing decreased back towards the spontaneous level, in the example shown starting after 1.5 s (Fig. 3 bottom). This trial illustrates the failure to maintain a stable short-term memory state, even when no distractor is applied.

In Fig. 3 the transitions to the incorrect activity states are caused by statistical fluctuations in the spiking activity of the integrate-and-fire neurons. We hypothesize that the stability of the high firing rate

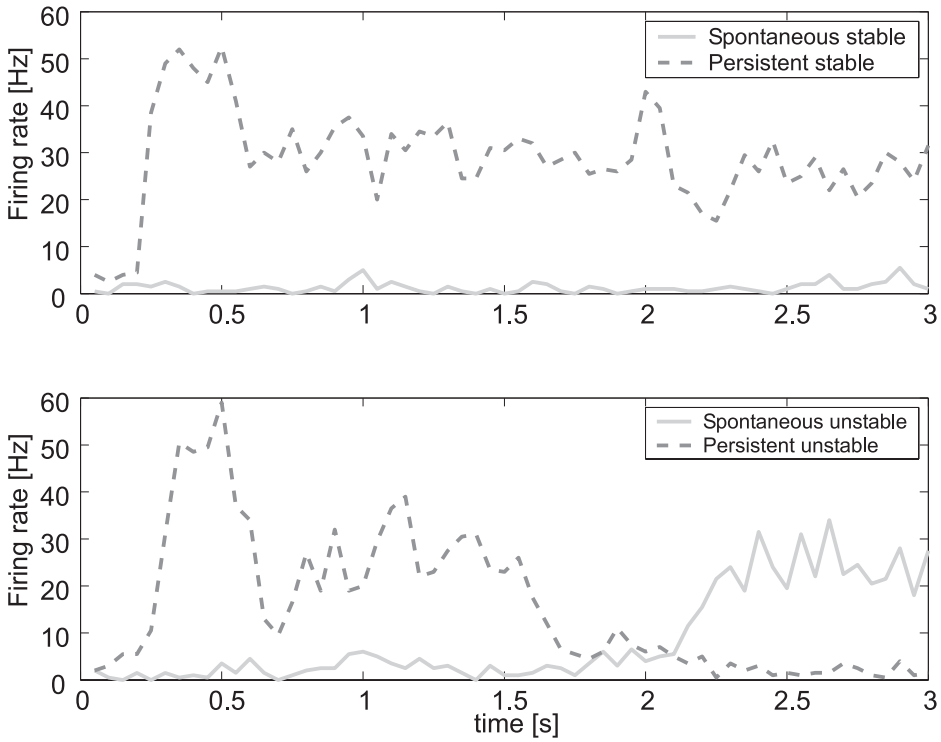

FIG. 3. Example trials of the integrate-and-fire attractor network simulations. The average firing rate of all the neurons in the S1 pool is shown. Top: normal operation. On a trial in which a recall stimulus was applied to S1 at 0–500 ms, firing continued normally until the end of the trial in the 'persistent' simulation condition. On a trial on which no recall stimulus was applied to S1, spontaneous firing continued until the end of the trial in the 'spontaneous' simulation condition. Bottom: unstable operation. On this persistent condition trial, the firing decreased during the trial as the network fell out of the attractor because of the statistical fluctuations caused by the spiking dynamics. On the spontaneous condition trial, the firing increased during the trial because of the statistical fluctuations. In these simulations the network parameter was $w_+ = 2.1$.

© The Authors (2008). Journal Compilation © Federation of European Neuroscience Societies and Blackwell Publishing Ltd

'persistent' attractor state may be increased in OCD, and that in addition the spontaneous state may be less likely to remain low, but may jump into a high firing rate attractor state. We note that there are two sources of noise in the spiking networks that cause the statistical fluctuations: the randomly arriving external Poisson spike trains, and the statistical fluctuations caused by the spiking of the neurons in the finite-sized network. The magnitude of these fluctuations increases as the number of neurons in the network becomes smaller (Mattia & Del Giudice, 2004).

For our investigations, we selected $w_+ = 2.1$, which with the default values of the NMDA and GABA conductances yielded stable dynamics, that is, a stable spontaneous state if no retrieval cue was applied, and a stable state of persistent firing after a retrieval cue had been applied and removed. To investigate the effects of changes (modulations) in the NMDA, AMPA and GABA conductances, we chose for demonstration purposes increases of 3% for the NMDA, and 10% for the AMPA and GABA synapses between the neurons in the network shown in Fig. 2A, as these were found to be sufficient to alter the stability of the attractor network. A strength of our approach is that we show that even quite small increases in the synaptic currents can alter the global behaviour of the network, e.g. the stability of its attractors.

We assessed how the stability of both the spontaneous and persistent states changes when NMDA and AMPA efficacies are modulated. Specifically we ran multiple trial integrate-and-fire network simulations and counted how often the system maintained the spontaneous or persistent state, assessed by the firing rate in the last second of the simulation (2–3 s) of each 3-s trial. Figure 4 shows the stability of the spontaneous and persistent attractor. We plot the % of the simulation runs on which the network during the last second of the simulation was in the high firing rate attractor state. Figure 4 shows that for the persistent run simulations, in which the cue triggered the attractor into the high firing rate attractor state, the network was still in the high firing rate attractor state in the baseline condition on approximately 88% of the runs, and that this had increased to 98% when the NMDA conductances were increased by 3% (+NMDA). Thus, increasing the NMDA receptor-activated synaptic currents increased the stability of the network. The effect was highly significant, with the means ± SEM shown. Figure 4 also shows that increasing AMPA by 10% (+AMPA) could also increase the stability of the persistent high firing rate attractor state, as did the combination +NMDA +AMPA.

Figure 4 shows that in the baseline condition the spontaneous state was unstable on approximately 10% of the trials, that is, on 10% of the trials the spiking noise in the network caused the network run in the condition without any initial retrieval cue to end up in a high firing rate attractor state. This is of course an error that is related to the spiking noise in the network. In the +NMDA condition, the spontaneous state had jumped to the high firing rate attractor state on approximately 22% of the runs, that is the low firing rate spontaneous state was present at the end of a simulation on only approximately 78% of the runs. Thus, increasing NMDA receptor-activated currents can contribute to the network jumping from what should be a quiescent state of spontaneous activity into a high firing rate attractor state. We relate this to the symptoms of OCDs, in that the system can jump into a state with a

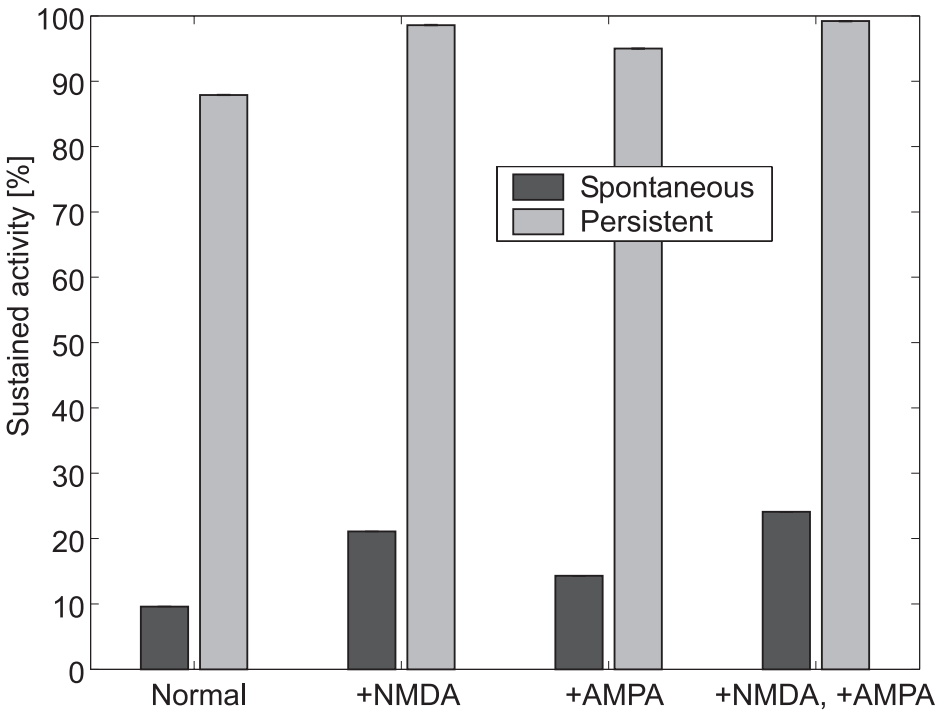

FIG. 4. The stability of the spontaneous and persistent attractor states. The percentage of the simulation runs on which the network during the last second of the 3-s simulation was in the high firing rate attractor state is shown on the ordinate. For the persistent run simulations, in which the cue triggered the attractor into the high firing rate attractor state, the network was still in the high firing rate attractor state in the baseline condition on approximately 88% of the runs, and this increased to nearly 100% when the *N*-methyl-D-aspartate (NMDA) conductances were increased by 3% (+NMDA). The effect was highly significant, with the means ± SEM shown. Increasing α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) by 10% (+AMPA) could also increase the stability of the persistent high firing rate attractor state, as did the combination +NMDA +AMPA. For the spontaneous state simulations, in the baseline condition the spontaneous state was unstable on approximately 10% of the trials, that is, on 10% of the trials the spiking noise in the network caused the network run in the condition without any initial retrieval cue to end up in a high firing rate attractor state. In the +NMDA condition, the spontaneous state had jumped to the high firing rate attractor state on 25% of the runs, that is the low firing rate spontaneous state was present at the end of a simulation on only approximately 75% of the runs. The +AMPA can also make the spontaneous state more likely to jump to a persistent high firing rate attractor state, as can the combination +NMDA +AMPA.

© The Authors (2008). Journal Compilation © Federation of European Neuroscience Societies and Blackwell Publishing Ltd

dominant memory (which might be an idea or concern or action) even when there is no initiating input. Figure 4 also shows that +AMPA can make the spontaneous state more likely to jump to a persistent high firing rate attractor state, as can the combination +NMDA +AMPA.

We next investigated to what extent alterations of the GABA receptor-mediated inhibition in the network could restore the system towards more normal activity even when NMDA and/or GABA conductances were high. Figure 5 shows that increasing the GABA currents by 10% when the NMDA currents are increased by 3% (+NMDA +GABA) can move the persistent state away from overstability back to the normal baseline state. That is, instead of the system ending up in the high firing rate attractor state in the persistent state simulations on 98% of the runs, the system ended up in the high firing rate attractor state on approximately 88% of the runs, the baseline level. Increasing GABA has a large effect on the stability of the spontaneous state, making it less likely to jump to a high firing rate attractor state. The combination +NMDA +GABA produced a spontaneous state in which the +GABA overcorrected for the effect of +NMDA. That is, in the +NMDA +GABA condition the network was very likely to stay in the spontaneous firing rate condition in which it was started or, equivalently, when tested in the spontaneous condition, the network was less likely than normally to jump to a high firing rate attractor. Increasing GABA thus corrected for the effect of increasing NMDA receptor-activated synaptic currents on the persistent type of run when there was an initiating stimulus; and overcorrected for the effect of increasing NMDA on the spontaneous state simulations when there was no initiating retrieval stimulus, and the network should remain in the low firing rate state until the end of the simulation run. The implications for symptoms are that agents that increase GABA conductances might reduce and normalize the tendency to remain locked into an idea or concern or action; and would make it much less likely that the quiescent resting state would be left by jumping because of the noisy spiking towards a state representing a dominant idea or concern or action. The effects of increasing GABA receptor-activated currents alone was to make the persistent simulations less stable (less likely to end in a high firing rate state) and the spontaneous simulations to be more stable (more likely to end up in the spontaneous state).

We next investigated how an increase of NMDA currents might make the system less distractible, and overstable in remaining in an attractor. This was investigated as shown in Fig. 2B by setting up a system with two high firing rate attractors, S1 and S2, then starting the network in an S1 attractor state with S1 applied at t = 0–0.5 s, and then applying a distractor S2 at time t = 1–1.5 s to investigate how strong S2 had to be to distract the network out of its S1 attractor. We assessed how often in 200 trials the average activity during the last second (2–3 s) stayed above 10 Hz in the S1 attractor (% sustained activity shown on the ordinate of Fig. 6). The strength of the distractor stimulus applied to S2 was an increase in firing rate above the 2.4 kHz background activity, which is distributed among 800 synapses per neuron. Figure 6 shows that increasing the NMDA receptor-activated currents by 5% (+NMDA) means that larger distractor currents must be applied to S2 to move the system away from S1 to S2. That is, the +NMDA condition makes the system more stable in its high firing rate attractor, and less able to be moved to another state by another stimulus (in this case S2). We relate this to the symptoms of OCD, in that once in an attractor state (which might reflect an idea or concern or action), it is very difficult to get the system to move to another state. Increasing AMPA receptor-activated synaptic currents (by 10%, +AMPA) produces similar, but smaller, effects.

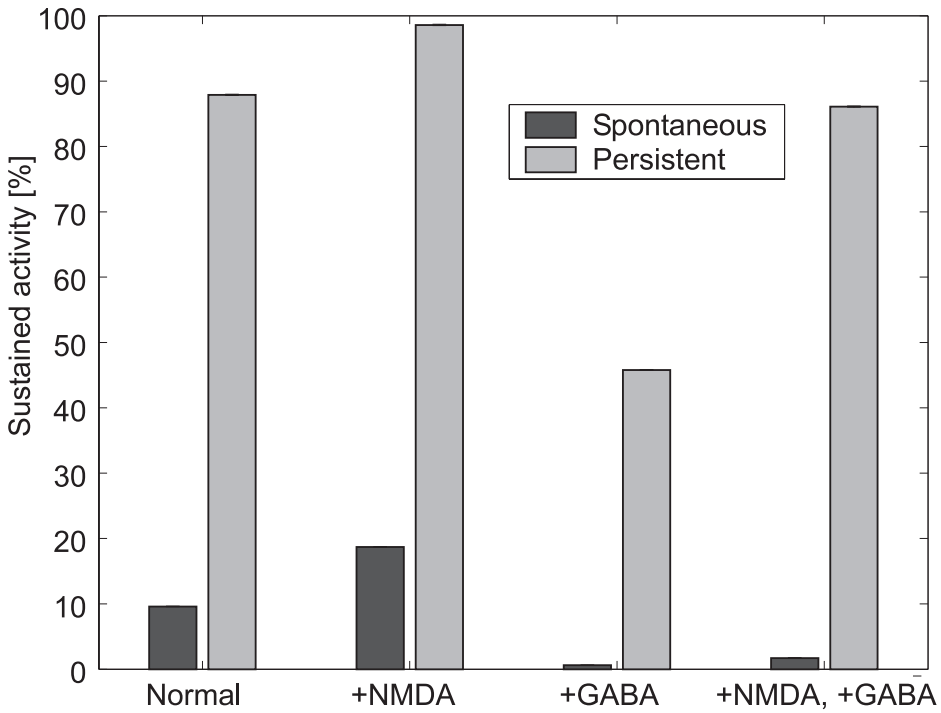

FIG. 5. The effect of increasing γ-aminobutyric acid (GABA)-receptor mediated synaptic conductances by 10% (+GABA) on the stability of the network. Conventions as in Fig. 4. Increasing the GABA currents by 10% when the *N*-methyl-D-aspartate (NMDA) currents are increased by 3% (+NMDA +GABA) moved the persistent state away from the overstability produced by +NMDA alone, and returned the persistent state to the normal baseline level. That is, instead of the system ending up in the high firing rate attractor state in the persistent state simulations on 98% of the runs in the +NMDA condition, the system ended up in the high firing rate attractor state on approximately 88% of the runs in the +NMDA +GABA condition. The combination +NMDA +GABA produced a spontaneous state that was less likely than the normal state to jump to a high firing rate attractor (see text).

© The Authors (2008). Journal Compilation © Federation of European Neuroscience Societies and Blackwell Publishing Ltd

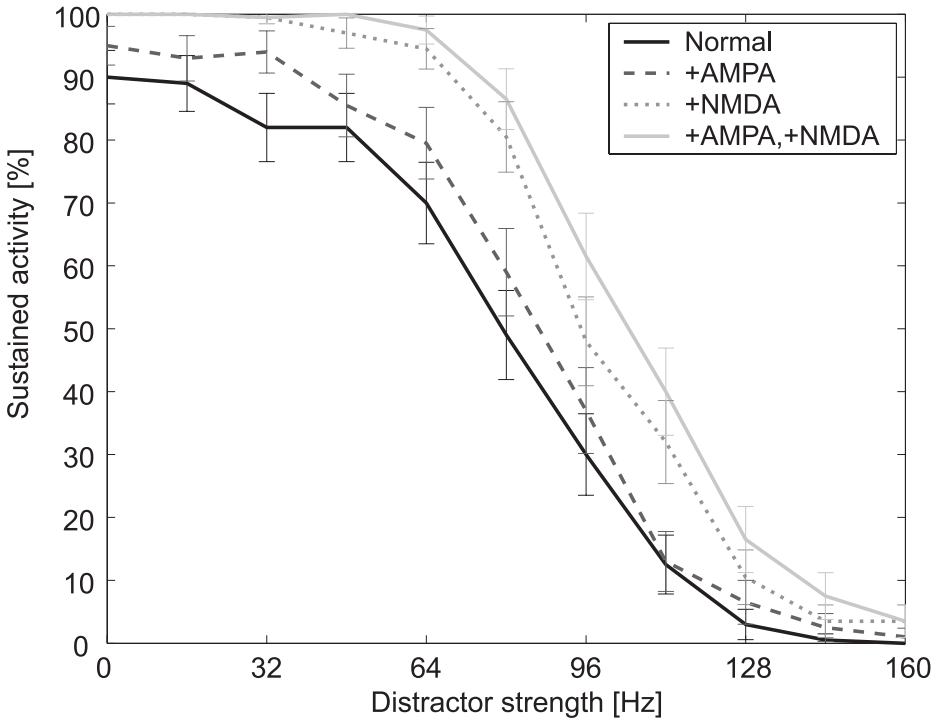

FIG. 6. Stability and distractibility as a function of the distractor strength and the synaptic efficacies. We assessed how often in 200 trials the average activity during the last second (2–3 s) stayed above 10 Hz in the S1 attractor, instead of being distracted by the S2 distractor. The modulation of the synapses +*N*-methyl-D-aspartate (NMDA) and +α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) corresponds to an increase of 5% and 10%, respectively. The strength of the distractor stimulus applied to S2 is an increase in firing rate above the 2.4 kHz background activity that is distributed among 800 synapses per neuron. The higher the sustained activity is in S1 the higher is the stability and the lower is the distractibility by S2. The standard deviations were approximated with the binomial distribution.

Figure 7 shows that increasing GABA-activated synaptic conductances (by 10%, +GABA) can partly normalize the overstability and decrease the distractibility that is produced by elevating NMDA receptor-activated synaptic conductances (by 3%, +NMDA). The investigation and conventions used the same protocol as for the simulations shown in Fig. 6. We assessed how often in 200 trials the average activity during the last second (2–3 s) stayed above 10 Hz in the S1 attractor, instead of being distracted by the S2 distractor. The strength of the distractor stimulus applied to S2 is an increase in firing rate above the 2.4 kHz background activity that is distributed among 800 synapses per neuron. The higher the sustained activity is in S1 the higher is the stability to S1 and the lower is the distractibility by S2. What was found (as shown in Fig. 7) is that the increase of GABA was able to move the curve in part back towards the normal state away from the overstable and less distractible state produced by increasing NMDA receptor-activated synaptic currents. The effect was particularly evident at low distractor currents.

Discussion

This is a new approach to the symptoms of OCD, for it deals with the symptoms in terms of overstability of attractor networks in the cerebral cortex. If the same generic change in stability was produced in different cortical areas, then we have indicated how different symptoms might arise. Of course, if these changes were more evident in some areas than in others in different patients, this would help to account for the different symptoms in different patients. Having proposed a generic hypothesis for the disorder, we recognize of course that the exact symptoms that arise if stability in some systems is increased will be subject to the exact effects that these will have in an individual patient, who may react to these effects, and produce explanatory accounts for the effects, and ways to deal with them, that may be quite different from individual to individual.

The integrate-and-fire simulations show that an increase of NMDA or AMPA synaptic currents can increase the stability of attractor networks. They can become so stable that the intrinsic stochastic noise caused by the spiking of the neurons is much less effective in moving the system from one state to another, whether spontaneously in the absence of a distracting, that is new, stimulus, or in the presence of a new (distracting) stimulus that would normally move the system from one attractor state to another. The simulations also show that the stability of the spontaneous (quiescent) firing rate state is reduced by increasing NMDA or AMPA receptor-activated synaptic currents. We relate this to the symptoms of OCD in that the system is more likely than normally, under the influence of the spiking stochastic noise caused by the neuronal spiking in the system, to jump into one of the dominant attractor states, which might be a recurring idea or concern or action.

This simulation evidence, that an increase of glutamatergic synaptic efficacy can increase the stability of attractor networks and thus potentially provide an account for some of the symptoms of OCD, is consistent with evidence that glutamatergic function may be increased in some brain systems in OCD (Rosenberg *et al.*, 2000, 2001, 2004; Pittenger *et al.*, 2006), and that cerebrospinal fluid glutamate levels are elevated (Chakrabarty *et al.*, 2005). Consistent with this, agents with anti-glutamatergic activity such as riluzole, which can decrease glutamate transmitter release, may be efficacious in OCD (Pittenger *et al.*, 2006; Bhattacharyya & Chakraborty, 2007). Further evidence for a link between glutamate as a neurotransmitter and OCD comes from genetic studies. There is evidence for a significant association

© The Authors (2008). Journal Compilation © Federation of European Neuroscience Societies and Blackwell Publishing Ltd

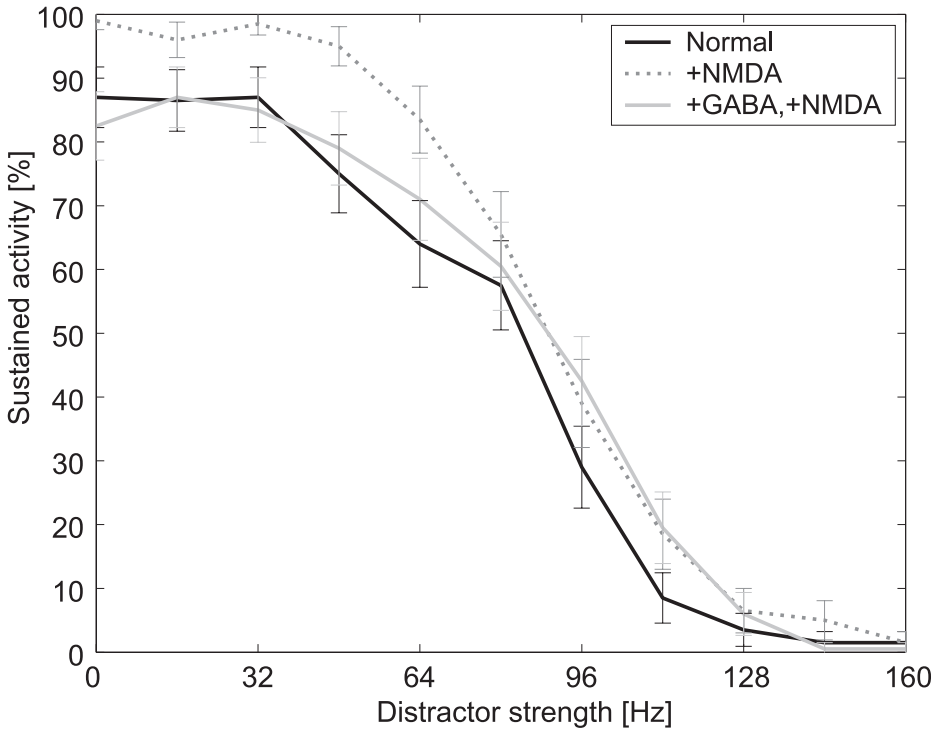

FIG. 7. Stability and distractibility: the effect of increasing γ-aminobutyric acid (GABA) by 10% (+GABA) on the overstability to S1/decrease in distractibility by S2 produced by increasing *N*-methyl-D-aspartate (NMDA) synaptic conductances by 3% (+NMDA). We assessed how often in 200 trials the average activity during the last second (2–3 s) stayed above 10 Hz in the S1 attractor, instead of being distracted to the S2 attractor state by the S2 distractor. The modulation of the synapses +NMDA and +GABA corresponds to an increase of 3% and 10%, respectively. The strength of the distractor stimulus applied to S2 is an increase in firing rate above the 2.4 kHz background activity that is distributed among 800 synapses per neuron. The higher the sustained activity is in S1 the higher is the stability to S1 and the lower is the distractibility by S2. The standard deviations were approximated with the binomial distribution.

between the SLC1A1 glutamate transporter gene and OCD (Stewart *et al.*, 2007). This transporter is crucial in terminating the action of glutamate as an excitatory neurotransmitter and in maintaining extracellular glutamate concentrations within a normal range (Bhattacharyya & Chakraborty, 2007). In addition, Arnold *et al.* (2004) postulated that NMDA receptors were involved in OCD, and specifically that polymorphisms in the 3′ untranslated region of GRIN2B (glutamate receptor, ionotropic, NMDA 2B) were associated with OCD in affected families.

We have described results with a single attractor network. It is natural to think of such a network as being implemented by associatively modifiable recurrent collateral synaptic connections between the pyramidal cells within a few mm of each other in a cortical area. Such circuitry is characteristic of the cerebral neocortex, and endows it with many remarkable and fundamental properties, including the ability to implement short-term memories, which are fundamental to cognitive processes such as top-down attention, and planning (Rolls, 2008b). However, when two cortical areas are connected by forward and backward connections, which are likely to be associatively modifiable, then this architecture, again characteristic of the cerebral neocortex (Rolls, 2008b), also provides the basis for the implementation of an attractor network, but with now a different set of glutamatergic synapses in which overactivity could be relevant to the theory described here. Further, there are likely to be interactions between connected attractor networks (Rolls & Deco, 2002; Deco & Rolls, 2003, 2006; Rolls, 2008b), and these are also relevant to the way in which the theory described here would be implemented in the brain.

While patients with OCD exhibit a wide variety of obsessions and compulsions, the symptoms tend to fall into specific clusters. Common patterns include obsessions of contamination, with accompanying cleaning compulsions; obsessions with symmetry or order, with accompanying ordering behaviours; obsessions of saving, with accompanying hoarding; somatic obsessions; aggressive obsessions with checking compulsions; and sexual and religious obsessions (Pittenger *et al.*, 2006). Although it is not part of the current theory, which addresses how cortical states of activity may become too stable in OCD, to account for the tendency for certain types of obsessive state to be common, we note the following. The types of obsessive state that are common may, in at least some cases, reflect states that would be useful to have in normal amounts as they have biological utility. For example, a tendency to avoid contamination, to hoard food and other goods, and to check whether there are dangers in the vicinity could all have biological utility in an appropriate amount, and selection could have produced predispositions to have these concerns or goals (Rolls, 2005). These concerns share in common that they may emerge when there are not high priority calls for action, such as the immediate need to find food to eat because of hunger, or to flee from or fight an aggressor. Given that these concerns then may have a chance of being expressed when there are not high-priority calls for immediate action, it seems possible that if cortical states do become too stable, then it is concerns of this type that might become too prominent, leading to obsessive thoughts, and thus to compulsive behaviour. The suggestion is that because these states about which obsessions are likely are not part of actions being performed to meet immediate environmental needs, then these states do not have anything particular to displace them, and can thus remain present for considerable periods even under normal circumstances. An exacerbation of attractor network stability when there is no particular

© The Authors (2008). Journal Compilation © Federation of European Neuroscience Societies and Blackwell Publishing Ltd

environmental need for immediate alternative action may in this way lead to this type of state becoming obsessional.

One of the interesting aspects of the present model of OCD is that it is in a sense the opposite of a recent model of schizophrenia (Loh *et al.*, 2007). In that model, we suggest that decreased NMDA synaptic currents may, by decreasing the firing rates and effectively the synaptic strengths in cortical attractor networks, decrease the stability of cortical attractor networks. This, if effected in the dorsolateral prefrontal cortex, might produce some of the cognitive symptoms of schizophrenia, including distractibility, poor attention and poor short-term memory. If the same generic effect were present in the temporal lobe semantic memory networks, this might make thoughts move too freely from one thought to another loosely associated thought, producing bizarre thoughts and some of the positive symptoms of schizophrenia (providing an alternative to the parasitic or spurious attractor hypothesis; Hoffman & McGlashan, 2001). If implemented in the orbitofrontal cortex, the decrease of NMDA synaptic currents would produce lower neuronal firing, and thus less affect, which is a negative symptom of schizophrenia. These two models, of OCD and schizophrenia, in being formal opposites of each other may underline the importance of attractor networks to cortical function (Rolls, 2008b), and emphasize that any alteration in their stability, whether towards too little stability or too much, will have major consequences for the operation of the cerebral cortex.

The theory about how the symptoms of OCD could arise in relation to the increased stability of cortical attractor networks has implications for possible pharmaceutical approaches to treatment. One is that treatments that reduce glutamatergic activity, for example by partially blocking NMDA receptors, might be useful. Another is that increasing the inhibition in the cortical system, for example by increasing GABA receptor-activated synaptic currents, might be useful, both by bringing the system from a state where it was locked into an attractor back to the normal level, and by making the spontaneous state more stable, so that it would be less likely to jump to an attractor state (which might represent a dominant idea or concern or action; see Fig. 5). However, we emphasize that the way in which the network effects we consider produce the symptoms in individual patients will be complex, and will depend on the way in which each person may deal cognitively with the effects. In this respect, we emphasize that what we describe is a theory of OCD, that the theory must be considered in the light of empirical evidence yet to be obtained, and that cognitive behaviour therapy makes an important contribution to the treatment of such patients. We hope that the theory will stimulate further thinking and research in this area.

Supplementary material

The following supplementary material may be found on http://www.blackwell-synergy.com

Appendix S1. Supporting information: neural and synaptic dynamics.

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the correspondence author for the article.

Acknowledgements

M.L. was supported by the Boehringer Ingelheim Fonds. Support was also provided by the Oxford McDonnell Centre for Cognitive Neuroscience.

Abbreviations

AMPA, α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid; ED, extra-dimensional; GABA, γ-aminobutyric acid; ID, intradimensional; NMDA, *N*-methyl-D-aspartate; OCD, obsessive–compulsive disorder; SSRT, stop-signal reaction time.

References

Abeles, M. (1991) *Corticonics – Neural Circuits of the Cerebral Cortex*. Cambridge University Press, New York.

Amit, D.J. (1989) *Modeling Brain Function*. Cambridge University Press, Cambridge.

Arnold, P.D., Rosenberg, D.R., Mundo, E., Tharmalingam, S., Kennedy, J.L. & Richter, M.A. (2004) Association of a glutamate (NMDA) subunit receptor gene (GRIN2B) with obsessive–compulsive disorder: a preliminary study. *Psychopharmacology*, **174**, 530–538.

Bannon, S., Gonsalvez, C.J., Croft, R.J. & Boyce, P.M. (2002) Response inhibition deficits in obsessive–compulsive disorder. *Psychiatry Res.*, **110**, 165–174.

Bannon, S., Gonsalvez, C.J., Croft, R.J. & Boyce, P.M. (2006) Executive functions in obsessive–compulsive disorder: state or trait deficits? *Aust. N. Z. J. Psychiatry*, **40**, 1031–1038.

Berg, E. (1948) A simple objective technique for measuring flexibility in thinking. *J. Gen. Psychol.*, **39**, 15–22.

Bhattacharyya, S. & Chakraborty, K. (2007) Glutamatergic dysfunction – newer targets for anti-obsessional drugs. *Recent Patents CNS Drug Discov.*, **2**, 47–55.

Blackerby, R.F. (1993) *Application of Chaos Theory to Psychological Models*. Performance Strategies Publications, Austin, TX, USA.

Braitenberg, V. & Schütz, A. (1991) *Anatomy of the Cortex*. Springer-Verlag, Berlin.

Brunel, N. & Wang, X.J. (2001) Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J. Comput. Neurosci.*, **11**, 63–85.

Chakrabarty, K., Bhattacharyya, S., Christopher, R. & Khanna, S. (2005) Glutamatergic dysfunction in OCD. *Neuropsychopharmacology*, **30**, 1735–1740.

Chamberlain, S.R., Fineberg, N.A., Blackwell, A.D., Robbins, T.W. & Sahakian, B.J. (2006) Motor inhibition and cognitive flexibility in obsessive–compulsive disorder and trichotillomania. *Am. J. Psychiatry*, **163**, 1282–1284.

Chamberlain, S.R., Fineberg, N.A., Menzies, L.A., Blackwell, A.D., Bullmore, E.T., Robbins, T.W. & Sahakian, B.J. (2007) Impaired cognitive flexibility and motor inhibition in unaffected first-degree relatives of patients with obsessive–compulsive disorder. *Am. J. Psychiatry*, **164**, 335–338.

Compte, A., Brunel, N., Goldman-Rakic, P.S. & Wang, X.J. (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb. Cortex*, **10**, 910–923.

Coyle, J.T. (2006) Glutamate and schizophrenia: beyond the dopamine hypothesis. *Cell. Mol. Neurobiol.*, **26**, 365–384.

Deco, G. & Rolls, E.T. (2003) Attention and working memory: a dynamical model of neuronal activity in the prefrontal cortex. *Eur. J. Neurosci.*, **18**, 2374–2390.

Deco, G. & Rolls, E.T. (2005) Neurodynamics of biased competition and co-operation for attention: a model with spiking neurons. *J. Neurophysiol.*, **94**, 295–313.

Deco, G. & Rolls, E.T. (2006) Decision-making and Weber's Law: a neurophysiological model. *Eur. J. Neurosci.*, **24**, 901–916.

Delorme, R., Gousse, V., Roy, I., Trandafir, A., Mathieu, F., Mouren-Simeoni, M.C., Betancur, C. & Leboyer, M. (2007) Shared executive dysfunctions in unaffected relatives of patients with autism and obsessive–compulsive disorder. *Eur. Psychiatry*, **22**, 32–38.

Durstewitz, D. & Seamans, J.K. (2002) The computational role of dopamine D1 receptors in working memory. *Neural Netw.*, **15**, 561–572.

Durstewitz, D., Kelc, M. & Gunturkun, O. (1999) A neurocomputational theory of the dopaminergic modulation of working memory functions. *J. Neurosci.*, **19**, 2807–2822.

Durstewitz, D., Seamans, J.K. & Sejnowski, T.J. (2000a) Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *J. Neurophysiol.*, **83**, 1733–1750.

Durstewitz, D., Seamans, J.K. & Sejnowski, T.J. (2000b) Neurocomputational models of working memory. *Nat. Neurosci.*, **3**(Suppl.), 1184–1191.

Faisal, A.A., Selen, L.P. & Wolpert, D.M. (2008) Noise in the nervous system. *Nat. Rev.*, **9**, 292–303.

© The Authors (2008). Journal Compilation © Federation of European Neuroscience Societies and Blackwell Publishing Ltd

Hampshire, A. & Owen, A.M. (2006) Fractionating attentional control using event-related fMRI. *Cereb. Cortex*, **16**, 1679–1689.
Hartston, H.J. & Swerdlow, N.R. (1999) Visuospatial priming and Stroop performance in patients with obsessive compulsive disorder. *Neuropsychology*, **13**, 447–457.
Hebb, D.O. (1949) *The Organization of Behavior: A Neuropsychological Theory*. Wiley, New York.
Heinrichs, D.W. (2005) Antidepressants and the chaotic brain: implications for the respectful treatment of selves. *Philos. Psychiatr. Psychol.*, **12**, 215–227.
Hertz, J., Krogh, A. & Palmer, R.G. (1991) *An Introduction to the Theory of Neural Computation*. Addison-Wesley, Wokingham.
van den Heuvel, O.A., Veltman, D.J., Groenewegen, H.J., Cath, D.C., van Balkom, A.J., van Hartskamp, J., Barkhof, F. & van Dyck, R. (2005) Frontal-striatal dysfunction during planning in obsessive–compulsive disorder. *Arch. Gen. Psychiatry*, **62**, 301–309.
Hoffman, R.E. & McGlashan, T.H. (2001) Neural network models of schizophrenia. *Neuroscientist*, **7**, 441–454.
Hopfield, J.J. (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl Acad. Sci. USA*, **79**, 2554–2558.
Hornak, J., Rolls, E.T. & Wade, D. (1996) Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal lobe damage. *Neuropsychologia*, **34**, 247–261.
Hornak, J., Bramham, J., Rolls, E.T., Morris, R.G., O'Doherty, J., Bullock, P.R. & Polkey, C.E. (2003) Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain*, **126**, 1691–1712.
Jahr, C.E. & Stevens, C.F. (1990) Voltage dependence of NMDA-activated macroscopic conductances predicted by single-channel kinetics. *J. Neurosci.*, **10**, 3178–3182.
Karno, M., Golding, J.M., Sorenson, S.B. & Burnam, M.A. (1988) The epidemiology of obsessive–compulsive disorder in five US communities. *Arch. Gen. Psychiatry*, **45**, 1094–1099.
Khalil, H.K. 1996. *Nonlinear Systems*. Prentice Hall, Upper Saddle River, NJ.
Koch, K.W. & Fuster, J.M. (1989) Unit activity in monkey parietal cortex related to haptic perception and temporary memory. *Exp. Brain Res.*, **76**, 292–306.
Kringelbach, M.L. & Rolls, E.T. (2004) The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog. Neurobiol.*, **72**, 341–372.
Lewis, M.D. (2005) Bridging emotion theory and neurobiology through dynamic systems modeling. *Behav. Brain sci.*, **28**, 169–194; discussion 194–245.
Loh, M., Rolls, E.T. & Deco, G. (2007) A dynamical systems hypothesis of schizophrenia. *PLoS Comput. Biol.*, **3**, e228. doi:210.1371/journal.pcbi.0030228.
Mattia, M. & Del Giudice, P. (2004) Finite-size dynamics of inhibitory and excitatory interacting spiking neurons. *Phys. Rev.*, **70**, 052903.
Menzies, L., Chamberlain, S.R., Laird, A.R., Thelen, S.M., Sahakian, B.J. & Bullmore, E.T. (2008) Integrating evidence from neuroimaging and neuropsychological studies of obsessive–compulsive disorder: The orbitofronto-striatal model revisited. *Neurosci. Biobehav. Rev.*, **32**, 525–549.
Nagahama, Y., Okada, T., Katsumi, Y., Hayashi, T., Yamauchi, H., Oyanagi, C., Konishi, J., Fukuyama, H. & Shibasaki, H. (2001) Dissociable mechanisms of attentional control within the human prefrontal cortex. *Cereb. Cortex*, **11**, 85–92.
Pantelis, C., Barber, F.Z., Barnes, T.R., Nelson, H.E., Owen, A.M. & Robbins, T.W. (1999) Comparison of set-shifting ability in patients with chronic schizophrenia and frontal lobe damage. *Schizophr. Res.*, **37**, 251–270.
Peled, A. (2004) From plasticity to complexity: a new diagnostic method for psychiatry. *Med. Hypotheses*, **63**, 110–114.
Penades, R., Catalan, R., Andres, S., Salamero, M. & Gasto, C. (2005) Executive function and nonverbal memory in obsessive–compulsive disorder. *Psychiatry Res.*, **133**, 81–90.
Penades, R., Catalan, R., Rubia, K., Andres, S., Salamero, M. & Gasto, C. (2007) Impaired response inhibition in obsessive compulsive disorder. *Eur. Psychiatry*, **22**, 404–410.
Pittenger, C., Krystal, J.H. & Coric, V. (2006) Glutamate-modulating drugs as novel pharmacotherapeutic agents in the treatment of obsessive–compulsive disorder. *NeuroRx*, **3**, 69–81.
Riley, M.A. & Turvey, M.T. (2001) The self-organizing dynamics of intentions and actions. *Am. J. Psychol.*, **114**, 160–169.
Robins, L.N., Helzer, J.E., Weissman, M.M., Orvaschel, H., Gruenberg, E., Burke, J.D. Jr & Regier, D.A. (1984) Lifetime prevalence of specific psychiatric disorders in three sites. *Arch. Gen. Psychiatry*, **41**, 949–958.
Rogers, R.D., Andrews, T.C., Grasby, P.M., Brooks, D.J. & Robbins, T.W. (2000) Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. *J. Cogn. Neurosci.*, **12**, 142–162.
Rolls, E.T. (2004) The functions of the orbitofrontal cortex. *Brain Cogn.*, **55**, 11–29.
Rolls, E.T. (2005) *Emotion Explained*. Oxford University Press, Oxford.
Rolls, E.T. (2008a) The anterior and midcingulate cortices and reward. In Vogt, B.A. (ed.), *Cingulate Neurobiology & Disease*. Oxford University Press, Oxford, in press.
Rolls, E.T. (2008b) *Memory, Attention, and Decision-Making: A Unifying Computational Neuroscience Approach*. Oxford University Press, Oxford.
Rolls, E.T. & Deco, G. (2002) *Computational Neuroscience of Vision*. Oxford University Press, Oxford.
Rolls, E.T. & Treves, A. (1998) *Neural Networks and Brain Function*. Oxford University Press, Oxford.
Rolls, E.T., Hornak, J., Wade, D. & McGrath, J. (1994) Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *J. Neurol. Neurosurg. Psychiatry*, **57**, 1518–1524.
Rosenberg, D.R., MacMaster, F.P., Keshavan, M.S., Fitzgerald, K.D., Stewart, C.M. & Moore, G.J. (2000) Decrease in caudate glutamatergic concentrations in pediatric obsessive–compulsive disorder patients taking paroxetine. *J. Am. Acad. Child Adolesc. Psychiatry*, **39**, 1096–1103.
Rosenberg, D.R., MacMillan, S.N. & Moore, G.J. (2001) Brain anatomy and chemistry may predict treatment response in paediatric obsessive–compulsive disorder. *Int. J. Neuropsychopharmacol.*, **4**, 179–190.
Rosenberg, D.R., Mirza, Y., Russell, A., Tang, J., Smith, J.M., Banerjee, S.P., Bhandari, R., Rose, M., Ivey, J., Boyd, C. & Moore, G.J. (2004) Reduced anterior cingulate glutamatergic concentrations in childhood OCD and major depression versus healthy controls. *J. Am. Acad. Child Adolesc. Psychiatry*, **43**, 1146–1153.
Rushworth, M.F., Behrens, T.E., Rudebeck, P.H. & Walton, M.E. (2007a) Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends Cogn. Sci*, **11**, 168–176.
Rushworth, M.F., Buckley, M.J., Behrens, T.E., Walton, M.E. & Bannerman, D.M. (2007b) Functional organization of the medial frontal cortex. *Curr. Opin. Neurobiol.*, **17**, 220–227.
Seamans, J.K. & Yang, C.R. (2004) The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog. Neurobiol.*, **74**, 1–58.
Stewart, S.E., Fagerness, J.A., Platko, J., Smoller, J.W., Scharf, J.M., Illmann, C., Jenike, E., Chabane, N., Leboyer, M., Delorme, R., Jenike, M.A. & Pauls, D.L. (2007) Association of the SLC1A1 glutamate transporter gene and obsessive–compulsive disorder. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **144**, 1027–1033.
Treves, A. (1991) Are spin-glass effects relevant to understanding realistic auto-associative networks. *J. Physics A*, **24**, 2645–2654.
Treves, A. & Rolls, E.T. (1991) What determines the capacity of autoassociative memories in the brain? *Network*, **2**, 371–397.
Tuckwell, H. (1988) *Introduction to Theoretical Neurobiology*. Cambridge University Press, Cambridge.
Veale, D.M., Sahakian, B.J., Owen, A.M. & Marks, I.M. (1996) Specific cognitive deficits in tests sensitive to frontal lobe dysfunction in obsessive–compulsive disorder. *Psychol. Med.*, **26**, 1261–1269.
Wang, X.J. (1999) Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *J. Neurosci.*, **19**, 9587–9603.
Wang, X.J. (2001) Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.*, **24**, 455–463.
Wang, X.J. (2002) Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, **36**, 955–968.
Watkins, L.H., Sahakian, B.J., Robertson, M.M., Veale, D.M., Rogers, R.D., Pickard, K.M., Aitken, M.R. & Robbins, T.W. (2005) Executive function in Tourette's syndrome and obsessive–compulsive disorder. *Psychol. Med.*, **35**, 571–582.
Weissman, M.M., Bland, R.C., Canino, G.J., Greenwald, S., Hwu, H.G., Lee, C.K., Newman, S.C., Oakley-Browne, M.A., Rubio-Stipec, M., Wickramaratne, P.J., Wittchen, H. & Yeh, E.K. (1994) The cross national epidemiology of obsessive compulsive disorder. The Cross National Collaborative Group. *J. Clin. Psychiatry*, **55**(Suppl.), 5–10.
Wilson, F.A., O'Scalaidhe, S.P. & Goldman-Rakic, P.S. (1994) Functional synergism between putative gamma-aminobutyrate-containing neurons and pyramidal neurons in prefrontal cortex. *Proc. Natl Acad. Sci. USA*, **91**, 4009–4013.
Winterer, G. & Weinberger, D.R. (2004) Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. *Trends Neurosci.*, **27**, 683–690.

© The Authors (2008). Journal Compilation © Federation of European Neuroscience Societies and Blackwell Publishing Ltd