WARWICK
THE UNIVERSITY OF WARWICK

Original citation:
Rolls, Edmund T. (2016) A non-reward attractor theory of depression. Neuroscience & Biobehavioral Reviews, 68. pp. 47-58.

Permanent WRAP URL:
http://wrap.warwick.ac.uk/80548

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:
© 2016, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International http://creativecommons.org/licenses/by-nc-nd/4.0/

A note on versions:
The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP URL' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications

A non-reward attractor theory of depression

Edmund T. Rolls

Oxford Centre for Computational Neuroscience, Oxford, UK and
Department of Computer Science, University of Warwick, Coventry, UK

Please address correspondence to Professor E T Rolls
Email: Edmund.Rolls@oxcns.org
http://www.oxcns.org

Running head: Non-reward attractors and depression

Key words: orbitofrontal cortex; attractor networks; insula; non-reward; reward; anterior cingulate cortex; behavioural inhibition; reward reversal

\papers\Emotion\depression\depression16b.docx

Summary

A non-reward attractor theory of depression is proposed based on the operation of the lateral orbitofrontal cortex and supracallosal cingulate cortex. The orbitofrontal cortex contains error neurons that respond to non-reward for many seconds in an attractor state that maintains a memory of the non-reward. The human lateral orbitofrontal cortex is activated by non-reward during reward reversal, and by a signal to stop a response that is now incorrect. Damage to the human orbitofrontal cortex impairs reward reversal learning. Not receiving reward can produce depression. The theory proposed is that in depression, this lateral orbitofrontal cortex non-reward system is more easily triggered, and maintains its attractor-related firing for longer. This triggers negative cognitive states, which in turn have positive feedback top-down effects on the orbitofrontal cortex non-reward system. Treatments for depression, including ketamine, may act in part by quashing this attractor. The mania of bipolar disorder is hypothesized to be associated with oversensitivity and overactivity in the reciprocally related reward system in the medial orbitofrontal cortex and pregenual cingulate cortex.

1. Introduction

1.1 Background

Major depressive disorder is ranked by the World Health Organization as the leading cause of years-of-life lived with disability (Drevets, 2007; Gotlib and Hammen, 2009; Hamilton et al., 2013). Major depressive episodes, found in both major depressive disorder and bipolar disorder are pathological mood states characterized by persistently sad or depressed mood. Major depressive disorders are generally accompanied by: (a) altered incentive and reward processing, evidenced by amotivation, apathy, and anhedonia; (b) impaired modulation of anxiety and worry, manifested by generalized, social and panic anxiety, and oversensitivity to negative feedback; (c) inflexibility of thought and behavior in association with changing reinforcement contingencies, apparent as ruminative thoughts of self-reproach, pessimism, and guilt, and inertia toward initiating goal-directed behavior; (d) altered integration of sensory and social information, as evidenced by mood-congruent processing biases; (e) impaired attention and memory, shown as performance deficits on tests of attention set-shifting and maintenance, and autobiographical and short-term memory; and (f) visceral disturbances, including altered weight, appetite, sleep, and endocrine and autonomic function (Drevets, 2007; Gotlib and Hammen, 2009).

This paper describes a new theory of some of the brain mechanisms that are related to depression and mania, and has implications for treatment. The theory is related to attractor network dynamics, which enable states to be maintained by continuing firing within a population of neurons with strong excitatory synaptic inter-connections (Rolls, 2016a; Rolls and Deco, 2010) (see Appendix), and which are evident for non-reward in the lateral orbitofrontal cortex, as described next. The approach taken in this paper is to provide a systems-level neuroscience framework for understanding depression, bipolar disorder, and mania that is based on a wide range of evidence on the functions of different brain areas in emotion (Rolls, 2014a). Moreover, the theory extends hypotheses of depression to the concept that over-active attractor networks (Deco et al., 2013; Rolls, 2016a; Rolls and Deco, 2010) in some brain regions related to emotion (Rolls, 2014a) are involved in depression.

However, to provide further background, it is useful to refer to neuroimaging studies on depression (Rigucci et al., 2010) which lead to hypotheses of the following type. Patients with depression show impairments in the coordinated activity of some brain regions considered to be important for several domains of mental functioning (Iwabuchi et al., 2015) such as emotional processing (amygdala and subgenual anterior cingulate cortex) (Disner et al., 2011; Sheline et al., 2010); self-referential processes (medial prefrontal cortex (MPFC), precuneus and posterior cingulate cortex) (Kuhn and Gallinat, 2013; Price and Drevets, 2010; Sheline et al., 2010); cognitive functions such as memory (hippocampus, parahippocampal cortex) (Lorenzetti et al., 2009); visual processing (fusiform gyrus, lingual gyrus and lateral temporal cortex) (Veer et al., 2010); and attention processing (dorsolateral prefrontal cortex and anterior cingulate cortex) (Hamilton et al., 2012).

The remainder of Section 1 describes some of the evidence on the basis of which the theory has been developed. Section 2 describes the new theory of depression. Section 3 describes evidence consistent with the theory. Section 4 describes how the new theory relates to other brain systems than the orbitofrontal cortex implicated in depression. Section 5 considers some implications for treatments of depression. Section 6 describes a possible extension of the theory to mania and bipolar disorder.

1.2 Neurophysiological evidence for a non-reward system in the orbitofrontal cortex

First, neurophysiological evidence indicates that neurons in the orbitofrontal cortex respond to non-reward. The orbitofrontal cortex contains a population of error neurons that respond to non-reward and maintain their firing for many seconds after the non-reward, providing evidence that they have entered an attractor state that maintains a memory of the non-reward (Fig. 1) (Rolls, 2014b; Thorpe et al., 1983).

An example is shown in Fig. 1. On reversal trials, when saline (S) was obtained if a lick (indicated by a double dot) was made indicating that reversal should occur, the neuron started to respond approximately 1 s after the unexpected saline was obtained, and continued to respond as shown for many seconds, and indeed was still responding before the start of the next trial 8-10 s later, as shown. These neurons did not respond just to the taste of saline, but only when the saline indicated that a taste reward was not being obtained in that trial, and that because of the non-reward, reversal behavior should occur. The continuing firing for many seconds of these neurons provides evidence that they are part of a population that has entered a high firing rate attractor state.

Attractor networks once triggered can maintain their high firing rate because of the strong recurrent excitatory synaptic connections between the neuronal sub-population. The operation of attractor networks is summarized in the Appendix and described in more detail elsewhere (Rolls, 2016a) (with Appendix B available at http://www.oxcns.org). This short-term memory of a recent negatively reinforcing event is an essential component of a system that must change its operation after non-reward in a rule-based way, and this has been modelled (Deco and Rolls, 2005b). The change of behavior in a reversal task must be rule-based, for after a single trial in which a reward was not received to one stimulus, the choice switches to the other stimulus, even though the most recent reinforcement association of the second stimulus is with punishment. This is clearly illustrated by both reversals in Fig. 1 (Thorpe et al., 1983). This is a non-reward detection system which operates on a timescale of 1-10 s (as is evident in Fig. 1), which is an appropriate timescale for controlling behavior when a non-reward event occurs that relates to an action just taken or not taken. Further, the non-reward neuronal firing is probably computed in the orbitofrontal cortex, for this region contains the expected reward value and reward outcome value neurons required for the computation that an expected reward has not been obtained. Further, this type of rule-based reversal is impaired by damage to the orbitofrontal cortex in primates including humans, as described in Section 1.4.

This neuron-level evidence is strongly supported by functional neuroimaging evidence, which indicates that the macaque lateral orbitofrontal cortex is activated by non-reward during a reversal task, and that the focus of the activation is the lateral orbitofrontal cortex (Chau et al., 2015).

1.3 Neuroimaging evidence for a non-reward system in the lateral orbitofrontal cortex

Second, human functional neuroimaging evidence indicates that the lateral orbitofrontal cortex is activated by non-reward, and also by punishers that affect emotion-related behavior similarly (Rolls, 2014a). The activation of the human lateral orbitofrontal cortex during reward reversal is illustrated in Fig. 2a, which shows activations on reversal trials, that is when the human subject chose one person's face, and did not obtain the expected reward (Kringelbach and Rolls, 2003). Activations in the lateral orbitofrontal cortex are also produced by a signal to stop a response that is now incorrect, which is another situation in which behavior must change in order to be correct (Fig. 2b) (Deng et al., 2016). Orbitofrontal cortex activations in the stop-signal task have further been related to how impulsive the behavior is (Whelan et al., 2012). In this context, it has been suggested that impulsiveness may reflect how sensitive an individual is to non-reward or punishment (Rolls, 2014a), and indeed we have shown that people with orbitofrontal cortex damage become more impulsive (Berlin et al., 2005; Berlin et al., 2004). The lateral orbitofrontal cortex also responds to many punishing, unpleasant, stimuli (Grabenhorst and Rolls, 2011; Rolls, 2014b) including bad odor (Rolls et al., 2003) and losing money (O'Doherty et al., 2001), as shown in Fig. 3. The computations involved in non-reward referred to below are more complicated than those involved in representing punishers, but both types of representation are present in the orbitofrontal cortex, are implemented by different neurons, and both are involved in changing behavior to no longer choose the now non-rewarded stimulus, or the punisher (Rolls, 2014a, 2016a; Thorpe et al., 1983). The lateral orbitofrontal cortex projects to the supracallosal anterior cingulate cortex, in which non-rewards and punishers are also represented (Grabenhorst and Rolls, 2011; Rolls, 2014a; Rolls and Grabenhorst, 2008) (Fig. 3a) in this region implicated in action-outcome

learning (Grabenhorst and Rolls, 2011; Noonan et al., 2011; Rolls, 2014a). The effects of reward reversal on this supracallosal system are illustrated in Fig. 2a (green circle).

Consistent with this evidence for humans, functional neuroimaging in macaques reveals that the macaque lateral orbitofrontal cortex is activated by non-reward during a reversal task (Chau et al., 2015) (Fig. 2c).

1.4 Effects of damage to the orbitofrontal cortex

Third, lesion evidence also shows that the orbitofrontal cortex is involved in changing rewarded behavior when non-reward is detected, in that damage to the human orbitofrontal cortex impairs reward reversal learning, with the previously rewarded stimulus still being chosen during reversal even when no reward is being obtained (Fellows, 2011; Fellows and Farah, 2003; Hornak et al., 2004; Rolls et al., 1994). There is consistent evidence in macaques (Rolls, 2016a) with orbitofrontal cortex damage impairing performance on Go/NoGo task performance, in that they Go on the NoGo trials (Iversen and Mishkin, 1970), and in an object-reversal task in that they respond to the object that was formerly rewarded with food, and in extinction in that they continue to respond to an object that is no longer rewarded (Butter, 1969; Jones and Mishkin, 1972; Meunier et al., 1997). The visual discrimination reversal learning deficit shown by monkeys with orbitofrontal cortex damage (Murray and Izquierdo, 2007) may be due at least in part to the tendency of these monkeys not to withhold choosing non-rewarded stimuli (Jones and Mishkin, 1972) including objects that were previously rewarded during reversal (Rudebeck and Murray, 2011), and including foods that are not normally accepted (Baylis and Gaffan, 1991; Butter et al., 1969). Consistently, orbitofrontal cortex (but not amygdala) lesions impaired instrumental extinction (Murray and Izquierdo, 2007).

In an evolutionary context, this very rapid, rule-based, reversal of stimulus-reward associations possible in primates including humans may be an important adaptation to allow rapid and flexible changes of behaviour when reinforcement contingencies change, and is likely to be very important in social interaction (Rolls, 2014a, 2016a). The detection of this non-reward may be computed in the orbitofrontal cortex using reciprocally inhibiting Reward and Non-Reward attractor neuronal populations in a single network (Rolls and Deco, 2016). The computational model with integrate-and-fire neurons shows how adaptation in the Reward attractor neurons after an expected reward input is received can lead to activation of the Non-Reward attractor neurons unless a reward outcome has been received which maintains the Reward attractor state. This thus provides a model of how the responses of lateral orbitofrontal cortex non-reward neurons are produced (Rolls and Deco, 2016). Importantly, the model is based on attractor networks of the type described in this paper that may be over-sensitive or persistent in depression. The orbitofrontal cortex is highly developed in primates compared to rodents (Passingham and Wise, 2012; Wise, 2008), and this may be one function made possible by the great development in evolution of the orbitofrontal cortex (Rolls, 2016a, b). The theory of depression described here relates to this primate including human non-reward system in the orbitofrontal cortex.

2. A non-reward attractor theory of depression

It is well established that not receiving expected reward, or receiving unpleasant stimuli or events, can produce depression (Beck, 2008; Drevets, 2007; Eshel and Roiser, 2010; Harmer and Cowen, 2013; Price and Drevets, 2012; Pryce et al., 2011). More formally, in terms of learning theory, the omission or termination of a reward can give rise to sadness or depression, depending on the magnitude of the reward that is lost, if there is no action that can be taken to restore the reward (Rolls, 2013b, 2014a). If an action can be taken, then frustration and anger may arise for the same reinforcement contingency (Rolls, 2014a). This relates the current approach to the learned helplessness approach to depression, in which depression arises because no actions are being taken to restore rewards (Forgeard et al., 2011; Pryce et al., 2011). The sadness or depression may be short lasting if it is a minor non-reward. The

depression may be longer lasting if for example a member of the family dies, for every time that we remember the person we are aware of the loss of the reward of being with them, and this contributes to the longer-term depression that may arise.

On the basis of the evidence just described on the brain mechanisms involved in non-reward, and the non-reward triggers for depression, the theory is now proposed that in depression, the lateral orbitofrontal cortex non-reward / punishment attractor network system is more easily triggered and/or is very strongly triggered, and is therefore activated more often and maintains its attractor-related firing for longer. The greater attractor-related firing of the orbitofrontal cortex non-reward / punishment system then triggers negative cognitive states held on-line in other cortical systems such as the dorsolateral prefrontal cortex which is implicated in attentional control and which in turn has top-down effects on the orbitofrontal non-reward system that help to bias it in a negative direction and thus to increase its sensitivity to non-reward and maintain its overactivity (Rolls, 2013a). It is proposed that the interaction of two different brain systems of this type contributes to the long-lasting ruminating and continuing depressive thoughts which occur as a result of a positive feedback attractor cycle between these types of brain system.

Indeed, we have shown that cognitive states can have top-down effects on affective representations in the orbitofrontal cortex (de Araujo et al., 2005; Grabenhorst et al., 2008; McCabe et al., 2008; Rolls, 2013a). Further, top-down selective attention can also influence affective representations in the orbitofrontal cortex (Ge et al., 2012; Grabenhorst and Rolls, 2008; Grabenhorst and Rolls, 2010; Luo et al., 2013; Rolls, 2013a; Rolls et al., 2008a), and paying attention to depressive symptoms when depressed may in this way exacerbate the problems in a positive feedback way. (Top-down attention refers to the process whereby an area such as the prefrontal cortex can hold in short-term memory what it is that attention should enhance, and can then bias 'lower' brain areas to respond more to some properties of what they respond to (Deco and Rolls, 2005a; Desimone and Duncan, 1995; Rolls, 2013a, 2016a; Rolls and Deco, 2002).)

More generally, the presence of the cognitive ability to think ahead and see the implications of recent events that is afforded by language may be a computational and evolutionary development in the brain that exacerbates the vulnerability of the human brain to depression (Rolls, 2014a). The circuitry that may implement this is illustrated in Fig. 4. Further, whenever a memory is retrieved from the hippocampal and related systems the emotional component reactivates the orbitofrontal cortex emotional system (Rolls, 2015b), contributing to the long-lasting nature of depression.

The theory is that one way in which depression could result from over-activity in this lateral orbitofrontal cortex non-reward system is if there is a major negatively reinforcing life event that produces reactive depression and activates this system, which then becomes self-re-exciting based on the cycle between the lateral orbitofrontal cortex non-reward / punishment attractor system and the cognitive system, which together operate as a systems-level attractor (Fig. 4). The theory is that a second way in which depression might arise is if this lateral orbitofrontal cortex non-reward / punishment system is especially sensitive in some individuals. This might be related for example to genetic predisposition; or to the effects of chronic stress which influences cortical regions including the orbitofrontal cortex (Gold, 2015; Radley et al., 2015). In this case, the orbitofrontal system would over-react to normal levels of non-reward or punishment, or even become active in the absence of a stimulus as described in Section 5, and start the local attractor circuit in the lateral orbitofrontal cortex, which in turn would activate the cognitive system, which would feed back to the over-reactive lateral orbitofrontal cortex system to maintain now a systems-level attractor with ruminating thoughts (Fig. 4). In the theory, an oversensitive or over-responding short-term non-reward system in the lateral orbitofrontal cortex can produce long-lasting depression (a) because it activates the dorsolateral prefrontal cortex and related cognitive including language systems that continue thinking about the non-reward and re-excite the orbitofrontal cortex by top-down influences (see illustration in Fig. 4); and (b) because the orbitofrontal cortex non-reward system is activated whenever the human memory system retrieves a memory associated with a sad event (Rolls, 2015b), re-activating the positive feedback

system between the orbitofrontal and dorsolateral prefrontal cortex. In that the lateral orbitofrontal cortex connects to the supracallosal part of the anterior cingulate cortex, this is also expected to be overactive or to have increased functional connectivity in depression.

3. Evidence consistent with the theory

There is some evidence for altered structure and function of the lateral orbitofrontal cortex in depression (Drevets, 2007; Ma, 2015; Price and Drevets, 2012). For example, reductions of grey-matter volume and cortex thickness have been demonstrated specifically in the posterolateral orbitofrontal cortex / ventrolateral prefrontal cortex (BA 47, caudal BA 11 and the adjoining BA 45), and also in the subgenual cingulate cortex (BA 24, 25) (Drevets, 2007; Nugent et al., 2006). In depression, there is increased cerebral blood flow in areas that include the ventrolateral orbitofrontal cortex (which is a prediction of the theory), and also in regions such as the subgenual cingulate cortex and amygdala, and these increases appear to be related to the mood change, in that they become more normal when the mood state remits (Drevets, 2007).

Because the lateral orbitofrontal cortex responds to many punishing and non-rewarding stimuli (Grabenhorst and Rolls, 2011; Rolls, 2014a, b) that are likely to elicit autonomic/visceral responses, as does the supracallosal anterior cingulate cortex, and in view of connections from these areas to the anterior insula which is implicated in autonomic/visceral function (Critchley and Harrison, 2013; Rolls, 2015a), the anterior insula would also be expected to be overactive in depression, and that prediction is confirmed (Drevets, 2007; Hamilton et al., 2013; Ma, 2015).

Evidence from the first brain-wide voxel-level resting state functional-connectivity neuroimaging analysis of depression with 421 patients with major depressive disorder and 488 controls (Cheng, Rolls et al (2016)) provides support for and helps to refine the theory of depression. Resting state functional connectivity between different voxels reflects correlations of activity between those voxels and is a fundamental tool in helping to understand the brain regions with altered connectivity and function in depression.

One major circuit with altered functional connectivity involved the medial orbitofrontal cortex BA 13, which is implicated in reward, and which had reduced functional connectivity in depression with memory systems in the parahippocampal gyrus and medial temporal lobe (Fig. 5). The lateral orbitofrontal cortex BA 47/12, involved in non-reward and punishing events, did not have this reduced functional connectivity with memory systems, so that there is an imbalance in depression towards decreased reward-related memory system functionality. The reduced functional connectivity of the medial orbitofrontal cortex, implicated in reward, with memory systems (relative to the lateral orbitofrontal cortex) provides a new way of understanding how memory systems may be biased away from pleasant events in depression (Cheng et al., 2016).

Second, the lateral orbitofrontal cortex BA 47/12 had increased functional connectivity with the precuneus, the angular gyrus, and the temporal visual cortex BA 21 (Fig. 5). This enhanced functional connectivity of the non-reward/punishment system (BA 47/12) with the precuneus (involved in the sense of self and agency), and the angular gyrus (involved in language) may it is suggested be related to the explicit affectively negative sense of the self, and of self-esteem, in depression. The increased functional connectivity of the lateral orbitofrontal cortex, implicated in non-reward and punishment, with areas of the brain implicated in representing the self, language, and inputs from face and related perceptual systems (Fig. 5) (Cheng et al., 2016) provides a new way of understanding how unpleasant events and thoughts, and lowered self-esteem, may be exacerbated in depression. The increased connectivity between the lateral orbitofrontal cortex and the angular gyrus system involved in language (Fig. 5) directly supports the hypothesis illustrated in Fig. 4.

Treatments that can reduce depression such as a single dose of ketamine (Iadarola et al., 2015) may act in part by quashing the attractor state in the lateral orbitofrontal cortex / ventrolateral prefrontal cortex at least temporarily. Evidence consistent with this is that glucose metabolism in the ventrolateral

prefrontal cortex is decreased by a single dose of ketamine that ameliorates depression (Carlson et al., 2013), and in the lateral orbitofrontal cortex the decrease is related to the increase in hedonia produced by the ketamine (Fig. 6) (Lally et al., 2015). This NMDA receptor blocker may act at least in part by decreasing the high firing rate state of attractor networks by reducing transmission in the recurrent collateral excitatory connections between the neurons (Deco et al., 2013; Rolls, 2012; Rolls and Deco, 2010, 2015; Rolls et al., 2008c). Another NMDA receptor blocker, nitrous oxide, has also been shown to have an antidepressant effect, though the therapeutic use of nitrous oxide is not recommended because it produces vitamin B12 depletion (Nagele et al., 2015). Electroconvulsive therapy (ECT), which may have antidepressant effects, may also knock the non-reward system out of its attractor state, and this may contribute to any antidepressant effect. It has been shown that successful ECT for major depressive disorder is associated with reduced activation of the orbitofrontal cortex in emotional tasks (Beall et al., 2012).

Electrical stimulation of the brain that may relieve depression (Hamani et al., 2009; Hamani et al., 2011; Lujan et al., 2013) may act in part by providing reward that reciprocally inhibits the non-reward system, and/or by knocking the lateral orbitofrontal cortex and connected systems out of their attractor state. Treatment with antidepressant drugs decreases the activity of this lateral orbitofrontal cortex system (Ma, 2015).

Antidepressant drugs such as Selective Serotonin Reuptake Inhibitors (SSRIs) may treat depression by producing positive biases in the processing of emotional stimuli (Harmer and Cowen, 2013), increasing brain responses to positive stimuli and decreasing responses to negative stimuli (Ma, 2015). The reward and non-reward systems are likely to operate reciprocally, so that facilitating the reward system, or providing rewards, and thus activating the medial orbitofrontal cortex (Grabenhorst and Rolls, 2011; Rolls, 2014a) (Fig. 3), may operate in part by inhibiting the overactivity in the lateral orbitofrontal cortex non-reward / punishment system. Reciprocal activations of the medial orbitofrontal cortex reward systems and lateral orbitofrontal cortex non-reward systems are evident in a monetary reward/loss task (O'Doherty et al., 2001). The generally decreased functional connectivity of the medial orbitofrontal cortex in depression and the increased functional connectivity of the lateral orbitofrontal cortex illustrated in Fig. 5 (Cheng et al., 2016) also provides support for the theory that these systems operate in different directions in depression.

4. Relation to other brain systems implicated in depression

In this section, I consider how the lateral orbitofrontal cortex with its functions in non-reward is related to other brain systems implicated in depression, using a systems-level neuroscience approach to help assess the functions of different brain systems in depression based on each of their particular functions in emotion (Rolls, 2014a). The aim of this section is to provide a framework for understanding the contributions of different brain systems to depression.

The human medial orbitofrontal cortex has activations related to many rewarding and subjectively pleasant stimuli (Grabenhorst and Rolls, 2011; Rolls, 2014a; Rolls and Grabenhorst, 2008) (Fig. 3a). In the sense that reward vs non-reward and punishment are reciprocally related in their effects in the medial vs lateral orbitofrontal cortex respectively (O'Doherty et al., 2001; Rolls, 2014a), the anhedonia of depression can also be related to decreased effects of pleasant rewarding stimuli in the medial orbitofrontal cortex during depression, effects that can be restored by antidepressants (Ma, 2015). The lateral orbitofrontal cortex / inferior frontal gyrus region that responds to signals to inhibit a response in the stop-signal task (Deng et al., 2016) (Fig. 2b) is implicated in impulsive behavior, with damage in this region increasing impulsive behavior (Aron et al., 2014). This is of potential importance, for treatment with antidepressants, which would be expected to reduce the overactivity in this ventrolateral prefrontal cortex region, might thereby increase impulsiveness relative to that in the depressed state. Indeed, it is an interesting hypothesis that impulsiveness might reflect underactivity in this ventrolateral prefrontal cortex region, and that depression produced by oversensitivity to non-reward and punishment

might reflect overactivity in this ventrolateral prefrontal cortex region. In a certain sense, these types of behavior might reflect opposite ends of a continuum of non-reward/punishment sensitivity, with impulsiveness related to underactivity, and depression to overactivity in the ventrolateral prefrontal cortex.

The supracallosal cingulate cortex is activated by many aversive stimuli, and the pregenual cingulate cortex by many pleasant stimuli (Fig. 3a) (Grabenhorst and Rolls, 2011; Rolls, 2014a). These cingulate regions receive inputs from the lateral orbitofrontal cortex non-reward / punishment system and from the medial orbitofrontal cortex reward system, which are value-representing systems (Grabenhorst and Rolls, 2011; Rolls, 2014a). However, the anterior cingulate cortex appears to be involved beyond value in action-outcome learning, where the outcome refers to the reward or punisher for which an action is being learned (Camille et al., 2011; Grabenhorst and Rolls, 2011; Rolls, 2014b; Rudebeck et al., 2008; Rushworth et al., 2012; Rushworth et al., 2011). In contrast, the medial orbitofrontal cortex is implicated in reward-related processing and learning, and the lateral orbitofrontal cortex in non-reward and punishment-related processing and learning (Rolls, 2014a). These involve stimulus-stimulus associations, where the second stimulus is a reward (or its omission), or a punisher (Rolls, 2014a). Now given that emotions can be considered as states elicited by rewarding and punishing stimuli, and that moods such as depression can arise from prolonged non-reward or punishment (Rolls, 2014a), the part of the brain that processes these stimulus-stimulus associations, the orbitofrontal cortex, is more likely to be involved in depression than the action-related parts of the cingulate cortex.

In addition to the lateral orbitofrontal cortex and two regions that receive inputs from it, the insula implicated in autonomic functions (Rolls, 2015a), and the supracallosal anterior cingulate cortex, which are all activated by unpleasant stimuli (Grabenhorst and Rolls, 2011; Rolls, 2014a), parts of the amygdala are activated by unpleasant stimuli, and parts by pleasant stimuli (Rolls, 2014a), and amygdala activity has been related to depression (Harmer and Cowen, 2013; Ma, 2015; Price and Drevets, 2012). However, the amygdala is less involved in non-reward, especially the rule-based reversal of which stimuli are classified as rewarding that is required in a rapid reward reversal task (Rolls, 2014a). The orbitofrontal cortex is special in this, because the evidence is that it has attractor states than can be activated by non-reward (Rolls, 2014a; Thorpe et al., 1983) (Fig. 1), and these attractor states provide a basis for biasing the correct populations of neurons in the orbitofrontal cortex to implement the rapid one-trial reversal (Deco and Rolls, 2005b; Rolls, 2014a). Because the lateral orbitofrontal cortex has recurrent collaterals that can maintain attractor states, it is more likely to be involved in maintaining attractor states elicited by non-reward, including depression, than the amygdala (Rolls, 2014a, 2016a). The amygdala may therefore because of its responsiveness to punishing stimuli be related to depression, but may not be a structure that independently (of the orbitofrontal cortex) maintains its activity in an attractor state after non-reward, and during the mood state of depression.

The subgenual cingulate cortex has also been implicated in depression, and electrical stimulation in that region may relieve depression (Hamani et al., 2009; Hamani et al., 2011; Laxton et al., 2013; Lozano et al., 2012; Lujan et al., 2013; Mayberg, 2003). Further, adolescents with depression showed increased activation to social rejection (a non-reward outcome) relative to controls in the subgenual anterior cingulate as well as the supracallosal cingulate non-reward/punishment area and the amygdala, anterior insula and nucleus accumbens (Silk et al., 2014). However, the subgenual cingulate cortex is also implicated in autonomic function (Gabbott et al., 2003), and its role as an autonomic output region receiving information from the orbitofrontal and cingulate cortices could be related to some of the effects found in the subgenual cortex that are related to depression. Indeed, it is sometimes stated that a core symptom of depression is a "pervasive negative interoceptive state", and this may relate directly to the roles of the subgenual cingulate cortex in depression. Whether the subgenual cingulate cortex is activated because of inputs from the orbitofrontal cortex, or performs separate computations is not yet clear. Indeed, the orbitofrontal cortex has the inputs and representations required to compute non-reward, namely representations of expected value, and reward and punishment outcome value (Rolls, 2014a; Rolls and Grabenhorst, 2008), and it is not clear that the subgenual cingulate cortex has the

information to perform that computation. (Similar arguments apply to the anterior insula, which is also implicated in autonomic function (Rolls, 2015a), though it has also been described as part of a "salience network" (Menon and Uddin, 2010). Insular as well as anterior cingulate metabolism have been found to predict the successful outcome of treatment in depression (Fu et al., 2013; McGrath et al., 2013).) Further, the possibility is considered that electrical stimulation of the subcallosal region, which includes parts of the ventromedial prefrontal cortex (Laxton et al., 2013), that may relieve depression, may do so at least in part by activating connections involving the orbitofrontal cortex, other parts of the anterior cingulate cortex, and the striatum (Hamani et al., 2009; Johansen-Berg et al., 2008; Lujan et al., 2013).

5. Implications for treatments

This non-reward / punishment attractor network sensitivity theory of depression has implications for treatments. These implications can be understood and further explored in the context of investigations of the factors that influence the stability of attractor neuronal networks with integrate-and-fire neurons with noise introduced by the close to Poisson spiking times of the neurons (Deco et al., 2013; Deco et al., 2009; Loh et al., 2007; Rolls, 2008, 2016a; Rolls and Deco, 2010, 2011, 2015; Rolls et al., 2010; Rolls et al., 2008b; Rolls et al., 2008c; Wang, 2002). One is that antianxiety drugs, by increasing inhibition, might reduce the stability of the high firing rate state of the non-reward attractor, thus acting to quash the depression-related attractor state. A second is that it might be possible to produce agents that decrease the efficacy of NMDA receptors in the lateral orbitofrontal cortex, thereby reducing the stability of the depression-related attractor state. The evidence that there are genes that are selective for NMDA receptors for the neurons in different neural populations is that there are separate knock-outs for NMDA receptors in the CA3 and CA1 regions of the hippocampus (Nakazawa et al., 2004; Nakazawa et al., 2002; Nakazawa et al., 2003; Tonegawa et al., 2003). The present theory suggests that searching for ways to influence the attractor networks in the lateral orbitofrontal cortex by decreasing excitatory transmission (for example by influencing NMDA receptors in the lateral orbitofrontal cortex) or increasing inhibitory transmission in this region may be of considerable interest. It should be noted that the present theory is a theory specifically of non-reward and punishment-related attractor networks in the lateral orbitofrontal cortex and related areas in relation to depression, and that alterations of attractor networks in other cortical areas may be related to other psychiatric disorders (Rolls, 2012, 2016a).

In terms of the implications of the attractor-based aspect of the present theory, an important point is that the attractor dynamics must be kept stable in the face of the randomness or noise introduced into the system by the almost Poisson firing times of neurons for a given mean firing rate (Rolls, 2012, 2016a; Rolls and Deco, 2010). Moreover, the spontaneous firing rate state of the non-reward attractor must be maintained stable when no non-reward inputs are present (or otherwise the non-reward attractor would jump into a high firing rate non-reward state for no external reason, contributing to depression). The inhibitory transmitter GABA may be important in maintaining this type of stability (Rolls, 2012, 2016a; Rolls and Deco, 2010). Moreover, the high firing rate state produced by non-reward must not reach too high a firing rate, as this would cause overstability of the non-reward / depression state (Rolls, 2016a; Rolls and Deco, 2010). In a complementary way, if the high firing rate attractor state is insufficiently high, then that attractor state might be unstable, and the individual might be relatively insensitive to non-reward, not depressed, and impulsive because of not responding sufficiently to non-reward or punishment. The excitatory transmitter glutamate acting at NMDA or AMPA receptors may be important in setting the stability of the high firing rate attractor state (Rolls, 2012, 2016a; Rolls and Deco, 2010). In this respect and in this sense, the tendency to become depressed or to be impulsive may be reciprocally related to each other. Predictions for treatments follow from understanding these noisy attractor-based dynamics (Rolls, 2012, 2016a; Rolls and Deco, 2010).

The whole concept of attractor states has many implications for the treatment of depression, for rewards and other environmental changes and activities that tend to compete with the non-reward

attractor state and quash it may be useful in the treatment of depression.

6. Mania and bipolar disorder

So far, we have been considering unipolar depression.

Bipolar disorder includes periods of mania in addition to periods of depression. The severity of the mania is greatest in bipolar I disorder, moderate in bipolar II disorder, and lower in cyclothymia. During depression, people feel helplessness, reduced energy, and risk aversion, while with mania behaviors include grandiosity, increased energy, less sleep, and risk preference / impulsivity (Nusslock et al., 2014).

What is the relation between mania and depression? Although bipolar disorder is often thought to be very different from depression, might an analogous approach be useful for bipolar disorder? Could it be that in mania, there is something that in terms of reward/non-reward systems, is almost the opposite of depression? Might there be in mania *increased sensitivity to reward, and decreased sensitivity to non-reward / punishment*? The latter might manifest itself as increased impulsiveness in mania. That is a suggestion that might be considered to be the opposite of what has been described for unipolar depression.

It turns out that there is support for this hypothesis. It indeed appears that the risk for mania is characterized by a hypersensitivity to goal- and reward-relevant cues (Nusslock et al., 2014). This hypersensitivity can lead to an excessive increase in approach-related affect and motivation during life events involving rewards or goal striving and attainment. In the extreme, this excessive increase in reward-related affect is reflected in manic symptoms, such as pursuit of rewarding activities without attention to risks, elevated or irritable mood, decreased need for sleep, increased psychomotor activation, and extreme self-confidence. Some evidence consistent with the hypothesis is that patients with bipolar I disorder and their relatives showed greater activation of the medial orbitofrontal cortex in response to reward delivery (Wessa et al., 2014). Also, reduced deactivation of the medial orbitofrontal cortex (where rewards are represented) during reward reversal might reflect a reduced error signal in bipolar disorder patients and their relatives in the lateral orbitofrontal cortex. (Activation of the lateral orbitofrontal cortex by non-reward in healthy individuals is illustrated in Fig. 2a (Kringelbach and Rolls, 2003).) This type of non-reward or punishment-related responsiveness has been found to be very different in mania, with apparently decreasing activations in the lateral orbitofrontal cortex during expectation of increasing loss (Bermpohl et al., 2010), the opposite of what is found in healthy participants (Grabenhorst and Rolls, 2011; O'Doherty et al., 2001). In this context, of potentially reduced sensitivity or even abnormal function of the lateral orbitofrontal cortex non-reward system in mania, it is relevant that manic bipolar patients continue to pursue immediate rewards despite negative consequences (Wessa et al., 2014). Further, impulsivity in mania is pervasive, encompassing deficits in attention and behavioral inhibition. In addition, impulsivity is greater if the illness is severe (with for example frequent episodes, substance use disorders, and suicide attempts) (Swann, 2009). The significance of this is that impulsivity may reflect decreased sensitivity to non-reward, reflected in decreased activations in the lateral orbitofrontal cortex, where non-reward is represented (see above).

Thus mania may reflect a state in which there is decreased sensitivity of non-reward systems and hence increased impulsiveness due to reduced sensitivity of the lateral orbitofrontal cortex, and at the same time, increased sensitivity to reward reflected in activations in the medial orbitofrontal cortex and pregenual cingulate cortex. Although these medial and lateral systems may show reciprocally related activations within an individual, with for example increasing activations in the medial orbitofrontal cortex to increasing monetary gains and decreasing activations in the lateral orbitofrontal cortex, and vice versa to increasing monetary loss (O'Doherty et al., 2001), the reward and non-reward systems could be, and indeed are likely to, have their sensitivity set by independent genes, providing a basis for some patients to be depressed, and others to show both mania and depression. Indeed, Rolls' theory of emotions (2014a) would go beyond this, and suggest that the sensitivity to many different rewards (e.g.

food when hungry, water when thirsty, pleasant touch, sensitivity to reputation), and correspondingly to many different non-rewards, may be set by genes somewhat independently. This provides a relation to personality (Rolls, 2014a), with the implication that people with depression may be particularly sensitive to certain non-rewards or punishers, and people with mania may be particularly sensitive to particular rewards. This has important implications for therapy, which might be well-directed towards particular sensitivities to particular non-rewards and particular rewards in different individuals.

The question then arises of the extent to which attractor network operations contribute to mania. In terms of responses to inputs that increase the expectancy of reward, a short-term attractor system, probably in the orbitofrontal cortex, is likely to be present, to bridge any temporal interval between the expected reward signal and the actual outcome (Rolls and Deco, 2016). This could in principle be oversensitive in mania. When the reward, the outcome, is delivered, it might also be useful to have a short-term attractor, to help reset a rule attractor for which stimulus is currently rewarding. However, it would be maladaptive if these reward-expectancy or reward-outcome attractors normally operated for more than perhaps 10 s, for this would tend to break the important contingency between input stimuli and outcomes.

In addition to these short-term attractors, there also needs to be a longer term attractor process to reflect mood state, which typically operates on a much longer time scale. This might again be an attractor (with separate competing attractors for different mood states), and this attractor might be re-activated by the longer loop through the language / planning cognitive system, which by recalling a recent reward might calculate the long-term benefits, helping to keep the mood state prolonged. This whole 'long-loop' attractor might also be more sensitive in mania. These are interesting ideas for future empirical exploration.

The evidence that a reward system is over-responsive in mania (Nusslock et al., 2014) provides elegant complementary support for approach taken here, that over-responsiveness or over-activation of the lateral orbitofrontal cortex non-reward system provides a neuroscience-based framework for understanding depression.

7. Concluding remarks

The theory is proposed that in depression, the lateral orbitofrontal cortex / ventrolateral prefrontal cortex non-reward / punishment attractor network system is more easily triggered, and maintains its attractor-related firing for longer.

The greater attractor-related firing of the non-reward system triggers negative cognitive states, which in turn have top-down effects on the orbitofrontal non-reward system that help to maintain its overactivity and account for ruminating and continuing depressive thoughts in a positive feedback cycle between the cognitive systems in for example the dorsolateral prefrontal cortex and language systems and orbitofrontal cortex systems (Fig. 4). Recent neuroimaging evidence is consistent with this, showing increased functional connectivity between the lateral orbitofrontal cortex BA 47/12 and language areas such as the angular gyrus, with areas such as the precuneus involved in the representation of the self which may be related to the low self-esteem in depression, and to the temporal lobe sensory areas which may be related to the negative interpretation of environmental stimuli in depression (Cheng et al., 2016). The reduced functional connectivity of the medial orbitofrontal cortex, implicated in reward, with parahippocampal memory-related areas may relate to the reduced feelings of hedonia in depression (Cheng et al., 2016).

Treatments that can reduce depression such as a single dose of ketamine and electroconvulsive therapy, may act in part by quashing the attractor state at least temporarily. Other treatments such as antidepressant drugs may facilitate the medial orbitofrontal cortex reward system, which by reciprocal inhibitory interactions quashes overactive attractor states in the lateral orbitofrontal cortex.

It is hypothesized that the medial orbitofrontal cortex reward system which projects to the pregenual cingulate cortex is more responsive to reward when patients with bipolar disorder are in a manic state.

The orbitofrontal cortex projects to a number of brain areas and this helps to provide a systems-level neuroscience framework for understanding depression. The orbitofrontal cortex projects to the supracallosal anterior cingulate cortex which is involved in learning and correcting actions in relation to non-reward and punishment; and to the anterior insula and subgenual cingulate cortex which are involved inter alia in autonomic responses (Rolls, 2015a) which become altered in depression. The orbitofrontal cortex also projects to the amygdala, and to the basal ganglia which may be involved in habit (stimulus-response) learning (Rolls, 2014a).

Acknowledgments

Figure Legends
Fig. 1. Evidence that there are non-reward error-related neurons that maintain their firing after non-reward is obtained. Error neuron: Responses of an orbitofrontal cortex neuron that responded only when the macaque licked to a visual stimulus during reversal, expecting to obtain fruit juice reward, but actually obtained the taste of aversive saline because it was the first trial of reversal (trials 3, 6, and 13). Each vertical line represents an action potential; each L indicates a lick response in the Go-NoGo visual discrimination task. The visual stimulus was shown at time 0 for 1 s. The neuron did not respond on most reward (R) or saline (S) trials, but did respond on the trials marked S x, which were the first or second trials after a reversal of the visual discrimination on which the monkey licked to obtain reward, but actually obtained saline because the task had been reversed. The two times at which the reward contingencies were reversed are indicated. After responding to non-reward, when the expected reward was not obtained, the neuron fired for many seconds, and was sometimes still firing at the start of the next trial. It is notable that after an expected reward was not obtained due to a reversal contingency being applied, on the very next trial the macaque selected the previously non-rewarded stimulus. This shows that rapid reversal can be performed by a non-associative process, and must be rule-based. (After Thorpe, Rolls and Maddison 1983.) (RasterOF204Lab.eps)

Fig. 2a. Evidence that the human lateral orbitofrontal cortex is activated by non-reward. Activation of the lateral orbitofrontal cortex in a visual discrimination reversal task on reversal trials, when a face was selected but the expected reward was not obtained, indicating that the subject should select the other face in future to obtain the reward. a) A ventral view of the human brain with indication of the location of the two coronal slices (b,c) and the transverse slice (d). The activations with the red circle in the lateral orbitofrontal cortex (OFC, peaks at [42 42 -8] and [-46 30 -8]) show the activation on reversal trials compared to the non-reversal trials. For comparison, the activations with the blue circle show the fusiform face area produced just by face expressions, not by reversal, which are also indicated in the coronal slice in (c). b) A coronal slice showing the activation in the right orbitofrontal cortex on reversal trials. Activation is also shown in the supracallosal anterior cingulate region (Cingulate, green circle) that is also known to be activated by many punishing, unpleasant, stimuli (see Grabenhorst and Rolls (2011)). (From NeuroImage 20 (2), Morten L. Kringelbach and Edmund T. Rolls, Neural correlates of rapid reversal learning in a simple model of human social interaction, pp. 1371--83, Copyright, 2003, with permission from Elsevier.)

Fig. 2b. Activations in the human lateral orbitofrontal cortex are related to a signal to change behavior in the stop-signal task. In the task, a left or right arrow on a screen indicates which button to touch. However on some trials, an up-arrow then appears, and the participant must change the behavior, and stop the response. There is a larger response on trials on which the participant successfully changes the behavior and stops the response, as shown by the contrast stop-success – stop-failure, in the ventrolateral prefrontal cortex in a region including the lateral orbitofrontal cortex, with peak at [-42 50 -2] indicated by the cross-hairs, measured in 1709 participants. There were corresponding effects in the right lateral orbitofrontal cortex [42 52 -4]. Some activation in the dorsolateral prefrontal cortex in an area implicated in attention is also shown. (After Deng, Rolls et al, 2016).

Fig. 2c. Bold signal in the macaque lateral orbitofrontal related to win-stay / lose-shift performance, that is, to reward reversal performance. (After Chau et al, 2015). (OFCfacereversalSSTmac.eps)

Fig. 3. Rewards tend to be represented medially in the human orbitofrontal cortex and in the pregenual cingulate cortex, and non-reward and punishment laterally in the orbitofrontal cortex and in the supracallosal anterior cingulate cortex. Maps of subjective pleasure in the human orbitofrontal cortex (ventral view) and anterior cingulate and ventromedial prefrontal cortex (sagittal

view). Yellow: sites where activations correlate with subjective pleasantness. White: sites where activations correlate with subjective unpleasantness. The numbers refer to effects found in specific studies. Taste: 1, 2; odor: 3-10; flavor: 11-16; oral texture: 17, 18; chocolate: 19; water: 20; wine: 21; oral temperature: 22, 23; somatosensory temperature: 24, 25; the sight of touch: 26, 27; facial attractiveness: 28, 29; erotic pictures: 30; laser-induced pain: 31. Consistent laterality effects are not generally observed in these investigations. (Reprinted from Trends in Cognitive Sciences, 15 (2), Fabian Grabenhorst and Edmund T. Rolls, Value, pleasure and choice in the ventral prefrontal cortex, pp. 56--67, Copyright, 2011, with permission from Elsevier.) (PleasureMaps.eps)

Fig. 4. Interaction of orbitofrontal cortex non-reward networks with language networks in depression. Illustration of how an overactive non-reward attractor network in the lateral orbitofrontal cortex could send information forward to networks for language and planning ahead which could in turn send top-down feedback back to the orbitofrontal non-reward network to maintain its over-activity. It is suggested that such a system contributes to the persistent ruminating thoughts in depression. (DepArchi.eps)

Fig. 5. Resting state functional connectivity in depression. A. Cluster functional connectivity matrix. The color bar shows the $-\log_{10}$ of the p value for the difference of the functional connectivity. Blue indicates reduced functional connectivity, and yellow/orange/red increased functional connectivity. The matrix contains rows and columns for all cases in which there were 10 or more significant voxels within a cluster. ACC - anterior cingulate cortex; MedTL - medial temporal lobe, including parts of the parahippocampal gyrus; Thal - thalamus. The abbreviations are from the AAL2 (Rolls et al., 2015) (see also (Cheng et al., 2016)). B. The medial and lateral orbitofrontal cortex networks that show different functional connectivity in patients with depression. A decrease in functional connectivity is shown in blue, and an increase in red. MedTL -- medial temporal lobe from the parahippocampal gyrus to the temporal pole; MidTG21R - middle temporal gyrus area 21 right; OFC13 - medial orbitofrontal cortex area 13; OFC47/12R - lateral orbitofrontal cortex area 47/12 right. The lateral orbitofrontal cortex cluster in OFC47/12 is visible on the ventral view of the brain anterior and lateral to the OFC13 clusters. (After Cheng, Rolls et al, 2016.) (BrainClusters4b.eps)

Fig. 6. Antidepressant (reduction of anhedonia) effects of a single dose of ketamine are associated with a decrease in metabolism in the lateral orbitofrontal cortex and inferior frontal gyrus. Percentage change in hedonia (measured with the Snaith-Hamilton Pleasure Scale, SHAPS) regressed against FDG PET difference images. Decreased metabolism in the orbitofrontal cortex and inferior frontal gyrus (upper) were associated with the greatest anti-anhedonic response (lower) (i.e. with the greatest increase in hedonia). (After Lally et al 2015.) (Lal3.eps)
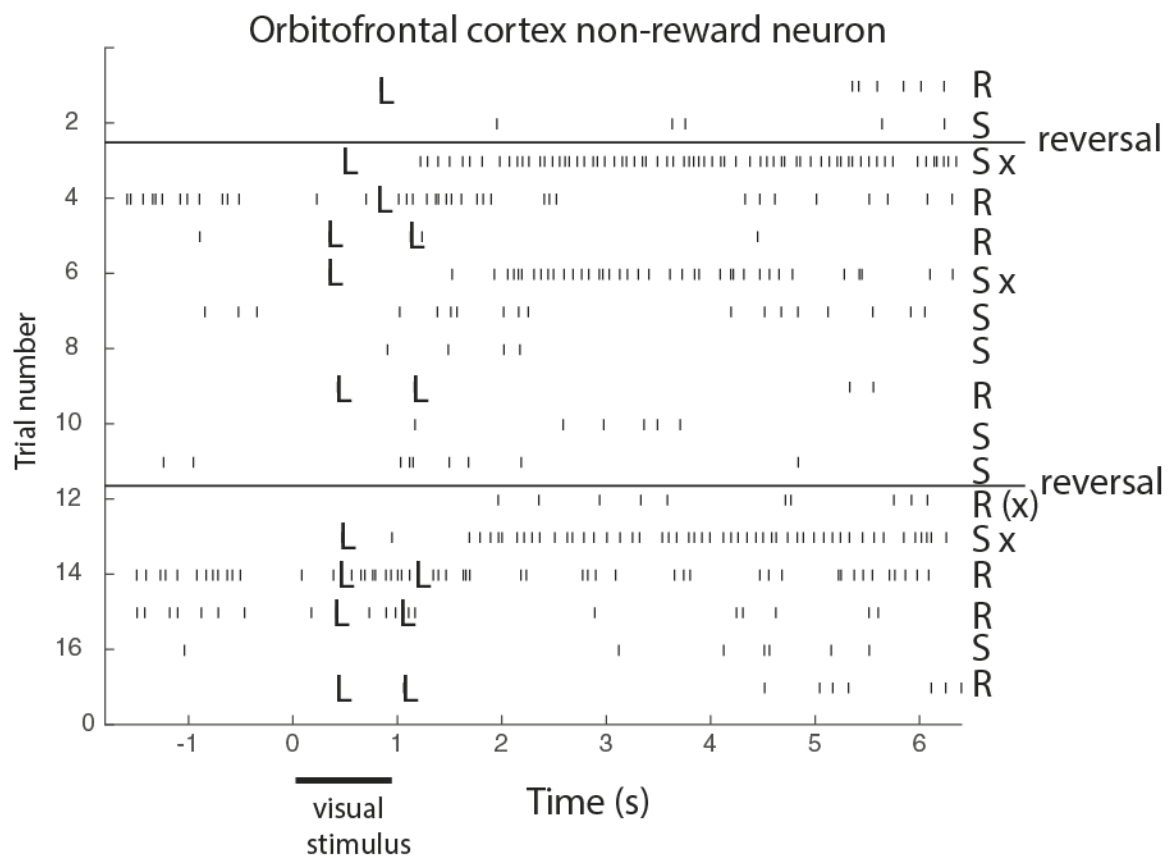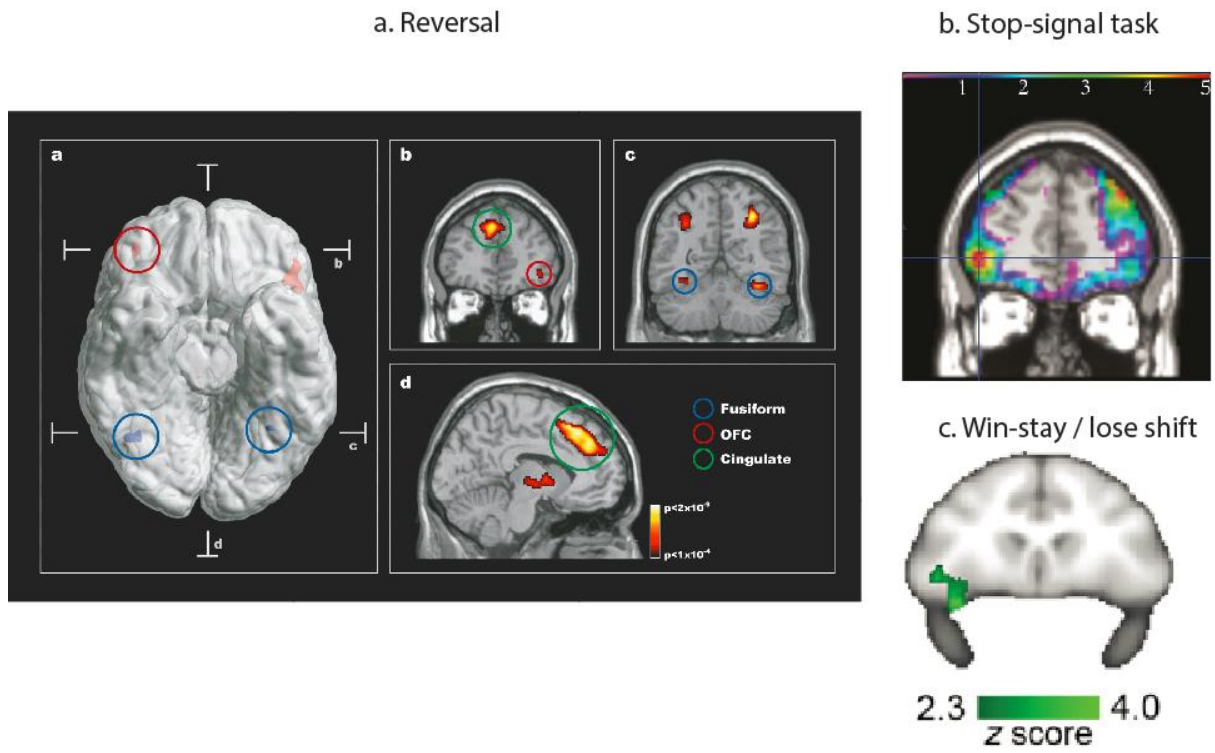
Fig. 1. (RasterOF204Lab.eps)

Orbitofrontal cortex non-reward neuron

Fig. 2. (OFCfacereversalSSTmac.eps)

a. Reversal

b. Stop-signal task

c. Win-stay / lose shift

Fig. 3. (PleasureMaps.eps)

Orbitofrontal cortex

Anterior cingulate and
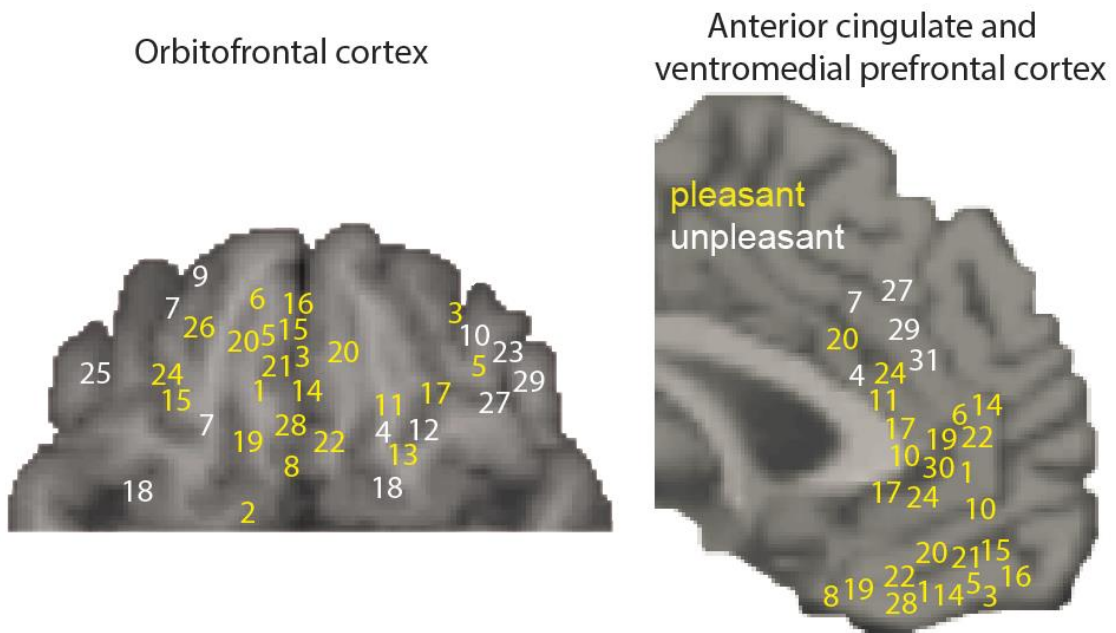ventromedial prefrontal cortex

Fig. 4. (DepArchi.eps)

Interaction of non-reward and language networks in depression

overactive lateral orbitofrontal cortex
non-reward attractor network

language networks for
planning and thinking ahead

Input

RC

RC

Output

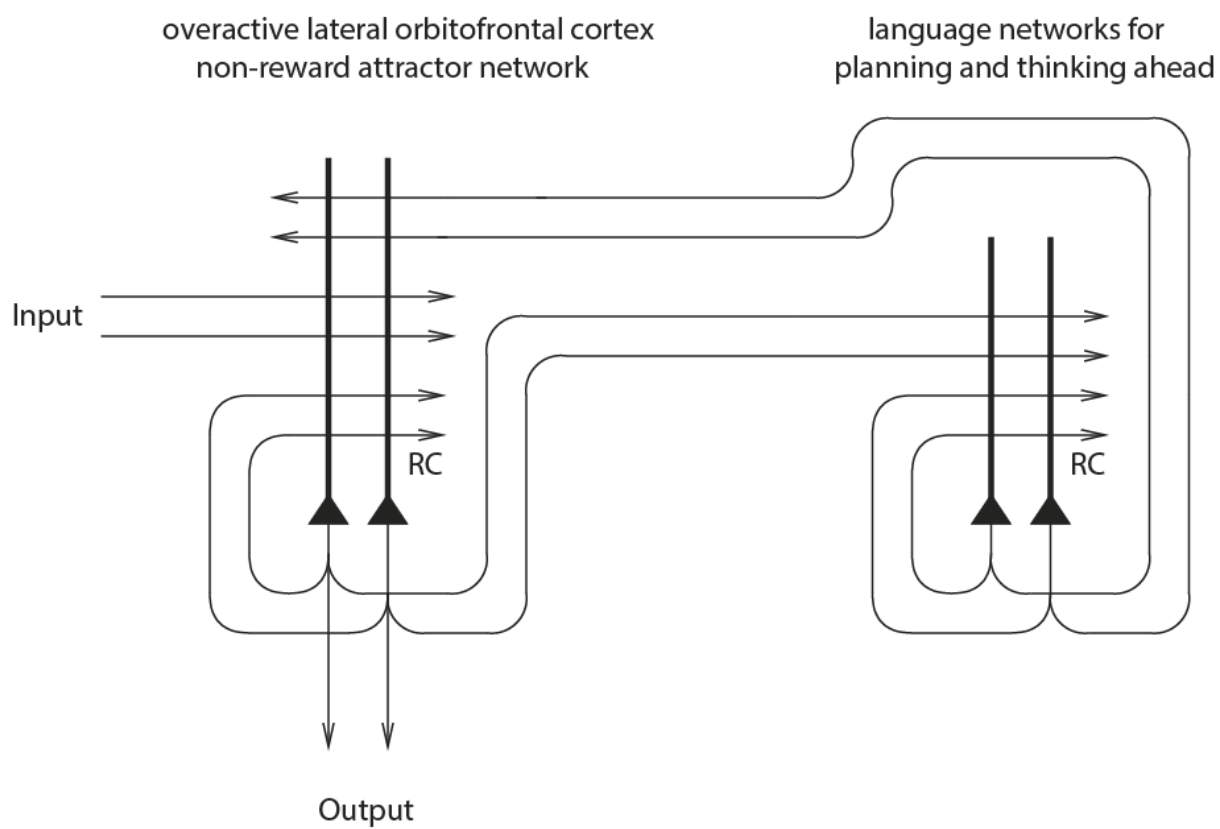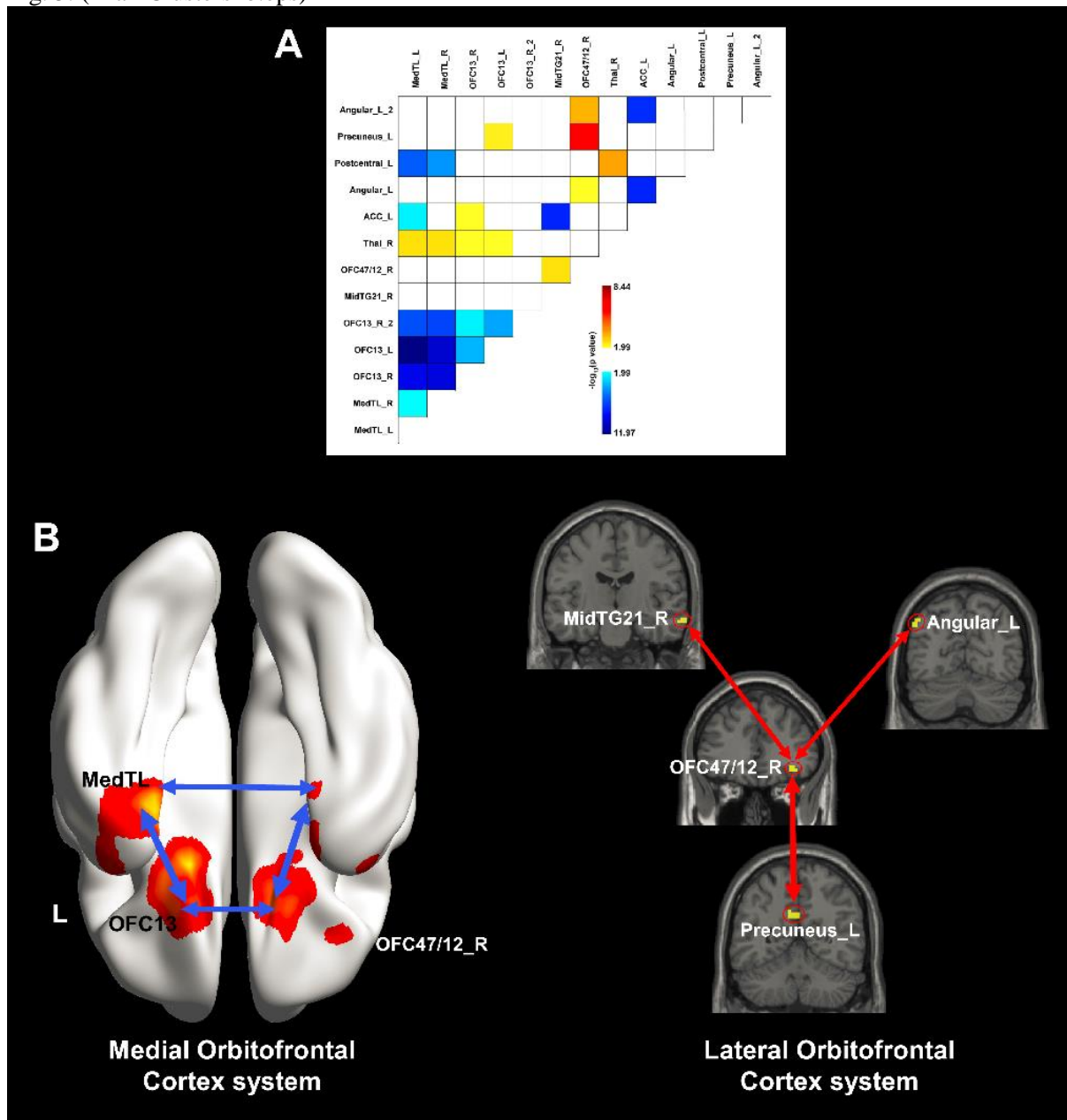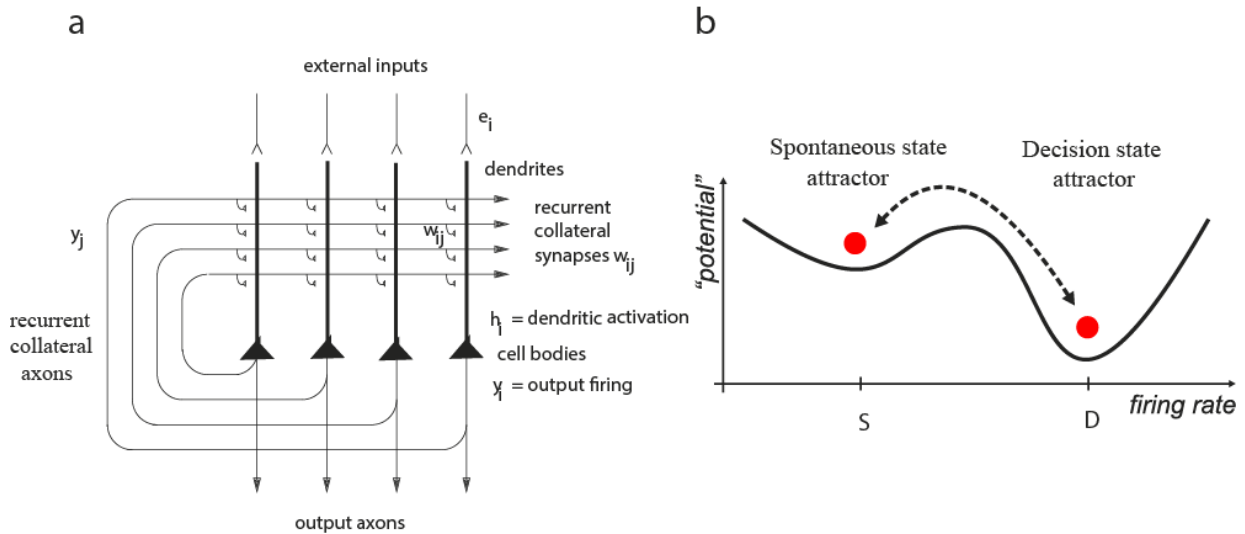Fig. 5. (BrainClusters4b.eps)

Appendix. Attractor networks.

a   b

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 i'th neuron. The neurons are connected by recurrent collateral synapses w_{ij}, where j refers to the j'th 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, or a decision) in an attractor network can be thought of as the local minimum in an energy landscape.

b. Energy landscape. Attractor networks are often conceptualized in terms of an energy landscape, in which a low energy, stable, state would be produced if neurons with high firing were connected by strong synapses (Hopfield, 1982; Rolls, 2016a; Rolls and Deco, 2010). 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 might represent a decision or a memory, and it is 'persistent' in that the neurons that implement it continue firing. The vertical axis of the 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 for the high firing attractor state are equally deep, making both states stable. In the situation that is illustrated, the high firing rate state is very deep, corresponding to a very stable and persistent attractor, such as one that might be present for non-reward or punishment in depression. In general, there will be many different high firing rate attractor basins, each corresponding to a different decision or memory. (autoarchi3Landscape.eps)

Neuronal network software written in Matlab to illustrate the operation of attractor networks is available in connection with *Cerebral Cortex: Principles of Operation* (E.T.Rolls, 2016, Oxford University Press) at
http://www.oxcns.org/NeuronalNetworkSimulationSoftware.html.

References

Aron, A.R., Robbins, T.W., Poldrack, R.A., 2014. Inhibition and the right inferior frontal cortex: one decade on. Trends Cogn. Sci. 18, 177-185.

Baylis, L.L., Gaffan, D., 1991. Amygdalectomy and ventromedial prefrontal ablation produce similar deficits in food choice and in simple object discrimination learning for an unseen reward. Exp. Brain Res. 86, 617-622.

Beall, E.B., Malone, D.A., Dale, R.M., Muzina, D.J., Koenig, K.A., Bhattacharrya, P.K., Jones, S.E., Phillips, M.D., Lowe, M.J., 2012. Effects of electroconvulsive therapy on brain functional activation and connectivity in depression. J. ECT 28, 234-241.

Beck, A.T., 2008. The evolution of the cognitive model of depression and its neurobiological correlates. Am. J. Psychiatry 165, 969-977.

Berlin, H., Rolls, E.T., Iversen, S.D., 2005. Borderline Personality Disorder, impulsivity and the orbitofrontal cortex. Am. J. Psychiatry 162, 2360-2373.

Berlin, H., Rolls, E.T., Kischka, U., 2004. Impulsivity, time perception, emotion, and reinforcement sensitivity in patients with orbitofrontal cortex lesions. Brain 127, 1108-1126.

Bermpohl, F., Kahnt, T., Dalanay, U., Hagele, C., Sajonz, B., Wegner, T., Stoy, M., Adli, M., Kruger, S., Wrase, J., Strohle, A., Bauer, M., Heinz, A., 2010. Altered representation of expected value in the orbitofrontal cortex in mania. Hum. Brain Mapp. 31, 958-969.

Butter, C.M., 1969. Perseveration in extinction and in discrimination reversal tasks following selective prefrontal ablations in Macaca mulatta. Physiol. Behav. 4, 163-171.

Butter, C.M., McDonald, J.A., Snyder, D.R., 1969. Orality, preference behavior, and reinforcement value of non-food objects in monkeys with orbital frontal lesions. Science 164, 1306-1307.

Camille, N., Tsuchida, A., Fellows, L.K., 2011. Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. J. Neurosci. 31, 15048-15052.

Carlson, P.J., Diazgranados, N., Nugent, A.C., Ibrahim, L., Luckenbaugh, D.A., Brutsche, N., Herscovitch, P., Manji, H.K., Zarate, C.A., Jr., Drevets, W.C., 2013. Neural correlates of rapid antidepressant response to ketamine in treatment-resistant unipolar depression: a preliminary positron emission tomography study. Biol. Psychiatry 73, 1213-1221.

Chau, B.K., Sallet, J., Papageorgiou, G.K., Noonan, M.P., Bell, A.H., Walton, M.E., Rushworth, M.F., 2015. Contrasting roles for orbitofrontal cortex and amygdala in credit assignment and learning in macaques. Neuron 87, 1106-1118.

Cheng, W., Rolls, E.T., Qiu, J., Liu, W., Chang, M., Huang, C.-C., Wang, X.-F., Zhang, J., Pu, J.-C., Tsai, S.-J., Lin, C.-P., Wang, F., Xie, P., Feng, J., 2016. Medial reward and lateral non-reward orbitofrontal cortex circuits change in opposite directions in depression.

Critchley, H.D., Harrison, N.A., 2013. Visceral influences on brain and behavior. Neuron 77, 624-638.

de Araujo, I.E.T., Rolls, E.T., Velazco, M.I., Margot, C., Cayeux, I., 2005. Cognitive modulation of olfactory processing. Neuron 46, 671-679.

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

Deco, G., Rolls, E.T., 2005b. Synaptic and spiking dynamics underlying reward reversal in orbitofrontal cortex. Cereb. Cortex 15, 15-30.

Deco, G., Rolls, E.T., Albantakis, L., Romo, R., 2013. Brain mechanisms for perceptual and reward-related decision-making. Prog. Neurobiol. 103, 194-213.

Deco, G., Rolls, E.T., Romo, R., 2009. Stochastic dynamics as a principle of brain function. Prog. Neurobiol. 88, 1-16.

Deng, W.L., Rolls, E.T., Ji, X., Banaschewski, T., Bodge, A., Bromberg, U., Buechel, C., Cattrell, A., Conrod, P., Flor, H., Frouin, V., Gallinat, J., Garavan, H., Gowland, P., Heinz, A., Ittermann, B., Lemaitre, H., Martinot, J.-L., Nees, F., Papadoulos Orfanos, D., Poustka, L., Smolka, M.N., Walter, H., Whelan, R., Schumann, G., Robbins, T.W., Feng, J., IMAGEN, C., 2016. Separate neural systems for behavioral change and for emotional responses to failure during behavioral inhibition.

Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. Annu. Rev. Neurosci. 18, 193-222.

Disner, S.G., Beevers, C.G., Haigh, E.A., Beck, A.T., 2011. Neural mechanisms of the cognitive model of depression. Nat. Rev. Neurosci. 12, 467-477.

Drevets, W.C., 2007. Orbitofrontal cortex function and structure in depression. Ann. N. Y. Acad. Sci. 1121, 499-527.

Eshel, N., Roiser, J.P., 2010. Reward and punishment processing in depression. Biol. Psychiatry 68, 118-124.

Fellows, L.K., 2011. Orbitofrontal contributions to value-based decision making: evidence from humans with frontal lobe damage. Ann. N. Y. Acad. Sci. 1239, 51-58.

Fellows, L.K., Farah, M.J., 2003. Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. Brain 126, 1830-1837.

Forgeard, M.J., Haigh, E.A., Beck, A.T., Davidson, R.J., Henn, F.A., Maier, S.F., Mayberg, H.S., Seligman, M.E., 2011. Beyond Depression: Towards a Process-Based Approach to Research, Diagnosis, and Treatment. Clin. Psychol. (New York) 18, 275-299.

Fu, C.H., Steiner, H., Costafreda, S.G., 2013. Predictive neural biomarkers of clinical response in depression: a meta-analysis of functional and structural neuroimaging studies of pharmacological and psychological therapies. Neurobiol. Dis. 52, 75-83.

Gabbott, P.L., Warner, T.A., Jays, P.R., Bacon, S.J., 2003. Areal and synaptic interconnectivity of prelimbic (area 32), infralimbic (area 25) and insular cortices in the rat. Brain Res. 993, 59-71.

Ge, T., Feng, J., Grabenhorst, F., Rolls, E.T., 2012. Componential Granger causality, and its application to identifying the source and mechanisms of the top-down biased activation that controls attention to affective vs sensory processing. Neuroimage 59, 1846-1858.

Gold, P.W., 2015. The organization of the stress system and its dysregulation in depressive illness. Mol.

Psychiatry 20, 32-47.

Gotlib, I.H., Hammen, C.L., 2009. Handbook of Depression. Guilford Press, New York.

Grabenhorst, F., Rolls, E.T., 2008. Selective attention to affective value alters how the brain processes taste stimuli. Eur. J. Neurosci. 27, 723-729.

Grabenhorst, F., Rolls, E.T., 2010. Attentional modulation of affective vs sensory processing: functional connectivity and a top-down biased activation theory of selective attention. J. Neurophysiol. 104, 1649-1660.

Grabenhorst, F., Rolls, E.T., 2011. Value, pleasure, and choice in the ventral prefrontal cortex. Trends Cogn. Sci. 15, 56-67.

Grabenhorst, F., Rolls, E.T., Bilderbeck, A., 2008. How cognition modulates affective responses to taste and flavor: top down influences on the orbitofrontal and pregenual cingulate cortices. Cereb. Cortex 18, 1549-1559.

Hamani, C., Mayberg, H., Snyder, B., Giacobbe, P., Kennedy, S., Lozano, A.M., 2009. Deep brain stimulation of the subcallosal cingulate gyrus for depression: anatomical location of active contacts in clinical responders and a suggested guideline for targeting. J. Neurosurg. 111, 1209-1215.

Hamani, C., Mayberg, H., Stone, S., Laxton, A., Haber, S., Lozano, A.M., 2011. The subcallosal cingulate gyrus in the context of major depression. Biol. Psychiatry 69, 301-308.

Hamilton, J.P., Chen, M.C., Gotlib, I.H., 2013. Neural systems approaches to understanding major depressive disorder: an intrinsic functional organization perspective. Neurobiol. Dis. 52, 4-11.

Hamilton, J.P., Etkin, A., Furman, D.J., Lemus, M.G., Johnson, R.F., Gotlib, I.H., 2012. Functional neuroimaging of major depressive disorder: a meta-analysis and new integration of base line activation and neural response data. Am. J. Psychiatry 169, 693-703.

Harmer, C.J., Cowen, P.J., 2013. 'It's the way that you look at it'--a cognitive neuropsychological account of SSRI action in depression. Philos. Trans. R. Soc. Lond. B Biol. Sci. 368, 20120407.

Hopfield, J.J., 1982. Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Science USA 79, 2554-2558.

Hornak, J., O'Doherty, J., Bramham, J., Rolls, E.T., Morris, R.G., Bullock, P.R., Polkey, C.E., 2004. Reward-related reversal learning after surgical excisions in orbitofrontal and dorsolateral prefrontal cortex in humans. J. Cogn. Neurosci. 16, 463-478.

Iadarola, N.D., Niciu, M.J., Richards, E.M., Vande Voort, J.L., Ballard, E.D., Lundin, N.B., Nugent, A.C., Machado-Vieira, R., Zarate, C.A., Jr., 2015. Ketamine and other N-methyl-D-aspartate receptor antagonists in the treatment of depression: a perspective review. Ther. Adv. Chronic Dis. 6, 97-114.

Iversen, S.D., Mishkin, M., 1970. Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. Exp. Brain Res. 11, 376-386.

Iwabuchi, S.J., Krishnadas, R., Li, C., Auer, D.P., Radua, J., Palaniyappan, L., 2015. Localized connectivity in depression: a meta-analysis of resting state functional imaging studies. Neurosci. Biobehav. Rev. 51, 77-86.

Johansen-Berg, H., Gutman, D.A., Behrens, T.E., Matthews, P.M., Rushworth, M.F., Katz, E., Lozano, A.M., Mayberg, H.S., 2008. Anatomical connectivity of the subgenual cingulate region targeted with deep brain stimulation for treatment-resistant depression. Cereb. Cortex 18, 1374-1383.

Jones, B., Mishkin, M., 1972. Limbic lesions and the problem of stimulus-reinforcement associations. Exp. Neurol. 36, 362-377.

Kringelbach, M.L., Rolls, E.T., 2003. Neural correlates of rapid reversal learning in a simple model of human social interaction. Neuroimage 20, 1371-1383.

Kuhn, S., Gallinat, J., 2013. Resting-state brain activity in schizophrenia and major depression: a quantitative meta-analysis. Schizophr. Bull. 39, 358-365.

Lally, N., Nugent, A.C., Luckenbaugh, D.A., Niciu, M.J., Roiser, J.P., Zarate, C.A., Jr., 2015. Neural correlates of change in major depressive disorder anhedonia following open-label ketamine. Journal of Psychopharmacology 29, 596-607.

Laxton, A.W., Neimat, J.S., Davis, K.D., Womelsdorf, T., Hutchison, W.D., Dostrovsky, J.O., Hamani, C., Mayberg, H.S., Lozano, A.M., 2013. Neuronal coding of implicit emotion categories in the subcallosal cortex in patients with depression. Biol. Psychiatry 74, 714-719.

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.

Lorenzetti, V., Allen, N.B., Fornito, A., Yucel, M., 2009. Structural brain abnormalities in major depressive disorder: a selective review of recent MRI studies. J. Affect. Disord. 117, 1-17.

Lozano, A.M., Giacobbe, P., Hamani, C., Rizvi, S.J., Kennedy, S.H., Kolivakis, T.T., Debonnel, G., Sadikot, A.F., Lam, R.W., Howard, A.K., Ilcewicz-Klimek, M., Honey, C.R., Mayberg, H.S., 2012. A multicenter pilot study of subcallosal cingulate area deep brain stimulation for treatment-resistant depression. J. Neurosurg. 116, 315-322.

Lujan, J.L., Chaturvedi, A., Choi, K.S., Holtzheimer, P.E., Gross, R.E., Mayberg, H.S., McIntyre, C.C., 2013. Tractography-activation models applied to subcallosal cingulate deep brain stimulation. Brain Stimul.

Luo, Q., Ge, T., Grabenhorst, F., Feng, J., Rolls, E.T., 2013. Attention-dependent modulation of cortical taste circuits revealed by Granger causality with signal-dependent noise PLoS Comput. Biol. 9, e1003265.

Ma, Y., 2015. Neuropsychological mechanism underlying antidepressant effect: a systematic meta-analysis. Mol. Psychiatry 20, 311-319.

Mayberg, H.S., 2003. Positron emission tomography imaging in depression: a neural systems perspective. Neuroimaging Clin. N. Am. 13, 805-815.

McCabe, C., Rolls, E.T., Bilderbeck, A., McGlone, F., 2008. Cognitive influences on the affective representation of touch and the sight of touch in the human brain. Social, Cognitive and Affective Neuroscience 3, 97-108.

McGrath, C.L., Kelley, M.E., Holtzheimer, P.E., Dunlop, B.W., Craighead, W.E., Franco, A.R., Craddock, R.C., Mayberg, H.S., 2013. Toward a neuroimaging treatment selection biomarker for major depressive disorder. JAMA Psychiatry 70, 821-829.

Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. Brain Struct Funct 214, 655-667.

Meunier, M., Bachevalier, J., Mishkin, M., 1997. Effects of orbital frontal and anterior cingulate lesions on object and spatial memory in rhesus monkeys. Neuropsychologia 35, 999-1015.

Murray, E.A., Izquierdo, A., 2007. Orbitofrontal cortex and amygdala contributions to affect and action in primates. Ann. N. Y. Acad. Sci. 1121, 273-296.

Nagele, P., Duma, A., Kopec, M., Gebara, M.A., Parsoei, A., Walker, M., Janski, A., Panagopoulos, V.N., Cristancho, P., Miller, J.P., Zorumski, C.F., Conway, C.R., 2015. Nitrous Oxide for Treatment-Resistant Major Depression: A Proof-of-Concept Trial. Biol. Psychiatry 78, 10-18.

Nakazawa, K., McHugh, T.J., Wilson, M.A., Tonegawa, S., 2004. NMDA receptors, place cells and hippocampal spatial memory. Nature Reviews Neuroscience 5, 361-372.

Nakazawa, K., Quirk, M.C., Chitwood, R.A., Watanabe, M., Yeckel, M.F., Sun, L.D., Kato, A., Carr, C.A., Johnston, D., Wilson, M.A., Tonegawa, S., 2002. Requirement for hippocampal CA3 NMDA receptors in associative memory recall. Science 297, 211-218.

Nakazawa, K., Sun, L.D., Quirk, M.C., Rondi-Reig, L., Wilson, M.A., Tonegawa, S., 2003. Hippocampal CA3 NMDA receptors are crucial for memory acquisition of one-time experience. Neuron 38, 305-315.

Noonan, M.P., Mars, R.B., Rushworth, M.F., 2011. Distinct roles of three frontal cortical areas in reward-guided behavior. J. Neurosci. 31, 14399-14412.

Nugent, A.C., Milham, M.P., Bain, E.E., Mah, L., Cannon, D.M., Marrett, S., Zarate, C.A., Pine, D.S., Price, J.L., Drevets, W.C., 2006. Cortical abnormalities in bipolar disorder investigated with MRI and voxel-based morphometry. Neuroimage 30, 485-497.

Nusslock, R., Young, C.B., Damme, K.S., 2014. Elevated reward-related neural activation as a unique biological marker of bipolar disorder: assessment and treatment implications. Behav. Res. Ther. 62, 74-87.

O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., Andrews, C., 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. Nat. Neurosci. 4, 95-102.

Passingham, R.E.P., Wise, S.P., 2012. The Neurobiology of the Prefrontal Cortex. Oxford University Press, Oxford.

Price, J.L., Drevets, W.C., 2010. Neurocircuitry of mood disorders. Neuropsychopharmacology 35, 192-216.

Price, J.L., Drevets, W.C., 2012. Neural circuits underlying the pathophysiology of mood disorders. Trends Cogn. Sci. 16, 61-71.

Pryce, C.R., Azzinnari, D., Spinelli, S., Seifritz, E., Tegethoff, M., Meinlschmidt, G., 2011. Helplessness: a systematic translational review of theory and evidence for its relevance to understanding and treating depression. Pharmacol. Ther. 132, 242-267.

Radley, J., Morilak, D., Viau, V., Campeau, S., 2015. Chronic stress and brain plasticity: Mechanisms underlying adaptive and maladaptive changes and implications for stress-related CNS disorders. Neurosci. Biobehav. Rev.

Rigucci, S., Serafini, G., Pompili, M., Kotzalidis, G.D., Tatarelli, R., 2010. Anatomical and functional correlates in major depressive disorder: the contribution of neuroimaging studies. World J. Biol. Psychiatry 11, 165-180.

Rolls, E.T., 2008. Memory, Attention, and Decision-Making: A Unifying Computational Neuroscience Approach. Oxford University Press, Oxford.

Rolls, E.T., 2012. Glutamate, obsessive-compulsive disorder, schizophrenia, and the stability of cortical attractor neuronal networks. Pharmacology, Biochemistry, and Behavior 100, 736-751.

Rolls, E.T., 2013a. A biased activation theory of the cognitive and attentional modulation of emotion. Front. Hum. Neurosci. 7, 74.

Rolls, E.T., 2013b. What are emotional states, and why do we have them? Emot. Rev. 5, 241-247.

Rolls, E.T., 2014a. Emotion and Decision-Making Explained. Oxford University Press, Oxford.

Rolls, E.T., 2014b. Emotion and Decision-Making Explained: Précis. Cortex 59, 185-193.

Rolls, E.T., 2015a. Functions of the anterior insula in taste, autonomic, and related functions. Brain Cogn., doi 10.1016/j.bandc.2015.1007.1002.

Rolls, E.T., 2015b. Limbic systems for emotion and for memory, but no single limbic system. Cortex 62, 119-157.

Rolls, E.T., 2016a. Cerebral Cortex: Principles of Operation. Oxford University Press, Oxford.

Rolls, E.T., 2016b. Evolution of the emotional brain, in: Watanabe, S., Hofman, M., Shimizu, T. (Eds.), Evolution of Brain, Cognition and Emotion in Vertebrates. Springer.

Rolls, E.T., Deco, G., 2002. Computational Neuroscience of Vision. Oxford University Press, Oxford.

Rolls, E.T., Deco, G., 2010. The Noisy Brain: Stochastic Dynamics as a Principle of Brain Function. Oxford University Press, Oxford.

Rolls, E.T., Deco, G., 2011. A computational neuroscience approach to schizophrenia and its onset. Neurosci. Biobehav. Rev. 35, 1644-1653.

Rolls, E.T., Deco, G., 2015. Stochastic cortical neurodynamics underlying the memory and cognitive changes in aging. Neurobiol. Learn. Mem. 118, 150-161.

Rolls, E.T., Deco, G., 2016. Non-reward neural mechanisms in the orbitofrontal cortex. Cortex.

Rolls, E.T., Grabenhorst, F., 2008. The orbitofrontal cortex and beyond: from affect to decision-making. Prog. Neurobiol. 86, 216-244.

Rolls, E.T., Grabenhorst, F., Deco, G., 2010. Decision-making, errors, and confidence in the brain. J. Neurophysiol. 104, 2359-2374.

Rolls, E.T., Grabenhorst, F., Margot, C., da Silva, M.A.A.P., Velazco, M.I., 2008a. Selective attention to affective value alters how the brain processes olfactory stimuli. J. Cogn. Neurosci. 20, 1815-1826.

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.

Rolls, E.T., Joliot, M., Tzourio-Mazoyer, N., 2015. Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. Neuroimage 122, 1-5.

Rolls, E.T., Kringelbach, M.L., de Araujo, I.E.T., 2003. Different representations of pleasant and unpleasant odors in the human brain. Eur. J. Neurosci. 18, 695-703.

Rolls, E.T., Loh, M., Deco, G., 2008b. An attractor hypothesis of obsessive-compulsive disorder. Eur. J. Neurosci. 28, 782-793.

Rolls, E.T., Loh, M., Deco, G., Winterer, G., 2008c. Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. Nature Reviews Neuroscience 9, 696-709.

Rudebeck, P.H., Behrens, T.E., Kennerley, S.W., Baxter, M.G., Buckley, M.J., Walton, M.E., Rushworth, M.F., 2008. Frontal cortex subregions play distinct roles in choices between actions and stimuli. J. Neurosci. 28, 13775-13785.

Rudebeck, P.H., Murray, E.A., 2011. Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. J. Neurosci. 31, 10569-10578.

Rushworth, M.F., Kolling, N., Sallet, J., Mars, R.B., 2012. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? Curr. Opin. Neurobiol. 22, 946-955.

Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E., Behrens, T.E., 2011. Frontal cortex and reward-guided learning and decision-making. Neuron 70, 1054-1069.

Sheline, Y.I., Price, J.L., Yan, Z., Mintun, M.A., 2010. Resting-state functional MRI in depression unmasks increased connectivity between networks via the dorsal nexus. Proc. Natl. Acad. Sci. U. S. A. 107, 11020-11025.

Silk, J.S., Siegle, G.J., Lee, K.H., Nelson, E.E., Stroud, L.R., Dahl, R.E., 2014. Increased neural response to peer rejection associated with adolescent depression and pubertal development. Soc. Cogn. Affect. Neurosci. 9, 1798-1807.

Swann, A.C., 2009. Impulsivity in mania. Curr Psychiatry Rep 11, 481-487.

Thorpe, S.J., Rolls, E.T., Maddison, S., 1983. Neuronal activity in the orbitofrontal cortex of the behaving monkey. Exp. Brain Res. 49, 93-115.

Tonegawa, S., Nakazawa, K., Wilson, M.A., 2003. Genetic neuroscience of mammalian learning and memory. Philos. Trans. R. Soc. Lond. B Biol. Sci. 358, 787-795.

Veer, I.M., Beckmann, C.F., van Tol, M.J., Ferrarini, L., Milles, J., Veltman, D.J., Aleman, A., van Buchem, M.A., van der Wee, N.J., Rombouts, S.A., 2010. Whole brain resting-state analysis reveals decreased functional connectivity in major depression. Front. Syst. Neurosci. 4.

Wang, X.J., 2002. Probabilistic decision making by slow reverberation in cortical circuits. Neuron 36, 955-968.

Wessa, M., Kanske, P., Linke, J., 2014. Bipolar disorder: a neural network perspective on a disorder of emotion and motivation. Restor. Neurol. Neurosci. 32, 51-62.

Whelan, R., Conrod, P.J., Poline, J.B., Lourdusamy, A., Banaschewski, T., Barker, G.J., Bellgrove, M.A., Buchel, C., Byrne, M., Cummins, T.D., Fauth-Buhler, M., Flor, H., Gallinat, J., Heinz, A., Ittermann, B., Mann, K., Martinot, J.L., Lalor, E.C., Lathrop, M., Loth, E., Nees, F., Paus, T., Rietschel, M., Smolka, M.N., Spanagel, R., Stephens, D.N., Struve, M., Thyreau, B., Vollstaedt-Klein, S., Robbins, T.W., Schumann, G., Garavan, H., Consortium, I., 2012. Adolescent impulsivity phenotypes characterized by distinct brain networks. Nat. Neurosci. 15, 920-925.

Wise, S.P., 2008. Forward frontal fields: phylogeny and fundamental function. Trends Neurosci. 31, 599-608.