Catecholamine function, brain state dynamics, and human cognition
Brink, R.L. van den

Citation
Brink, R. L. van den. (2017, November 7). *Catecholamine function, brain state dynamics, and human cognition*. Retrieved from https://hdl.handle.net/1887/54947

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/54947 holds various files of this Leiden University dissertation.

Author: Van den Brink R.L.
Title: Catecholamine function, brain state dynamics, and human cognition
Issue Date: 2017-11-07

Catecholamine function, brain state dynamics, and human cognition

Proefschrift ter verkijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus, Prof.mr. C.J.J.M. Stolker

volgens besluit van het College voor Promoties

te verdedigen op 7 November 2017

klokke: 13:45u

Door

Ruud Lucas van den Brink

geboren in 1987

te Delft

Leden van de promotiecomissie

Prof.dr. R.K. Ridderinkhof
Prof.dr. B.U. Forstmann
Dr. M.X. Cohen

Promotor

Prof.dr. S.T. Nieuwenhuis

Co-promotor

Dr. H. van Steenbergen

Table of contents

1. General introduction

1.1 Introduction

The locus coeruleus (LC) is a small nucleus that is located in the pontine tegmentum, and derives its name (literally meaning 'blue spot') from its color, which is a result of neuromelanin deposits within its cell bodies. The LC projects widely to the forebrain (Figure 1) where it releases norepinephrine (NE; also referred to as noradrenaline) (Aston-Jones et al., 1984; Berridge and Waterhouse, 2003).

Catecholamines such as NE do not have a unitary effect on their target neurons, but instead influence the function of other neurotransmitters, a process that is known known as neuromodulation. By virtue of the LC's wide projection profile and the neuromodulatory properties of NE, the LC-NE system profoundly influences neural firing characteristics and associated cognitive processes (Berridge and Waterhouse, 2003; Aston-Jones and Cohen, 2005; Bouret and Sara, 2005; Yu and Dayan, 2005).

In this introductory chapter, an overview of current findings and accounts pertaining to the LC-NE system and its relationship with 'brain state' (defined further below) and cognition is presented, followed by a summary of the chapters of this dissertation.
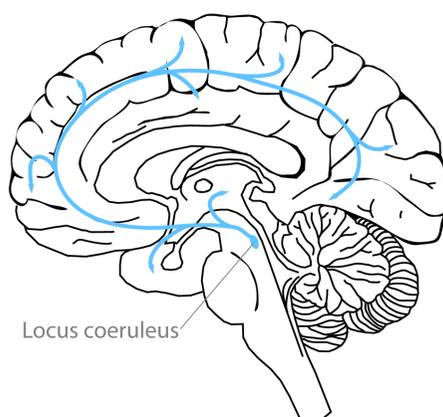
Figure 1. Schematic illustration of the locus ceuruleus and its projections.

1.2 Anatomical overview of the LC-NE system

The LC is situated directly anterior to the fourth ventricle, in the dorsal pons. In the healthy adult brain, the LC contains approximately 35,000 neurons in either hemisphere, which amount to a bilateral nucleus that is roughly the size of a grain of rice (Mouton et al., 1994). Despite its size, the LC sends wide, ascending, projections to the forebrain. For example, major innervation targets of the LC include the amygdala, hippocampus, thalamus, basal ganglia, cerebellum, spinal cord, and all

cortical lobes (Aston-Jones et al., 1984), Consequently, the LC is the dominant source of NE in the central nervous system.

While the projections that emanate from the LC have long been believed to be homogenously distributed across the brain, recent evidence suggests that distinct portions of the LC preferentially innervate select brain areas (Chandler et al., 2014; Schwarz and Luo, 2015; Schwarz et al., 2015). Moreover, recent evidence suggests that the LC does not solely supply the brain with NE, but, but may also release the catecholaminergic neuromodulator dopamine (DA) (Devoto et al., 2004; Kempadoo et al., 2016; Takeuchi et al., 2016). Although this introductory chapter is focused primarily on NE function, it should be noted that NE shares some of its functional properties with DA, such as its effect on neural gain (see below).

1.3 Functional overview of the LC-NE system and theories of LC-NE function in cognition

When released from noradrenergic terminals, NE acts on receptors that can be divided into three major classes: α1, α2, and β. Following release, NE is cleared from the synaptic cleft by the NE transporter (NET). Due to the cortical paucity of DA transporters, NET is also responsible for the reuptake of DA within the cortex. In chapters 3 and 4, NET is blocked pharmacologically in order to causally manipulate catecholamine levels in healthy human participants.

Direct measurements in monkeys have suggested that the LC has two distinct modes of operation (Aston-Jones and Cohen, 2005). In the 'phasic' mode, the LC fires rapidly and transiently in response to salient, novel, or otherwise behaviorally relevant stimuli. During bursts of phasic activity, LC neurons discharge *en masse* in a highly synchronized manner as a consequence of direct electrical coupling between individual neurons of the LC (Ishimatsu and Williams, 1996), although asynchronous firing has also been reported (Totah et al., 2017). Phasic bursts show a close temporal relationship with behavioral responses, suggesting that bursts occur to facilitate internally generated top-down decision processes.

Conversely, in the 'tonic' mode, the LC shows a sustained and regular pattern of firing, without an immediate temporal correspondence between LC discharges and behavioral responses. Across extended periods of time, however, the level of tonic LC activity cofluctuates with task performance, whereby periods of strong tonic activity are marked by distractible behavior, periods of weak tonic activity are marked by under-arousal or sleep, and periods of intermediate tonic activity are marked by (near-) optimal task performance (Aston-Jones and Cohen, 2005). Moreover, the strongest phasic LC activity occurs at time points of intermediate tonic firing.

At the synaptic level, NE can enhance the effect of both excitatory and inhibitory input (Moises et al., 1979; Rogawksi and Aghajanian, 1980). These and other findings have led to the view that NE boosts the efficacy of synaptic interactions between

neurons (Berridge and Waterhouse, 2003), a phenomenon that is known as gain modulation (Aston-Jones and Cohen, 2005). An increase in neural gain results in an increased difference in firing rates between strongly and weakly active neurons (Waterhouse et al., 1998), and consequently yields a system-wide facilitation of signal transmission (Servan-Schreiber et al., 1990; Aston-Jones and Cohen, 2005). In other words, an increase in neural gain does not necessarily modify the likelihood of a single neuron responding to its input, but at the system-level, increased signal propagation emerges and allows dominant neural firing patterns to prevail at the expense of less dominant firing patterns (Servan-Schreiber et al., 1990). The gain-regulating properties of the LC-NE system form a key ingredient of adaptive gain theory (Aston-Jones and Cohen, 2005), which is discussed further below.

In addition to the regulation of neural gain, work done primarily on crustaceans has revealed that NE has the ability to fundamentally reshape the firing properties of the combined set of its target neurons (Marder, 2012; Bargmann and Marder, 2013; Marder et al., 2014). For example, depending on the concentration of NE, the firing pattern in target neurons, and the presence of other neuromodulators, NE can elicit a shift towards rapid, synchronous bursting, or intermittent and asynchronous firing (Marder, 2012; Marder et al., 2014). These findings form the basis of another influential account of LC-NE function, 'network reset' (Bouret and Sara, 2005), also discussed further below.

The study of the LC-NE system in humans has been limited by methodological problems associated with its size and location. Consequently, almost everything we know about its function is based on animal work and computational modeling. Nevertheless there are several major theoretical accounts about how the LC-NE system affects brain state, cognition, and behavior, making it a rare example in cognitive neuroscience where theory outweighs data. The major theoretical accounts regarding LC-NE function will be discussed next.

1.3.1   Adaptive gain theory

In their adaptive gain theory, Aston-Jones and Cohen (2005) propose that the LC-NE system balances the trade-off between exploitation of the current task set and exploration of alternative task sets through NE's effect on neural gain. Because phasic LC activity predominantly occurs only in response to salient or motivationally relevant stimuli, phasic LC activity promotes immediately goal-relevant sensory information at the expense of goal-irrelevant and distracting information, and consequently, fast and accurate behavioral responses to goal-relevant information. Thus, timely bursts of LC activity result in the exploitation of reward from the current task set. Conversely, periods of sustained LC firing (tonic activity) provide a nonspecific and temporally less constrained amplification of incoming sensory information, and thus enable alternative, potentially rewarding, task sets to be explored. Thus, by balancing the trade-off between tonic and phasic activity, the LC

can orchestrate shifts in behavioral strategies in accordance with environmental demands.

1.3.2   The network reset account

The network reset account by Bouret and Sara (2005) describes NE as inducing large-scale neuronal reorganization to promote behavioral adaptation following environmental changes in behavioral requirements. In this context, phasic bursts of LC activity elicit a dynamic reorganization of the LC's target neuronal networks, thereby provoking or facilitating a cognitive shift in task set when such a shift is needed. In contrast to adaptive gain theory, the network reset account proposes that the distractible behavior that accompanies periods of strong tonic LC activation results from inappropriate and repeated cognitive shifts that are a consequence of the sustained and temporally nonspecific characteristics of tonic LC activity. However, in a broader sense, both adaptive gain theory and the network reset account converge on the notion that the LC facilitates goal-directed behavior via adaptation to environmental demands.

1.3.3   Unexpected uncertainty and GANE

In addition to the adaptive gain theory and network reset account discussed above, other accounts exist. The latter, however, will be discussed only briefly here because they are not central to this dissertation.

The Bayesian theory outlined by Yu and Dayan (2005) proposes that NE signals unforeseen changes in task demands. This 'unexpected uncertainty' reflects changes in environmental parameters that require an appropriate modification of predictions about the environment, and therefore a change in behavior. In this sense, prolonged heightened NE release prompts behavioral adaptation, akin to the heightened exploration resulting from increased tonic activity in adaptive gain theory. Moreover, the theory by Yu and Dayan (2005) proposes that NE functions (partially) antagonistically with another neuromodulator, acetylcholine, which in this framework signals known uncertainty about task contingencies.

Another, more recent, and neurobiologically based account (Mather et al., 2015) proposes that the LC-NE system promotes neural representations of goal-relevant information through the 'ignition' of local hotspots with locally concentrated pockets of the neurotransmitter glutamate. In this 'glutamate amplifies noradrenergic effects' (GANE) account, high-priority perceptual representations are favored over low-priority representations through the synergetic action of glutamate and phasically released NE.

The accounts outlined above broadly converge on the notion that NE prompts behavioral adaptation to the demands of the environment. Where these accounts differ lies mostly in how NE is proposed to orchestrate such behavioral adaptation

neurally. As discussed below, sensory information that informs an agent of the state of the environment is not processed neurally as a linear function of the stimulus, but instead interacts with ongoing, intrinsic, neural activity. Resent research indicates that NE may play a critical role in shaping the state of intrinsic neural activity and its interplay with external sensory information, offering new insights into the neural mechanisms by which LC-NE system dynamically regulates behavior.

1.4 The junction between brain state, neuromodulation, and cognition

Brain activity does not simply follow from external (sensory) input, but instead arises from a nonlinear interaction between sensory input and spontaneous - internally generated - brain activity (Luczak et al., 2009; Harris and Thiele, 2011). The state of such spontaneous activity, and the way it shapes cortical responses to sensory input, fluctuates dynamically over time.

A well-known example of fluctuations in dynamic brain state is the sleep-wake cycle. In the deep stages of sleep, neural activity alternates rhythmically between mass-synchronized spiking and near-complete quintessence. These low-frequency fluctuations form a stark contrast with the cortical state that is seen during alert wakefulness, in which neurons fire predominantly asynchronously (Pace-Schott and Hobson, 2002). More recently, less prominent fluctuations between such synchronous and asynchronous cortical firing states have been shown to occur within periods of wakefulness as well (Crochet and Petersen, 2006; Greenberg et al., 2008; Poulet and Petersen, 2008). The membrane potential of cortical neurons and their responsivity to input covary with fluctuations in cortical state (Zagha et al., 2013), leading to the view that the brain's repertoire of possible activity states – the joint set of parameters that are subject to rapid variation, such as gamma power, spiking correlation, and intracellular potentials – is determined by the brain's dynamic state (Okun and Lampl, 2008; Luczak et al., 2009; Harris and Thiele, 2011).

The ability to select and respond to the appropriate sensory information while ignoring irrelevant sensory information, known as top-down attention, shows similar neural characteristics as the desynchronized cortical state. In addition to the near-complete absence of neural and behavioral responsivity to sensory input during sleep, fluctuations in cortical state during wakefulness determine responsivity of cortical neurons to relevant sensory input (Reimer et al., 2014; McGinley et al., 2015b) as well as an animal's ability to respond appropriately to such input (McGinley et al., 2015a). Specifically, local desynchronization in neural population activity co-occurs with better signal detection performance at the behavioral level (McGinley et al., 2015a). Accordingly, attention to task-relevant stimuli has been proposed to rely on similar neural mechanisms as global cortical state change (Harris and Thiele, 2011).

Interestingly, the activity of the LC-NE system fluctuates as a function of brain state. For instance, the transition from wakefulness to the onset of sleep is marked by a progressive reduction of LC firing that continues until near-complete silence during paradoxical sleep (Aston-Jones and Bloom, 1981). Moreover, rapid changes in cortical activity state that occur during wakefulness co-occur with NE release within the cortex (Reimer et al., 2016), and with fluctuations in pupil diameter (Reimer et al., 2014). Fluctuations in pupil diameter in turn co-vary with activity in the LC (Aston-Jones and Cohen, 2005; Murphy et al., 2014b; Varazzani et al., 2015; Joshi et al., 2016). As mentioned earlier, the changes in brain state that accompany changes in pupil diameter and cortical NE release also co-vary with behavioral signal detection performance (McGinley et al., 2015a). Moreover, the magnitude of the pupil-linked attentional orienting response predicts the degree of behavioral adaptation following performance errors (Murphy et al., 2016).

The above indicates that cortical state, noradrenergic neuromodulation, and cognitive processes such as attention, are tightly intertwined. It is this junction that forms the central theme of this dissertation. Below, an overview of the chapters of this dissertation is presented, and each chapter is discussed within the context of brain state, neuromodulation, cognition, or a combination of these sub-themes.

1.5 An overview of the current dissertation

1.5.1 Chapter 2: Post-Error Slowing as a Consequence of Disturbed Low-Frequency Oscillatory Phase Entrainment

One of the most ubiquitous findings across reaction time (RT) tasks is that RTs slow down on trials following errors (Rabbitt, 1966; Laming, 1979). This phenomenon is known as post-error slowing (PES) and occurs across various task conditions and response modalities (Gehring and Fenscik, 2001b; Ridderinkhof, 2002; Endrass et al., 2005; Cavanagh et al., 2009a; Cohen et al., 2009; Dudschig and Jentzsch, 2009; Jentzsch and Dudschig, 2009; Eichele et al., 2010). PES has been suggested to reflect the strategic adjustment of behavior (Botvinick et al., 2001; Dutilh et al., 2012a) as well as a detrimental processing interference caused by the error (Jentzsch and Dudschig, 2009; Notebaert et al., 2009).

As discussed in the previous section, brain state influences our ability to select and respond to relevant sensory information. One line of literature suggests that, under conditions of rhythmic stimulus presentation, our brain may dynamically adjust its activity state in order to actively anticipate incoming stimuli by rhythmically aligning neural oscillations to the stimulus stream (Lakatos et al., 2008; Schroeder and Lakatos, 2009; Saleh et al., 2010; Stefanics et al., 2010b; Besle et al., 2011a; Henry and Obleser, 2012). Such 'entrainment' ensures that goal-relevant sensory

information is processed in the optimal neural context, and thus facilitates appropriate behavioral responses.

In chapter 2 of this dissertation, we test the novel hypothesis that PES may reflect a temporary perturbation of the entrained state. To test this hypothesis, we measured oscillatory EEG dynamics while human subjects performed a demanding discrimination task under time pressure. We show that brain state actively adjusts to the stimulus presentation rhythm by entraining low-frequency neuronal oscillations, and that the phase of these oscillations at stimulus onset predicts the speed of responding. Importantly, we show that entrainment is disrupted following errors, and that the degree of phase disturbance is closely related to the degree of PES on the subsequent trial.

Our results are consistent with the orienting account of PES, which proposes that errors, by virtue of being surprising events, result in the temporary reorientation of attention away from the current task, and as a consequence, longer RTs on the following trial (Notebaert et al., 2009). Interestingly, the LC-NE system is known to fire phasically in response to salient and surprising events, and theoretical accounts exist that link PES to the LC-NE system (Cohen et al., 2000; Nunez Castellar et al., 2010). Moreover, and in line with the concept of an orienting response, phasic NE release has also been proposed to act as a neural interrupt signal, whereby unexpected events (e.g., errors) lead to a reset and reorganization in target neuronal networks, and subsequent behavioral adaptation (Bouret and Sara, 2005; Dayan and Yu, 2006). We speculate that it is possible that the entrained brain state and consequent mode of behavioral responding are disrupted by an error-evoked orienting response in the LC-NE system (Nieuwenhuis et al., 2010; Nunez Castellar et al., 2010; Ullsperger et al., 2010). An orienting response (or interrupt signal) may aid task performance at longer intervals between errors and subsequent trials by facilitating the appropriate adjustment of behavior (Murphy et al., 2016).

1.5.2 Chapter 3: Catecholaminergic Neuromodulation Shapes Intrinsic MRI Functional Connectivity in the Human Brain

Spontaneously generated, ongoing, brain activity is correlated across brain regions (Biswal et al., 1995; Leopold et al., 2003; Fox and Raichle, 2007; Hiltunen et al., 2014). Moreover, the global structure of correlated activity changes dynamically with alterations in conscious state (Barttfeld et al., 2015) and task conditions (Nir et al., 2006; Sepulcre et al., 2010). In chapter 3 of this dissertation, we test the hypothesis that fluctuations in the strength of these intrinsic correlations are induced by the LC-NE system (Leopold et al., 2003; Drew et al., 2008; Schölvinck et al., 2010). Using a double-blind placebo-controlled cross-over design, we pharmacologically increase synaptic NE and DA levels by administering atomoxetine, a selective NET blocker, and examine the effects on the strength and spatial structure of 'resting-state' MRI functional connectivity.

As discussed earlier in this introductory chapter, NE increases neural gain, and as a result facilitates brain-wide signal transmission (Servan-Schreiber et al., 1990; Berridge and Waterhouse, 2003; Aston-Jones and Cohen, 2005). Computational modeling has indicated that such an increase in signal transmission should result in a brain-wide increase in the strength of both positively and negatively correlated activity, and the degree of clustering of that activity (Eldar et al., 2013). Eldar et al. (2013) accordingly showed that increased pupil diameter is indeed accompanied by such an increase in the strength of connectivity and clustering. Based on these findings, we predicted that the administration of atomoxetine should increase the strength and clustering of connectivity. Moreover, given the putative spatial aselectivity of the LC-NE system, we expected that an atomoxetine-induced increase in connectivity should be homogenous across the brain.

However, in contrast to an atomoxetine-induced increase in the strength of connectivity, we show that atomoxetine *reduced* the strength of inter-regional correlations across three levels of spatial organization. Furthermore, this modulatory effect on intrinsic correlations exhibited a substantial degree of spatial specificity: the decrease in functional connectivity showed an anterior-posterior gradient in the cortex, depended on the strength of baseline functional connectivity, and was strongest for connections between regions belonging to distinct intrinsic connectivity networks.

Our findings are the first to show that neuromodulation shapes the topography of intrinsic correlations in the human brain in a spatially specific manner. The unexpected reduction of the strength of connectivity indicates that neuromodulation may shape intrinsic correlations in a brain state-dependent manner, which dovetails with positron emission tomography findings (Coull et al., 1999) and theoretical proposals (Mather et al., 2015), but is difficult to account for by a global modulation of neural gain alone. Moreover, spatial specificity in the effect of atomoxetine on intrinsic correlations may be explained by recent findings that the projection profile of the LC-NE system is more heterogeneous than once thought (Chandler et al., 2014; Schwarz and Luo, 2015; Schwarz et al., 2015), and by the heterogeneous distribution of noradrenergic receptors (e.g. α2) across the cortex (Zilles and Amunts, 2009; Nahimi et al., 2015).

1.5.3   Chapter 4: Catecholamines Modulate Intrinsic Long-range Correlations in the Human Brain

In chapter 4, we test a prediction from the network reset account (Bouret and Sara, 2005): an increase in NE should lead to a reorganization of brain functional networks. To do so, we reanalyzed the dataset used in chapter 3: a double-blind placebo-controlled cross-over design in which we pharmacologically increase synaptic NE and DA levels by administering the selective NET blocker atomoxetine. We applied two complementary analysis approaches to examine the effect of NE on fine-grained

patterns of intrinsic functional connectivity patterns: 'dual regression' and 'spatial mode decomposition'. As opposed to chapter 3, in chapter 4 we examine if atomoxetine results in changes to the spatial structure (topology) of intrinsic fMRI correlations rather than a modulation of their strength alone.

Both analysis approaches provided converging evidence for an atomoxetine-related reduction in correlations between distributed brain regions, specifically sensory and motor-related networks. Additionally, spatial mode decomposition revealed a shift in dominance from left to right-lateralized frontoparietal network cofluctuations. Importantly, the pre-dominant effect of atomoxetine was a quantitative change to correlations within existing functional networks that left the spatial structure of these networks intact, rather than a reconfiguration of network topology.

Our findings are consistent with earlier work on primates (Guedj et al., 2016) which demonstrated similar connectivity atomoxetine-induced reductions in sensory and motor-related networks. However, we demonstrate that such reductions can be quantitative in nature, rather than necessarily stemming from a topological reconfiguration of network structure as would be predicted by the network reset account (Bouret and Sara, 2005). We conclude that catecholamines modulate dynamic changes in the strength of intrinsic inter-regional correlations, which may serve to coordinate flexible modulations of network interactions in order to facilitate goal-directed behavior.

1.5.4   Chapter 5: Pupil Diameter Tracks Lapses of Attention

Sustained attention, our ability to continuously monitor and respond to goal-relevant sensory information, is limited. Studies on the relationship between lapses of attention and psychophysiological markers of attentional state, such as pupil diameter, have yielded contradicting results. Adaptive gain theory (Aston-Jones and Cohen, 2005) predicts that baseline pupil diameter should show an inverted-U shaped relationship with attentional performance, whereby most lapses of attention occur in both periods where the pupil is relatively large, and where it is relatively small. In chapter 5, we test this prediction directly. Moreover, we explore additional markers of attentional state, based partially on recent research that showed a close relationship between the derivative of pupil diameter and brain state and attentional performance (Reimer et al., 2014; McGinley et al., 2015a).

We investigate the relationship between tonic fluctuations in pupil diameter and performance on a demanding sustained attention task. We found robust linear relationships between baseline pupil diameter and several measures of task performance, suggesting that attentional lapses tended to occur when pupil diameter was small. However, these observations were primarily driven by the joint effects of time-on-task on baseline pupil diameter and task performance. The linear relationships disappeared when we statistically controlled for time-on-task effects and were replaced by consistent inverted U-shaped relationships between baseline pupil

diameter and each of the task performance measures, such that most false alarms and the longest and most variable response times occurred when pupil diameter was both relatively small and large.

Finally, we observed strong linear relationships between the temporal derivative of pupil diameter and task performance measures, which were largely independent of time-on-task. Our results help to reconcile contradicting findings in the literature on pupil-linked changes in attentional state, and are consistent with the adaptive gain theory of LC-NE function. Moreover, our results suggest that the derivative of baseline pupil diameter is a potentially useful psychophysiological marker that could be used in the on-line prediction and prevention of attentional lapses.

1.5.5 Chapter 6: Task-free Spectral EEG Dynamics Track and Predict Patient Recovery From Severe Acquired Brain Injury

As previously discussed, our ability to process and respond to sensory information is dependent on the dynamic brain state. One prominent example in which brain state is fundamentally altered is that of disorders of consciousness resulting from brain injury. Some of these patients develop signs of awareness, while other patients remain in a state of unresponsiveness (Jennett and Plum, 1972; Laureys et al., 2004). At the neural level, the pathophysiological signatures of disorders of consciousness are reminiscent of hypoactivity in the LC-NE system that occurs during sleep or under-arousal, and concurrent mass-synchronization of cortical neurons. As shown in chapters 3 and 4 of this dissertation, brain activity that is synchronized across cortical areas in the absence of sensory input is susceptible to noradrenergic neuromodulation. In chapter 6 of this dissertation, we explore if the state (quantified as amplitude and connectivity) of such synchronized cortical activity in the absence of sensory input can be used to track and predict the level of awareness of patients with disorders of consciousness and their respective level of recovery. We analyze an existing dataset of patients who participated in an 'Early Intensive Neurorehabilitation Programme' (Eilander et al., 2005; Wijnen et al., 2007).

We show that compared to healthy control participants, patients showed a general 'slowing down' of cortical rhythms, whereby low-frequency (synchronized) cortical states are relatively dominant. Moreover, across the course of their recovery, patients exhibit nonlinear frequency band-specific changes in spectral amplitude and connectivity metrics, and these changes align well with the metrics' frequency band-specific diagnostic value. Remarkably, connectivity during a single task-free EEG measurement could predict the level of patient recovery approximately 3 months later with 75% accuracy.

Our findings show that amplitude and connectivity metrics of spectral brain state track patient recovery in a longitudinal fashion, and that these metrics are robust pathophysiological markers that can be used for the automated diagnosis and prognosis of disorders of consciousness. These metrics can be acquired

inexpensively at bedside, and are fully independent of the patient's neurocognitive abilities, which offers substantial improvements on existing methodologies. Lastly, our findings tentatively suggest that the relative preservation of ascending and recurrent interactions between the cortex and subcortical nuclei (speculatively, the thalamus or ascending arousal systems such as the LC), putatively responsible for desynchronized cortical states (Schiff, 2010; Schiff et al., 2014), may predict the later reemergence of awareness. Thus, our findings shed new light on the pathophysiological brain state-related processes that underlie disorders of consciousness.

2. Post-Error Slowing as a Consequence of Disturbed Low-Frequency Oscillatory Phase Entrainment

Abstract

A common finding across many reaction-time tasks is that people slow down on trials following errors, a phenomenon known as post-error slowing. In the present study we tested a novel hypothesis about the neural mechanism underlying post-error slowing. Recent research has shown that when task-relevant stimuli occur in a rhythmic stream, neuronal oscillations entrain to the task structure, thereby enhancing reaction speed. We hypothesized that under such circumstances post-error slowing results from an error-induced disturbance of this endogenous brain rhythm. To test this hypothesis, we measured oscillatory EEG dynamics while human subjects performed a demanding discrimination task under time pressure. We found that low-frequency neuronal oscillations entrained to the stimulus-presentation rhythm, and that low-frequency phase at stimulus onset predicted the speed of responding. Importantly, we found that this entrainment was disrupted following errors, and that the degree of phase disturbance was closely related to the degree of post-error slowing on the subsequent trial. These results describe a new mechanism underlying behavioral changes following errors.

Chapter 2

2.1 Introduction

One of the most common findings in empirical studies employing reaction time (RT) measures is that RTs slow down on trials following errors, a phenomenon known as post-error slowing (PES, Rabbitt, 1966; Laming, 1979). PES has been shown to occur across many different tasks and response modalities, including go/no-go (Cohen et al., 2009); flanker (Cavanagh et al., 2009b; Eichele et al., 2010); categorization (Dudschig and Jentzsch, 2009; Jentzsch and Dudschig, 2009); Stroop (Gehring and Fenscik, 2001a); Simon (Ridderinkhof, 2002); and saccade countermanding tasks (Endrass et al., 2005). Researchers have proposed several accounts of the cognitive mechanisms that are responsible for PES (Danielmeier and Ullsperger, 2011), suggesting that PES can arise from strategic adjustments in response caution (Botvinick et al., 2001; Dutilh et al., 2012a) as well as a detrimental processing interference caused by the error (Jentzsch and Dudschig, 2009; Notebaert et al., 2009). However, although studies have started to elucidate the neural mechanisms underlying error-related strategic adjustments (King et al., 2010; Danielmeier et al., 2011; Narayanan et al., 2013), the neural basis of error-related processing interference is still poorly understood.

In the present study, we tested the hypothesis that this component of PES results from an error-evoked disturbance of internally generated brain rhythmicity. It has been suggested that under circumstances of high external predictability, the timing of endogenous periodic processes is systematically aligned with the environmental rhythmicity (Large and Jones, 1999; Schroeder and Lakatos, 2009). Such temporal alignment can be implemented by neuronal oscillations, which are rhythmic fluctuations in the excitability of large-scale neuronal ensembles (Wang, 2010). This oscillatory phase locking, or entrainment, ensures that behaviorally relevant stimuli are processed in the optimal neuronal context, thereby optimizing speed and accuracy of behavioral responses. Indeed, low-frequency neuronal oscillations have been shown to align with stimulus-presentation rates, are modulated by expectations about stimulus onset, and predict the latency of behavioral responses (Lakatos et al., 2008; Schroeder and Lakatos, 2009; Saleh et al., 2010; Stefanics et al., 2010a; Besle et al., 2011b; Henry and Obleser, 2012).

To test our hypothesis that errors slow down subsequent responses through a disturbance of low-frequency EEG oscillatory phase entrainment, we examined the oscillatory dynamics of EEG signals from participants performing a modified Eriksen flanker task under high time pressure. We found that low-frequency neuronal oscillations entrained to the stimulus-presentation rhythm, and that low-frequency phase at stimulus onset predicted the speed of responding. Importantly, we found that this entrainment was disrupted following errors, and that the degree of phase disturbance was closely related to the degree of PES on the subsequent trial. These results support our novel hypothesis regarding the neural origin of PES.

2.2 Materials and Methods

Participants. Twenty-one participants (aged 17-29 years, 17 female, all right-handed) gave informed consent to take part in this study. Following EEG artifact rejection, one participant was excluded from further analysis due to an insufficient number of error trials (<30). All participants had normal or corrected-to-normal vision, and were free from any neurological or psychiatric disorders. Participants were recruited via the Leiden University Research Participation website and received €7,50/h or course credit. The experiment was approved by the Leiden University Institute of Psychology Ethics Committee.

Behavioral task. The participants performed a modified version of the flanker task (Eriksen and Eriksen, 1974), in which a target letter ('H', 'K', 'C' or 'S') was flanked by three identical flanker letters ('H', 'K', 'C' or 'S') on each side. Participants had to classify the target letter by giving one of two left-hand responses or one of two right-hand responses. The flanking letters were always incongruent with the target letter (e.g., SSSHSSS or KKKCKKK) and mapped to a finger of the hand opposite to the hand associated with the correct response. This ensured similar difficulty across trials, enabling cross-trial comparisons of RTs, while retaining a sufficiently high error rate.

Stimuli were presented in black on a white background for 200 ms at 2.77° horizontal visual angle and at a viewing distance of 120 cm. Stimulus onset asynchronies (SOAs) were randomly selected from a uniform distribution with a mean of 1350 ms, and varying between 1200 ms and 1500 ms with 50 ms increments (Figure 1A). During the inter-trial interval, a black fixation cross was shown. Participants were instructed to keep their eyes fixated on the cross at all times. Between task blocks participants received RT and accuracy feedback, and were pressed for speed. In total there were 10 blocks of 104 trials each. Participants practiced the task beforehand (120 trials), to ensure they understood the task.

Because several studies have found that PES is exclusively found for errors of which participants are aware (Hughes and Yeung, 2011; Murphy et al., 2012), we conducted a pilot experiment (N=4) to verify whether participants detected the errors made in our challenging version of the flanker task. If they detected an error, participants were to press an error-signaling button (space bar) with their thumb, immediately after the error, and before the next trial started (an additional task on top of the already demanding flanker task). All other methods were identical to those described for the main experiment. Participants correctly signaled over 76% of their errors, which provides a lower bound on the percentage of aware errors, while misclassifications of correct responses as errors were rare (<2%). This suggests that participants were aware of the large majority of their errors.

Behavioral data acquisition and analysis. The experiment was run on a personal computer with an 18-inch monitor. Stimulus presentation and the recording of responses were performed using E-Prime (v2.0; Psychology Software Tools, Inc.).

In order to quantify PES, researchers usually subtract the mean reaction time (MRT) on post-error trials from the MRT on post-correct trials (MRT$_{post-error}$ − MRT$_{post-correct}$). This will hereafter be referred to as PES$_{traditional}$ (Dutilh et al., 2012b). However, Dutilh et al. (2012b) have shown that this measure of PES is vulnerable to confounds related to global performance fluctuations. They therefore proposed an alternative measure of PES, dubbed PES$_{robust}$. Following this method, we conducted a pairwise comparison of correct trials around each error (RT$_{post-error}$ − RT$_{pre-error}$), resulting in single-trial values of PES. For the calculation of PES$_{robust}$, we included error trials that were both preceded and followed by at least one correct trial. To test for group-level significance of PES$_{traditional}$ and PES$_{robust}$, MRTs on post-error trials were compared to post-correct and pre-error trials, respectively, with one-tailed paired-sample *t*-tests. Additionally, Pearson's correlation was used to determine whether PES$_{traditional}$ and PES$_{robust}$ were correlated.

EEG recording and preprocessing. EEG data were recorded using a BioSemi ActiveTwo system from 18 electrodes placed according to the international 10/20 system: F3; Fz; F4; C3; Cz; C4; P3; Pz; P4; PO7; PO3; POz; PO4; PO8; O1; Oz; O2; and Iz. Additionally, a reference electrode was placed on each earlobe, and bipolar electro-oculogram (EOG) recordings were obtained from electrodes placed approximately 1 cm lateral of the outer canthi (horizontal EOG) and from electrodes placed approximately 1 cm above and below the left eye (vertical EOG). During acquisition, impedances were kept below 30 kΩ. The EEG signal was pre-amplified at the electrode to improve the signal-to-noise ratio with a gain of 16×, and digitized at 24-bit resolution with a sampling rate of 1024 Hz. Each active electrode was measured online with respect to a common mode sense (CMS) active electrode producing a monopolar (non-differential) channel.

All EEG data were analyzed in MATLAB 2011b, using the EEGLAB toolbox (Delorme and Makeig, 2004a) and custom in-house code. First, EEG data were down sampled to 512 Hz and re-referenced off-line to the average of the earlobe electrodes. Next, to remove drifts, the continuous EEG data were high-pass filtered offline at 0.5 Hz with a zero-phase-shift, two-way, least-squares finite impulse response filter (constructed using MATLAB's 'fir1' function which smooths the filter kernel using a Hamming window by default). Zero-phase-shift filters prevent the introduction of filter artefacts that could distort real oscillatory phase information. Filtering the continuous time course rather than epoched data also prevented edge artifacts from contaminating the data. Nevertheless, to verify that high-pass filtering did not influence low-frequency phase estimates, we also ran the relevant analyses on data to which no high-pass filter was applied, and found that it did not influence the results.

Additionally, the EEG data were notch-filtered at 50 Hz to remove line-noise. Following filtering, the continuous data were segmented into epochs ranging from -1 to 3s centered on stimulus onset and baseline-corrected by subtracting the average offset during the -400 to -100 ms pre-stimulus window. Next, the following types of trials were excluded from further analysis: trials in which the participant failed to

respond, trials which were part of a sequence of more than three consecutive errors, and trials with an RT exceeding 1200 ms. Trials with transient artifacts and eye movements were manually rejected. Following trial rejection, per participant on average 609 correct trials, and 78 error trials remained (393 and 47 respective lower limits) and 693 post-correct trials and 94 post-error trials remained (501 and 51 respective lower limits). Next, eye blinks and continuous electromyogram (EMG) artifacts were identified using JADE independent component analysis as implemented in the EEGLAB toolbox (Delorme and Makeig, 2004a) and the corresponding components were removed from the data.

ERP analysis. To confirm that our task showed the event-related potential (ERP) components that are typically found during flanker tasks—the error-related negativity and error positivity—we computed response-locked ERPs for correct and error trials, and baseline-corrected them by subtracting the average offset during the -100 to 0 ms pre-response window. These ERPs were then compared on each time-point with two-tailed *t*-tests with a *p*-threshold of 0.001 or less, and cluster-corrected for multiple comparisons across time-points (Maris and Oostenveld, 2007).

Time-frequency analyses. Spectral power and phase dynamics were extracted via Morlet wavelet decomposition. Wavelet decomposition involves convolving the data with a set of Gaussian-windowed complex sine waves, here defined as:

$$\psi_{f_w} = e^{2i\pi f_w t} \cdot e^{\frac{-t^2}{2s_w^2}}$$

where f_w denotes frequency, which ranged from 0.5 to 30 Hz with 30 logarithmically-spaced steps. Time is denoted by t, and s_w determines the width of the Gaussian window, and thus the tradeoff between time and frequency precision. A wider Gaussian will result in a wider wavelet, and thus results in more temporal smearing of instantaneous power/phase estimates. Because we were primarily interested in lower frequencies, and the lower frequencies inherently span a wider temporal range, we preferred to minimize temporal smearing at the expense of frequency resolution in the lower frequency bands. We therefore linearly increased the Gaussian width with frequency, such that:

$$s_w = \frac{c_w}{2\pi f_w}$$

where c_w denotes the number of wavelet cycles, which ranged from 3 to 12.

To accommodate the large wavelet width at the lower frequencies, prior to running wavelet decomposition we expanded each data epoch with a mirror (time-reversed) image of itself such that each epoch consisted of a mirror image of the epoch, then the 'true' epoch, and another mirror image following it. This increased the effective epoch size from 4 s to 12 s while preserving data continuity, thereby preventing the introduction of transients in the signal that could cause edge artifacts (Cohen, 2014b).

After convolution, frequency-specific instantaneous power (P) is given by:

$$P(f_w, t) = ReM(f_w, t)^2 + iM(f_w, t)^2$$

where *ReM* and *iM* denote the magnitude of the real and imaginary components of the convolution result, respectively. To enable comparisons across frequency bands, power was converted to decibel (dB) scale via: $10 \cdot \log_{10}[P(t,f)/P(t_{baseline}, f)]$ where $t_{baseline}$ baseline ranged from -300 to -100 ms. Frequency-specific instantaneous phase (ϕ) is given by the angle of the convolution result (arctangent of *iM* over *ReM*) at time *t*.

To demonstrate entrainment of EEG oscillations to the task rhythm, we computed inter-trial phase coherence (ITPC), a measure of the consistency of phase across trials. This measure ranges between 0 (no consistency) and 1 (perfect consistency), and can be computed using:

$$\text{ITPC}(f_w, t) = \left| n^{-1} \sum_{N=1}^{n} e^{i\phi(f_w, t)} \right|$$

where *n* is the number of trials.

Statistical analyses of time-frequency data. As randomly distributed phases across trials produce a known (Rayleigh) distribution, the significance of observed ITPC can be assessed by comparing it to ITPC under the null hypothesis. Conversely, a critical value (ITPC$_{crit}$) at which ITPC significantly deviates from randomly distributed phases can be calculated, given a *p*-value (0.001 in our case), via:

$$\text{ITPC}_{\text{crit}} = \sqrt{-\log(p) \cdot n^{-1}}$$

where *n* is the number of trials (Zar, 1999; Cohen, 2014b). We thus calculated ITPC$_{crit}$, and for each time-frequency point values higher than ITPC$_{crit}$ were considered significant.

We used non-parametric permutation testing to assess whether power deviated from the -300 to -100 ms pre-stimulus baseline window, for each time-frequency point. Permutation testing is robust against violations of assumptions about data distributions that can occur with conventional parametric statistics (Maris and Oostenveld, 2007). First, within each frequency, the assignment of evoked power to 'baseline' or 'power' distributions was shuffled, and *t*-statistics specific to each time-frequency point were computed iteratively 1000 times. The resulting *t*-value distribution was then Z-scored, and the standardized value of the non-shuffled comparison (actual baseline vs. power) was then computed per time-frequency point. This procedure generated a time-frequency matrix with a Z-value for each time-frequency point that indicated the statistical likelihood of finding the experimentally obtained power values, given that the null hypothesis (no difference in power compared to baseline) is true. To correct for multiple comparisons, the threshold for the time-frequency matrix was set as Z-scores corresponding to *p*-values of 0.001 or less. In a second step, a distribution of maximum cluster sizes (the number of contiguous significant time-frequency points) under the null hypothesis was

computed. The cluster size corresponding to the 95th percentile of this distribution was then taken as the lower bound for cluster correction of the time-frequency matrix, resulting in a cluster-level threshold of $p<0.05$ (c.f. Cohen et al., 2012).

To test whether phase distributions differed from a uniform distribution, we used Rayleigh's test for uniformity (Fisher, 1993).

Fixed-effects analysis. Circular-linear correlations (i.e correlations between single-trial phase and behavioral measures) were performed using a similar non-parametric permutation testing approach as described above. Correlation distributions under the null hypothesis in this case were created by shuffling the assignment of pooled behavioral data (i.e. RT or PES$_{robust}$) to phase information at the single-trial level. The test statistic which was subsequently Z-scored for each time-frequency point was the correlation coefficient resulting from circular to linear correlation as implemented in the circular statistics (CircStat) toolbox (Berens, 2009).

To assess differences in average phase angle between conditions (i.e. correct vs. error; post-correct vs. post-error; and post-correct vs. large and small PES bins), we again used a similar procedure, but instead shuffled the assignment of pooled single-trial phase to trial types to compute a standardized distribution of angle differences under the null hypothesis. We used the Watson-Williams test for angular means to obtain the test-statistic that was informative of the angle differences between conditions. This test statistic was subsequently standardized, similar to the analyses described above. Because there were more correct and post-correct trials than error and post-error trials, respectively, we matched these trial numbers between conditions by selecting a random subset of trials from the larger condition prior to computing differences between them in the average phase angle.

Random-effects analysis. Because the fixed-effects procedure described above does not take into account the between-subjects variance in effect size, we repeated these analyses for a number of select time-frequency points (stimulus onset, and 600 ms post-stimulus /response), but without assuming constant between-subject variance. Specifically, we computed phase-RT and phase-PES correlations for each individual participant to obtain a distribution of correlation coefficients. Because circular-to-linear correlations are bound between zero and one, correlation coefficients under the null hypothesis are unlikely to be exactly zero. We therefore computed the correlations again but with shuffled phase-behavior assignments for 1000 permutations. This resulted in a distribution of correlation coefficients under the null hypothesis. To assess the group-level significance of the observed distribution of correlation coefficients, we used a paired-sample t-test to compare them to the distribution under the null hypothesis averaged across permutations.

To test for differences in phase angle between conditions, we first computed the average phase angle and resultant vector length across trials for each participant and each condition. Then, group-level differences in phase angle between conditions were assessed using the parametric Hotelling paired-sample test for angular means (Zar, 1999). The extension of this test described by Zar (1999) takes into account the

resultant vector lengths of the phase distributions of individual cases (participants, in our case), which carries meaningful information regarding the consistency of the phase angle difference between the average distributions. It is therefore more suited to be used in second-level, random-effects analyses than the Watson-Williams test.

2.3 Results

Behavioral and ERP results. The average RT on correct trials was 608 ms (SD 69 ms); average RT on error trials was 604 ms (SD 87 ms; p = 0.53). The average percentage of correct responses was 86.7% (SD 5.8%). To confirm that participants were sensitive to the average stimulus-presentation rate (one stimulus every 1350 ms), we binned correct trials by preceding SOA, and expected the shortest RTs in the 1350-ms bin. A repeated-measures ANOVA revealed that correct RT had an inverted U-shaped relationship with the preceding SOA (Figure 1B; quadratic trend: F(1,19) = 12.73, p < 0.01), indicating that participants were good at estimating the average SOA of 1350 ms. No such trend was found for accuracy (F(1,19) = 0.8, p = 0.38), ruling out preceding SOA as a confounding factor in our comparisons between post-error and post-correct trials.

As a next step we wanted to confirm that PES occurred in our flanker task with only incongruent stimuli. In line with our expectations, both PES$_{traditional}$ (mean 27 ms, SD 25 ms) and PES$_{robust}$ (mean 44 ms, SD 24 ms) were significantly larger than zero (p < 0.001; Figure 1C). Although PES$_{robust}$ was significantly larger (p < 0.001), the two measures of PES were highly correlated (p < 0.001, Figure 1D), suggesting that PES$_{robust}$, although more precise (Dutilh et al., 2012b), does not provide radically different values for PES than the more traditional measure. Post-error accuracy (mean: 84.1%; SD 12.0%) was somewhat lower than post-correct accuracy (mean: 87.7; SD 4.2%), but this difference was not significant, p = 0.24.

Figure 1E shows that the error-related negativity and error positivity, two well-known error-related ERP components, were present in our EEG data.

Functional entrainment of oscillations. To replicate previous results indicating that low-frequency oscillations entrain to the stimulus-presentation rhythm (Lakatos et al., 2008; Schroeder and Lakatos, 2009; Stefanics et al., 2010a; Henry and Obleser, 2012; Ng et al., 2012), we computed ITPC, a measure of the consistency of oscillatory phase across trials. We calculated ITPC averaged across all trial types and electrodes to assess which frequencies showed the highest degree of consistency, and expected greatest consistency in low (< 2Hz) frequencies, overlapping with our task rhythm. In line with this expectation, Figure 2A shows that ITPC was significantly higher than expected by chance in the lowest frequency range. Additionally, the cluster of time-frequency points that showed significant ITPC extended into the higher frequencies.
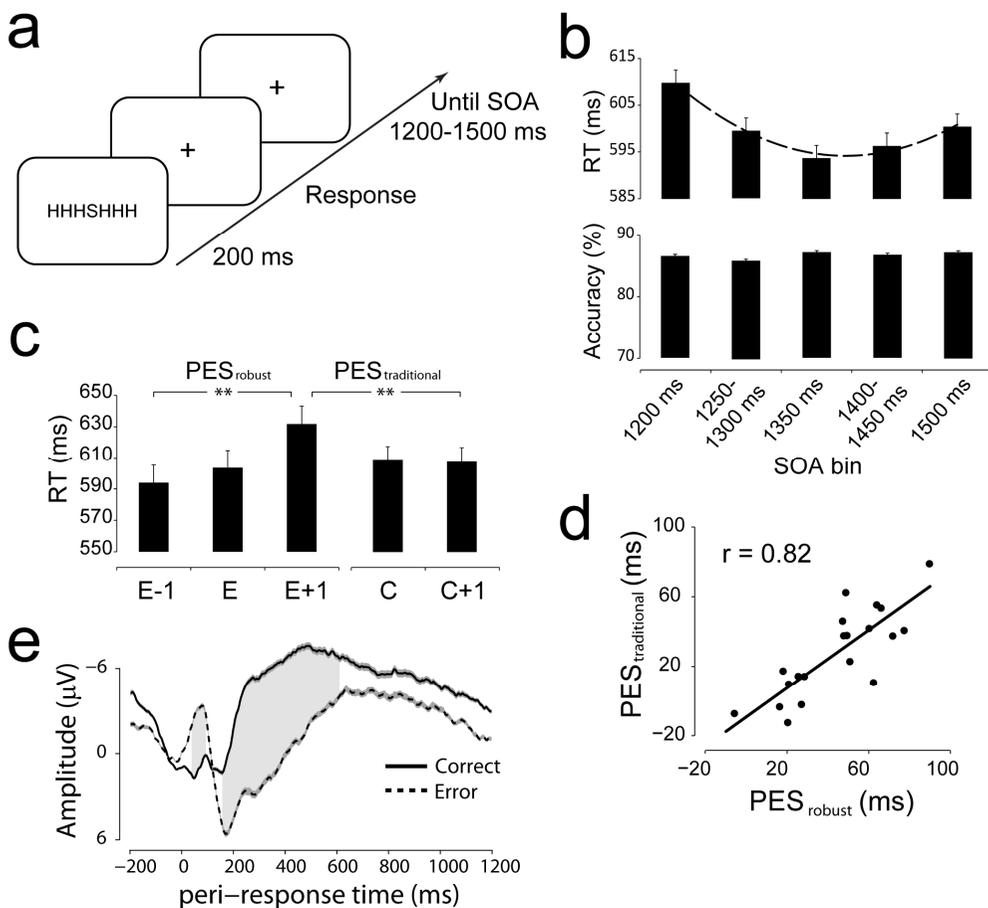
Figure 1. Task, behavioral and ERP results. **a)** Timing of trial events. SOA: stimulus onset asynchrony. **b)** Response time (RT) and accuracy as a function of (binned) preceding SOA. Intermediate bins are averaged for display purposes. The dashed line in the upper panel is a fitted quadratic curve. **c)** Behavioral results showing post-error slowing, as measured by both the traditional and robust methods of quantifying PES. Asterisks indicate significant ($p < 0.001$) differences. **d)** PESrobust and PEStraditional correlated significantly across subjects ($p < 0.01$). **e)** Response-locked event-related potentials for correct and error trials, averaged across electrodes Fz and Cz. Shaded area indicates a significant ($p < 0.001$) difference between conditions. All error bars show the standard error of the mean.

To confirm the dependence of ITPC on the task rhythm, we ran a control experiment (N = 4) with the same task, except that the stimulus-presentation rhythm was manipulated in a block-wise fashion (0.5 and 0.85 Hz). All other task parameters
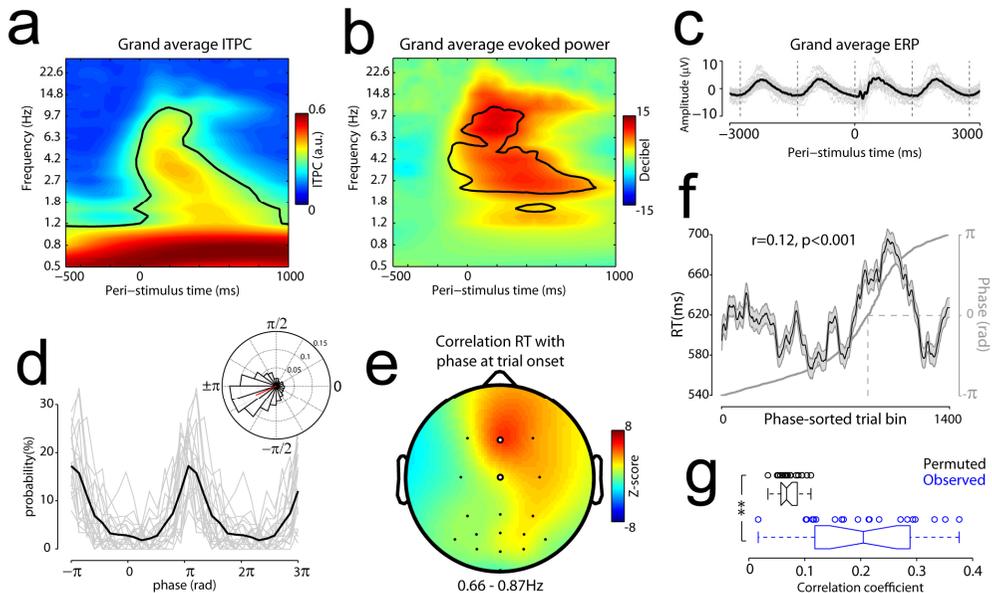
Figure 2. Functional entrainment of oscillations to the task rhythm. **a)** Low-frequency oscillations entrain to the task rhythm, as indicated by inter-trial phase coherence (ITPC) averaged across all trials and all channels. **b**) Evoked (ERP) power averaged across all conditions and all channels. Only frequencies above the task rhythm show significant increases in evoked power, suggesting that elevated low-frequency ITPC reflects entrainment of endogenous oscillations. Significant (p<0.001) regions are outlined in black using MATLAB's contourf function. **c)** Channel- and condition-averaged ERP showing oscillations entrained to the stimulus-presentation rhythm. The solid gray lines show the ERPs of individual participants. The solid black line shows the average. The vertical dashed gray lines show the time of average stimulus onsets. **d)** Low-frequency (0.76Hz) phase distribution at correct trial onset for the average of channels Fz and Cz. The solid black line shows the average. Gray lines show individual participants. The inset shows a rose histogram, where the radial extent of the bars indicates the probability of a given phase occurring on a single trial. The red line is the average vector of the histogram. **e)** Topographical distribution of Z-scored correlation of correct RTs with low-frequency phase at trial onset. The two highlighted channels are Fz and Cz, the average of which was used in all further reported analyses. **f)** Correct RTs sorted by low-frequency (0.76Hz) phase at trial onset, smoothed with a moving average of 100 trials for display purposes. The non-linearity in the curve arises due to a correlation of phase with RT. The error bars represent the standard error of the mean and are the result of smoothing. The light gray line and corresponding y-axis on the right-hand side represent phase for each of the sorted trials. **g)** Box plots of the correlation coefficients for permuted (black) and observed (blue) correlations between correct RT and 0.76Hz phase at trial onset for the average of channel Fz and Cz. A 95% confidence interval around the median of each distribution is indicated by the notches in the boxes. The whiskers extend to the most extreme values of each distribution. The circles show the correlation coefficients of individual participants. ** p < 0.001.

were as in the main experiment. For each of the four participants, we found clearly dissociable peaks in ITPC (at 0.5 and 0.85 Hz) that differed between blocks, indicating that oscillatory phase locking was highly dependent on the stimulus-presentation rhythm.

Elevated ITPC can come about in two ways (Tallon-Baudry and Bertrand, 1999; Donner and Siegel, 2011). The first possibility is that ITPC arises due to cross-trial phase alignment of true, endogenously driven oscillatory activity. The second possibility is that ITPC is a result of stimulus-evoked activity, meaning simple stimulus-evoked EEG components cause consistency in phase across trials. To exclude the possibility that stimulus-evoked activity was the driving force behind consistency in phase across trials, we computed oscillatory power of stimulus-locked trial-averaged (ERP) data. As Figure 2B shows, no significant evoked power was found below ~2Hz, indicating that ITPC in these low frequencies was likely driven by endogenous oscillatory activity. Above ~2Hz, however, we found significant evoked power, indicating that elevated ITPC in those frequencies was most likely stimulus-evoked. In sum, lower frequencies showed significant ITPC, which likely reflected entrainment of endogenous neuronal oscillations to the stimulus-presentation rhythm, in line with our prediction. These entrained oscillations are readily visible in the stimulus-locked grand-average ERP (Figure 2C). Moreover, the phase angle of our frequency of interest (0.76Hz) at stimulus onset was highly consistent across participants (Figure 2D; deviation from uniform distribution: Rayleigh's Z = 11.3, $p <$ 0.0001).

Next, to show that this entrainment is directly linked to task performance, we correlated correct RTs with single-trial instantaneous phase at trial onset, pooling the trials of all participants, and focusing on the frequency corresponding with the average stimulus-presentation rate (0.76 Hz). The relationship between phase and RT was strongest over fronto-central scalp regions (Figure 2E), in accord with previous findings (Stefanics et al., 2010a; Henry and Obleser, 2012; Ng et al., 2012). Accordingly, all subsequent analyses were performed on the average of the signals recorded at Fz and Cz. The significant phase-RT relationship (r = 0.12, $p <$ 0.001) with trials pooled across participants is shown in more detail in Figure 2F. This correlation was also significant ($p <$ 0.001) at the group level with a random-effects analysis approach (Figure 2G). Thus, the entrainment of oscillations is functional, in that it predicts the latency of responding, and this effect is consistent across participants.

Error-induced disturbance in phase entrainment. To test our hypothesis that errors result in a disturbance of phase entrainment, we calculated the difference in phase angle between correct and error trials (response-locked, i.e., aligned with the commission of errors / correct responses), and post-correct and post-error trials (stimulus-locked, i.e. aligned with the onset of stimuli following errors / correct responses). If the average phase angle on correct responses facilitates response speed, then PES may reflect a temporary deviation from this phase angle. Note that
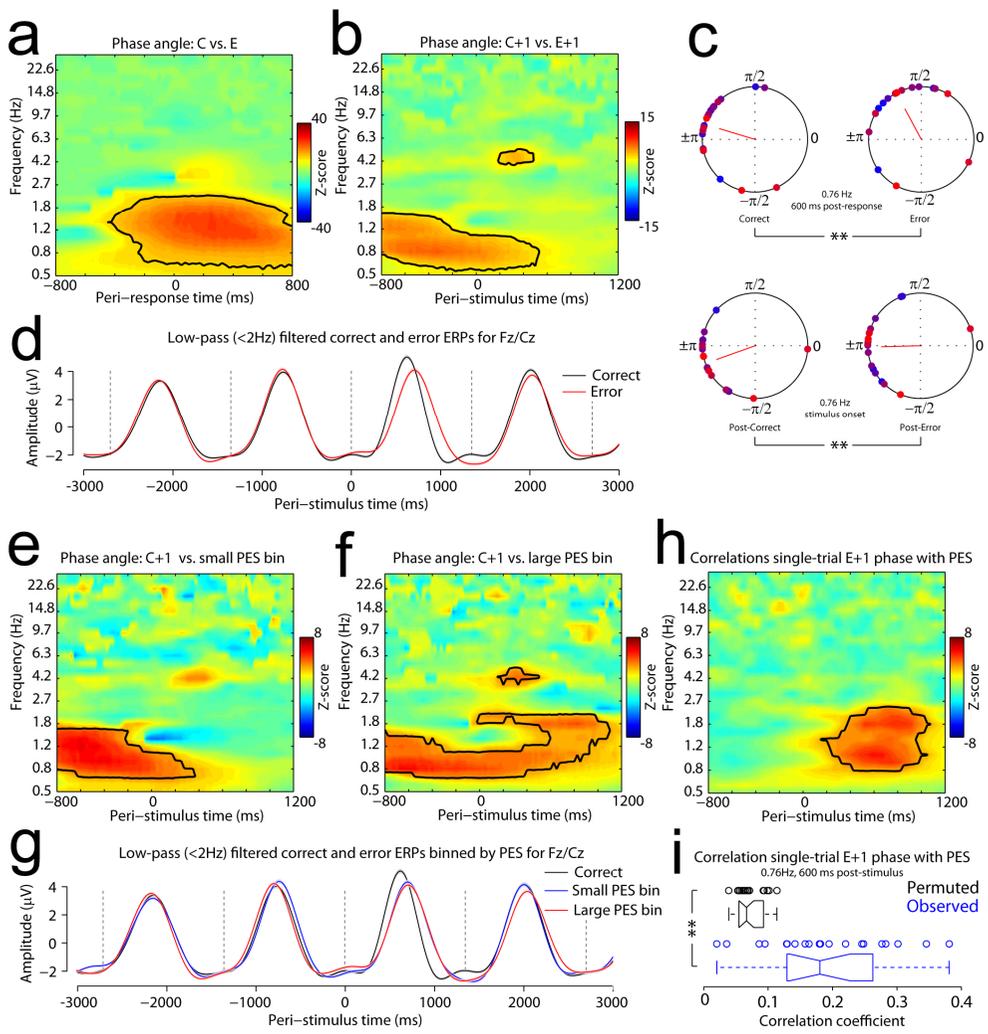
Figure 3. Relation between phase angle and post-error slowing. **a)** Z-scored differences in absolute phase angle between correct (C) and error (E) trials, pooled across participants. Red colors indicate phase on error trials significantly lagging behind correct trials. Time zero indicates the time of a response. **b)** Z-scored differences in absolute phase angle between post-correct (C+1) and post-error (E+1) trials, pooled across participants. Red colors indicate phase on post-error trials significantly lagging behind post-correct trials. Time zero indicates the onset of the stimulus. **c)** Low-frequency (0.76 Hz) phase distributions for post-correct and post-error trials for individual subjects. Each participant is denoted by a uniquely colored dot on the circle. The average vector of the distribution is shown in red. ** p < 0.001. **d)** Low-pass (<2Hz) filtered ERPs for the average of channel Fz and Cz. Error bars denote the standard error of the mean. Dashed vertical lines show the average stimulus onset times. **e)** Z-scored phase angle differences between trials that show a low degree of PES (small PES bin) and post-correct trials. **f)** Z-scored phase angle differences between trials that show a high degree of PES (large PES bin) and post-correct trials. **g)** Same as d), but error trials are binned according to PES$_{robust}$. **h)** Z-scored correlations of [caption continues on next page]

[continued caption Figure 3] single-trial phase on post-error trials with single-trial PES$_{robust}$. In all time-frequency plots, significant ($p < 0.001$) regions are outlined in black using MATLAB's contourf function. **i)** Box plots of the correlation coefficients for permuted (black) and observed (blue) correlations between single-trial 0.76Hz phase at 600 ms post-stimulus on post-error trials and single-trial PES$_{robust}$. A 95% confidence interval around the median of each distribution is indicated by the notches in the boxes. The whiskers extend to the most extreme values of each distribution. The circles show the correlations coefficients of individual participants. ** $p < 0.001$.

if errors result in a *systematic* disturbance of phase entrainment, this will become apparent as a difference in phase angle following correct and erroneous responses rather than as a difference in phase-locking strength.

As Figure 3A shows, following the response, low-frequency phase angle on error trials significantly differs from that of correct trials. Although the difference in phase angle between correct and error trials is already evident before the response, this pre-response difference is likely the result of temporal smearing of the effect that is inherent to wavelet decomposition. To confirm this, we re-ran our wavelet decomposition with a lower number of wavelet cycles, thereby increasing the temporal resolution at the expense of frequency resolution. We found that this eliminated all pre-response differences in phase angle between errors and correct trials, while post-response differences were still present (results not shown). This indicates that the disturbance in phase entrainment following the response is likely to be error-induced. The difference in phase angle following correct and error responses was highly consistent across participants (Figure 3C, upper panel; F(2,18) = 19.3, *p* < 0.001).

Figure 3B shows that this error-induced disturbance in phase entrainment extends into the post-error trial, where it could potentially influence RT, thereby causing slowing on post-error trials. The difference in phase angle at trial onset for post-correct and post-error trials was highly consistent across participants (Figure 3C, lower panel; F(2,18) = 6.04, *p* < 0.001). The error-induced disturbance in entrainment is also visible in the low-pass filtered ERPs. Figure 3D shows that around the average response time, the phase of the error-trial ERP starts to lag behind that of the correct-trial ERP. This phase difference extends into the subsequent trial and then dissolves again.

If the error-related disturbance in phase entrainment relates to how much PES occurs, phase angle should differ between trials that show a large degree of PES and trials with a small degree of PES. To test this prediction we divided up the post-error trials into three equally sized bins according to the amount of PES$_{robust}$, and compared, for each time and frequency point, the phase angle of these bins to the phase angle on post-correct trials. We expected to see a more pronounced disturbance in phase angle on trials with a large degree of PES. Consistent with our prediction, Figures 3E and 3F show that the disturbance in phase entrainment

27
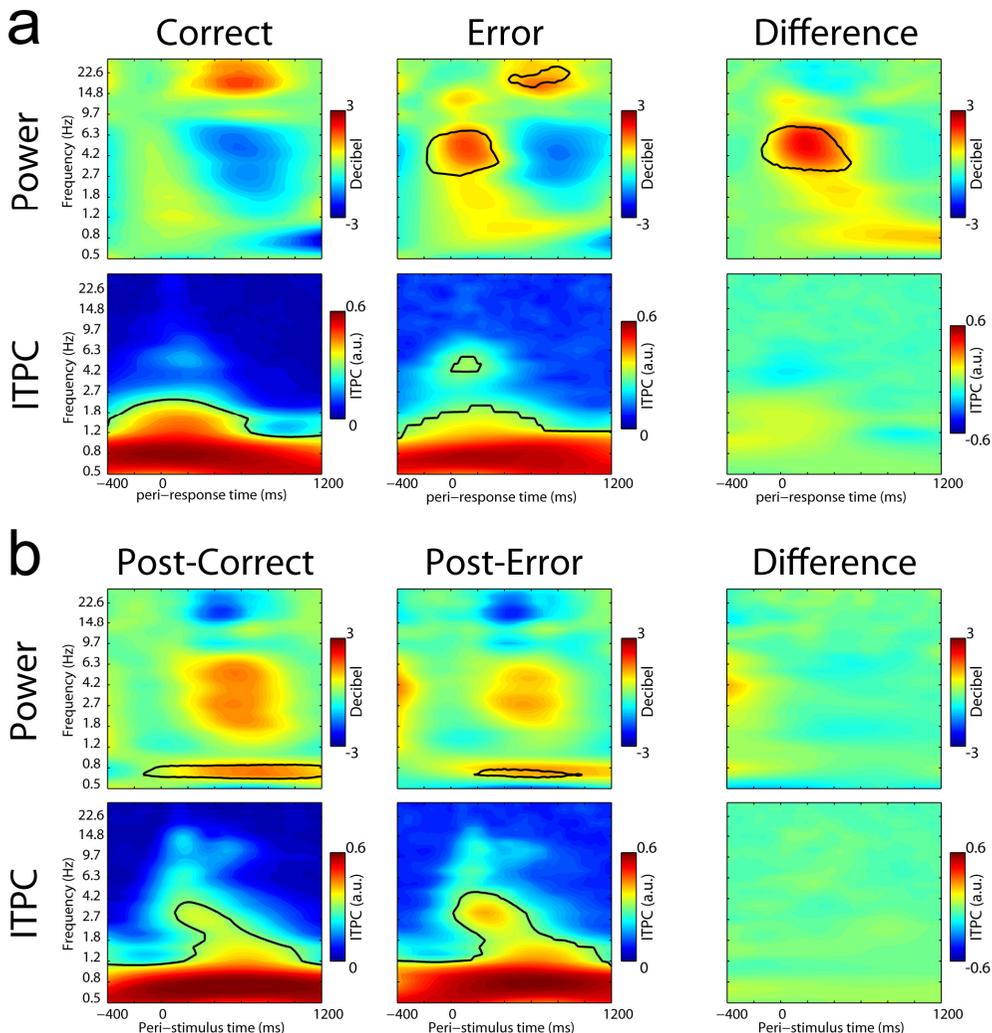
Figure 4. Power and inter-trial phase coherence (ITPC) comparisons for the average of channel Fz and Cz, between **a)** correct and error trials, and **b)** post-correct and post-error trials. Significant (*p* < 0.001) regions are outlined in black using MATLAB's contourf function.

extended for a longer period of time on trials with a large degree of PES. This indicates that the longer the disturbance in phase entrainment lasts following an erroneous response, the slower the participant is in responding on the subsequent trial. This strongly suggests a link between PES and disturbed phase entrainment.

The prolonged disturbance during trials that show a large degree of PES is also visible in the ERPs (Figure 3G). Whereas the ERP of the small PES bin rejoins the ERP of correct trials before the trial N+1 response period, the ERP of the large PES bin is still out of phase with respect to correct trials during that period. The difference

in 0.76Hz phase angle between post-error trials with large PES and post-correct trials was highly consistent across participants (at stimulus onset: $F(2,18) = 7.32$, $p = 0.0047$; at 600 ms post-stimulus: $F(2,18) = 3.74$, $p = 0.0053$). The difference in phase angle between post-error trials with small PES and post-correct trials was not significant ($F(2,18) = 2.40$, $p = 0.12$), conceivably due to the fact that binning resulted in too few trials for an accurate estimate of average phase angle for each participant separately.

Next, to test directly whether PES can be predicted by the level of phase disturbance, we correlated phase on post-error trials with PES_{robust} at the single-trial level. Figure 3H shows that on post-error trials, low-frequency phase predicted PES, demonstrating a trial-by-trial relationship between disturbed phase entrainment and PES. Interestingly, the time-window in which phase predicted PES overlapped with the period during which trials with large PES (Figure 3F) but not trials with small PES (Figure 3E) differed in phase angle from post-correct trials. This suggests that post-error trials were generally characterized by some degree of phase disturbance, and that the amount of PES on a given trial was mainly dependent on the duration of the error-induced phase disturbance. The correlation between post-error phase around the time of the response (~600 ms) and single-trial PES was highly consistent across participants (Figure 3I).

Finally, in order to confirm that the disturbance is specific to the entrained low-frequency oscillations rather than reflecting a broader spectral perturbation, we computed power and ITPC for correct- and error trials, and post-correct and post-error trials, as well as respective differences between them. As can be seen in Figure 4, there are no significant differences in power or ITPC between conditions, except for the well-documented error-related increase in theta-band power (Narayanan et al., 2013). This indicates that the phase angle differences between correct and error trials, and between post-correct and post-error trials are unlikely to result from a broad-band perturbation. Instead, the results are in line with our hypothesis of a specific phase disturbance in the task-entrained frequency.

2.4 Discussion

In the present study, we tested the hypothesis that PES results from a disturbance of internally generated brain rhythmicity. Taken together, our results show that: 1) endogenous low-frequency neuronal oscillations entrain to the stimulus-presentation rhythm; 2) the entrainment facilitates speeded responding; 3) entrainment is disturbed following the commission of an error; and importantly, 4) the level of disturbance of entrainment predicts how much slowing occurs on the following trial. These results support our novel hypothesis about the neural origin of PES.

What might be the mechanism underlying this error-related disturbance of phase entrainment? One possibility is that errors evoke a transient process that temporarily

distracts from the current task and therefore results in slower responding on the subsequent trial (Jentzsch and Dudschig, 2009). This possibility is suggested by the orienting account of PES (Notebaert et al., 2009), which posits that errors, due to their infrequent occurrence, automatically draw attention away from the ongoing task, much like other surprising events tend to do. The orienting account is supported by the finding that slowing occurs following correct trials instead of errors when correct responses are infrequent (Notebaert et al., 2009; Nunez Castellar et al., 2010); and by the observation that participants who make fewer errors (i.e. for whom errors are more unexpected) show larger PES (Steinborn et al., 2012). Our findings suggest that this hypothesized orienting response may cause the phase angle of entrained oscillations to deviate from the more advantageous phase angle observed following correct trials, thus leading to slower responding. This account is consistent with recent studies that have linked the phase of ongoing delta oscillations to the rate of perceptual evidence accumulation (Wyart et al., 2012; Cravo et al., 2013).

Other evidence in support of a limited-duration orienting response during which attention is distracted from the task is provided by studies that examined the relationship between post-error performance and the interval between the error and the subsequent trial (Dudschig and Jentzsch, 2009; Jentzsch and Dudschig, 2009; Danielmeier and Ullsperger, 2011). By varying the response-stimulus interval (RSI), these studies found that short RSIs (<500 ms) are typically associated with large PES and a post-error decrease in accuracy. As the RSI increases up to 1000 ms and larger, PES strongly diminishes and the post-error decrease becomes a post-error *increase* in accuracy. We found an intermediate pattern of results (modest PES, slightly decreased accuracy after errors) that seems consistent with the intermediate length of our average RSI (~750 ms). The orienting account can explain these findings by claiming that at short RSIs, attentional reorientation is the dominant cause for the impaired performance on the subsequent trial. As the RSI increases, it becomes increasingly likely that the system has recovered from the orienting response by the time the next stimulus is presented. The timing of the error-induced disturbance in phase entrainment observed here, which partially overlapped with the subsequent trial (and more so on trials characterized by large PES), is consistent with the notion that at intermediate RSIs there is a substantial probability of overlap between the orienting response and the subsequent trial. Of course, due to the correlative nature of our methods, we cannot unequivocally conclude that disturbed entrainment is the *causative* mechanism underlying PES. This issue can be addressed by manipulating entrainment directly, for example using transcranial magnetic stimulation.

A disturbance in phase entrainment is unlikely to be the only cause of PES. Dutilh et al. (2012a), using drift-diffusion modelling of performance on a lexical decision task, found that when the RSI is relatively long (1000 ms), PES can be attributed almost exclusively to increased response caution. This suggests that a strategic change in response threshold (Botvinick et al., 2001) can also contribute to PES. Another study

using drift-diffusion model analysis found that PES could be explained by a change in several model parameters, including an increased response threshold and a decreased rate of evidence accumulation, consistent with disturbed phase entrainment (White et al., 2010). However, in this study participants received trial-to-trial feedback, so post-error effects may have been contaminated by the possibly distracting presence of error feedback. Additionally, several studies have found PES with inter-trial intervals of several seconds (Hajcak et al., 2003; Marco-Pallarés et al., 2008; King et al., 2010; Danielmeier et al., 2011), precluding a contribution of a transient disturbance of phase entrainment. However, these studies either measured PES$_{traditional}$, which is confounded by global fluctuations in motivation and task performance (Dutilh et al., 2012b), which are often present in tasks with long RSIs; or they used PES$_{robust}$ but did not control for differences in trial type (e.g., congruent vs incongruent) between pre- and post-error trials—a plausible source of confound (Steinhauser and Yeung, 2012). Therefore, it is uncertain if PES truly occurred in these studies. Nonetheless, taken together, the literature suggests that in contrast to the intermediate RSI effects discussed in the previous paragraph, at longer RSIs PES is mainly determined by a time-consuming strategic change in speed-accuracy trade-off. This notion is corroborated by the finding that post-error accuracy is increased following errors only at longer RSIs (Jentzsch and Dudschig, 2009).

An important question is what neural mechanism orchestrates the alignment of neuronal oscillations to environmental rhythmicity. The widespread cortical topography and slow temporal dynamics of the entrained oscillations observed here suggest a possible neuromodulatory involvement. Accordingly, the norepinephrine-producing neurons of the locus coeruleus (LC) are phase-locked in firing to ongoing cortical slow-wave oscillations during sleep (Eschenko et al., 2012; Sara and Bouret, 2012). Specifically, neurons of the prefrontal cortex and the LC fire in phasic opposition, suggesting a mutual excitatory drive. Additionally, one concurrent EEG-fMRI study has provided tentative evidence of temporal alignment of human LC activity with slow-wave oscillations (Dang-Vu et al., 2008). These findings have led to the suggestion that the LC facilitates transitions from the down- to the up-state of slow cortical oscillations (Eschenko et al., 2012; Sara and Bouret, 2012), which would provide a plausible mechanism for temporal alignment of slow oscillations with environmental rhythmicity. In line with the concept of an orienting response, phasic norepinephrine release has also been proposed to act as a neural interrupt signal, whereby unexpected events (e.g., errors) lead to a reset and reorganization in target neuronal networks, and subsequent behavioral adaptation (Bouret and Sara, 2005; Dayan and Yu, 2006). Interestingly, several proposals have linked PES to noradrenergic activity (Cohen et al., 2000; Nunez Castellar et al., 2010), and preliminary evidence suggests that PES is partly determined by a genetic marker of norepinephrine synthesis (Colzato et al., 2013). Thus, it is possible that the noradrenergic system is involved in the entrainment of cortical oscillations, and that this entrainment and consequent mode of behavioral responding are disrupted by an

error-evoked orienting response (or interrupt signal) in the noradrenergic system (Nieuwenhuis et al., 2010; Nunez Castellar et al., 2010; Ullsperger et al., 2010).

A remaining question is whether our findings will generalize to other task designs. Here, phase entrainment followed the low-frequency stimulus-presentation rhythm. However, phase entrainment does not always occur in the lower frequency range (e.g. Maltseva et al., 2000), raising the question whether an error-induced disturbance might also occur in higher frequencies. Indeed, because we used only one stimulus-presentation frequency, it is conceivable that the error-induced disturbance in phase angle occurred independent of endogenous entrainment. This possibility seems unlikely given that the choice of stimulus-presentation rhythm was motivated by earlier work on entrainment (Stefanics et al., 2010a; Lakatos et al., 2013) rather than a specific hypothesis about the involvement of 0.76-Hz oscillations in error processing. Furthermore, others have shown attention-related phase shifts of entrained oscillations in substantially higher frequencies (Lakatos et al., 2008; Besle et al., 2011b). Nevertheless, the specificity of the error-related disturbance in phase angle to the task frequency warrants further investigation using multiple stimulus-presentation rhythms.

Furthermore, in our experiment the task rhythm was determined mainly by the highly predictable SOA, whereas in many other studies of PES the RSI instead of SOA is the predictable time interval. So an interesting question is whether in these studies PES is also related to a disturbance of phase entrainment. Previous studies have shown that participants can use various temporal cues to facilitate entrainment (Stefanics et al., 2010a). However, it remains to be investigated whether response time can function as a temporal cue to entrain oscillations to subsequent stimulus onset. Alternatively, to the extent that variability in RTs is modest and the RSI and other intervals are fixed, the SOA will be relatively constant across trials. The entrainment in the current study was robust to a certain amount of temporal variability (SOA 1200-1500 ms), suggesting that even in tasks in which RSI is the predictable factor, neuronal oscillations may entrain to the SOA.

In conclusion, our brain exploits the innate periodicity in the environment by means of oscillatory entrainment. We have shown that when an error is made, a temporary perturbation occurs such that the entrained oscillations are 'out of sync' with the current task, the degree of which predicts the magnitude of slowing of the subsequent behavioral response.

3. Catecholaminergic Neuromodulation Shapes Intrinsic MRI Functional Connectivity in the Human Brain

Abstract

The brain commonly exhibits spontaneous (i.e., in the absence of a task) fluctuations in neural activity that are correlated across brain regions. It has been established that the spatial structure, or topography, of these intrinsic correlations is in part determined by the fixed anatomical connectivity between regions. However, it remains unclear which factors dynamically sculpt this topography as a function of brain state. Potential candidate factors are subcortical catecholaminergic neuromodulatory systems, such as the locus coeruleus-norepinephrine (LC-NE) system, which send diffuse projections to most parts of the forebrain. Here, we systematically characterized the effects of endogenous central neuromodulation on correlated fluctuations during rest in the human brain. Using a double-blind placebo-controlled cross-over design, we pharmacologically increased synaptic catecholamine levels by administering atomoxetine, an NE transporter blocker, and examined the effects on the strength and spatial structure of resting-state MRI functional connectivity. First, atomoxetine reduced the strength of inter-regional correlations across three levels of spatial organization, indicating that catecholamines reduce the strength of functional interactions during rest. Second, this modulatory effect on intrinsic correlations exhibited a substantial degree of spatial specificity: the decrease in functional connectivity showed an anterior-posterior gradient in the cortex, depended on the strength of baseline functional connectivity, and was strongest for connections between regions belonging to distinct resting-state networks. Thus, catecholamines reduce intrinsic correlations in a spatially heterogeneous fashion. We conclude that neuromodulation is an important factor shaping the topography of intrinsic functional connectivity.

3.1 Introduction

The resting-state – here defined as periods during which a participant is not engaged in a complex explicit task – is characterized by fluctuations in neural activity that are correlated across brain regions (Biswal et al., 1995; Leopold et al., 2003; Fox and Raichle, 2007; Hiltunen et al., 2014). Such spontaneous, correlated fluctuations exhibit a rich spatial (Yeo et al., 2011) and temporal (Allen et al., 2014; Zalesky et al., 2014) structure that is reflective of the brain's functional organization (Tavor et al., 2016). The strength and spatial distribution of these correlated fluctuations are predictive of behavior and pathological conditions (Greicius et al., 2004; De Luca et al., 2005). Moreover, the global structure, or topography, of correlated activity changes dynamically with alterations in conscious state (Barttfeld et al., 2015) and task conditions (Nir et al., 2006; Sepulcre et al., 2010). While the existence and overall spatio-temporal structure of the spontaneous inter-regional correlations are well-established (Fox and Raichle, 2007), uncertainty remains regarding the underlying physiological mechanisms. It has been proposed that correlations across distant brain regions could be induced by brainstem neuromodulatory systems – and in particular the locus coeruleus-norepinephrine (LC-NE) system, which sends diffuse, ascending projections to the forebrain (Leopold et al., 2003; Drew et al., 2008; Schölvinck et al., 2010), where noradrenergic terminals co-release dopamine (DA; Devoto and Flore, 2006). Here, we examined if and how the catecholaminergic neuromodulators NE and DA shape correlated fluctuations during rest in the human brain.

A number of observations suggest that catecholamines should generally increase the strength of functional connectivity. Both iontophoretic NE application and DA agonism enhance neuronal responses to excitatory synaptic input (Rogawksi and Aghajanian, 1980; Seamans et al., 2001b; Wang and O'Donnell, 2001). Furthermore, NE and DA can amplify synaptic GABAergic inhibition (Moises et al., 1979; Seamans et al., 2001a). These and other findings have led to the view that catecholamines boost the efficacy of synaptic interactions between neurons (Berridge and Waterhouse, 2003; Winterer and Weinberger, 2004), resulting in an increased difference in firing rates between strongly and weakly active neurons. Such signal amplification yields a system-wide facilitation of signal transmission (Waterhouse et al., 1998). Recent computational work suggests that this effect of catecholamines should boost both positive and negative temporal correlations between the activities of local groups of neurons, resulting in stronger and increasingly clustered network connectivity (Donner and Nieuwenhuis, 2013; Eldar et al., 2013). Putative behavioral and pupillary indices of heightened NE activity have accordingly been shown to co-occur with stronger functional coupling throughout the brain (Eldar et al., 2013). A first consideration of the anatomy of the LC-NE system suggests that these changes in functional connectivity might show little spatial specificity. LC neurons exhibit tightly synchronous firing and collateralize broadly, resulting in largely homogeneous catecholaminergic

innervation throughout the brain (Swanson and Hartman, 1975; Aston-Jones et al., 1984; Ishimatsu and Williams, 1996; Berridge and Waterhouse, 2003).

In the present study we systematically characterized catecholamine effects on the strength and spatial structure of resting-state inter-regional correlations, measured with functional magnetic resonance imaging (fMRI). Using a double-blind placebo-controlled cross-over design, we manipulated catecholamine activity by administering a single dose of atomoxetine, a selective NE transporter (NET) blocker. Within the cortex NET is also responsible for DA reuptake, due to the cortical paucity of DA transporters (Devoto and Flore, 2006). Thus, NET blockers increase both central NE and cortical DA availability (Bymaster et al., 2002; Devoto et al., 2004; Swanson et al., 2006; Koda et al., 2010). We systematically quantified catecholamine effects on functional connectivity—globally, between brain networks, and at the level of individual connections between brain regions. In contrast to the notion of a catecholamine-induced homogeneous increase in functional connectivity, we found that atomoxetine *reduced* correlations across most pairs of brain regions. Most remarkably, atomoxetine altered the strength of inter-regional correlations in a highly spatially specific manner. These results have important ramifications for our understanding of resting-state activity and central catecholaminergic function.

3.2 Materials and Methods

Participants. Neurologically healthy right-handed individuals (N=24, age 19-26, 5 male) were recruited and medically screened by a physician for physical health and drug contraindications. Exclusion criteria included: standard contraindications for MRI; current use of psychoactive or cardiovascular medication; a history of psychiatric illness or head trauma; cardiovascular disease; renal failure; hepatic insufficiency; glaucoma; hypertension; drug or alcohol abuse; learning disabilities; poor eyesight (myopia ≤ -6 diopters); smoking more than 5 cigarettes a day; and pregnancy. All participants gave written informed consent prior to the experiment and screening, and were compensated with €135 or course credit.

Design and functional MRI data. We used a double-blind placebo-controlled cross-over design. In each of two sessions, scheduled one week apart at the same time of day, participants received either a single oral dose of atomoxetine (40 mg) or placebo (125 mg of lactose monohydrate with 1% magnesium stearate, visually identical to the drug). Elsewhere we report data showing that the atomoxetine treatment significantly increased salivary levels of cortisol and alpha amylase, reliable markers of sympathetic nervous system and hypothalamus-pituitary-adrenal axis activation, respectively (Warren et al., in preparation), thus confirming drug uptake. In both sessions participants were scanned once before pill ingestion (t = -20 min) and once at t = 90 min, when approximate peak-plasma levels are reached. The interaction contrast (post atomoxetine – pre atomoxetine) minus (post placebo – pre placebo)

allowed us to examine the effects of atomoxetine while controlling for other session-related differences. Each scan comprised 8 minutes of eyes-open resting-state fMRI. During scanning the room was dark, and participants fixated on a black fixation cross presented on a gray background.

MRI data collection and preprocessing. All MRI data were collected with a Philips 3T MRI scanner. In each of the scanning sessions we collected a T2*-weighted echo planar imaging (EPI) resting-state image (echo time 30 ms, repetition time 2.2 s, flip angle 80º, field of view 80 x 80 x 38 voxels of size 2.75 mm isotropic, and 216 volumes). To allow magnetic equilibrium to be reached, the first 5 volumes were automatically discarded.

In addition, each time the participant entered the scanner we collected a B_0 field inhomogeneity scan (echo time 3.2 ms, repetition time 200 ms, flip angle 30º, and field of view 256 x 256 x 80 voxels with a reconstructed size of 0.86 x 0.86 mm with 3 mm thick slices). Finally, at the start of the first session we collected a high-resolution anatomical T1 image (echo time 4.6 ms, repetition time 9.77 ms, flip angle 8º, and field of view 256 x 256 x 140 voxels with size 0.88 x 0.88 mm with 1.2 mm thick slices).

We used tools from the FMRIB Software Library (FSL) for preprocessing of the MRI data (Smith et al., 2004; Jenkinson et al., 2012). EPI scans were first realigned using MCFLIRT motion correction and skull-stripped using BET brain extraction. We used B_0 unwarping to control for potential differences in head position each time the participant entered the scanner and resulting differences in geometric distortions in the magnetic field. The B_0 scans were first reconstructed into an unwrapped phase angle and magnitude image. The phase image was then converted to units radians per second and median-filtered, and the magnitude image was skull-stripped. We then used FEAT to unwarp the EPI images in the y-direction with a 10% signal loss threshold and an effective echo spacing of 0.332656505.

The unwarped EPI images were then high-pass filtered at 100 s, prewhitened, smoothed at 5 mm FWHM, and co-registered with the anatomical T1 to 2 mm isotropic MNI space (degrees of freedom: EPI to T1, 3; T1/EPI to MNI, 12). Any remaining artifacts (e.g., motion residual, susceptibility-motion interaction, cardiac and sinus artifacts) were removed using FMRIB's ICA-based X-noiseifier (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014) with pre-trained weights (Standard.RData). Noise classification performance was checked afterwards, by manually classifying components as either 'signal', 'noise', or 'unknown'. Then, the accuracy of the automated artifact detection algorithm was quantified as the percentage of components that had the label 'noise' in both classifications. The accuracy was found to be 96.4% correct. All subsequent analyses were conducted in MATLAB 2012a.

Physiological recordings and correction. We recorded heart rate using a pulse oximeter and breath rate using a pneumatic belt at 500 Hz during acquisition of each EPI scan. We used these time series for retrospective image correction [RETROICOR; (Glover et al., 2000)]. This method assigns cardiac and respiratory phases to each volume in each individual EPI time series which can then be removed

from the data. The physiological time series were first down-sampled to 100 Hz. Next, the pulse oximetry data were band-pass filtered between 0.6 and 2 Hz, and the respiration data were low-pass filtered at 1 Hz, using a two-way FIR filter. We then extracted peaks in each time series corresponding to maximum blood oxygenation and maximum diaphragm expansion. The inter-peak intervals were then converted to phase-time by linearly interpolating across the intervals to between 0 and 2π. Next, we used these phase-time series to extract the sine- and co-sine components of the dominant and first harmonic Fourier series of each signal. After down-sampling to the EPI sample rate this yielded 8 regressors (4 cardiac and 4 respiratory) that could then be used to remove cardiac and respiratory effects from the blood-oxygen-dependent (BOLD) time series using multiple linear regression. The findings reported here were based on non-corrected data, but we replicated all of our results using the RETROICOR-corrected data (see Results).

Pupillometry. Pupil size was measured from the right eye at 500 Hz with an MRI-compatible Eyelink 1000 eye tracker. Blinks and other artifacts were interpolated offline using shape-preserving piecewise cubic interpolation. Pupil data were low-pass filtered at 5 Hz to remove high-frequency noise and Z-scored across conditions. Five participants were excluded from pupil-related analyses due to poor signal quality (>50% of continuous time series interpolated) or missing data. Of the remaining participants, on average 20% (SD 9%) of the data were interpolated.

Brain parcellation and connectivity. Time series of brain regions were extracted for the 90 regions of the Automated Anatomical Labeling [AAL, (Tzourio-Mazoyer et al., 2002)] atlas (Fig. 1a). We did not include the cerebellum because it was not fully inside the field of view for all participants. Following averaging across voxels within each brain region, time series (M) for each run i were Z-scored and correlation matrices (R) were computed between them via:

$$R_i = \frac{M_i{}' \cdot M_i}{nTR - 1}$$

where $'$ denotes transposition and *nTR* is the number of volumes (211). Because positive and negative correlations jointly determine a network's functional organization (Fox et al., 2005), many prior studies have used the absolute value of the correlation coefficient to describe functional interactions (Achard and Bullmore, 2007; Eldar et al., 2013; Li et al., 2013). Moreover, computational work suggests that catecholamines should boost temporal correlations regardless of their sign (Donner and Nieuwenhuis, 2013; Eldar et al., 2013). We therefore used the absolute correlation coefficient as our measure of connectivity strength. The signed and absolute matrices were very similar, because anti-correlations were rare (mean 3.4% of all connections, SD 3.5%), as is common when no global signal regression has been performed. In the group- and condition-averaged correlation matrix, 0.28% were anti-correlations (11 out of 4005 unique connections, Fig. 1b). To facilitate comparisons of values across participants, we range-normalized each participant's absolute correlation matrices between 0 and

1 across the 4 conditions. This procedure discarded the between-participant variance while leaving the spatial structure and between-condition variance intact.

In addition, for the post atomoxetine condition time-resolved connectivity (Allen et al., 2014) was computed for 189 tapered windows w of length nw (22 volumes) via:

$$R_{wi} = \frac{M_{wi}' \cdot M_{wi}}{nw - 1}$$

The taper was created by convolving a Gaussian (SD 3 TRs) with a rectangle. R_{wi} was Fisher-transformed to stabilize variance across windows. We then again used the absolute value as our measure of connectivity strength. An identical sliding window was applied to the pupil diameter data in the post atomoxetine condition such that for each window in R_{wi} there was a corresponding value of pupil size during that window. Then, we divided up pupil size into 3 equal-sized bins, and averaged the corresponding values in R_{wi} for each pupil bin separately. To rule out the possibility that the results depended on the choice of bin size, we also tried alternate bin sizes (2, 5, and 7 bins) and found similar effects.

Graph-theoretical analysis of global correlation structure. For each condition, we constructed a binary undirected (adjacency) matrix A. We did this by first concatenating the correlation matrices across participants such that for each condition we had a brain region by brain region by N (90 by 90 by 24) matrix of connectivity. We then assessed with a t-test across the participant dimension for each element y,x in the connectivity matrix whether its value differed significantly from the average of its row y or column x (Hipp et al., 2012). In other words, for each connection we obtained a distribution across participants of weighted values, and two distributions corresponding to the mean weighted values of each brain region that was linked by that particular connection. The connection distribution was then compared to each of the brain region distributions with a t-test. If either of the two comparisons was significant, the connection was scored as 1, and otherwise it was scored as 0. The alpha level was set to 0.01, Bonferroni-corrected for two comparisons to 0.005 (Hipp et al., 2012).

Note that this procedure – as opposed to simply applying a fixed-percentage threshold – results in adjacency matrices that can differ in the number of connections between conditions, and therefore allows the assessment of correlation structure, or degree. We thus quantified the global degree k in each condition as the average across the adjacency matrix (Hipp et al., 2012) via:

$$k = n^{-1} \sum_{x=1}^{n} n^{-1} \sum_{y=1}^{n} A(x, y)$$

where n is the number of brain regions in the AAL atlas.

To test the prediction that increased catecholamine levels should result in stronger functional connectivity, we used k as our measure of connectivity strength rather than relying on the mean weighted values (i.e. the average of R_i). The binarization of weighted graphs is common in functional network analysis (Achard and Bullmore,

2007; Rubinov and Sporns, 2010; Hipp et al., 2012; Li et al., 2013) and is intended to preserve only the strongest (most probable) connections. This ensures that weak edges, which are more likely to be spurious (Rubinov and Sporns, 2010), do not convolute the global mean. Given that these edges are less likely to reflect true neurophysiological interactions, they are less likely to be sensitive to any experimental manipulation that is specifically intended to alter neurophysiology (in our case, drug intake). Thus, excluding these connections decreases the likelihood of false negatives in between-condition comparisons of the global mean. In addition, by treating each connection equally (either present or absent) the global mean is not disproportionally influenced by extremely strong connections that are more likely to decrease in strength after an experimental manipulation by virtue of regression towards the mean.

Furthermore, by defining adjacency matrices using a statistical test across participants, each connection that is present in the adjacency matrix is ensured to be reliably expressed across the group of participants for a given condition. Thus, the adjacency matrices are representative of the group-level topography of connectivity. We used two measures of clustering, defined using these group-level adjacency matrices, to test the prediction that an increase in central catecholamine levels should be accompanied by more strongly clustered network connectivity. The clustering coefficient C was quantified as the average fraction of triangles τ around a node, the latter given by:

$$\tau_x = 2^{-1}\sum A(x,y)A(x,z)A(y,z) \ , \ \text{where} \ y,z \in N$$

and N represents the total set of nodes. C was then given by:

$$C = n^{-1}\sum_{y=1}^{n} n^{-1}\sum_{x=1}^{n}\frac{2\tau_x}{k_x(k_x-1)}$$

The clustering coefficient here is equivalent to the average proportion of the node's neighbors that are in turn neighbors to each other (Watts and Strogatz, 1998; Rubinov and Sporns, 2010). Thus, the clustering coefficient represents the mean fraction of clustering around each node.

Because C is normalized by degree (k) individually per node, it may be biased by nodes with a relatively low k. We therefore also included a measure of clustering that is normalized by k collectively and hence does not suffer from the same potential bias. This measure is known as transitivity (T), and is given by:

$$T = \frac{\displaystyle\sum_{x=1}^{n} 2\tau_x}{\displaystyle\sum_{x=1}^{n} k_x(k_x-1)}$$

Note that this is equivalent to the ratio of triangles to triplets in the network. Both clustering coefficient and transitivity capture the extent to which the network is segregated in terms of processing, because a large number of triangles implies

functional clustering. These two measures were computed using the Brain Connectivity Toolbox (Rubinov and Sporns, 2010). Note that both clustering and transitivity are (partially) dependent on global degree (van Wijk et al., 2010).

To test statistically whether degree, clustering coefficient and transitivity differed between conditions, we used non-parametric permutation testing. We shuffled the condition labels for each participant prior to computing the adjacency matrices and then computed the graph-theoretical measures. This was done for 10,000 iterations to produce a null distribution. We then derived a p value for each contrast by dividing the number of null observations more extreme than the observed contrast by the total number of null observations, and subtracting this value from 1.

Network identification via community detection. We used the Louvain method for community detection optimized for stability (Blondel et al., 2008; Le Martelot and Hankin, 2013) to classify each brain region as belonging to a particular network, or module. This method works by maximizing the number of within-group connections (edges) while minimizing the number of between-group connections via greedy optimization. We first defined an adjacency matrix A_s by concatenating the condition-averaged correlation matrices across participants, and then statistically comparing each element y,x to the average of its row y or column x, similar as described above. However, to accurately classify networks we needed to retain only those connections that were most informative about community structure. We therefore promoted sparsity in the condition-averaged adjacency matrix by defining it using a one-tailed t-test with a conventional alpha level (0.05) and a correction for multiple comparisons using the false discovery rate (FDR). This preserved only those connections that were consistently the strongest across participants (16.9% of all possible connections). We then submitted this sparse condition-averaged adjacency matrix to the Louvain community detection algorithm. The optimization procedure (Le Martelot and Hankin, 2013) ensured a stable solution across multiple runs of the algorithm. In the optimization procedure, the Markov time acts as a resolution parameter that determines the community scale, and thus the number of modules that the algorithm will return. This parameter was set to 0.9, resulting in 6 separate modules. We set the number of modules to be detected to 6 because, given the relatively coarse anatomical layout of the AAL atlas, this number yielded a relatively reliable modular organization. The community detection and optimization resulted in a 'module number' for each AAL brain region indicating to which module it belonged, and a single Q-value indicating the strength of modularity.

We first verified whether the Q-value was significantly higher than chance. To do so, we generated 10,000 randomized null networks with an identical size, density and degree distribution as A_s (Maslov and Sneppen, 2002), and submitted them to Louvain community detection and optimization to produce a null-distribution of Q-values. We then derived a p-value for the observed modularity by dividing the number of null Q-values more extreme than the empirical Q-value by the total number of null Q-values, and subtracting it from 1.

The observed Q-value of 0.46 was significantly higher than chance ($p < 0.001$), showing that group-average connectivity was strongly modular. We then visualized the modular structure by re-arranging the condition-averaged correlation matrix by module. The assignment of brain regions to modules corresponded closely to a number of well-characterized intrinsic connectivity networks, indicating that the modular structure reflected a functionally meaningful grouping of brain regions.

Graph-theoretical analysis of network structure. The procedure described above allowed us to group brain regions into modules of intrinsically coupled AAL brain regions. We could then use these modules to assess changes in the structure of intrinsic correlations at the within- and between-network level, rather than as a function of the system in its entirety. To do this, we first re-arranged the condition-specific adjacency matrices by their module number, and computed average degree of elements within and between modules via:

$$k_m = n_a^{-1} \sum_{x_a=1}^{n_a} n_b^{-1} \sum_{y_b=1}^{n_b} A_s(x_a, y_b)$$

where n_a is the number of brain regions belonging to module *a* and n_b is the number of brain regions in module *b.* This yielded, for each condition, a symmetric and module-by-module matrix of continuous average degree values, in which values on the diagonal indicated the average number of connections within each module, and each value around the diagonal indicated the average number of connections between a combination of modules.

We could then use these 'module matrices' to test for atomoxetine-related changes in degree of the connections within modules, and the connections linking different modules. This allowed us to characterize changes in connectivity in a spatially more specific way than for global degree. We again used non-parametric permutation testing, similar as described for global degree, except that it was done for individual elements within the module matrices.

Control analyses using an alternate atlas and multiple thresholds. To rule out the possibility that our results were specific to the use of the AAL atlas, we repeated all of our key analyses using the atlas made available by Craddock et al. (2012), which comprised 87 distinct regions after excluding the cerebellum, and found similar effects in terms of both direction and significance. Moreover, to verify that our results were independent of the statistical threshold used to define the adjacency matrices, we conducted a control analysis in which a range of adjacency matrices was created per condition with varying condition-averaged connection densities (40-75%). This was done by progressively raising / lowering the alpha level of the *t*-test that was used to determine whether a connection is present or absent (see above). Then, for each threshold we computed the graph-theoretical measures, and for each condition and measure separately calculated the area under the curve (AUC) across thresholds. This allowed us to compare the AUC between conditions with permutation testing

(10,000 iterations). For all measures the critical interaction contrast was significant and in the same direction as our original findings (see Results).

Controlling for regression towards the mean. The correlation between baseline coupling strength and the atomoxetine-related change in coupling strength (Fig. 3e, see Results) is confounded by regression towards the mean. That is, if two particular brain regions show strong baseline coupling, then simply by chance they are more likely to show a reduction under atomoxetine, and so a negative correlation is likely to occur. We therefore controlled for regression towards the mean using permutation testing. For 10,000 permutations we shuffled the condition labels across participants prior to computing the atomoxetine-related change in coupling strength. We then computed the correlation between baseline coupling and atomoxetine-related change in coupling to produce a distribution of correlation coefficients under the null hypothesis of regression towards the mean. Finally, we derived a *p* value for the empirical correlation coefficient by dividing the number of null observations more extreme than the correlation coefficient by the total number of null observations, and subtracting this value from 1. This *p* value indicated the significance of the observed correlation coefficient beyond regression towards the mean.

Analysis of BOLD signal variance. We calculated for each participant and each AAL brain region the fractional amplitude (i.e., variance) of low-frequency fluctuations in the non-Z-scored BOLD time series [fALFF (Zou et al., 2008)]. This measure indexes the relative contribution of low-frequency (0.01-0.08 Hz) fluctuations to the total amplitude spectrum. We compared fALFF between conditions using repeated measures analysis of variance. Additionally, for each participant we correlated the atomoxetine-related change in fALFF with the atomoxetine-related change in inter-regional correlation strength across AAL brain regions. We then compared the distribution of Fisher-transformed correlation coefficients to zero using a two-tailed *t*-test. Very similar results were obtained using alternative measures of variance (e.g., average 0.01-0.08 Hz amplitude or the signal standard deviation rather than fractional amplitude).

3.3 Results

Atomoxetine reduces global degree and clustering. In a first set of analyses, we examined the effect of atomoxetine on graph-theoretical summary measures of functional connectivity strength. We parcellated each participant's brain into 90 separate regions according to the Automated Anatomical Labeling (AAL) atlas [(Tzourio-Mazoyer et al., 2002), Fig. 1a] and computed the correlation between the Z-scored time series of all pairs of regions (Fig. 1b). We then took the absolute correlation coefficient as our measure of functional connectivity strength (see Materials and Methods). In general, functional connectivity was strongest between
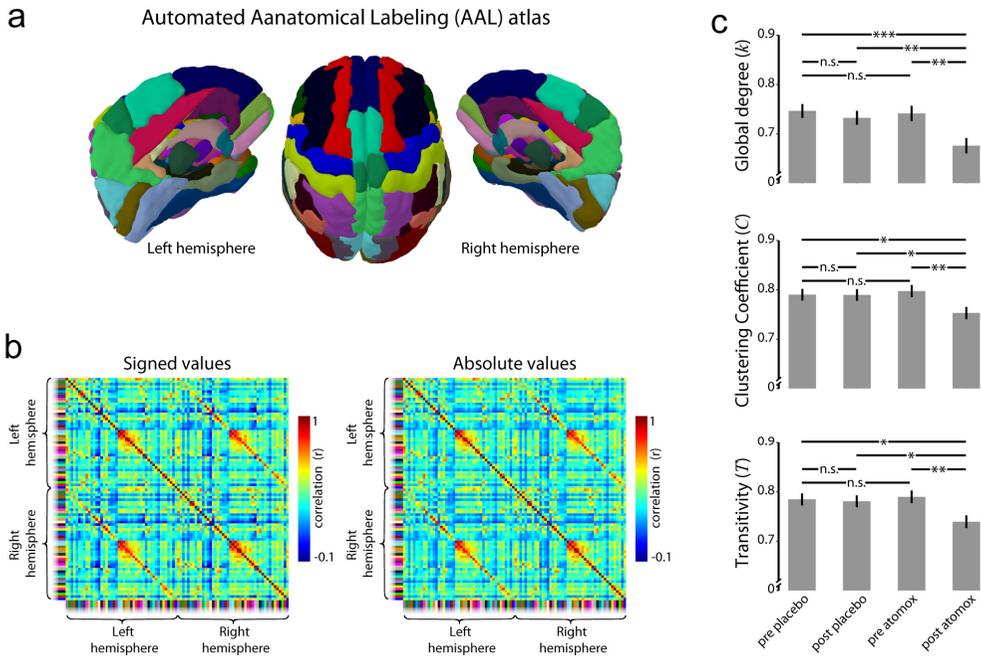
Figure 1. Inter-regional correlation and global graph-theoretical results. **a)** Topography of the AAL atlas. Each brain region within hemispheres has a unique color. **b)** Condition-averaged inter-regional correlation. Both the signed and absolute values are shown. Color labels on the left and bottom axes correspond to brain regions in a). **c)** Atomoxetine effects on global graph-theoretical measures. Error bars represent the SD of the bootstrapped null-distribution. n.s.: non-significant; *: $p < 0.05$; **: $p < 0.01$, ***: $p < 0.001$.

visual cortical areas and between homologue areas in both hemispheres (Fig. 1b), consistent with a host of previous work (Fox and Raichle, 2007).

For each condition (pre placebo, post placebo, pre atomoxetine, post atomoxetine), we constructed a binary matrix of connections (*edges*) between pairs of brain regions that consistently differed in strength across participants from the average of other connections involving either of the two brain regions [c.f. (Hipp et al., 2012)]. Graph theory allowed us to capture different properties of these matrices of intrinsic correlations in a small number of diagnostic scalar quantities (Bullmore and Sporns, 2009; Rubinov and Sporns, 2010). Specifically, we assessed three such measures: the *global degree,* which indexes the number of strongly correlated regions (above a certain threshold; see Materials and Methods) in the network, and two descriptors of the extent to which network connectivity is clustered in segregated local groups of brain regions: *clustering coefficient* and *transitivity*, both of which are (partially) dependent on the strength of connectivity (van Wijk et al., 2010; Eldar et al., 2013). If catecholamines increase global functional connectivity, then atomoxetine should increase all three measures.

43

Figure 1c shows that atomoxetine significantly reduced the number of strong correlations present in the network, as indicated by lower global degree. This was reflected in a significant interaction between treatment and time (p = 0.039). A similar pattern of results was found for the two measures of clustering, both of which decreased in magnitude (Fig. 1c): clustering coefficient (p = 0.043) and transitivity (p = 0.048). Thus, atomoxetine reduced the number of strongly correlated brain regions, as well as the extent to which correlated brain regions formed local functional ensembles. Together, these result show that atomoxetine decreases, rather than increases, overall inter-regional correlations in the brain at rest.

Atomoxetine reduces inter-network degree. Many studies of resting-state activity in humans have revealed a consistent set of groups or *modules* of brain regions that are characterized by strong coupling between brain regions belonging to the same module, and weaker coupling between brain regions belonging to different modules (Bullmore and Sporns, 2009). These modules are often referred to as "intrinsic functional connectivity networks" (Fox and Raichle, 2007). In a next set of analyses, we investigated atomoxetine-related changes in the strength of functional connectivity within and between these networks.

To do this we arranged the connectivity matrix by network [Fig. 2a; see (Blondel et al., 2008)]. This resulted in 6 functional networks that correspond closely to previously reported resting-state networks (Yeo et al., 2011; Zalesky et al., 2014). Based on their topography (Fig. 2b) we termed them: default; somato-motor; visual; sub-cortical; inferior-frontal; and fronto-parietal networks. We then calculated the average number of connections within and between these networks, resulting in a 6 by 6 *network degree* matrix for each condition (Fig. 2c,d). Finally, we examined atomoxetine-related changes in within- and between-network degree using permutation testing. This allowed us to explore if changes in functional connectivity occurred in intra- or inter-network connections. Note that the atomoxetine-related reduction in global degree (Fig. 1c) is visible in the network degree matrices as an overall increase in 'brightness' in the right panel of Fig. 2d. Consistent with the decrease in global degree reported above, we observed only atomoxetine-related *reductions* in network degree (Fig. 2e). The interaction between treatment and time was significant for the connections between the visual and somato-motor networks (p < 0.001); between the visual and fronto-parietal networks (p = 0.044); and between the fronto-parietal and default networks (p < 0.001). After using the FDR (q = 0.05) to correct for multiple comparisons, all connections except the connection between the visual and fronto-parietal networks remained significant. However, when comparing the area under the curve across a range of thresholds (see Materials and Methods), all connections remained significant after FDR correction.

Interestingly, all significant reductions in network degree were in connections between (as opposed to within) functional networks. Thus, the most robust decreases in functional coupling occurred for connections linking functionally dissociable groups of brain regions. These results corroborate the conclusion drawn above, that
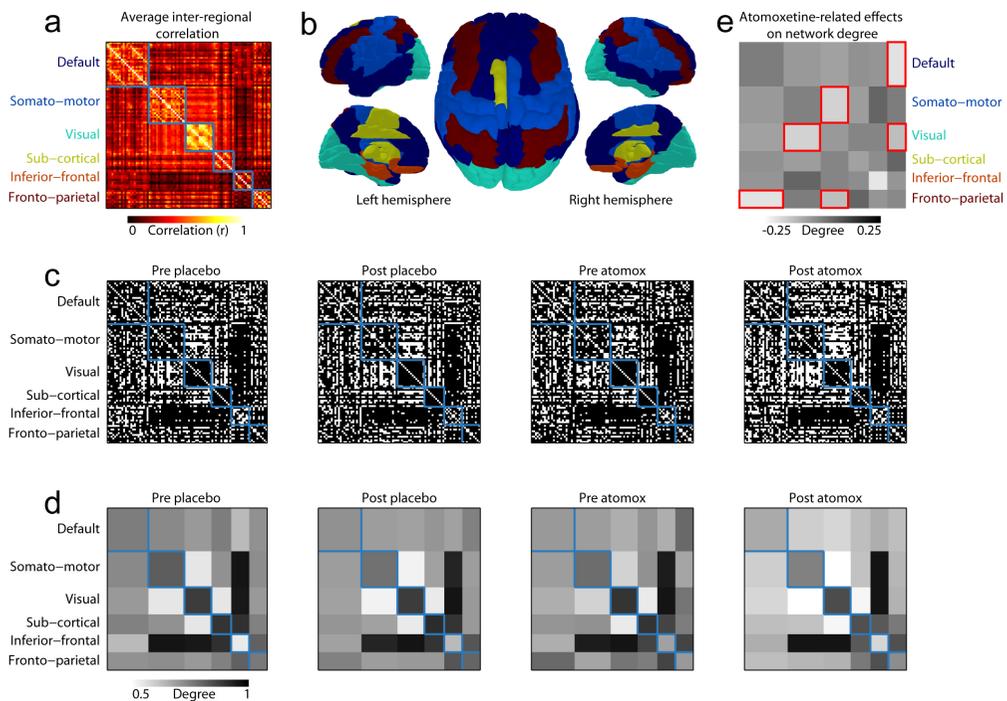
Figure 2. Intrinsic connectivity networks and changes in graph-theoretical measures of network structure. **a)** Condition-averaged inter-regional correlation arranged by network. The networks are outlined in blue. **b)** Topography of functional networks. Colors correspond to the labels in a). **c)** Condition-specific adjacency matrices arranged by network. Black elements indicate that a connection is present. **d)** Average degree for within- and between-network connections. To facilitate visual comparison, the size of each network is the same as in c). Note, however, that all statistical comparisons were conducted on symmetrically sized matrices in which each network contributed equally to the global mean. **e)** Atomoxetine-induced changes in degree for connections within and between networks. Shades of grey reflect the value of the interaction contrast (post atomoxetine - pre atomoxetine) minus (post placebo - pre placebo). Significant ($p < 0.05$) changes in degree are outlined in red.

atomoxetine decreased inter-regional correlations in the brain at rest. More importantly, these results provide a first indication that this reduction in inter-regional correlations is not spatially homogeneous across the brain. In the following, we further characterize the spatial heterogeneity of the atomoxetine-induced reductions in inter-regional correlations.

Regionally-specific reductions and baseline-dependent changes in connectivity with atomoxetine. Having assessed the topographical changes induced by atomoxetine at the global level and the level of functional networks, we next assessed differences in the strength of inter-regional correlations at the level of individual connections between brain regions, using the absolute correlation coefficient. We
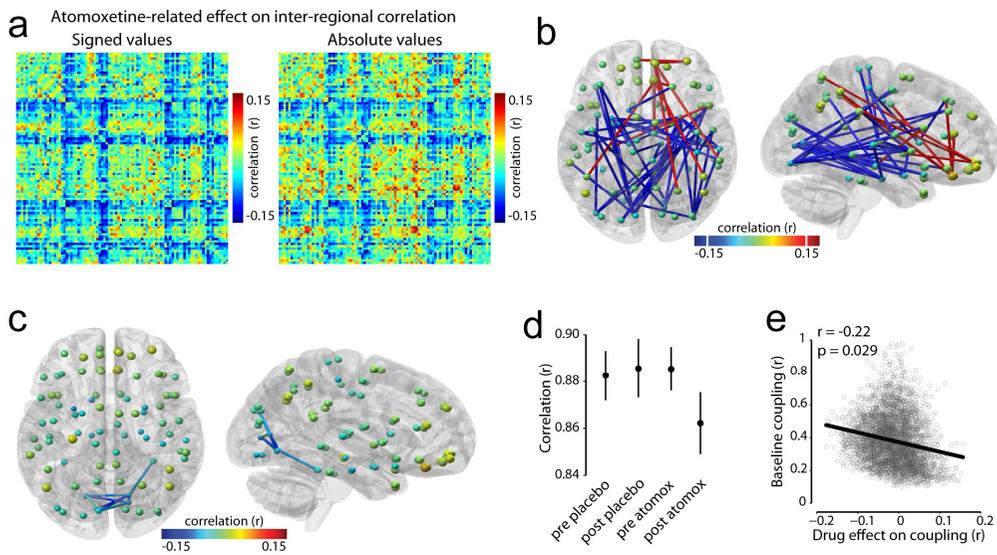
Figure 3. Atomoxetine-related effects on inter-regional correlation. **a)** Region-by-region matrix of atomoxetine-related changes in inter-regional correlation strength. Colors represent the value of the interaction contrast (post atomoxetine - pre atomoxetine) minus (post placebo - pre placebo). Blue colors indicate reduced correlation following atomoxetine. The matrices are organized following Figure 1b. **b)** Atomoxetine-related effect on the absolute inter-regional correlation coefficient, rendered in 3D with an arbitrary threshold applied. The threshold is indicated by the white dashes in the color bar. Spheres are placed in the center of mass of their respective AAL atlas regions. Both the size and color indicate the average atomoxetine-related effect on coupling [i.e the average across rows or columns in a)]. **c)** Transverse (top is anterior) and sagittal (right is anterior) view on 3D rendering of significant correlation changes, resulting from the whole-brain two-step analysis. Individual connections that changed significantly with atomoxetine are plotted as cylinders between the corresponding regions. **d)** Inter-regional correlation in each condition, averaged across the significant connections shown in c). Error bars represent the SEM. **e)** Correlation between baseline inter-regional correlation strength (collapsed across pre placebo and pre atomoxetine) and change with atomoxetine. Each dot represents a unique region-by-region connection. Self-connections were excluded.

found that atomoxetine altered correlation strength in a strikingly structured fashion (Fig. 3a): In general, connectivity was reduced by atomoxetine, especially in posterior brain regions (Fig. 3b). These observations align with our findings of reductions in inter-network degree involving the visual system. To quantify these effects, we used a two-step procedure. Specifically, we first derived a set of data-driven hypotheses by identifying, in the first half of the fMRI volumes, the limited number of individual connections that exhibited an atomoxetine-related change in connectivity that was reliable across participants ($p < 0.05$, using a two-tailed t-test), thereby reducing the number of comparisons for the subsequent step. We then re-tested those connections

using the (independent) second half of the volumes, and selected those that again showed a systematic atomoxetine-related change in correlation strength ($p < 0.005$, two-tailed). Atomoxetine significantly lowered correlation strength in a cluster of occipital brain regions (Fig. 3c,d), specifically correlations between left calcarine cortex and right calcarine cortex / bilateral lingual gyrus; between left cuneus and right calcarine cortex / lingual gyrus; between left lingual gyrus and right calcarine cortex / lingual gyrus; and between right lingual gyrus and right calcarine cortex / right fusiform gyrus. These contiguous connections remained significant after applying a highly conservative cluster size threshold ($p < 0.0001$), obtained by generating a distribution of maximum cluster sizes under the null hypothesis with permutation testing (Nichols and Holmes, 2001). Thus, the cluster involved significantly more connections than would be expected by chance.

Note that we did not find significant changes in connectivity between structures of the basal ganglia, which have been widely studied in relation to catecholaminergic drug effects (Sulzer et al., 2016). This lack of an atomoxetine-related effect in the human basal ganglia is consistent with the observation that the basal ganglia receive relatively sparse noradrenergic innervation (Aston-Jones et al., 1984), and with findings that atomoxetine has little effect on DA levels within the basal ganglia of rodents (Bymaster et al., 2002). Indeed, unlike in the cortex, in the basal ganglia there is an abundance of DA transporter (Sulzer et al., 2016), so DA reuptake is not dependent on the NE transporter. Thus, our finding that atomoxetine reduced the strength of inter-regional correlations in (predominantly visual) cortical areas is consistent with the specific effect of atomoxetine on synaptic catecholamine levels within the cortex.

In sum, atomoxetine lowered the strength of correlations between visual cortical brain regions (Fig. 3c), regions that on average showed strong connectivity (Fig. 1b, 2a). This raises the question whether the atomoxetine-induced change in connectivity was dependent on the baseline level of connectivity between any pair of brain regions. To address this question we correlated weighted coupling strength collapsed across the pre placebo and pre atomoxetine conditions (i.e., baseline connectivity) with the coupling change following atomoxetine (controlling for regression towards the mean with permutation testing). The observed correlation between baseline connectivity strength and the change with atomoxetine was significant ($r = -0.22$, $p = 0.029$, Fig. 3e). This indicates that the strongest functional connections tended to show the largest connectivity reductions after atomoxetine, and vice versa.

Atomoxetine induces de-coupling of early visual cortex from the rest of the brain. As noted above, atomoxetine reduced correlations between occipital brain regions. To establish whether these occipital regions also showed reduced coupling to the rest of the brain, as suggested by visual inspection (Fig. 3b), we computed a summary statistic (median) of correlation strength between early visual (pericalcarine) cortex and all other AAL atlas regions. There were no differences between left and right early visual cortex, so we collapsed the data across hemispheres.
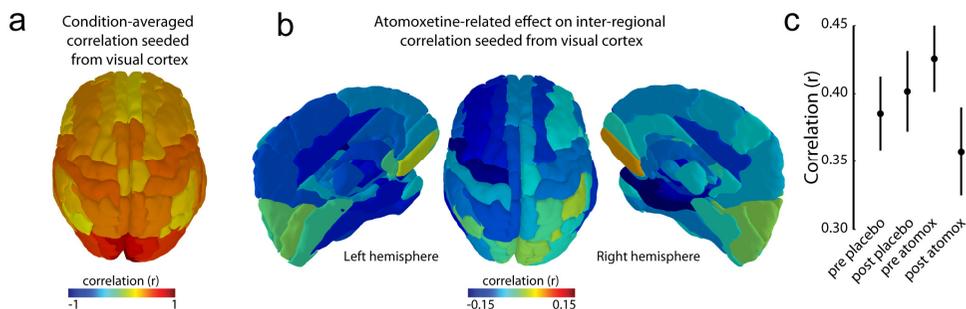
Figure 4. Atomoxetine reduces correlation strength between early visual (pericalcarine) cortex and the rest of the brain. **a)** Topography of condition-averaged correlation seeded from left and right early visual cortex. **b)** Topography of atomoxetine-related effects on correlation seeded from left and right early visual cortex. Colors represent the value of the interaction (post atomoxetine - pre atomoxetine) minus (post placebo - pre placebo). **c)** Median correlation values across the brain seeded from left and right early visual cortex. Error bars represent the SEM.

Average connectivity seeded from early visual cortex is shown in Figure 4a. Atomoxetine reduced connectivity between early visual cortex and the rest of the brain as reflected by a significant interaction between treatment and time (F(1,23) = 5.31, p = 0.031, Fig. 4b,c). The only significant pairwise comparison was post atomoxetine versus pre atomoxetine (t(23) = 2.34, p = 0.028). Together these results suggest that the early visual cortical areas not only de-coupled from each other following atomoxetine (Fig. 3c) but also from the rest of the brain.

The results of our analyses at the level of individual connections between brain regions converge with those at the global level and at the level of networks of brain regions, showing that atomoxetine decreased functional connectivity. In addition, the results show that atomoxetine modulated functional connectivity in a highly regionally specific fashion, with more robust changes in visual cortex than in other brain areas. *Excluding alternative explanations.* In five sets of control analyses, we ruled out the possibility that the atomoxetine-related changes in inter-regional correlations were driven either by local changes in BOLD variance, by retinal effects due to pupil dilation associated with atomoxetine, by head motion, by saccade-related retinal transients, or by atomoxetine-induced changes in physiology (heart rate and breath rate). First, the correlation coefficient between two signals is their covariance normalized by the signals' variances. Thus, it is possible that the observed changes in inter-regional correlations are caused by local changes in variance alone (Haynes et al., 2005; Freeman et al., 2011), rather than by changes in covariance (i.e., the degree to which the BOLD signals in two regions fluctuated together). If this is the case, then the atomoxetine-related change in average inter-regional correlation and the atomoxetine-related change in BOLD signal variance should be negatively correlated across brain regions.

Figure 5. Spectral BOLD characteristics and the relation with inter-regional correlations. (Left) atomoxetine-induced changes in spectral amplitude for AAL brain regions that showed an atomoxetine-induced increase (red) and decrease (blue) in inter-regional correlation strength. Brain-region and condition-averaged amplitude is shown in black. (Right) Mean correlation between the region-averaged atomoxetine-induced change in coupling strength and fractional amplitude of low-frequency BOLD fluctuations. Error bars show the SEM. **: $p < 0.01$.
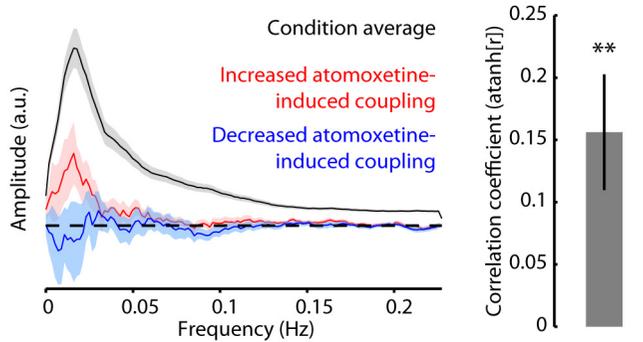
Instead, we found a positive relationship between changes in inter-regional correlation and changes in BOLD variance, which was consistent across participants ($t(23)$ = 3.36, p = 0.003, Fig. 5), ruling out variance as a confound. Moreover, there was no interaction between treatment and time in overall BOLD variance ($F(1,23)$ = 0.71, p = 0.40), or in variance for only the occipital brain regions that showed reduced atomoxetine-related inter-regional correlation ($F(1,23)$ = 0.41, p = 0.53).

Second, because atomoxetine increased the size of the pupil (Fig. 6a), it is conceivable that this peripheral effect, rather than the effect of atomoxetine on central catecholamine levels, was driving the changes in inter-regional correlation in visual cortex (Haynes et al., 2004). To examine this potential confound, we binned inter-regional correlation by pupil size in the post atomoxetine condition, focusing on those correlations that showed a significant reduction under atomoxetine. If larger pupil size is responsible for the reduction in correlations, then time

Figure 6. a) Atomoxetine effect on pupil diameter. **b,c)** Correlation strength in the post atomoxetine condition binned by pupil size, only for connections that showed an atomoxetine-related reduction in inter-regional correlation. Error bars represent the SEM. n.s.: non-significant; *: $p < 0.05$.
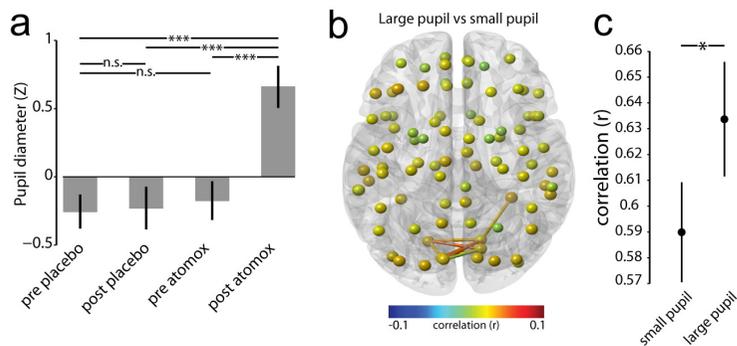
periods during which the pupil is large should be associated with weaker correlations than time periods during which the pupil is small. Interestingly, we found the opposite pattern (Fig. 6b, c): stronger correlations for large pupil ($t(18) = 2,84$, $p = 0.010$), ruling out an interpretation in terms of pupil size.

Third, head motion can have a strong influence on the strength of inter-regional correlations (Van Dijk et al., 2012). To rule out the possibility that our key finding of atomoxetine-related changes in inter-regional correlation was driven by head motion, we first compared head motion between conditions. Neither mean head motion nor mean absolute head motion differed between conditions (all ps > 0.05). No participant's head motion exceeded 2 mm, indicating that overall there was little head motion. However, general mild head motion tends to increase correlations between proximate areas and decrease connectivity between distant areas (Van Dijk et al., 2012). Thus head motion can potentially lead to spatially heterogeneous effects on connectivity in a manner that is related to the distance between brain areas. To rule out the possibility that the spatial structure of atomoxetine-related changes in connectivity was driven by subtle (non-significant) differences in head motion between conditions, we correlated Euclidean distance between the center of mass of each pair of AAL brain areas and the strength of functional connectivity between those areas, for each participant and each condition. We then compared the distribution of Fisher-transformed correlation coefficients between conditions. If head motion is responsible for the observed change in connectivity between conditions, then the correlation between Euclidean distance and strength of connectivity should also differ between conditions. However, we did not find any differences between conditions (all ps > 0.05), ruling out head motion as a confound.

Fourth, it is possible that that the atomoxetine-related reduction in the strength of correlation between visual cortical areas occurred due to differences between conditions in saccade-related retinal transients. In order to rule out this possibility, we extracted several eye movement metrics from the eye tracker gaze position data using the EYE-EEG toolbox (Dimigen et al., 2011). There was no interaction between treatment and time for any of the metrics: the number of saccades ($F(1,18) = 0.47$, $p = 0.50$), median saccade amplitude ($F(1,18) = 0.45$, $p = 0.51$), median saccade duration ($F(1,18) = 0.11$, $p = 0.74$), or median saccade peak velocity ($F(1,18) = 3.32$, $p = 0.085$). This latter trend was driven by a numeric difference between the pre placebo and post placebo conditions. Pre and post atomoxetine did not differ significantly in saccade peak velocity ($t(18) = -0.43$, $p = 0.67$). Together, these results show that our key result of an atomoxetine-related reduction in the strength of correlation between visual cortical regions was unlikely to be driven by saccade-related retinal transients.

Finally, atomoxetine significantly increased breath rate ($F(1,23) = 8.96$, $p = 0.007$) and heart rate ($F(1,23) = 4.66$, $p = 0.041$), as reflected by a significant interactions between treatment and time. We therefore corrected the BOLD time series using the RETROICOR method [see Materials and Methods; (Glover et al., 2000)]. The average

R^2 of the physiology regressors was relatively low (0.034), indicating that physiology accounted for a small proportion of the total BOLD variance [which was likely the result of artifact removal by FMRIB's ICA-based X-noiseifier (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014)]. Nevertheless, to conclusively rule out atomoxetine-related changes in physiology as confounds, we repeated the key analyses on the physiology-corrected data. All three global graph-theoretical measures remained significant and in the same direction as reported above (all p's < 0.05). We also found significant reductions in network degree in the same inter-network connections (all p's < 0.05). Lastly, we observed a similar contiguous cluster of significantly reduced inter-regional correlations within visual cortex (all p's < 0.005, and cluster-corrected at $p < 0.0001$). Thus, our key results were unlikely to be driven by atomoxetine-related changes in physiology.

3.4 Discussion

Using a pharmacological manipulation, we examined the effects of increased extracellular levels of the catecholamines NE and DA on resting-state fMRI connectivity in the human brain. First, we found that our manipulation reduced the strength of inter-regional correlations across three levels of spatial organization, indicating that catecholamines reduce the strength of functional interactions during rest. Second, this modulatory effect on the structure of resting-state correlations exhibited a substantial degree of spatial specificity, indicating that catecholamines differentially reduce spontaneous correlations between select brain regions. These two key findings are surprising in light of the common understanding of the neurophysiology and computational function of catecholaminergic systems. They also identify catecholaminergic neuromodulation as an important factor shaping the spatial structure and strength of intrinsic functional connectivity in the human brain.

Our first key finding is that atomoxetine, a selective NET blocker that increases synaptic NE and DA levels (Bymaster et al., 2002; Devoto et al., 2004; Invernizzi and Garattini, 2004; Swanson et al., 2006; Koda et al., 2010), reduced the strength of inter-regional correlations. Specifically, atomoxetine reduced the strength of connectivity globally (Fig. 1c), between nodes belonging to distinct intrinsic connectivity networks (Fig. 2e), and between individual brain regions within the visual system (Fig. 3c). This consistent pattern of results seems to be at odds with the notion of a facilitative effect of catecholamines on brain-wide signal transmission (Aston-Jones and Cohen, 2005; Eldar et al., 2013). One possible explanation for this discrepancy lies in the fact that in our experiment participants did not actively respond to incoming sensory information. According to a recent theory, the effects of NE on neural activity strongly depend on interactions with local glutamate release (Mather et al., 2015). Accordingly, enhanced NE may have qualitatively different effects during task processing, associated with relatively high glutamate activity, than during states of relative cortical

quintessence (i.e., at rest), associated with relatively low glutamate activity. In line with this possibility, Coull *et al* showed dissociable effects of the α2 adrenergic agonist clonidine on positron emission tomography effective connectivity obtained during task performance and during rest (Coull et al., 1999). Whereas during task performance clonidine increased connectivity between frontal and parietal cortical regions, during rest clonidine reduced connectivity from frontal cortex to thalamus, and in connections to and from visual cortex. Two other studies that employed NE drugs also provided evidence for regional reductions in connectivity strength during rest (McCabe and Mishor, 2011; Metzger et al., 2015). These studies, however, only used a small number of seed regions to assess connectivity, and hence did not examine large-scale topographical changes.

Our second key finding is that atomoxetine resulted in spatially heterogeneous changes in inter-regional correlations. For example, atomoxetine caused a reduction in the number of strongly correlated brain regions between (but not within) distinct resting-state networks (Fig. 2e). Furthermore, the effect of atomoxetine on inter-regional correlations was dependent on the baseline level of coupling: the strongest functional connections tended to show the largest connectivity reductions after atomoxetine (Fig. 3d). How can such spatially structured effects of catecholamines come about? First, recent anatomical tracing work has suggested that the projection profile of the LC is more heterogeneous than once thought (Schwarz and Luo, 2015). For example, even though on the whole there is broad collateralization within the LC-NE system, sub-populations of LC neurons selectively innervate distinct brain regions (Chandler et al., 2014; Schwarz et al., 2015). Moreover, sub-populations of LC neurons that differ in their afferent projection profile also show marked differences in their firing characteristics (Chandler et al., 2014). The firing modes of LC neurons in turn have differentiable effects on neuronal synchronization within the cortex (Safaai et al., 2015). Importantly, LC neurons have been reported to co-release DA (Devoto and Flore, 2006). Thus, spatially selective effects of catecholamines on correlated fluctuations in the brain can be achieved via a heterogeneous cortical innervation by the LC.

Second, heterogeneity in the effect of catecholamines on inter-regional correlations could be achieved by regional differences in the expression of different receptor types. For example, expression of the α2 receptor roughly follows an anterior to posterior gradient (Nahimi et al., 2015), with particularly strong expression in primary visual cortex (Zilles and Amunts, 2009). Interestingly, we observed an anterior to posterior gradient in the effect of atomoxetine on the strength of correlations (Fig. 3b). Moreover, we found a pronounced reduction in the strength of correlations between regions within visual cortex, and between early visual cortex and the rest of the brain (Fig. 4). The similarity between the spatial distributions of α2 receptors and the effects of atomoxetine thus warrants further investigation into the relationship between specific NE receptor types and their influence on correlated activity across the brain.

A number of limitations of the present study should be acknowledged. First, we examined the effects of only one dose (40 mg) of atomoxetine. Dose-dependent pharmacological effects of catecholaminergic drugs on neural function are not uncommon (Berridge and Waterhouse, 2003). Future work on the neurochemical basis of functional connectivity will need to examine dose-dependent effects of atomoxetine, and other catecholaminergic drugs, with different pharmacokinetic profiles. Second, we do not know whether atomoxetine would have similar effects on functional connectivity in clinical populations characterized by disturbed catecholaminergic function (e.g., attention deficit hyperactivity disorder and depression). Third, although we used BOLD activity as a proxy for neural activity, the link between neuronal interactions and BOLD activity is not entirely clear (Logothetis, 2008). Models of catecholamine function make predictions about how NE and DA should affect neural communication. However, the translation of these predictions to BOLD correlations is not straightforward. Lastly, we used an atlas-based brain parcellation to investigate inter-regional correlations. Thus, the spatial resolution of our analyses was restricted by the resolution of the atlas. Future work, using voxel-level approaches, is needed to investigate more fine-grained spatial effects of catecholamine levels on functional connectivity.

The synaptic effects of catecholamines have been relatively well charted (Berridge and Waterhouse, 2003; Winterer and Weinberger, 2004). However, there is considerable uncertainty about how these low-level effects translate to system-wide functional interactions. Recently, a study by Safaai *et al* provided an important first glimpse into how the LC-NE system modulates spontaneous cortical activity and how this modulation in turn affects sensory processing in anesthetized rats (Safaai et al., 2015). Specifically, they showed that LC bursts can both attenuate and enhance processing of sensory stimuli depending on their timing relative to the stimulus and the cortical activity state. However, the effects of catecholamines on the large-scale communication *between* distant brain areas and their neurophysiological underpinnings remain exceedingly unexplored. Our finding that atomoxetine reduced inter-regional correlations in a spatially structured manner thus calls for novel work on the neural mechanisms that produce such effects.

Theory and evidence indicate that the topography of intrinsic fMRI correlations is dictated to an important extent by the fixed anatomical connectivity of each brain region (Deco et al., 2011; Deco et al., 2013). That is, brain regions that are anatomically strongly connected are more likely to show strong functional coupling than those that are connected weakly or only indirectly. However, within the constraints of physical connectivity there is substantial room for state-dependent movement in functional topological space (Allen et al., 2014; Zalesky et al., 2014; Barttfeld et al., 2015). Our results identify NE and DA as important factors driving these movements, and thus suggest that spontaneous fluctuations of catecholamine levels can serve to flexibly alter the structure of spontaneous correlations both globally and in specific brain regions, around the anatomical backbone.

4. Catecholamines Modulate Intrinsic Long-Range Correlations in the Human Brain

Abstract

Brain activity fluctuates intrinsically, even in the absence of changes in sensory input and motor output. These fluctuations are correlated across large-scale networks of brain regions, and their strength and topography changes dynamically. Such dynamic changes in functional connectivity may be induced by brainstem neuromodulatory systems: in particular the locus coeruleus, which projects widely to the forebrain where it co-releases the catecholamines norepinephrine and dopamine. In the current study we examined whether catecholamines change the strength or the spatial structure (topology) of intrinsic long-range correlations, or both. Using a double-blind placebo-controlled crossover design, we pharmacologically increased central catecholamine levels in healthy human participants by administering atomoxetine. We used two complementary analysis approaches to examine the effect of catecholamines on fine-grained strength and topology of intrinsic functional connectivity patterns: 'dual regression' and 'spatial mode decomposition'. Both approaches provided converging evidence for an atomoxetine-related reduction in correlation strength between distributed brain regions. Importantly, the pre-dominant effects of the drug were quantitative changes of correlations within existing functional networks that left the spatial structure of these networks intact, rather than reconfigurations of the topology of these networks. We conclude that catecholamines modulate dynamic changes in the strength of intrinsic inter-regional correlations.

Chapter 4

4.1 Introduction

In the absence of changes in sensory input and motor output, brain activity fluctuates in intrinsically organized correlated networks (Biswal et al., 1995; Fox and Raichle, 2007). The strength and spatial structure of these intrinsic correlations predict task-based brain activation (Cole et al., 2016; Tavor et al., 2016), behavior (De Luca et al., 2005; Seeley et al., 2007), and are useful to study neural dysfunction in clinical populations (Greicius et al., 2004; De Luca et al., 2005; Rombouts et al., 2005; Di Perri et al., 2016). The topology of this 'functional connectivity' is constrained by the (largely) fixed structural connectivity between brain regions (Deco et al., 2011; Deco et al., 2013), which determines the anatomical backbone along which functional connectivity patterns can change dynamically (Allen et al., 2014; Zalesky et al., 2014; Barttfeld et al., 2015). It has been proposed that such changes in functional connectivity patterns may be induced by brainstem neuromodulatory systems (Leopold et al., 2003; Drew et al., 2008; Schölvinck et al., 2010). An important example is the locus coeruleus, which sends diffuse, ascending projections to the forebrain, where noradrenergic terminals release the catecholamines norepinephrine (NE) and dopamine (DA) (Devoto and Flore, 2006).

Several lines of evidence suggest that catecholamines might shape intrinsic correlations in activity between brain regions, possibly in diverse ways, changing either the strength or the topology of these correlation patterns, or both. First, at the single-cell level, catecholamines enhance neuronal responses to excitatory synaptic input (Rogawksi and Aghajanian, 1980; Seamans et al., 2001b; Wang and O'Donnell, 2001) and can amplify GABAergic inhibition (Moises et al., 1979; Seamans et al., 2001a). Such enhanced synaptic efficacy results in system-level signal amplification (Berridge and Waterhouse, 2003). Second, direct pharmacological manipulations of synaptic catecholamine levels have been shown to alter the global strength of inter-regional cofluctuations (Guedj et al., 2016; van den Brink et al., 2016; Warren et al., 2016). Third, evidence from small-scale circuits in crustaceans suggests that (catecholaminergic) neuromodulation can dynamically reconfigure functional networks, despite a constant structural connectome (Marder, 2012; Bargmann and Marder, 2013; Marder et al., 2014). Analogous effects have been suggested to underlie fast "resets" of brain network dynamics in the mammalian brain (Bouret and Sara, 2005), but direct evidence for catecholaminergic reconfiguration of cortical networks in humans has been lacking so far.

Here, we investigated catecholaminergic modulations of large-scale patterns of intrinsic fMRI signal correlations in the human brain at 'rest'. To this end, we re-analyzed data from a double-blind placebo-controlled crossover study (van den Brink et al., 2016) of central catecholamine effects using atomoxetine, a selective NE transporter blocker. Our previous study quantified atomoxetine-induced modulations of the global strength of intrinsic correlations (van den Brink et al., 2016). Here, by contrast, we examined finer-grained patterns of intrinsic correlations, in order to test

for possible atomoxetine-induced quantitative changes in existing correlation patterns versus reconfiguration of correlation patterns.

We used two complementary analysis approaches. The first approach, known as 'dual regression', has been widely used to study the effects of pharmacological manipulations on fMRI signal correlations during the resting state (Beckmann, 2009; Filippini et al., 2009). The dual regression approach first delineates patterns of intrinsically correlated brain regions, so-called functional networks, and then compares voxel-level cofluctuation strength with those networks between conditions or groups. This method has proven useful for elucidating pharmacological effects on fMRI functional connectivity (Chamberlain et al., 2007; Klumpers et al., 2012; Cole et al., 2013; Klaassens et al., 2015; Guedj et al., 2016; Schrantee et al., 2016; Klaassens et al., 2017). The second approach was linear decomposition of intrinsic signal correlation matrices into so-called 'spatial modes', again constituting patterns (or 'networks') of cofluctuations in brain activity (Mitra and Pesaran, 1999; Friston and Büchel, 2004; Donner et al., 2013). We used a generalization of the spatial mode decomposition that, different from dual regression, directly delineated networks showing the strongest drug-related changes in correlations, without prior selection of certain candidate networks (Friston & Büchel, 2004; Donner et al, 2013).

Both approaches provided converging evidence for the notion that catecholamines reduce the strength of fine-grained cofluctuation between diverse brain regions (including sensory- and motor-related networks). In addition, spatial mode decomposition revealed an atomoxetine-related shift from left to right-lateralized frontoparietal dominance in cofluctuation strength. Importantly, the predominant changes of correlation patterns we detected, all reflected quantitative changes in existing correlations, rather than a qualitative reconfiguration of network topology.

4.2 Materials and Methods

Design and MRI preprocessing. We reanalyzed data from van den Brink et al. (2016). This dataset comprised eyes open 'resting-state' (blank fixation) fMRI scans of 28 participants who received either placebo or atomoxetine (40 mg) on two separate sessions, scheduled one week apart. Atomoxetine is a selective NE transporter blocker that increases synaptic catecholamine levels (Bymaster et al., 2002; Devoto et al., 2004; Swanson et al., 2006; Koda et al., 2010). The study had a double-blind placebo-controlled crossover design, and was approved by the Leiden University Medical Ethics Committee. All participants gave written informed consent before the experiment, in accordance with the declaration of Helsinki.

Salivary markers of central catecholamine levels confirmed drug uptake (Warren et al., 2017). A full description of scan parameters and preprocessing details can be found in van den Brink et al. (2016). In brief, we applied the following preprocessing steps to the fMRI data (TR = 2.2 s; voxel size = 2.75 mm isotropic): realignment and

motion correction; B0 unwarping; high-pass filtering at 100 s; prewhitening; smoothing at 5 mm FWHM; coregistration of the functional scans with an anatomical T1 scan to 2 mm isotropic MNI space; artifact removal using FMRIB's ICA-based X-noiseifier (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014); and retrospective image correction to account for differences in heart and breath rate between the atomoxetine and placebo conditions (Glover et al., 2000). In the current article, we focus on the runs following atomoxetine / placebo ingestion.

Dual regression analysis. We estimated a set of independent components (ICs) that were representative of the combined set of resting-state runs (i.e., runs from all participants and both the atomoxetine and placebo conditions) by applying a spatial independent component analysis (ICA) to all temporally concatenated data using FSL's MELODIC. The number of ICs to be detected (51) was automatically estimated from the data. Each IC represented a statistical parametric map and corresponding time series of consistent spatio-temporal dynamics. Next, we spatially correlated each IC spatial map with the 10 intrinsic connectivity networks reported by Smith et al. (2009) and selected the ICs that showed the highest correlation coefficient. The selected components showed an average correlation coefficient of 0.48 (range: 0.28 - 0.70), which indicated that the ICs as expressed in our data corresponded relatively well to previously reported intrinsic connectivity networks (Smith et al., 2009).

The 10 selected ICs represented spatial maps of ICs that were reliably expressed across the combined set of resting-state runs. They were thus representative of group-level spatiotemporal dynamics, but did not necessarily represent spatiotemporal dynamics within individual runs. To produce a time series and a spatial map for the individual resting-state runs and for each IC, we used dual regression (Beckmann, 2009; Filippini et al., 2009). Figure 1 shows a schematic overview of this analysis
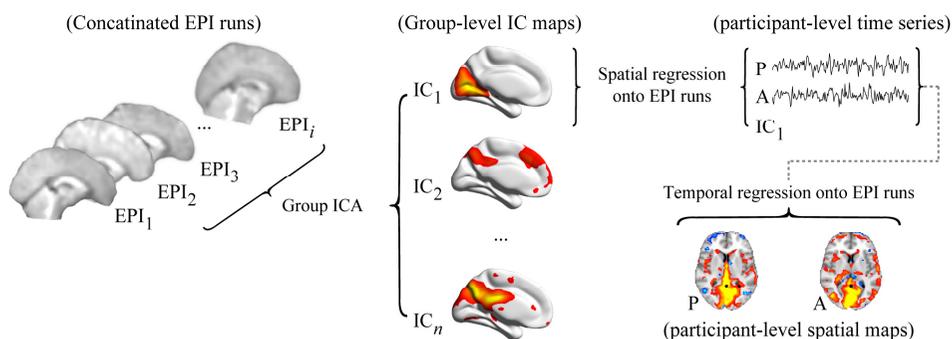
Figure 1. Schematic overview of the dual regression method. First, a group-level independent component analysis is run to produce spatial maps. A selection of these maps is subsequently regressed onto the individual runs to produce participant- and component-specific time series. Finally, these time series are used in temporal regression to produce participant- and component-specific spatial maps, which can then be compared between conditions. P: placebo; A: atomoxetine; EPI: echo planar imaging; IC: independent component

approach. In a first step, we used the group-level IC spatial maps in multiple spatial regression onto the individual runs. This produced a time series for each IC as expressed within the individual runs. Then, in a second step, we used the participant-level IC time series as temporal regressors to produce spatial maps of regression coefficients for each IC and each run. Thus, this two-stage regression approach resulted in a spatial map for each participant, condition, and IC, that indicated the degree of covariation between individual voxels and the IC time series.

Finally, we collected the IC spatial maps of the individual runs into single 4D files (one per condition, per IC). This allowed us to compare these spatial maps to zero across participants to examine which brain regions cofluctuated with the IC time series, and compare them between conditions to assess which voxels displayed changes in cofluctuation strength with the IC, using non-parametric permutation testing (10,000 iterations) as implemented in FSL's Randomise. The α level was set at 0.05, family-wise error (FWE) corrected for multiple comparisons using threshold-free cluster enhancement.

Brain parcellation and inter-regional covariance analysis. We extracted the fMRI time series of individual brain regions using the Automated Anatomical Labeling (AAL; Tzourio-Mazoyer et al., 2002) atlas, which contained 90 regions (cf. van den Brink et al., 2016). In addition, we used an alternate atlas that was based on a functional parcellation (Craddock et al., 2012). This atlas contained 140 individual brain regions. After averaging across voxels within each brain region (for each atlas separately), we Z-scored the multivariate time series (*M,* with dimensionality imaging volumes by brain regions) for each run *i* and then computed the group-averaged covariance matrices (C) for the placebo and atomoxetine conditions (subscript P and A, respectively) via the following:

$$(1) \quad C_P = N^{-1} \sum_{iP=1}^{N} \frac{M_{iP}^{T} \cdot M_{iP}}{nTR-1} \; , \; C_A = N^{-1} \sum_{iA=1}^{N} \frac{M_{iA}^{T} \cdot M_{iA}}{nTR-1}$$

where *nTR* is the number of volumes (211), N is the number of participants (24), and ^{T} denotes a matrix transposition. The matrices C_P and C_A represented the covariance between the BOLD time series of all brain regions, averaged across participants. Note that by variance normalizing (Z-scoring) the time series, the units of C (covariance) are equivalent to the Pearson correlation coefficient.

Singular value decomposition of covariance matrices. The dual regression analysis described at the beginning of this section relies on a linear decomposition of the data (ICA). An alternative linear decomposition, eigenvalue decomposition (SVD) as used in principal component analysis, can be extended to comparisons of correlation patterns between two conditions. Singular value decomposition (SVD) is a multivariate linear decomposition that identifies spatial modes of signal cofluctuations; each of these spatial modes can be conceptualized as a 'network' of correlated (or anti-correlated) brain regions (Mitra and Pesaran, 1999; Friston and Büchel, 2004; Donner et al., 2013). The decomposition can be generalized to extract spatial modes

that are more strongly expressed in one experimental condition than in the other, in other words: maximize the ratios of explained variance between conditions (Friston and Büchel, 2004; Donner and Nieuwenhuis, 2013).

Our first objective was to determine if SVD (specifically: principal component analysis) identified similar correlation patterns as ICA. We thus submitted the covariance matrices (C) of the placebo and atomoxetine conditions to SVD:

$$(2) \quad C = V \cdot \lambda \cdot V^{T}$$

where ^{T} denotes transposition, λ is an *n*-by-*n* matrix with eigenvalues on its diagonal, and V is an *n*-by-*n* matrix of corresponding eigenvectors in which rows are brain regions (*n* = 90) and columns define individual modes (*p*).
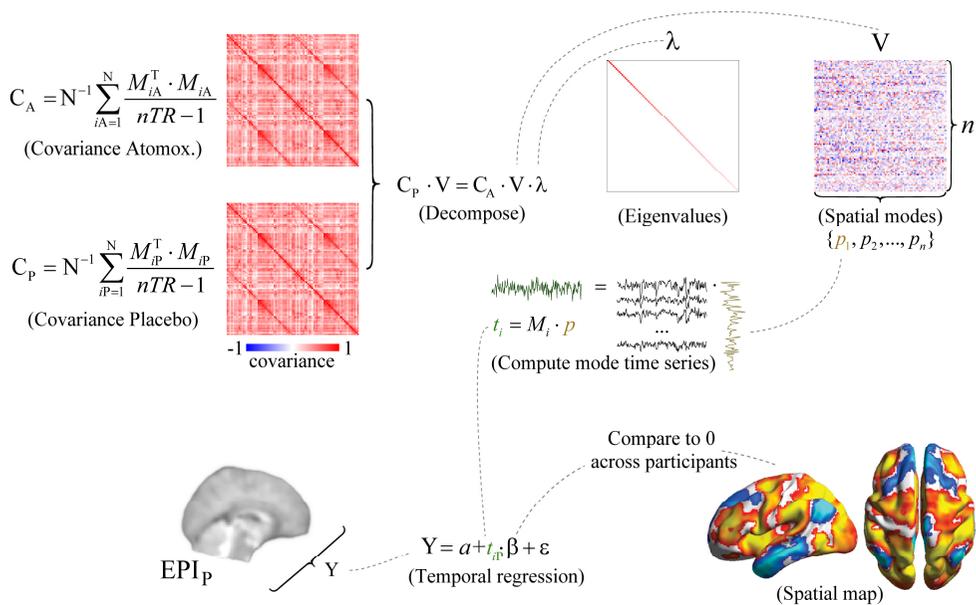
Figure 2. Schematic overview of the spatial mode decomposition method. The covariance matrices C_A and C_P are submitted to generalized eigenvalue decomposition to produce a matrix of eigenvalues (λ) and eigenvectors (V). The decomposition equation as given here delineated modes that were more strongly expressed in the placebo condition than in the atomoxetine condition. To identify modes that were more strongly expressed in the atomoxetine condition, the covariance matrices C_A and C_P were swapped. After decomposition, the participant-level time series (*t*) corresponding to each individual spatial mode (p) can be computed for each run *i* by projecting the mode onto the data (*M*). The number of brain regions in the parcellation scheme is denoted by *n*. A spatial map of brain regions that consistently covaried with the mode time series is computed by regressing the spatial mode time series for the atomoxetine (A) and placebo (P) conditions onto the voxel-level fMRI time series, and comparing the regression coefficients to zero across participants.

The overall sign of the elements in *p* is arbitrary but the sign of one element with respect to another indicates their relative co-variation, with equal signs indicating positive correlation and unequal signs indicating negative correlation.

For each run *i*, separately for the atomoxetine and placebo condition, we calculated participant-level time series *t* corresponding to each mode by projecting the mode onto the participant-level multivariate time series *M* via:

$$(3)\quad t_i = M_i \cdot p$$

The so-computed *t* described the time-varying strength of the expression of the spatial mode (functional network) in each individual participant's data. Next, we obtained voxel-level spatial maps for each mode and each run by regressing the vectors t_i onto the corresponding voxel-level BOLD data using multiple linear regresson. We then selected modes based on maximal spatial correlation with the 10 intrinsic connectivity networks reported by Smith et al. (2009), similar to the selection of ICA components described above. For the placebo condition, the average correlation coefficient was 0.41 (SD 0.12, min 0.16, max 0.56), and for the atomoxetine condition the average correlation coefficient was 0.40 (SD 0.12, min 0.15, max 0.53), indicating that SVD was able to identify networks of intrinsically cofluctuating activity reasonably well (Figure S1 and S2). Similar results were obtained with the Craddock atlas. Next, we describe the generalization of SVD to extract modes that are more strongly expressed in one condition relative to the other.

Generalized eigenvalue decomposition of covariance matrices. We used generalized eigenvalue decomposition to decompose the covariance matrices into spatial modes that maximized the ratio of explained variance in the placebo condition relative to the atomoxetine condition (Mitra and Pesaran, 1999; Friston and Büchel, 2004; Donner et al., 2013). Figure 2 shows a schematic overview of this analysis approach. We refer to previous work for experimental validation of generalized eigenvalue decomposition for use on fMRI data (Donner et al, 2013). For simplicity, we here refer to this method as 'spatial mode decomposition'. Using the 'eig' function in MATLAB 2012a, we decomposed the participant-averaged atomoxetine covariance matrix C_A and placebo covariance matrix C_P by solving the equation:

$$(4)\quad C_P \cdot V = C_A \cdot V \cdot \lambda$$

where λ is an *n*-by-*n* matrix with generalized eigenvalues on its diagonal, and V is an *n*-by-*n* matrix of corresponding eigenvectors in which rows are brain regions (*n* = 90 for the AAL atlas, and *n* = 140 for the Craddock atlas) and columns define individual modes (*p*). Here, *p* were spatial patterns that maximized the variance accounted for in one condition relative to the other (as measured by the corresponding λ_p). The above equation identified spatial modes that were more strongly expressed in the placebo condition than in the atomoxetine condition. To identify spatial modes that were more strongly expressed in the atomoxetine condition, the covariance matrices C_A and C_P were swapped. We arranged V and λ such that their first entries

corresponded to the modes that explained most variance. In other words, we sorted λ in descending order and then sorted V by λ.

We next calculated participant-level time series t corresponding to p for each individual run i via:

$$(5) \quad t_i = M_i \cdot p$$

Here, t_i was a vector with length 211 (the number of volumes), and M_i was a matrix of Z-scored fMRI time series from the run, with size 211 by n (volumes by brain regions).

To examine the spatial distribution of each mode, we used the corresponding time series in multiple temporal regression. Specifically, for each participant and condition separately, we regressed the mode time series onto the single-voxel time series from the corresponding run. This yielded one spatial map of regression coefficients per participant, condition, and mode. For each mode and for each condition, we could then compare the regression coefficients to zero using non-parametric permutation testing (10,000 iterations). The α level was set at 0.05, FWE-corrected for multiple comparisons using threshold-free cluster enhancement. The resulting statistical parametric maps indicated which voxels (if any) covaried with the mode time series consistently across participants, and were thus indicative of the spatial distribution of the modes.

Quantifying the across-subject consistency and reliability of spatial modes. The spatial modes were computed such they explained more variance in the group-average data, in the atomoxetine condition than in the placebo condition (or the converse). We aimed to quantify, in a cross-validated fashion, how consistently the fluctuation strength of these group-average spatial modes distinguished between conditions within individual subjects. The fluctuation amplitude s_i corresponding to each mode's time series in each individual run from each participant quantified the amount of variance that the mode explained in the data, and was calculated via:

$$(6) \quad s_i = t_i^{\mathrm{T}} \cdot t_i$$

Note that this is equivalent to:

$$(7) \quad p^{\mathrm{T}} \cdot M_i^{\mathrm{T}} \cdot M_i \cdot p = p^{\mathrm{T}} \cdot C_i \cdot p = s_i$$

We then divided s_i by the sum of eigenvalues (λ) to convert it to units of percentage variance explained. In contrast to the eigenvalues, which capture the group-level mode's ratio of explained variance between conditions, s_i captured the amount of variance that the mode captured in the condition-specific runs at the individual participant-level. For cross-validation, we defined modes (using eq. 4) based on the group-average covariance matrices C_A and C_P that were generated from the first half of volumes in M_i (using eq. 1). Then, each mode was projected onto the (independent) remaining half of volumes in M_i as described above (eq. 5) and their corresponding fluctuation amplitudes were calculated (via eq. 6). We then used the second half of volumes to define the modes and projected them onto the first half, and averaged the two values of s_i. The percentage variance explained by each mode could then be

compared between conditions with non-parametric permutation testing (10,000 iterations).

We used receiver operating characteristic (ROC) analysis to quantify the reliability of the spatial modes in discriminating between experimental conditions, at the level of short segments (25% of volumes, ~114 s) of individual fMRI runs. ROC analysis performs more accurately with densely populated distributions of measurements. Thus, we defined spatial modes based on the group average covariance matrices calculated from a smaller subset of volumes (25%), as described above (using eq. 1 and eq. 4). We subdivided the remainder of volumes into 20 equal-sized bins, and computed s_i for each of them. We cross-validated the fluctuation amplitude calculation by computing modes and projecting them onto the remaining data four times such that eventually all data were used to define the modes. This yielded four distributions of s_i per condition and participant that were submitted to ROC analysis, resulting in four ROC-curves per participant. We calculated the area under the ROC-curve (AUC) and averaged the resulting AUC values across the four ROC-curves of each participant. This AUC value could then be interpreted as the probability with which we could predict the condition from the mode's fluctuation strength in a given data segment. The AUC values were tested for significance by comparing them to chance level (0.5) using non-parametric permutation testing (10,000 iterations). In order to exclude the possibility that the significance of the ROC results depended on the number (25%) of volumes on which the mode was defined, we repeated the ROC analyses for modes defined on ~14%, 20%, and ~33% of the data, and found identical results in terms of direction and significance.

Correlation between mode spatial maps and independent components. To determine if the mode spatial maps depended on the parcellation scheme, we used spatial correlation. Specifically, for each individual participant and condition, we correlated the (unthresholded) spatial maps of regression coefficients of the modes that were generated with the AAL atlas, and those that were generated with the Craddock atlas. We then compared the distribution of Fisher-transformed correlation coefficients to zero using a two-tailed *t*-test. Similarly, we characterized the correspondence in mode spatial maps between the individual conditions by correlating the unthresholded spatial maps at the individual participant level, and comparing the resulting distribution of Fisher-transformed correlation coefficients to zero using a two-tailed *t*-test.

To characterize correspondence between the mode spatial maps and well-characterized intrinsic connectivity networks, we first created a mode spatial map by temporally concatenating the mode time-series of the atomoxetine and placebo conditions, and regressing this concatenated time series onto the temporally concatenated BOLD time-series data for each participant. The purpose of this concatenation procedure was to create spatial maps that were independent of drug condition, similar to ICA components that were identified in temporally concatenated EPI data. We could then correlate these condition-invariant unthresholded participant-

level mode spatial maps with the IC spatial maps that were selected for dual regression analysis, and compare the distribution of Fisher-transformed regression coefficients to zero using a two-tailed *t*-test. In all cases where we report average correlation coefficients, we applied Fisher's *r*-to-Z transform prior to averaging, and subsequently applied the Z-to-*r* transform.

4.3 Results

Our first aim was to characterize atomoxetine-induced changes in fine-grained (voxel-level) cofluctuation strength with a set of 10 well-characterized intrinsic connectivity networks (Smith et al., 2009) using conventional methods for the analysis of pharmacological resting-state fMRI: dual regression (Figure 1) (Beckmann, 2009; Filippini et al., 2009). Second, we report the results of an alternative analysis approach that is targeted at finding spatial patterns ("modes") of correlated activity that maximize the ratio of explained variance between conditions in a fully data-driven manner (Figure 2). Thus, instead of testing if any of an *a priori* selection of networks showed
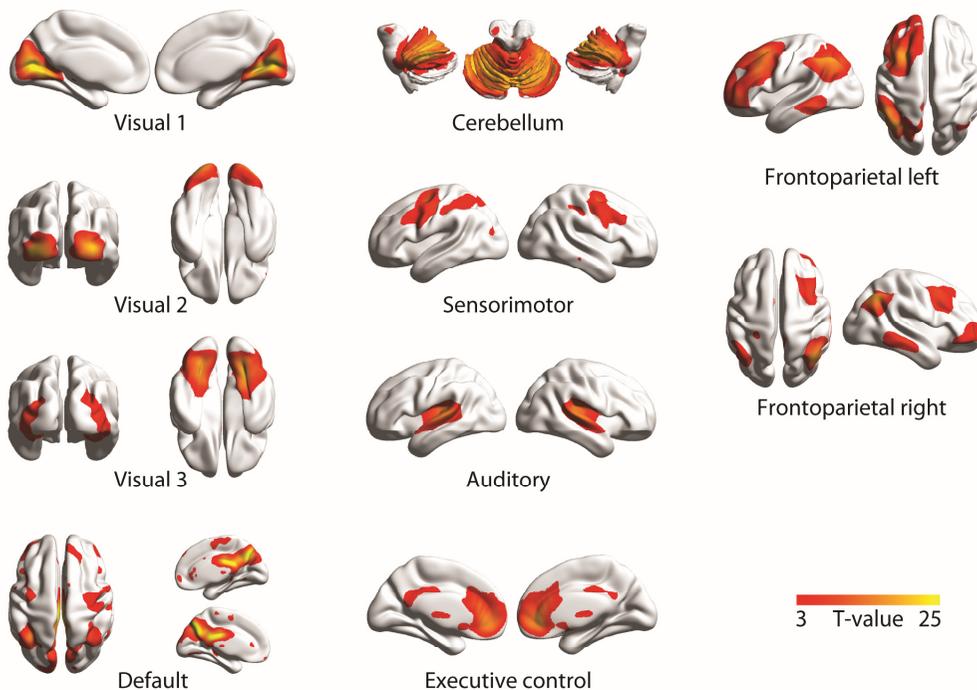
Figure 3. Spatial maps of the independent components that were selected for dual regression analysis. Components were selected based on spatial correlation with the 10 canonical resting-state networks presented by Smith et al. (2009).

drug-induced changes in the cofluctuation strength (as with dual regression), spatial mode decomposition directly yielded the networks that exhibited drug-induced changes incofluctuations, in terms of cofluctuation strength or spatial pattern, or both. In contrast to singular value decomposition of condition-level data (see Materials and Methods) or linear decomposition of the data using ICA, spatial mode decomposition directly reveals those patterns that cofluctuate more/less in one condition than in the other. Moreover, this analysis allows us to characterize to what extent atomoxetine-
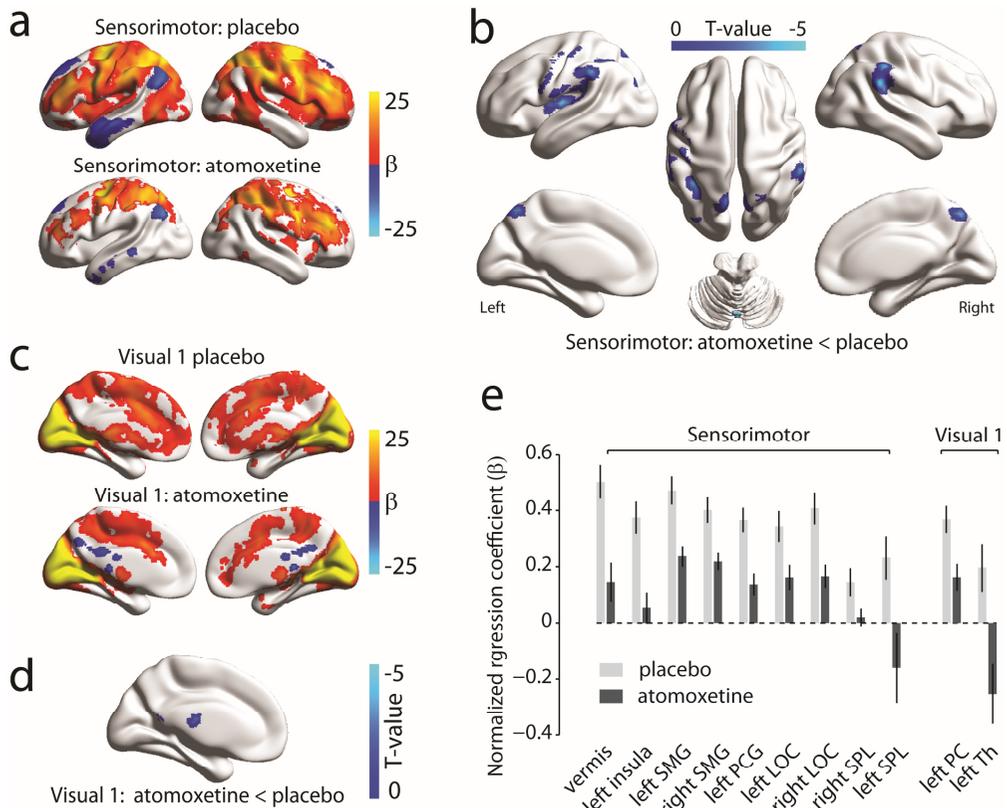
Figure 4. Results of the dual regression analysis. **a)** Brain areas that significantly ($p < 0.05$, FWE-corrected) cofluctuated with the sensorimotor network. **b)** Atomoxetine-induced changes in cofluctuation strength with the sensorimotor network. **c)** Brain areas that significantly cofluctuated with the visual 1 network. **d)** Atomoxetine-induced changes in cofluctuation strength with the visual 1 network. Blue colors indicate reduced cofluctuation strength following atomoxetine compared to placebo. **e)** Cofluctuation strength (range-normalized across participants to between -1 and 1 for illustrative purposes only) for each condition and each significant ($p < 0.05$, FWE-corrected) cluster of atomoxetine-induced changes in cofluctuation strenth. Error bars show the SEM. Abbreviations: PC: precuneous cortex; Th: thalamus; SMG: supramarginal gyrus; PCG: precentral gyrus; LOC: lateral occipital cortex; SPL: superior parietal lobule.

related modulations of covariance reflect reconfigurations in the topological structure of cofluctuations by directly comparing the mode's spatial structure between the atomoxetine and placebo conditions.

Results dual regression analysis. We first computed group-level ICA spatial maps and then submitted a selection to dual regression analysis (Figure 1). The spatial maps of the included components are shown in Figure 3.

The sensorimotor network and first visual network both showed significant ($p <$ 0.05, FWE-corrected) atomoxetine-induced changes in cofluctuation strength (Figure 4a-d). In all significant clusters, atomoxetine reduced the strength of cofluctuation, consistent with our earlier findings obtained at coarser levels of spatial granularity (van den Brink et al., 2016). All clusters that showed a significant atomoxetine-related reduction in cofluctuation strength also covaried positively with the component time series in the placebo condition (Figure S3), and were thus functionally linked to the networks. In most cases, atomoxetine moved cofluctuations from positive towards zero, and in some cases cofluctuations reversed polarity, at least numerically (Figure 4e; Figure S3). Cofluctuations with the component time series in the clusters that numerically reversed polarity, however, were not significantly negative in the atomoxetine condition (Figure S4). Thus, the primary effect of atomoxetine on cofluctuation strength was to move positive cofluctuations towards zero. The MNI coordinates and peak T-statistics of all significant clusters are summarized in Table 1. Together, these results suggest that atomoxetine attenuated voxel-level cofluctuation strength between brain regions that cofluctuated positively with sensory- and motor-related networks, and the fluctuations of those networks.

Table 1. Clusters that showed an atomoxetine-induced change in cofluctuation strength with resting-state networks.

Component	Location	MNI Coordinates (x y z)	Peak T-statistic
Visual 1	Left PC	-4 -56 12	-4.91
	Left Th	0 -22 12	-4.21
Sensorimotor	Vermis	0 -72 -14	-6.04
	Left insula	-40 -12 8	-4.65
	Left SMG	-66 -34 34	-5.71
	Right SMG	58 -44 28	-4.57
	Left PCG	-50 -10 56	-4.85
	Left LOC	-34 -82 26	-4.36
	Right LOC	18 -66 72	-4.73
	Right SPL	38 -54 68	-3.30
	Left SPL	-34 -42 70	-3.43

Cluster locations were assessed using the Harvard-Oxford structural atlas. Peak MNI coordinates are indicated in mm. Abbreviations: PC: precuneous cortex; Th: thalamus; SMG: supramarginal gyrus; PCG: precentral gyrus; LOC: lateral occipital cortex; SPL: superior parietal lobule.

Spatial modes that are less strongly expressed in the atomoxetine condition relative to placebo. The above reported atomoxetine-related reductions in cofluctuation strength with sensory- and motor-related networks resulted from dual regression analysis. In the following, we report the results of an alternative analysis approach that directly reveals those patterns that cofluctuate more/less in one condition than in the other, and that allows us to characterize to what extent atomoxetine-related modulations of covariance reflect reconfigurations in the topological structure of cofluctuations. Given that dual regression analysis identified only atomoxetine-reductions in the strength of cofluctuations, we first focus on spatial modes that reflected an atomoxetine-related reduction in cofluctuations.

The eigenvalues of the modes that were less strongly expressed in the atomoxetine condition are shown in Figure 5a. We focused on mode number 1 because it had the largest eigenvalue and thus accounted for most variance in the data, and because mode orthogonality can obscure the interpretation of modes with higher ranks (c.f. Donner et al., 2013).

This first spatial mode robustly differed in its fluctuation strength (i.e., variance explained) between the atomoxetine and placebo conditions. We first tested, using cross-validation, if the first spatial mode consistently explained less variance in the atomoxetine condition than in the placebo condition across subjects: we computed the mode based on covariance in each half of the volumes in each participant's runs, projected the mode onto the remaining half of the volumes, and calculated the mode's proportion of explained variance in each condition (see Materials and Methods). Indeed, the first spatial mode accounted for significantly less variance in the atomoxetine condition than in the placebo condition, for both parcellation schemes (AAL: $p = 0.003$; Craddock: $p < 0.001$; Figure 5b). Further, ROC analysis showed that even at the level of short individual data segments (~114 s), the first spatial mode's fluctuation strength reliably discriminated between drug and placebo conditions, with AUC-values larger than 0.6 for both parcellation schemes (Figure 5c). This indicates that the spatial mode identified by our analysis reflected a robust effect of the pharmacological intervention on brain-wide intrinsic correlations.

We next compared the spatial distributions of the expressions of the first spatial mode, between different parcellation schemes and with the spatial maps obtained from the dual regression analysis. The unthresholded spatial map of mode 1 (reflecting voxel-level covariation with the mode's time series) is shown in Figure 5d, separately per condition and for modes that were generated using the AAL atlas, and for modes that were generated using the Craddock atlas. Figure S5 shows thresholded ($p < 0.05$, FWE-corrected) mode spatial maps. Despite using parcellation schemes that differed both in the number of brain regions and in the way the brain regions were defined (anatomical parcellation and functional clustering, respectively), the mode spatial maps generated with the two atlases corresponded robustly across
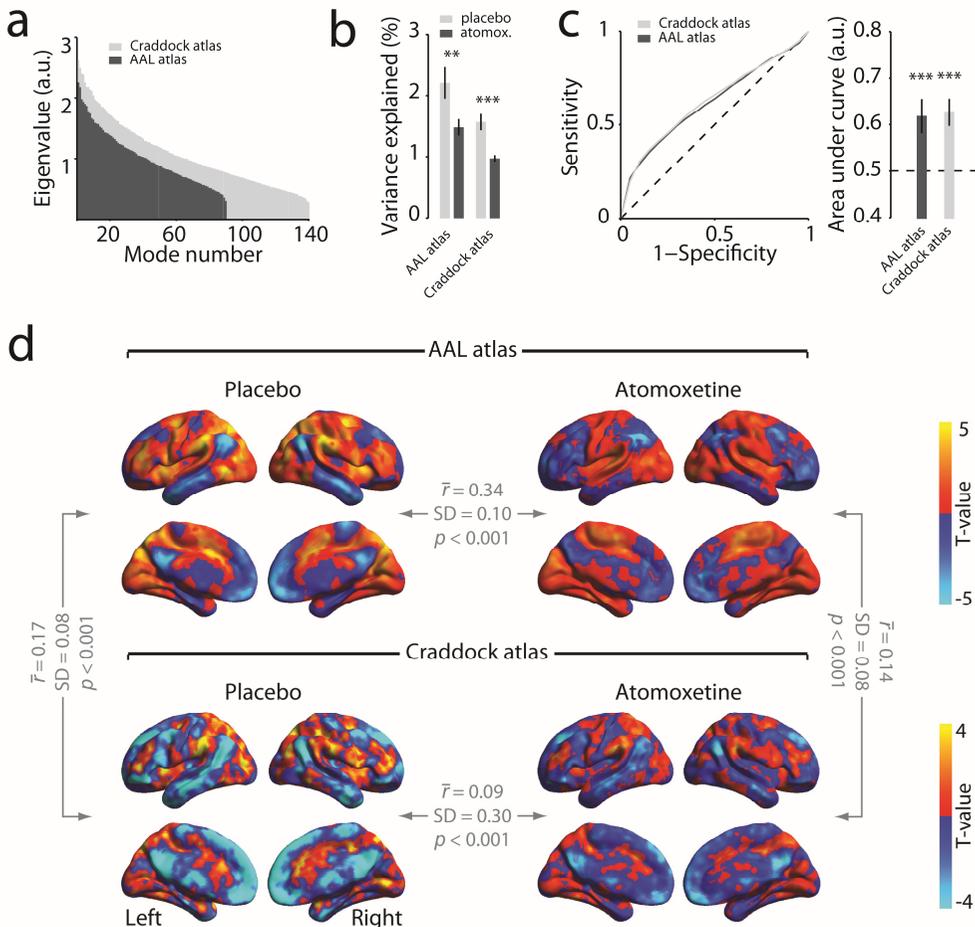
Figure 5. Spatial modes that are less strongly expressed in the atomoxetine condition. **a)** Eigenvalues of all modes. **b)** A comparison between conditions of the percentage of variance explained by the first mode. **c)** ROC curves to distinguish conditions based on the fluctuation amplitude of the first mode. **d)** Spatial map of the first mode. Colored regions show covariation with the mode time series. Error bars show the SEM. **: $p < 0.01$; ***: $p < 0.001$. The r values indicate the average correlation coefficient across participants.

participants (placebo: $t(23) = 10.43$, $p < 0.001$; atomoxetine: $t(23) = 9.54$, $p < 0.001$; Figure 5d).

To determine if the mode corresponded to any of the intrinsic connectivity networks that were used for dual regression analysis, we correlated the mode spatial map with the ICA component spatial maps at the individual participant level. The spatial map of mode 1 that was generated with the AAL atlas correlated most strongly with the left-lateralized frontoparietal ICA component (mean r = -0.15, SD 0.05; $t(23)$ = -16.33, $p < 0.001$). The spatial map of mode 1 that was generated with the Craddock atlas also correlated significantly across participants with the left-lateralized

frontoparietal component (mean r = -0.07, SD 0.04; $t(23)$ = -8.68, p < 0.001). Moreover, for both atlases the spatial map of mode 1 correlated significantly with the sensorimotor component (AAL: mean r = 0.13, SD 0.04; $t(23)$ = 17.41, p < 0.001; Craddock: mean r = 0.07, SD 0.03; $t(23)$ = 14.36, p < 0.001), suggesting that atomoxetine reduced the strength of cofluctuations in a network that resembled the ICA-identified sensorimotor network (Figure S6). If this is indeed the case, then the regions that showed an atomoxetine-related reduction in cofluctuations in the dual regression analysis (Figure 4) should show a similar sign in the mode spatial map (i.e. be part of the same cofluctuating network). We therefore masked the thresholded spatial map of mode 1 in the placebo condition (Figure S5) with the significant clusters in Figure 4. All clusters showed the same sign (Figure S7), indicating that the spatial mode reflected a reduction in cofluctuation strength across brain regions that showed similar reductions in the dual regression analyses.

Spatial correlation also enabled us to examine if the mode reflected a reconfiguration of the spatial structure of cofluctuations or if it reflected a quantitative change in strength that left the structure of cofluctuations intact. To this end, we operationally defined reconfiguration as a change in mode topology, implying a spatial mode that was only expressed in one condition, but not in the other. By contrast, quantitative changes would entail the spatial mode to be expressed in both conditions, only to a different degree. Note that both scenarios might lead to a robust spatial mode maximizing the ratio between variance accounted for in both conditions.

We correlated the mode spatial map of the placebo condition with that in the atomoxetine condition (Figure 5d). Inconsistent with the notion of an atomoxetine-related reconfiguration of cofluctuation structure, the mode spatial map correlated between conditions for both atlases (AAL: $t(23)$ = 15.57, p < 0.001; Craddock: $t(23)$ = 14.89, p < 0.001). In other words, the spatial distribution of the expression of the first spatial mode that most discriminated between conditions, was in fact similar between conditions. This indicates that the predominant effect of atomoxetine was a quantitative reduction of the strength of cofluctuations rather than a topological reconfiguration of intrinsic functional connectivity networks (see Discussion).

Spatial modes that are more strongly expressed in the atomoxetine condition relative to placebo. Thus far we have shown, using dual regression analysis, that atomoxetine reduced cofluctuation strength between distributed brain regions (including sensory- and motor-related networks). Spatial mode decomposition revealed similar reductions in cofluctuation strength, and further indicated that these changes in cofluctuation strength left the spatial structure of cofluctuations intact. We now turn to spatial modes that reflected an atomoxetine-related increase in cofluctuations (see Materials and Methods), which may have not been identified by dual regression analysis.

The eigenvalues of the modes that were more strongly expressed in the atomoxetine condition are shown in Figure 6a. Again, we selected mode 1 because it accounted for most variance in the data. Similar to the above reported analysis of
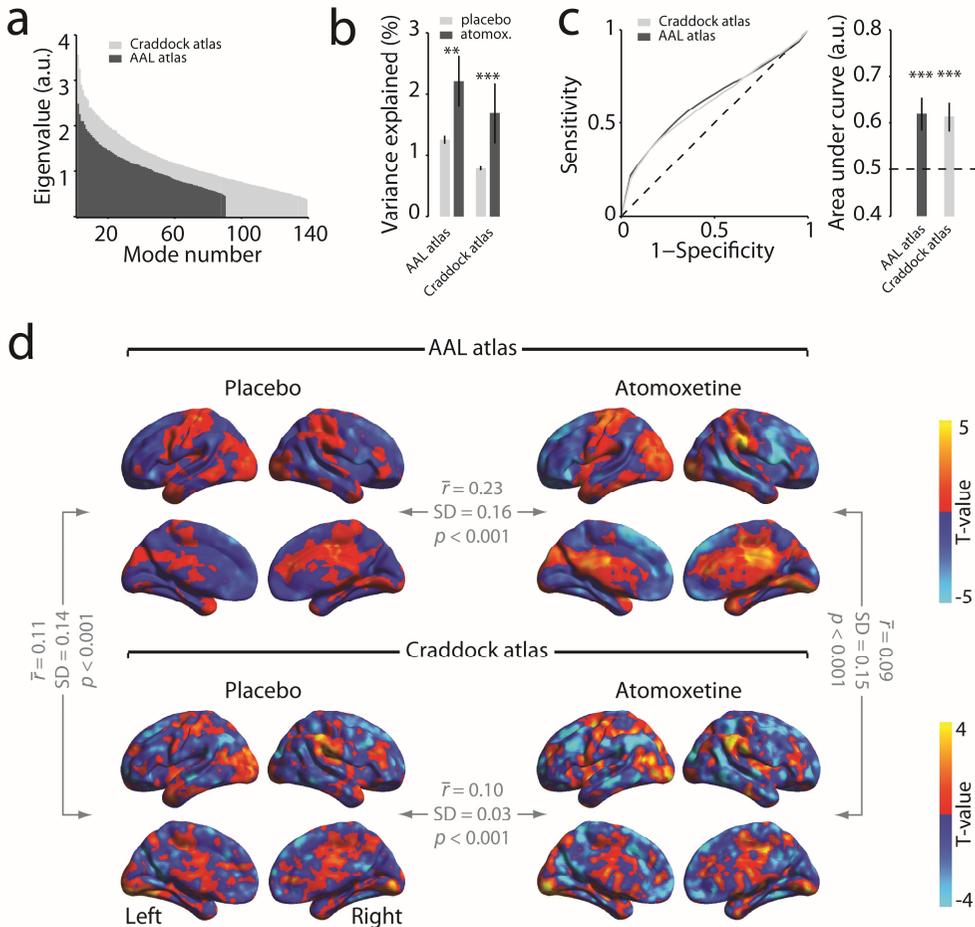
Figure 6. Spatial modes that are more strongly expressed in the atomoxetine condition. **a)** Eigenvalues of all modes. **b)** A comparison between conditions of the percentage of variance explained by the first mode. **c)** ROC curves to distinguish conditions based on the fluctuation amplitude of the first mode. **d)** Spatial map of the first mode. Colored regions show covariation with the mode time series. Error bars show the SEM. **: $p < 0.01$; ***: $p < 0.001$. The r values indicate the average correlation coefficient across participants.

mode variance, we computed the mode based on covariance in each half of the volumes, and projected it onto the remaining half. For both atlases, the mode explained significantly more variance in the atomoxetine condition than in the placebo condition (AAL: $p = 0.002$; Craddock: $p < 0.001$; Figure 6b), and consistently throughout the resting-state runs (ROC values >0.6; $p < 0.001$; Figure 6c). Thus, the mode reflected a pattern of brain regions in which activity cofluctuated more strongly following atomoxetine than following placebo.

Figure 6d shows the (unthresholded) spatial map of mode 1, separately per condition, for modes that were generated using the AAL atlas, and for modes that

were generated using the Craddock atlas. Figure S8 shows thresholded ($p < 0.05$, FWE-corrected) mode spatial maps. Again, the spatial map of the modes generated with the two atlases corresponded robustly across participants (placebo: $t(23) = 3.96$, $p < 0.001$; atomoxetine: $t(23) = 3.98$, $p < 0.001$).

The spatial map of mode 1 correlated most strongly with the right-lateralized frontoparietal component (AAL atlas: mean $r = -0.05$, SD 0.03; $t(23) = -7.98$, $p < 0.001$; Craddock atlas: mean $r = -0.09$, SD 0.03; $t(23) = -14.44$, $p < 0.001$). Together, these results suggest that atomoxetine increased the strength of cofluctuations in a distributed network that resembled the right-lateralized frontoparietal network.

Next, we again examined if the mode reflected a change in cofluctuation network structure, or if it reflected a modulation of cofluctuation strength alone. We thus correlated the mode spatial map in the atomoxetine condition and in the placebo condition, and found robust correlations (AAL atlas: $t(23) = 6.93$, $p < 0.001$; Craddock atlas: $t(23) = 14.89$, $p < 0.001$). Thus, the spatial structure of the first mode was similar between conditions, and therefore the most prominent atomoxetine-related increases in cofluctuations can be interpreted as a quantitative increase in the strength of those cofluctuations rather than a more profound reconfiguration of network topology.

In sum, dual regression analysis and spatial mode decomposition converge on the conclusion that atomoxetine weakens cofluctuation strength between distributed brain regions, including sensory- and motor-related networks. In addition, spatial mode decomposition revealed a shift from left- to right-lateralized frontoparietal network dominance, as confirmed by a significant interaction in the strength of correlation between mode polarity (atomoxetine-induced increase versus decrease) and component (frontoparietal left versus right) (repeated-measures ANOVA; AAL: $F(1,23) = 163.14$, $p < 0.001$; Craddock: $F(1,23) = 56.15$, $p < 0.001$). Lastly, spatial mode analysis revealed that the predominant effect of atomoxetine was the quantitative scaling of preexisting cofluctuation patterns whereby the overall spatial structure of these cofluctuating networks was left intact.

4.4 Discussion

In the present study, we examined the effect of the selective NE transporter blocker atomoxetine on the fine-grained spatial structure of resting-state fMRI cofluctuations using dual regression and spatial mode decomposition. First, dual regression analysis revealed that atomoxetine reduced cofluctuation strength between a distributed set of brain regions that included sensory- and motor-related networks. Second, spatial mode decomposition provided converging evidence for such a reduction in sensory- and motor related coupling. Third, spatial mode decomposition revealed an atomoxetine-related shift in the dominance from left-lateralized to right-lateralized frontoparietal network activity. Importantly, spatial mode decomposition indicated that the most prominent atomoxetine-related changes in

cofluctuations did not alter the topology of the networks in which these changes occurred, but instead reflected quantitative modulations within these networks that left the overall cofluctuation structure intact.

The study of small neural circuits has revealed dynamical reconfigurations of functional networks through neuromodulators, including catecholamines (Marder, 2012; Bargmann and Marder, 2013; Marder et al., 2014). Yet, our results show that the total landscape of cofluctuation changes is dominated to a greater extent by quantitative catecholamine-related changes (i.e. mode 1 accounted for most variance relative to other modes, and it reflected quantitative changes). It is worth noting that more subtle catecholamine-related reconfigurations may have occurred, but were not detected by our current analyses (e.g. may have been captured by modes that accounted for less variance, which we did not examine). Additionally, our findings leave open the possibility that the rapid and transient (i.e. phasic) release of catecholamines has a more profound influence on the topological organization of intrinsic cofluctuations, given that such phasic catecholamine release can have qualitatively different effects on neural conductance properties (Rodgers et al., 2011b; Rodgers et al., 2011a) and behavioral performance (de Gee et al., 2017) than changes in tonic levels. Moreover, the dynamical structure of time-varying changes in network topology may be more susceptible to influence by neuromodulatory tone than the stationary network topology, given that such time-varying topological changes have been reported to covary with behavioral performance and pupillary indices of neuromodulation (Shine et al., 2016).

Our findings are broadly consistent with an earlier study (Guedj et al., 2016) that examined the effect of atomoxetine on resting-state cofluctuations in rhesus macaques using dual regression, and found widespread atomoxetine-related cofluctuation reductions within and between networks, including the somatomotor, somatosensory, (peripheral) visual, and a bilateral frontoparietal network. Interestingly, the authors report an atomoxetine-related reduction in cofluctuations between the somatomotor network and bilateral clusters that strongly resemble the bilateral clusters in the supramarginal gyrus reported here (Figure 4, Table 1), suggesting that the supramarginal gyrus is a particularly prominent target region of catecholaminergic neuromodulation. Moreover, the general correspondence between the findings reported by Guedj et al. (2016) and those reported here suggests that the modulation of cofluctuating networks, mediated by catecholamines, is a mechanism that occurs consistently across species. However, in contrast to the interpretation by Guedj et al. (2016) our findings suggest that these altered cofluctuation dynamics may reflect quantitative changes rather than broad network reconfigurations.

The finding that atomoxetine reduced cofluctuation strength with the sensorimotor network is noteworthy in light of the therapeutic profile of atomoxetine. Atomoxetine has been shown to improve inhibitory motor control in rats (Robinson et al., 2008), patients with ADHD (Chamberlain et al., 2007), and healthy humans (Chamberlain et al., 2006b). Our findings suggest that these response inhibition-enhancing effects of

atomoxetine may be the result of catecholaminergic action within the sensorimotor system. Nevertheless, the potential link between the here reported reduction in cofluctuation strength with the sensorimotor network and the response inhibition-enhancing effects of atomoxetine awaits further investigation.

Furthermore, we found that atomoxetine caused a shift in the dominance from left- to right-lateralized frontoparietal network dominance. Frontoparietal regions in the right hemisphere have been implicated in attentional reorientation and the regulation of goal-directed stimulus selection (Corbetta and Shulman, 2002; Corbetta et al., 2008; Thiebaut de Schotten et al., 2011). Interestingly, right-lateralized frontoparietal regions have also been suggested to be particularly susceptible to noradrenergic influences (Corbetta and Shulman, 2002; Corbetta et al., 2008), and atomoxetine has been reported to improve the precision of neural representations of stimuli (Warren et al., 2016). It is tempting to speculate that the here observed atomoxetine-related shift from left- to right-lateralized frontoparietal network dominance may indicate a shift towards goal-oriented stimulus processing. While our participants were not engaged in a task (other than active fixation), this speculation provides an interesting avenue for future research.

Our findings show the utility of spatial mode decomposition for the analysis of pharmacological resting-state fMRI data. One of its primary advantages over dual regression analysis is that it does not require an *a priori* selection of functional networks, but instead automatically yields the networks (spatial modes) that show the strongest drug-related effects. Thus, it reduces the chances of overlooking prominent drug-related changes in inter-regional cofluctuations, as evidenced by the atomoxetine-related increases in covariance that were identified by spatial mode decomposition, but not by dual regression. Moreover, spatial mode decomposition is computationally inexpensive when used in combination with an anatomical atlas, as we have done here. We should note, however, that even though our results demonstrate some robustness of the method to the particular parcellation scheme, it is not a certainty that the resulting networks will generalize to other parcellation schemes, in particular those of radically different densities. In addition, we only examined the first modes because they explained the largest amount of variance in the data, but modes with higher rank numbers may contain information regarding relevant changes in connectivity as well. Examining these, however, would require additional statistical corrections that could increase the false negative rate. Moreover, the interpretability of modes with higher ranks may be hindered by mode orthogonality. Lastly, the decomposition can only be used to compare two separate conditions (or groups), which limits its applicability in complex (e.g., longitudinal) study designs. Nevertheless, spatial mode decomposition offers a thorough characterization of drug-related changes in the structure of cofluctuating activity.

In sum, we have shown that dual regression and spatial mode decomposition converge on the conclusion that catecholamines reduce cofluctuation strength within and between distributed systems, including sensory- and motor-related networks. In

addition, spatial mode decomposition revealed an atomoxetine-related shift from left to right-lateralized frontoparietal network dominance. Importantly, however, these quantitative changes left the overall spatial structure of cofluctuations intact, suggesting that the predominant effect of increased synaptic catecholamine levels was to quantitatively scale cofluctuations in preexisting networks. Lastly, our findings lend support to the notion that catecholamines modulate dynamic changes in the strength of intrinsic inter-regional coupling, possibly to coordinate flexible modulations of network interactions to facilitate goal-directed behavior.

4.5 Supplementary Materials

Ruling out confounding artifacts in the global signal. Recent findings have suggested that the global MRI signal may contain artifacts that are related to various non-neural sources, and these artifacts are not effectively removed by standard preprocessing techniques (Power et al., 2017). While the independent components that were used for dual regression analysis by definition do not contain such global artifacts (due to the spatial independence of components), these artifacts may have caused spurious differences between conditions in the structure of inter-regional covariance. We therefore applied global signal (the mean of all regional time series) regression to the regional BOLD time series prior to computing covariance matrices, and repeated our key spatial mode decomposition analyses.

For the decomposition placebo < atomoxetine, the percentage variance explained of mode 1 differed between conditions and in the expected direction (AAL: $t(23)$ = 4.45, $p < 0.001$, area under ROC curve = 0.64, $t(23)$ = 6.88, $p < 0.001$; Craddock: $t(23)$ = 4.55, $p < 0.001$, area under ROC curve = 0.69, $t(23)$ = 7.54, $p < 0.001$). Similarly, for the decomposition placebo > atomoxetine the percentage variance explained of mode 1 also differed between conditions and in the expected direction (AAL: $t(23)$ = -5.15, $p < 0.001$, area under ROC curve = 0.63, $t(23)$ = 8.97, $p < 0.001$; Craddock: $t(23)$ = -6.23, $p < 0.001$, area under ROC curve = 0.63 $t(23)$ = 7.06, $p < 0.001$). Moreover, the spatial distribution of mode 1 computed without global signal regression correlated significantly with the spatial distribution of mode 1 computed on global signal-removed time series, for both atlases and for the decomposition in both directions (all p values < 0.001). Thus, the findings as presented in the main text were unlikely to be driven by spurious differences between conditions relating to artifacts in the global signal.
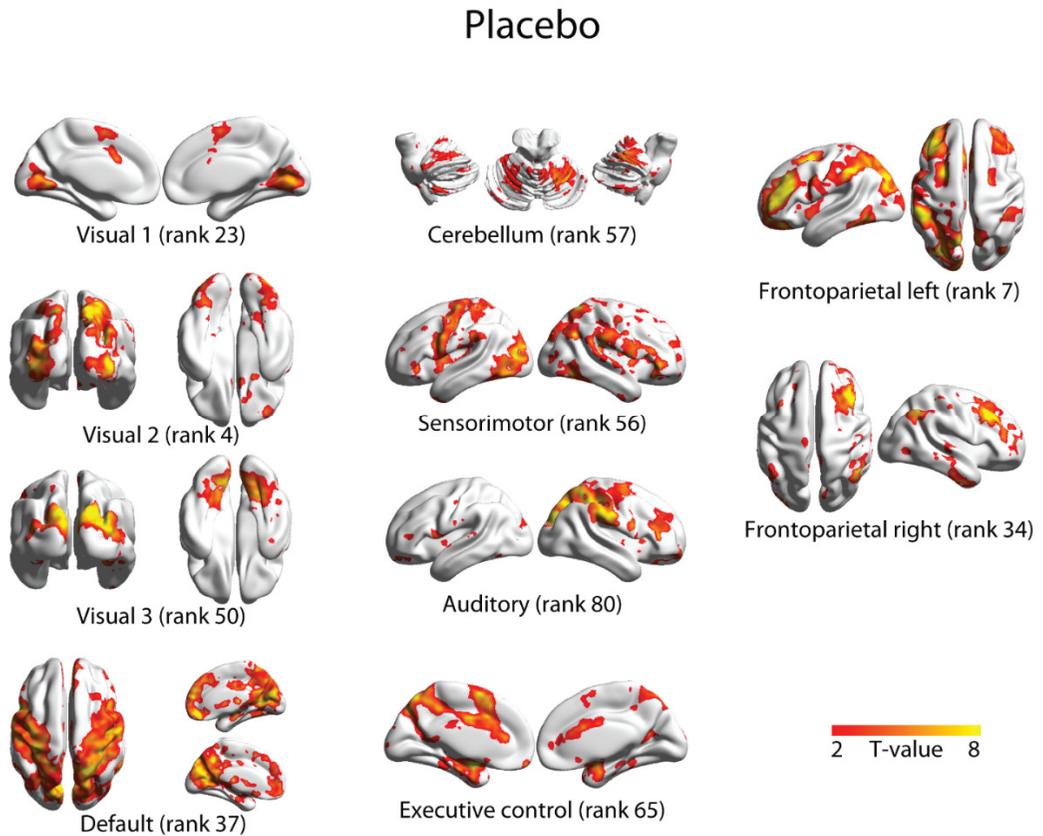
Placebo

Figure S1. Modes resulting from singular value decomposition of AAL atlas covariance in the placebo condition. The modes were selected based on maximal spatial correlation with the independent component topographies presented by Smith et al. (2009). The rank number indicates the relative proportion of explained variance of each mode, where lower rank numbers account for relatively more variance in the data than high rank numbers.

Atomoxetine

Visual 1 (rank 23)

Cerebellum (rank 78)

Frontoparietal left (rank 35)

Visual 2 (rank 4)

Sensorimotor (rank 56)

Frontoparietal right (rank 9)

Visual 3 (rank 50)

Auditory (rank 30)

2   T-value   8

Default (rank 37)

Executive control (rank 65)
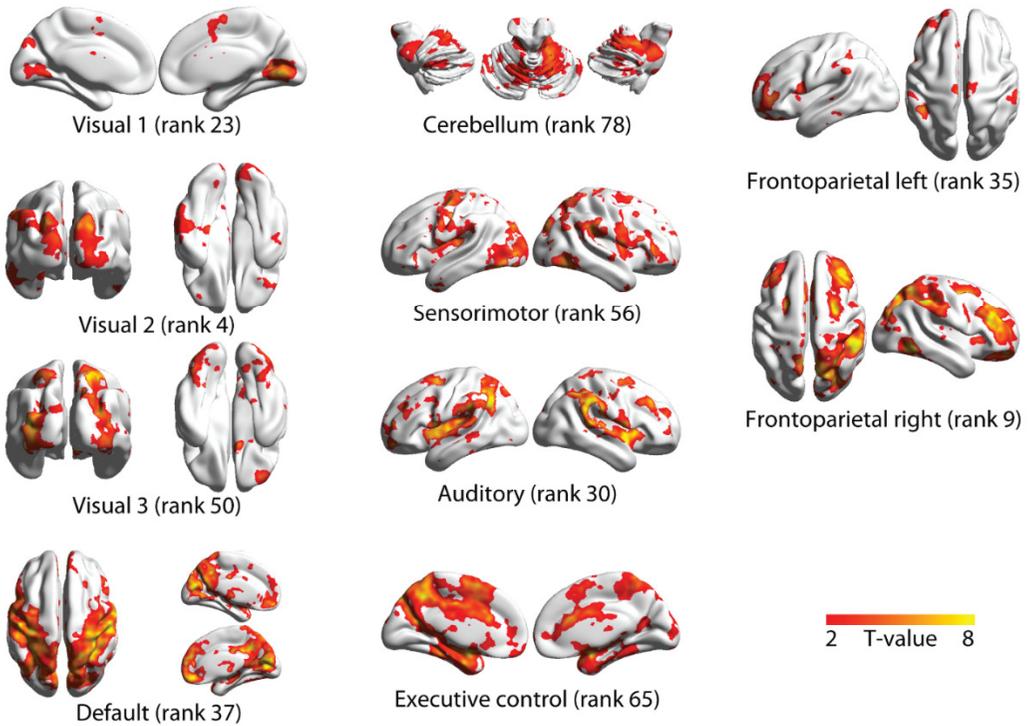
Figure S2. Modes resulting from singular value decomposition of AAL atlas covariance in the atomoxetine condition. The modes were selected based on maximal spatial correlation with the independent component topographies presented by Smith et al. (2009). The rank number indicates the relative proportion of explained variance of each mode, where lower rank numbers account for relatively more variance in the data than high rank numbers.
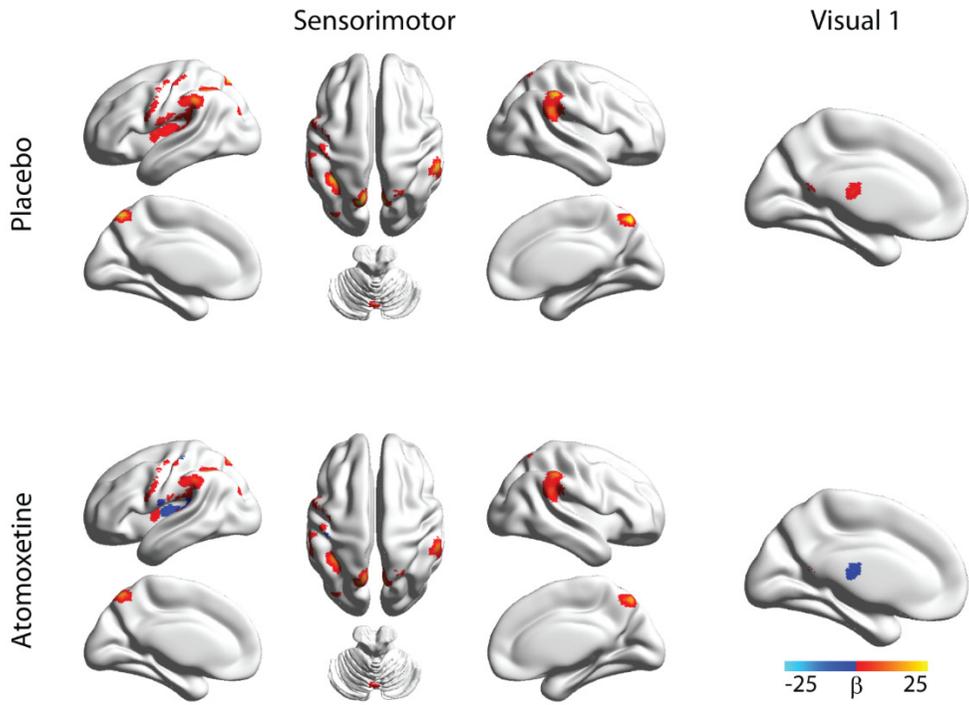
Figure S3. Covariation with component time-series in the individual conditions, only for clusters that showed a significant ($p < 0.05$, FWE-corrected) atomoxetine-induced reduction in coupling.
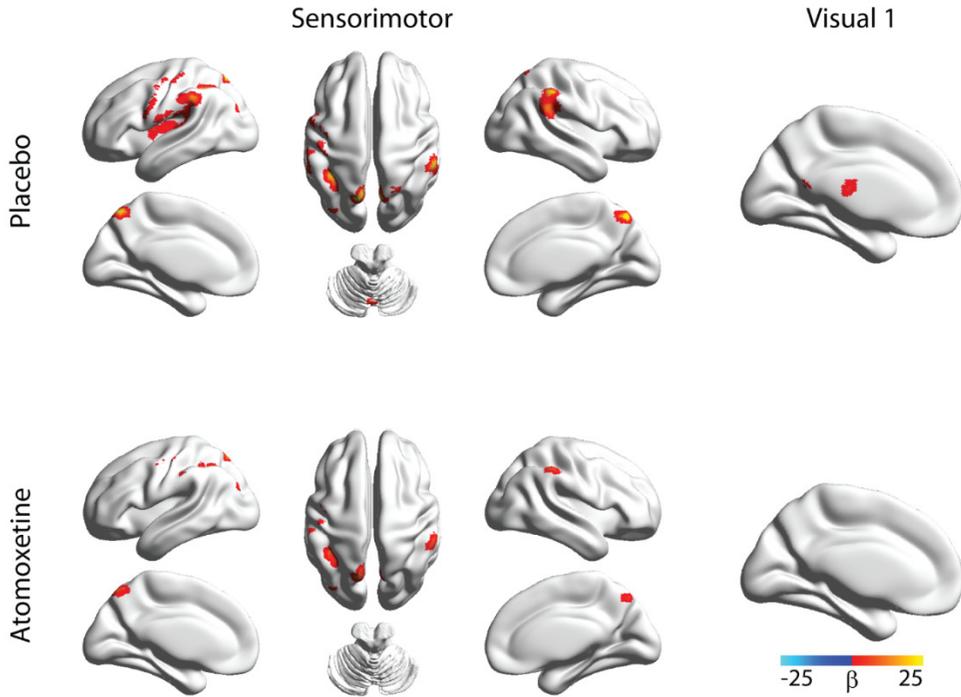
Figure S4. Covariation with component time-series in the individual conditions, only for clusters that both showed a significant ($p < 0.05$, FWE-corrected) atomoxetine-induced reduction in coupling, and significant ($p < 0.05$, FWE-corrected) covariation with the component time series. Note that brain regions that showed (numerically) an atomoxetine-induced polarity reversal are not significant.
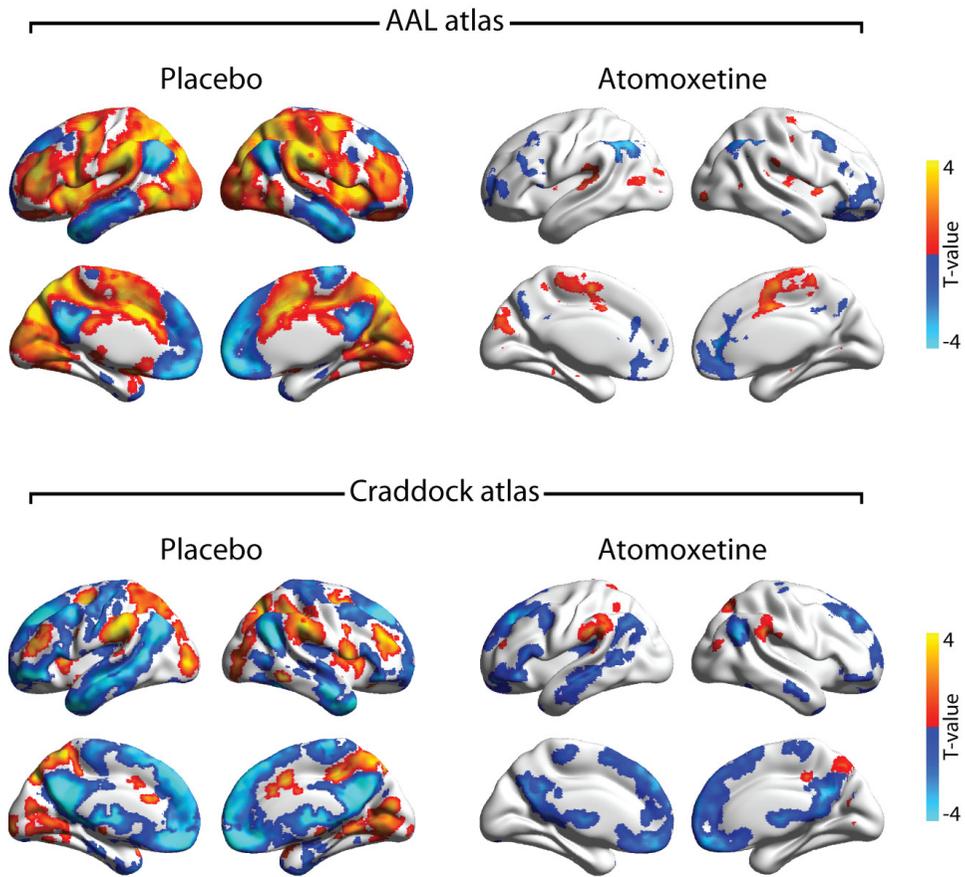
Figure S5. Spatial mode 1 for the decomposition placebo > atomoxetine, with FWE-corrected threshold of p < 0.05.
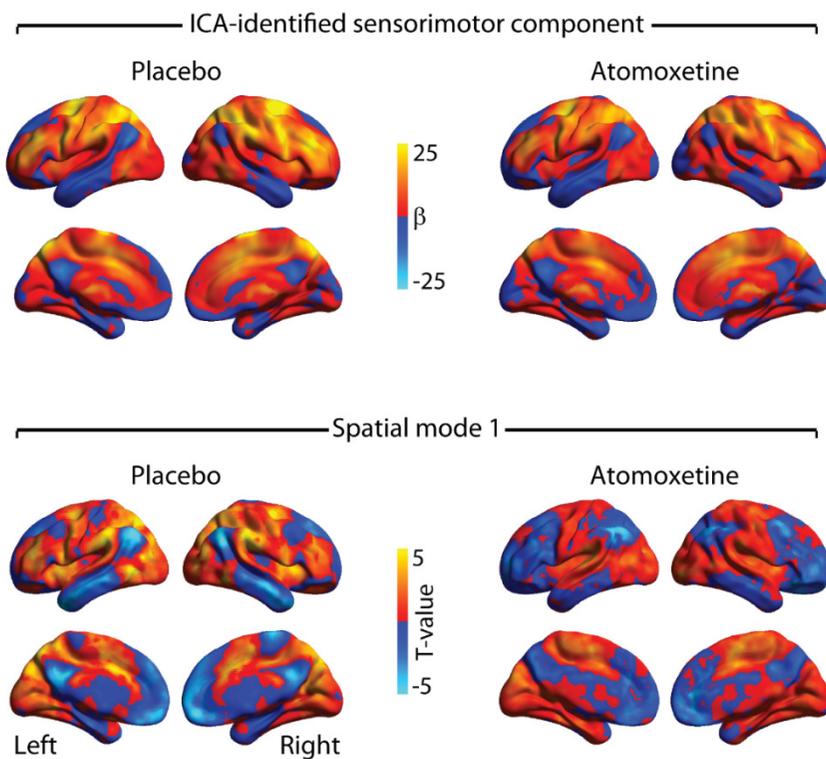
Figure S6. Unthresholded spatial maps of average regression coefficients of the ICA-identified sensorimotor component, and spatial maps of mode 1 generated using the AAL atlas, and for the decomposition placebo > atomoxetine.
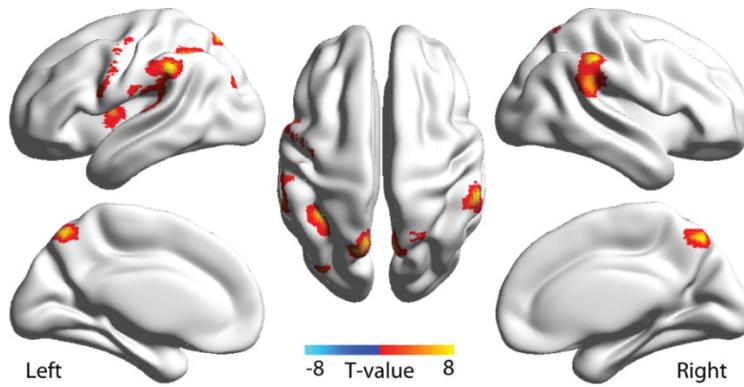
Figure S7. Spatial mode 1 in the placebo condition generated with the AAL atlas (top left panel of Figure S5), masked with the regions that showed a significant reduction in connectivity with the sensorimotor network in the dual regression analysis.
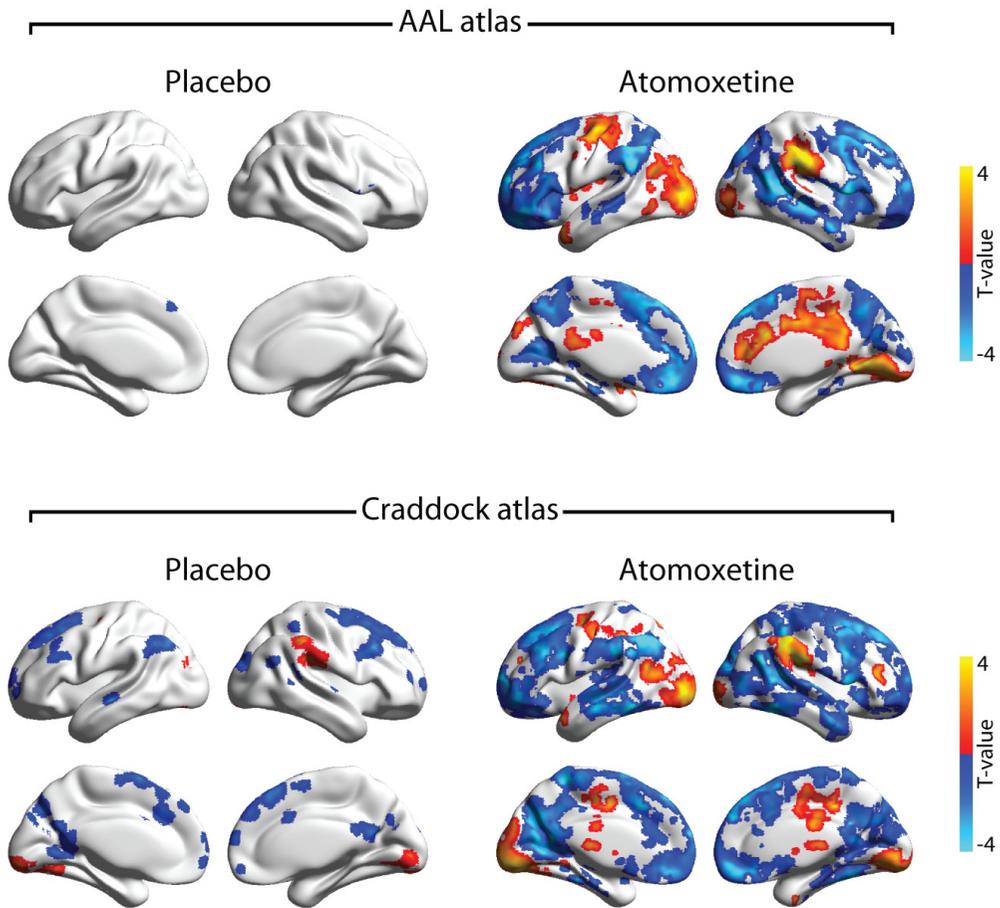
Figure S8. Spatial mode 1 for the decomposition placebo > atomoxetine, with FWE-corrected threshold of $p < 0.05$.

5. Pupil Diameter Tracks Lapses of Attention

Abstract

Our ability to sustain attention for prolonged periods of time is limited. Studies on the relationship between lapses of attention and psychophysiological markers of attentional state, such as pupil diameter, have yielded contradicting results. Here, we investigated the relationship between tonic fluctuations in pupil diameter and performance on a demanding sustained attention task. We found robust linear relationships between baseline pupil diameter and several measures of task performance, suggesting that attentional lapses tended to occur when pupil diameter was small. However, these observations were primarily driven by the joint effects of time-on-task on baseline pupil diameter and task performance. The linear relationships disappeared when we statistically controlled for time-on-task effects and were replaced by consistent inverted U-shaped relationships between baseline pupil diameter and each of the task performance measures, such that most false alarms and the longest and most variable response times occurred when pupil diameter was both relatively small and large. Finally, we observed strong linear relationships between the temporal derivative of pupil diameter and task performance measures, which were largely independent of time-on-task. Our results help to reconcile contradicting findings in the literature on pupil-linked changes in attentional state, and are consistent with the adaptive gain theory of locus coeruleus-norepinephrine function. Moreover, they suggest that the derivative of baseline pupil diameter is a potentially useful psychophysiological marker that could be used in the on-line prediction and prevention of attentional lapses.

Chapter 5

5.1 Introduction

The ability to sustain attention for prolonged periods of time is essential for normal functioning in everyday life. Lapses of attention can have dramatic consequences, such as when a car driver is absent-minded and brakes too late in response to an unexpected traffic backup, or when an air traffic controller fails to spot that two aircraft are about to cross paths. Physiological markers that indicate when such lapses of attention are more likely to occur could yield insight into the cognitive mechanisms that underlie attentional lapses, as well as provide preventative measures. A potentially useful physiological marker for detecting lapses of attention might be pupil diameter. The diameter of the pupil has long been known as a marker of cognitive load and attentional performance (Kahneman and Beatty, 1966, 1967). More recently, some researchers have considered endogenous ('baseline') variability in pupil diameter as an indicator of fluctuations in attentional control state (e.g., (Aston-Jones and Cohen, 2005; Gilzenrat et al., 2010; Jepma and Nieuwenhuis, 2011)).

Despite the potential utility of pupil diameter as a marker of attentional engagement, the available studies in which the relationship between baseline pupil diameter and sustained attentional performance has been investigated display a remarkable lack of consistency. Some researchers have reported that moments of poor task performance or off-task thought are associated with larger baseline diameter (Smallwood et al., 2011; Smallwood et al., 2012; Franklin et al., 2013; Unsworth and Robison, 2016). Others have reported that poor task performance is associated with *smaller* baseline diameter (Van Orden et al., 2000; Kristjansson et al., 2009; Grandchamp et al., 2014; Mittner et al., 2014; Hopstaken et al., 2015b), or is preceded by a progressive decline in pupil diameter (Murphy et al., 2011; Grandchamp et al., 2014). Finally, some studies have found that poor task performance can be accompanied by both relatively small and relatively large baseline diameter (Murphy et al., 2011; Smallwood et al., 2012; McGinley et al., 2015a; Unsworth and Robison, 2016). Several theoretical and methodological factors may be responsible for this discrepancy. For instance, the studies reviewed here differed considerably with regard to the measures they used to assess attentional performance: response time (RT; (Smallwood et al., 2011; Smallwood et al., 2012)); a proportion of slowest responses (Kristjansson et al., 2009; Unsworth and Robison, 2016); variability in RTs (Murphy et al., 2011); perceptual sensitivity (Beatty, 1982; Hopstaken et al., 2015b, a); and self-reported off-task thought (Franklin et al., 2013; Grandchamp et al., 2014; Mittner et al., 2014).

Another factor that may contribute to the lack of consistency in this literature concerns time-on-task effects. Prolonged task performance often results in decrements in attentional performance due to reduced vigilance (Mackworth, 1948, 1968), and concurrent changes in pupil diameter (Van Orden et al., 2000; Hopstaken et al., 2015b; Unsworth and Robison, 2016). For instance, Hopstaken and colleagues found a progressive decrease in both baseline diameter and perceptual sensitivity with

prolonged performance of an N-back task (Hopstaken et al., 2015b). A similar decline in both pupil diameter and performance was reported by Van Orden and colleagues, using a sustained attention task (Van Orden et al., 2000). However, in other studies time-on-task has been reported to lead to contrasting effects on pupil diameter and task performance. For instance, Murphy and colleagues found a progressive *increase* over time in baseline diameter during an oddball task and a trend towards poorer performance over time (Murphy et al., 2011). Beatty reported a decrement over time in perceptual sensitivity during an oddball task, but no change in baseline diameter (Beatty, 1982). These time-on-task effects are often not taken into account when assessing the relationship between pupil diameter and performance (but see Kristjansson et al. (2009) and Mathôt et al. (Mathôt et al., 2015) for notable exceptions). Thus, depending on the behavioral task and context, shared effects of time-on-task could in principle impose a relationship between diameter and task performance, or obscure a more nuanced relationship.

An example of such a nuanced relationship is the Yerkes-Dodson law, the phenomenon that performance often varies as an inverted-U function of arousal, such that both under- and over-arousal are associated with poor performance (Yerkes and Dodson, 1908). Aston-Jones and Cohen (2005), in their adaptive gain theory, proposed that this relationship reflects the effects of neuromodulation originating from the locus coeruleus-norepinephrine (LC-NE) system. The LC is a small nucleus in the pontine tegmentum that collateralizes broadly and supplies NE to almost the entire brain (Berridge and Waterhouse, 2003; Aston-Jones and Cohen, 2005). Over longer time periods, the level of baseline activity of LC neurons fluctuates with task performance. Intermediate levels of baseline LC activity are associated with (near-)optimal performance, whereas shifts toward either end of the baseline activity continuum are associated with declining performance (Coull et al., 1995; Aston-Jones and Cohen, 2005; Chamberlain et al., 2006a; Brown et al., 2015). Notably, activity in the LC has been reported to correlate with the size of the pupil (Aston-Jones and Cohen, 2005; Murphy et al., 2014b; Varazzani et al., 2015; Joshi et al., 2016). Thus, taken together, this framework predicts that both periods of small baseline diameter and periods of large baseline diameter should be associated with impaired attentional performance. Unfortunately, most studies so far have been confined to categorical comparisons of pupil diameter between on-task and off-task thought, or fast and slow response times, without taking nonlinear relationships into consideration.

Thus, in the present study we carried out a detailed investigation of the interrelationships between performance on a sustained attention task, slow baseline fluctuations in the diameter of the pupil, and the effects of time-on-task on both these variables. In contrast to some previous studies, we assessed these relationships at a within-participant, moment-by-moment level, using multiple measures of attentional state. This approach requires large numbers of trials, which in many studies is made difficult by the fact that short intertrial intervals can lead to contamination of pre-trial baseline pupil measurements by task-related pupil dilations on the previous trial. Here,

we overcame this challenge by using a fast-paced, isoluminant, gradual-onset continuous performance task (Esterman et al., 2013) to minimize stimulus-evoked pupil dilations, and by regressing out the remaining task-related transient pupil dilations.

Our results show that attentional performance and baseline diameter progressively declined over time, resulting in strong linear relationships between these variables. However, when we controlled for time-on-task, the relationships between task performance and pupil diameter became U-shaped, consistent with the Yerkes-Dodson law and the adaptive gain theory of LC function (Aston-Jones and Cohen, 2005). Moreover, we explored the relationship between performance and changes in pupil size quantified as the temporal derivative of baseline diameter. This measure was inspired by prior work in rodents, showing that the derivative of pupil diameter tracks changes in cortical state and signal detection performance (Reimer et al., 2014; McGinley et al., 2015a). As opposed to baseline diameter, its derivative showed a linear relationship with behavioral performance that was robust to the effect of time-on-task.

5.2 Materials and Methods

Participants. A total of 30 right-handed individuals took part in the study. Two participants were excluded due to technical difficulties with the eye tracker, resulting in a final N of 28 (mean age: 20.9; SD 2.5; min/max 18-26; 6 male). Exclusion criteria included a history of psychiatric disorders or wearing glasses. All participants gave written informed consent prior to the experiment and were compensated with €7,50 or course credit. The study was approved by the Leiden University Institutional Review Board (IRB).

Task. We used a modified version of the gradual continuous performance task (gradCPT) described by Esterman *et al.* (Esterman et al., 2013). Participants were asked to respond to images of cities by pressing the space bar and withhold a response when presented with an image of a mountain (Figure 1a). City trials were more frequent (90% of trials) than mountain trials (10%). The images subtended approximately 6 degrees of visual angle, were isoluminant, grayscale, and were presented on a black background. The images linearly and continuously morphed from one into the next, with an 800-ms interval between 100% coherence levels (stimulus onset asynchrony, SOA). This was done to provide a task context in which the participant had to continuously monitor the stimulus stream, and thus could not take 'mini breaks' in between trials. To allow the pupil to normalize, the first and last seven images in each block were scrambled. On these trials the participant did not respond and these trials were not included in any of the analyses. Participants were first familiarized with the environment and task by passively viewing all images, without continuous transitions. Then, they practiced the task for 34 trials at ~45% of

the normal speed, and then for another 75 trials at the regular speed. Each participant performed a total of 3 blocks of 600 trials (~8 minutes each) per block. Participants took a forced break of at least 5 minutes between blocks and were offered a small snack (chocolate chip cookie) during this interval. The total duration of the experiment was approximately 40 minutes.

Measures of task performance. Hits (responses to cities) and correct rejections (withheld responses to mountains) were considered as correct trials. Misses constituted withheld responses to cities. Within the context of continuous performance tasks, lapses of attention are usually defined as false alarms (Robertson et al., 1997; Esterman et al., 2013). In our study false alarms corresponded to responses to mountains. However, as noted in the introduction, a variety of other measures have been used to infer attentional state. Therefore, besides false alarm rate, we included three additional performance metrics: 1) the proportion of trials that fell within the slowest quintile of RTs within the block; 2) average RT; and 3) the RT coefficient of variation (RTCV)—that is, the standard deviation of RT divided by the block mean RT.

RTs were measured relative to the onset of each stimulus. For example, an RT of 640 ms (i.e., 80% of the SOA) indicated that the participant responded when the displayed image consisted of 80% trial *n*, and 20% trial *n*-1. An iterative algorithm assigned responses to trials in the case of multiple responses, or unusually fast responses (before 70% coherence of trial *n*) and unusually slow responses (after 40% coherence of trial *n*+1). First, the number of correct responses was optimized. Then, ambiguous responses were assigned to a neighboring trial if either of them had no response. If both had a response, it was assigned to the closest city (go) trial. Lastly, if a trial was assigned multiple responses, the fastest response was selected. This procedure was identical to the one described by Esterman *et al.* (Esterman et al., 2013), and is unlikely to have substantially influenced the results, given that ambiguous responses were relatively rare (<4% of the trials).

Pupillometry. Participants were seated in a dimly lit room with their head stabilized by a chin rest. During the task participants were asked to keep their eyes focused on a small white fixation dot in the center of the image. We measured the diameter of the right pupil at a sampling rate of 1 kHz with an EyeLink 1000 eye tracker. Prior to the start of each block the eye tracker was calibrated and validated with a 9-point fixation routine.

Moments when the eye tracker received no pupil signal (e.g., during blinks) were marked automatically during data acquisition by the manufacturer's blink detection algorithm. Afterwards, an iterative algorithm detected additional moments of poor signal quality (e.g., due to partial occlusion of the pupil by the eyelashes). For 200 iterations over the entire signal time series for a given participant and block, any sample for which the difference in pupil diameter compared to the previous sample exceeded a threshold was marked as 0. The default threshold was set to 25 pixels, but the threshold was individually-tailored for participants for whom the algorithm failed to identify sharp spikes in the data or inappropriately marked clean sections of data.
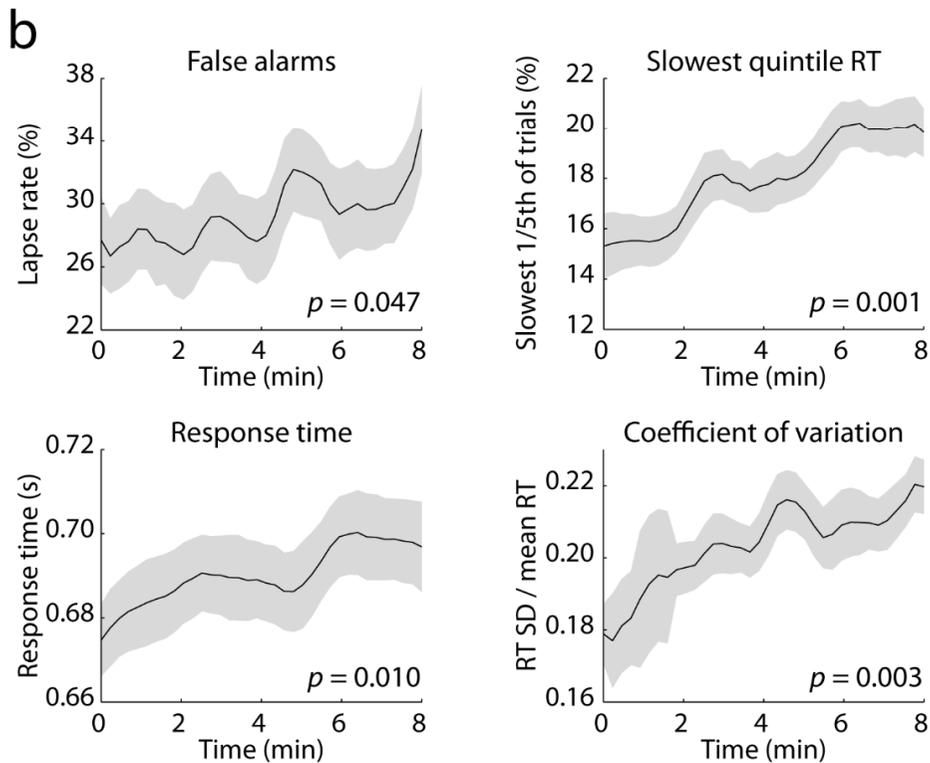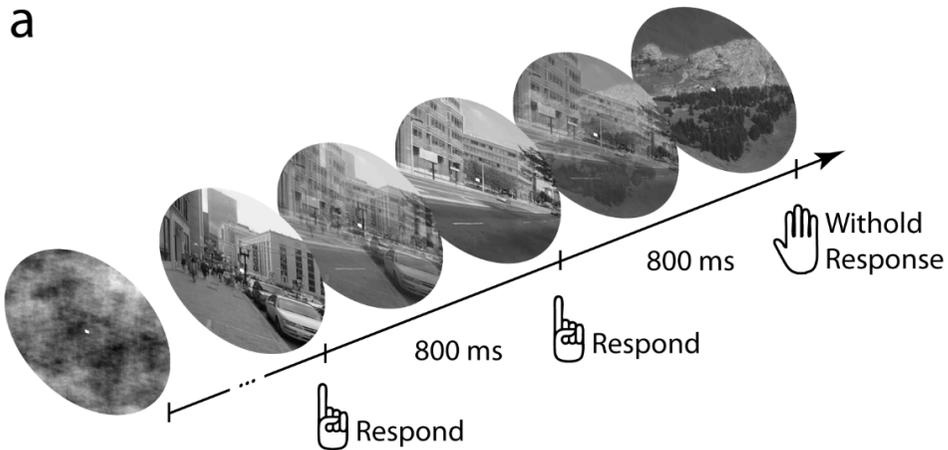
Figure 1. Task and behavioral results. **a)** Gradual continuous performance task. Each block started and ended with 7 scrambled images. The participant was asked to respond to city scenes but not to mountain scenes. Each image continuously morphed into the next, with an 800-ms interval between 100% coherence levels. **b)** Behavioral results. Data are smoothed for display only. All measures showed a significant linear increase with time-on-task, *p*-values are listed in the lower right corner of each panel. Error bars represent the SEM.

All marked data sections were then interpolated across using shape-preserving piecewise cubic interpolation. On average 6.8% (SD 4.6, min/max 0.2/18.7) of the data points were interpolated. After interpolation each pupil time series was low-pass filtered at 6 Hz to remove any residual high-frequency noise.

We were primarily interested in the relationship between tonic (endogenous) variations in pupil diameter and behavioral performance. Due to the short SOA of the gradCPT (800 ms), it is possible that stimulus-related pupil dilations precluded a reliable estimation of tonic pupil fluctuations. However, we first note that stimulus-related pupil responses accounted on average for only ~8% (SD 4%) of total pupil fluctuations, indicating that tonic fluctuations were a far more dominant source of variance in the observed pupil time series. Moreover, we reduced this already small contribution of trial-related pupil responses by employing linear regression to calculate residualized pupil time series for each participant and block that represented fluctuations in pupil diameter that were independent of the phasic pupil dilations evoked by task stimuli and their associated behavioral responses. The measured pupil time series were segmented around the onset of each stimulus, distinguishing between the four trial types (hits, misses, correct rejections, and false alarms), and around response onset. For each participant, we then computed average stimulus-locked and response-locked pupil waveforms, and extracted the peak amplitude in a 0 to 5 s post-event window, relative to a 200-ms pre-event baseline. This resulted in an estimate of the amplitude of phasic pupil dilations for each participant and type of event. Next, for each participant we created separate stick functions for each type of event in which the latency of the sticks corresponded to stimulus onsets and the participant's RTs, and the amplitude corresponded to the estimated amplitude of the phasic pupil dilation for that participant and type of event. We then convolved the stick functions with the canonical pupillary response function (h) presented by Hoeks and Levelt (Hoeks and Levelt, 1993):

$$h = s \cdot (t^n) \cdot e^{\left(\frac{-n \cdot t}{t_{max}}\right)}$$

where t is time, n is the number of layers (10.1), t_{max} corresponds to the latency of maximum dilatory response per participant and type of event, and s was a constant ($2.7569 \cdot 10^{-29}$) to scale the response function to unit height.

Finally, we used multiple linear regression to remove the stimulus- and response-related phasic dilations (Figure 2) from the unsegmented pupil time series. This procedure minimized the extent to which phasic pupil dilations convoluted the estimates of tonic variations in the diameter of the pupil. Note that this approach is highly similar to analysis of the first-stage general linear model of functional magnetic resonance imaging data, to correct for signal variance associated with trial-type-specific evoked responses (as implemented by e.g. (Esterman et al., 2013)).
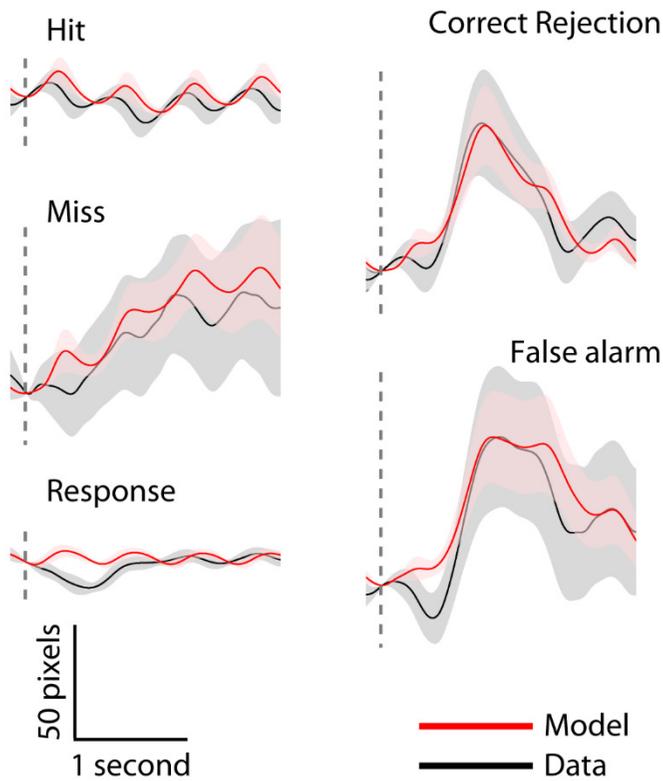
Figure 2. Phasic pupillary responses. Trial-averaged modeled (red) and empirical (black) stimulus-related pupil dilations. The vertical dashed line represents event (stimulus or response) onset. Error bars represent the SEM.

5.3 Results

Performance decrements with time-on-task. We first verified whether behavioral performance degraded over the course of a block, as is expected in demanding tasks like the gradCPT that require continually sustained attention (Robertson et al., 1997; Esterman et al., 2013). To do so, we calculated temporally resolved metrics of trial-averaged behavior by applying a sliding window to the behavioral data of each block of each participant. The window had a width of 50 trials (40 seconds duration) and was slid across the data in steps of 15 trials. For each of these windows we calculated several measures of task performance: 1) the proportion of false alarms; 2) the proportion of trials that fell within the slowest quintile of RTs within the block; 3) average RT; and 4) the RTCV (see Materials & Methods). For each of these measures this approach resulted in a continuous time series for each block. We then Z-scored the time series and fitted a straight line to them. The slopes of the fitted lines indicated whether the time series were on average increasing or decreasing (or not changing) over time. We averaged the slopes across blocks for each participant and tested if the distribution of slopes was larger than 0 using a one-tailed *t*-test. As expected, we found significant performance decrements for all behavioral measures. Over the course of a block, progressively more false alarms occurred ($t(27) = 1.74$, $p = 0.047$), and RTs

became longer (RT: $t(27) = 2.47$, $p = 0.010$; quintile: $t(27) = 3.31$, $p = 0.001$) and more variable ($t(27) = 3.06$, $p = 0.003$; Figure 1b). The proportion of misses also increased with time ($t(27) = 3.31$, $p = 0.001$), but misses were rare (0.2% of all trials) and will thus not be considered in any further analyses. In sum, over the course of a block performance deteriorated. For the sake of simplicity, we hereafter refer to the effect of time within blocks as 'time-on-task' effects.

The effects of time-on-task on tonic pupil fluctuations. Having established that behavioral performance on the task degraded over time, we next turned to the pupil data. We applied a sliding window to the unsegmented pupil data that was identical to the one applied to the behavioral data (a width of 50 trials and a step size of 15 trials). We extracted two measures: 1) the average pupil diameter in each window, hereafter referred to as 'baseline diameter'; and 2) the average temporal derivative of baseline diameter, which quantifies the extent to which the pupil tended to dilate or constrict within each window. The derivative measure was calculated as the average difference between each two consecutive samples within the window (using MATLAB's 'diff' function). This is equivalent to the difference in baseline diameter between the first and last sample of the window. For each of the pupillary measures this resulted in a time series that was identical in length to the time series of the behavioral measures.

As a direct follow-up on the behavioral analyses, we first examined whether the pupillary measures also showed time-on-task effects. To do so, we fitted a straight line to each pupil time series. The slope of the fitted line was informative of linear trends over time. We averaged the slopes across blocks for each participant and compared the distribution of slopes to 0 using a two-tailed t-test. We had no clear hypothesis regarding the direction of the time-on-task effect for the derivative of pupil diameter, so for this test we also used a two-tailed t-test. Both pupil measures showed
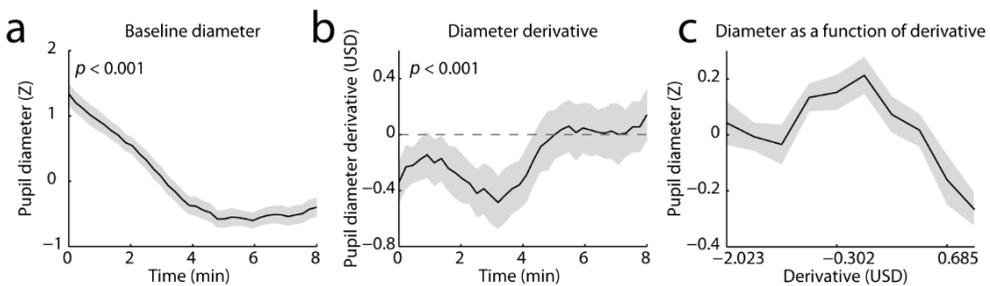
Figure 3. Baseline diameter and derivative. **a)** Time-on-task effect for baseline diameter, and **b)** for diameter derivative. *p*-values are listed in the top left corner of each panel. The derivative is shown as variance-normalized but without the mean removed. Values below the horizontal dotted line indicate that on average the pupil is constricting, whereas values above the line indicate that the pupil is dilating. **c)** The relationship between pupil diameter and its derivative. Baseline pupil diameter plotted as a function of the derivative. Diameter is smallest when the pupil is dilating the fastest. USD: Units standard deviation.

significant linear time-on-task effects. Over the course of a block baseline diameter became smaller ($t(27) = 8.10$, $p < 0.001$; Figure 3a). On average, its derivative was initially negative and became less negative over time ($t(27) = 4.40$, $p < 0.001$; Figure 3b), reflecting the fact that the pupil progressively decreased in diameter during the early-to-mid portions of a block and reached a relatively stable diameter thereafter.

Before examining the relationship between baseline diameter and its derivative vis-à-vis the behavioral performance measures, we wanted to make sure that the two pupil measures were not highly correlated with each other, so that they might be expected to explain unique variance in the behavioral measures. In order to clarify the relationship between baseline diameter and its derivative, we correlated their respective time series derived from the sliding-window approach, for each participant and each block, and compared the distribution of Fisher-transformed correlation coefficients averaged across blocks to zero using a t-test. Although the correlation was consistently negative across participants ($t(27) = -2.48$, $p = 0.020$; Figure 3c), the average correlation coefficient was rather small: -0.12. Thus, the two pupil measures only weakly co-varied ($R^2 < 1.5\%$) and their capacities to explain unique portions of the variance in behavior were high.

The relationship between tonic pupil fluctuations and behavior. We next used multiple regression to examine linear relationships between the time series of each of the Z-scored pupillary measures and each of the Z-scored behavioral measures, within participants and within blocks. A separate model was constructed for each of the pupillary/behavioral measure pairings. We also included quadratic regressors in these models, but only report the quadratic relationships between baseline diameter and the behavioral measures. The inclusion of quadratic regressors in the regression
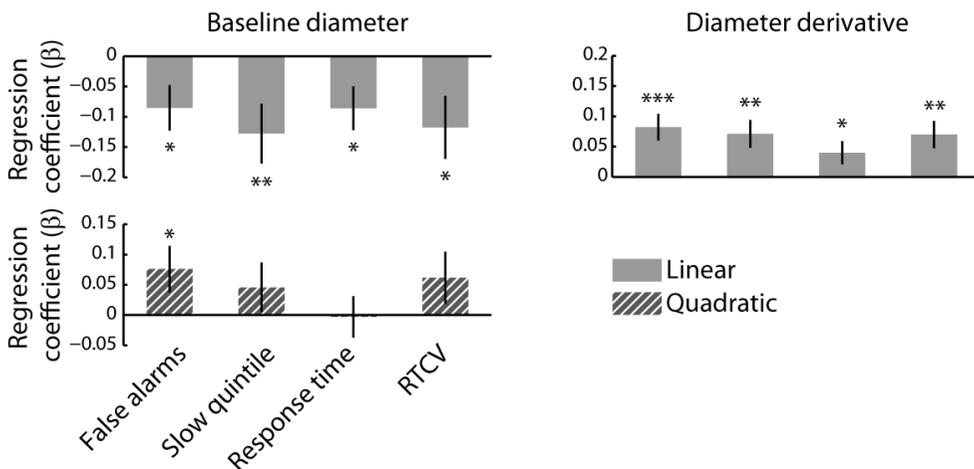
Figure 4. The relationship between pupil diameter and behavior. Regression coefficients are shown per pupil measure and behavioral measure. USD: Units standard deviation. Error bars represent the SEM. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

models for the derivative did not affect the direction and significance of the linear regression coefficients. We averaged the resulting regression coefficients across blocks for each participant and compared the distribution of regression coefficients to 0 using t-tests. We expected the typical Yerkes-Dodson relationship between baseline diameter and behavior (but see below), and therefore used one-tailed t-tests to compare the quadratic regression coefficients to zero. Furthermore, because all analyses concerning the derivative of baseline diameter were exploratory, we used two-tailed t-tests in these analyses. The linear and quadratic relationships between the pupil measures and the behavioral measures are summarized in Figure 4.

We found a significant positive quadratic relationship between baseline diameter and false alarm rate ($t(27)$ = 1.99, p = 0.029), indicating that false alarm rate tended to increase at both the upper and lower extremes of baseline pupil diameter. This finding is consistent with the long-recognized inverted U-shaped relationship between arousal and task performance (Yerkes and Dodson, 1908). However, we found no such quadratic relationship for the other behavioral measures (all ps > 0.05).

Given the linear time-on-task effects on baseline diameter and each of the behavioral measures, it may be expected that baseline diameter be linearly related to false alarm rate, RT, RTCV, and the proportion of trials that fell within the slowest RT quintile. We thus used one-tailed t-tests to test this hypothesis. In line with this notion, all behavioral measures were negatively related to baseline diameter (false alarm rate: $t(27)$ = 2.28, p = 0.02; quintile: $t(27)$ = -2.60, p = 0.008; RT: $t(27)$ = -2.38, p = 0.012; RTCV $t(27)$ = -2.27, p = 0.016). Thus, more false alarms and longer and more variable RTs tended to occur when baseline diameter was smallest which, as shown earlier, also tended to coincide with the end of task blocks. The linear relationships between baseline diameter and each of the behavioral measures are shown in Figure 5a.

Interestingly, the derivative of pupil diameter showed a significant positive linear relationship with all behavioral measures (false alarm rate: $t(27)$ = 3.71, p = 0.001; quintile: $t(27)$ = 3.10, p = 0.005; RT: $t(27)$ = 2.10, p = 0.046; RTCV: $t(27)$ = 3.11, p = 0.005). The positive relationship indicated that periods during which the pupil was relatively stable or dilating (i.e., the value of the derivative was positive/least negative) were characterized by the most false alarms and the slowest and most variable RTs (Figure 5b). In other words, periods in which the pupil showed little change in size over time or tended to dilate slowly, were marked by the poorest behavioral performance.

In order to rule out the possibility that these results were dependent on the choice of window size, we repeated the regression analysis for a range of sliding window sizes (40 s to 4 min, and an 8-s difference in width between each consecutive window size). For each window size, we then computed the regression coefficients indicating linear and quadratic relationships between the time series of each of the Z-scored pupillary measures and each of the Z-scored behavioral measures. We then averaged
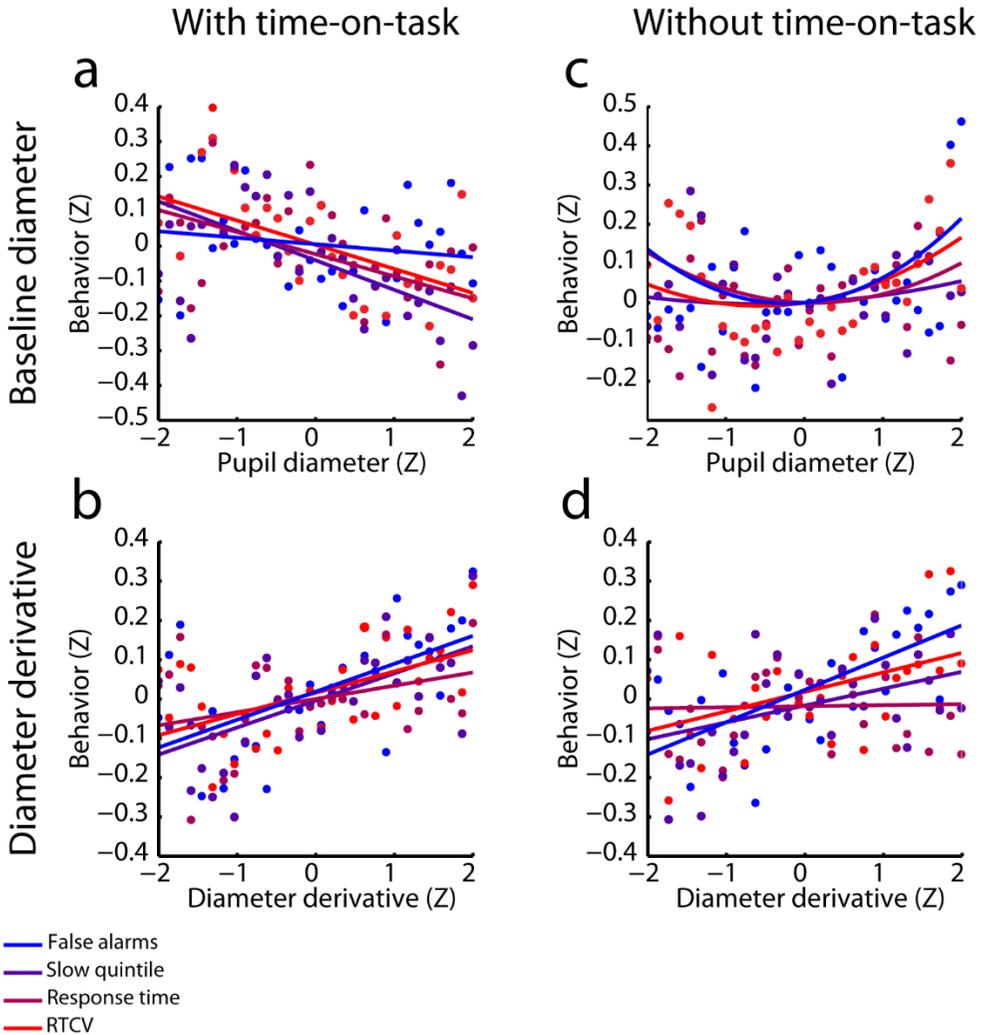
Figure 5. Relationship between pupillary measures and behavior, before **(a,b)** and after **(c,d)** regressing out time-on-task. Pupil data were z-scored within participants and blocks, aggregated across participants, and then divided up into 30 bins, and the behavioral data were sorted according to pupil diameter. Large positive values on the Y-axis indicate relatively poor behavioral performance. The initially linear relationship between baseline diameter and behavior becomes U-shaped after controlling for time-on-task, whereas the relationship between the derivative of baseline diameter and behavior remains linear after controlling for time-on-task. Straight lines are least squares regression lines, curved lines are fitted 2nd-order polynomials.

the resulting regression coefficients across blocks, and across the behavioral measures, and computed their area under the curve (AUC) across window sizes. This AUC summary statistic indicated whether on average the behavioral measures

showed a relationship (linear or quadratic) with the two pupil measures. Finally, we tested if the group-level distribution of AUCs differed from 0 using one-tailed *t*-tests. If the linear pupil-behavior relationships were not dependent on the choice of a single (arbitrary) window size, we expected the AUC of the linear regression coefficients to go in the same direction as the initial regression coefficients. That is, we would expect the AUC to be negative for diameter, and positive for the derivative. As expected, the linear AUCs were significantly different from 0 and in the predicted direction for both pupil measures (diameter: $t(27) = -2.62$, $p = 0.007$; derivative: $t(27) = 4.05$, $p < 0.001$). Also in line with our expectations, the quadratic AUC for baseline diameter did not differ from zero ($t(27) = 0.09$, $p = 0.47$). Altogether, these results show that periods during which the pupil was smallest and remained relatively stable or dilated again were marked by the poorest behavioral performance on the task. These effects were consistent across a range of time scales.

The relationship between tonic pupil fluctuations and behavior, controlled for time-on-task. It is possible that the relationships between baseline diameter and behavior reported above simply reflect the strong effects of time-on-task on these two types of variables, rather than a more intrinsic, time-invariant relationship. We therefore wondered whether shared effects of time-on-task on baseline diameter and behavior might be obscuring more subtle relationships between the associated measures. To address this possibility, we explored whether the relationship between the pupillary measures and behavior remained after statistically controlling for time-on-task. To do so, we performed similar regression analyses as before, except that we included a linearly increasing predictor that tracked time-on-task (i.e., the time elapsed within each block). As a result, the regression coefficients represented the relationship between the pupillary measures and behavior, independent of a linear time-on-task effect.

As can be seen in Figure 6, the initial linear relationships between baseline diameter and the RT measures became quadratic when time-on-task was taken into account. Both relatively small and large diameters were associated with an increased false alarm rate, and slower and more variable RTs (false alarm rate: $t(27) = 1.99$, $p = 0.028$; quintile: $t(27) = 1.45$, $p = 0.08$; RT: $t(27) = 2.06$, $p = 0.025$; RTCV: $t(27) = 2.79$, $p = 0.005$), whereas linear relationships between pupil size and these behavioral measures were no longer present (all $p > 0.2$). This suggests that a U-shaped relationship between baseline diameter and RT measures was indeed initially obscured by strong time-on-task effects (Figure 5c). In contrast, the linear relationships between the derivative and the behavioral measures that were evident in the original regression models were largely preserved in the model that statistically controlled for time-on-task (false alarm rate: $t(27) = 3.09$, $p = 0.005$; quintile: $t(27) = 2.13$, $p = 0.041$; RT: $t(27) = 0.93$, $p = 0.360$; RTCV: $t(27) = 3.07$, $p = 0.005$; Figure 5d). These effects indicate that periods in which linearly detrended pupil diameter was generally increasing were associated with relatively impaired performance.
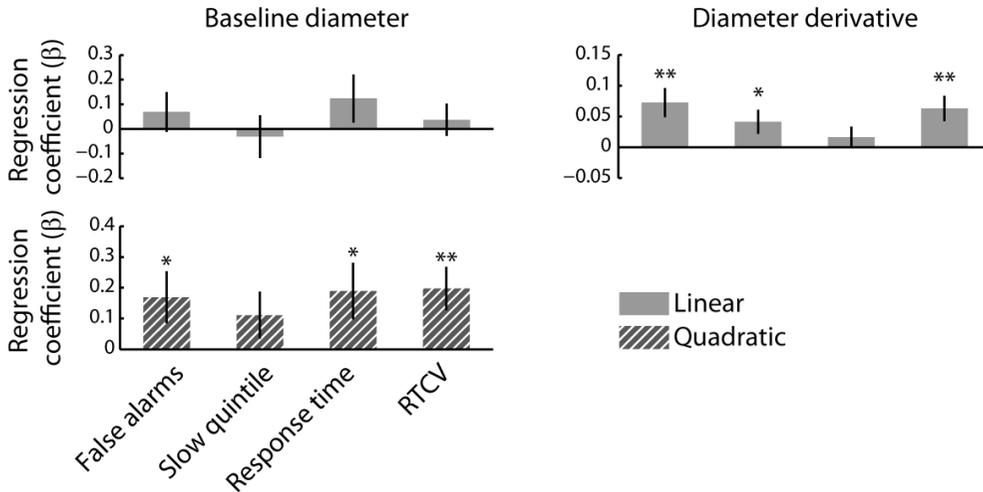
Figure 6. The relationship between pupil diameter and behavior, after statistically controlling for time-on-task. Regression coefficients per pupil measure and behavioral measure with time-on-task included as a variable of non-interest. Error bars represent the SEM. *: $p < 0.05$; **: $p < 0.01$.

Again, these results were not dependent on the choice of window size, because the AUC summary statistics across window sizes and behavioral measures showed a similar shift from linear to quadratic for baseline diameter after statistically controlling for time-on-task (linear AUC: $t(27) = 0.29$, $p = 0.39$; quadratic AUC: $t(27) = 3.08$, $p = 0.002$). The linear relationship between the derivative and behavior was also preserved in the AUC across window sizes (linear AUC: $t(27) = 3.08$, $p = 0.002$).

Together, these results suggest that time-on-task was driving the initially observed linear relationships between mean baseline diameter and task performance, and to some extent obscured latent quadratic relationships between these variables. In contrast, the linear relationship between task performance and the derivative of pupil diameter was mostly robust to controlling for time-on-task. As we discuss below, this relationship likely reflects the quadratic relationship between diameter and behavior that occurred independent of time-on-task.

5.4 Discussion

Using a fast-paced sustained attention task, we found robust linear relationships between baseline pupil diameter and several behavioral manifestations of attentional lapses. However, these linear relationships primarily reflected the joint effect of time-on-task on baseline pupil and behavior: as performance deteriorated over the course of a block (as indexed by increased false alarm rate and slower and more variable RTs), the pupil became progressively smaller. Importantly, when this effect of time-

on-task was statistically partialled out, the relationship between baseline diameter and behavior became U-shaped: more false alarms, and longer and more variable RTs occurred during periods of both relatively small and relatively large baseline diameter, a pattern that is consistent with the Yerkes-Dodson law of mental task performance (Yerkes and Dodson, 1908) and the adaptive gain theory of LC-NE function (Aston-Jones and Cohen, 2005).

Previous studies on the relationship between pupil diameter and attentional state have yielded contradicting results. Some studies have reported that moments of poor task performance or off-task thought are associated with larger baseline diameter (Smallwood et al., 2011; Smallwood et al., 2012; Franklin et al., 2013; Unsworth and Robison, 2016). Conversely, others have reported that poor task performance is associated with smaller baseline diameter (Van Orden et al., 2000; Kristjansson et al., 2009; Grandchamp et al., 2014; Mittner et al., 2014; Hopstaken et al., 2015b). Our research suggests three methodological reasons for these mixed results. First, a multitude of measures have been used to assess attentional state. Although the performance measures used in the current study generally showed similar relationships with pupil diameter, there were some differences between the measures. For example, as opposed to the RT measures, false alarm rate already displayed a U-shaped relationship with diameter before the effect of time-on-task was partialled out. This discrepancy may be explained by the possibility that slow and variable RTs primarily reflect a decrease in attentional focus (Weissman et al., 2006)—equivalent to a lower (e.g. (Nunez et al., 2015)) and/or more variable (Murphy et al., 2014a) rate of decision formation— whereas false alarms may reflect either a decrease in attentional focus or an inadvertent lowering of the response threshold (Forstmann et al., 2016). Thus these signatures of attentional lapses may have partially dissociable mechanistic bases. To make sure that key conclusions do not depend on the specific choice of measure, future studies should ideally use a range of performance measures, as we have done here. Second, the majority of previous studies have reported only categorical comparisons (e.g., on-task versus off-task thought (Franklin et al., 2013; Grandchamp et al., 2014; Kang et al., 2014); or normal versus slow RTs (Kristjansson et al., 2009; Unsworth and Robison, 2016)) to assess the relationship between pupil diameter and attentional state. However, such comparisons cannot reveal potential non-linear relationships between pupil and behavior. Thus, the manner in which the relationship between baseline diameter and attentional state is assessed restricts the conclusions that can be drawn from the data.

Finally, our results suggest that contradictory findings in the literature may also be due to differences between studies in the presence and nature of parallel effects of time-on-task on pupil diameter and behavior. In tasks that are demanding, such as our task, the dominant finding is that attentional lapses and mind wandering are associated with a smaller baseline pupil diameter than non-lapses or on-task thought (e.g. (Kristjansson et al., 2009; Mittner et al., 2014; Hopstaken et al., 2015b)). This pattern may simply be due to a progressive decrement in behavioral performance

along with a monotonic decline in pupil diameter over time, perhaps reflecting a shift from center to left on the Yerkes-Dodson curve and a corresponding abandonment of exploitative behavior (Aston-Jones and Cohen, 2005), or reduced top-down control of behavior (Mathôt et al., 2015). Such time-dependent shifts on the Yerkes-Dodson curve could be the consequence of depleted cognitive resources. As noted by Hopstaken et al. (2015b), there is substantial overlap between the behavioral consequences of mental fatigue and the characteristics of low-arousal states. Nevertheless, the mechanistic origin of simultaneous effects of time-on-task on pupil diameter and performance remains an interesting open question for future research. In less demanding tasks, by contrast, time-related performance decrements are often less severe, and pupil diameter has even been reported to increase over time in such settings (Murphy et al., 2011). Such an absence of shared time-on-task effects might in turn afford greater scope for revealing more nuanced relationships between pupil diameter and task performance in the observed data. We suggest that future studies should carefully distinguish between pupil-behavior relationships due to time-on-task and potentially more subtle relationships that operate on a faster time scale. As we have shown, this dissociation can be easily achieved via the implementation of appropriate statistical control.

Aside from yielding insight into the mechanisms underlying attentional lapses, an important long-term goal of studies such as ours is to establish psychophysiological markers that can be used in on-line biofeedback systems, aimed at predicting and preventing lapses of attention. Recently, deBettencourt et al. (deBettencourt et al., 2015) made an important step towards the realization of such a system. By providing participants with well-timed performance feedback based on the on-line analysis of brain imaging data, they could improve participants' performance on a sustained attention task. However, the involvement of brain imaging equipment imposes obvious restrictions on the real-world applicability of this technique. Our results, however, indicate that the pupil could potentially be used to predict when lapses of attention are likely to occur. Given the relatively non-invasive and cost-effective nature of eye-tracking, such a system would offer substantial advantages over neuroimaging-based systems. However, it should be noted that the average regression coefficients that captured the relationship between the dynamics of the pupil and the dynamics of behavior, although consistent across participants, were modest in size (between 0.1 and 0.2). Thus, future work is needed to establish the practical feasibility of using pupil diameter and its derivative as on-line markers of attentional lapses.

Our findings that the average derivative of the pupil diameter time series was linearly related to behavioral performance, and that this relationship was independent of time-on-task, indicate that the derivative of pupil diameter offers a potential marker of attentional performance. The robustness of the derivative to time-on-task compared to baseline diameter may be explained by the way we computed this measure. Specifically, the derivative reflected the difference in baseline diameter between the first and last time point in the sliding window. Thus, this measure was less affected by

block-wide trends in pupil diameter but instead captured changes at the temporal scale of the applied sliding window. Moreover, the derivative of a U-shaped signal is monotonically increasing ($f(x) = ax^2 + c \rightarrow f'(x) = 2ax$). Any quadratic relationship between a variable (e.g., baseline diameter) and another variable (e.g., behavior) will therefore be measurable as a linear relationship between the derivative of the first variable (baseline diameter derivative) and the second variable (behavior). This holds true even in the presence of a superimposed linear relationship between diameter and behavior (e.g., due to time-on-task effects), because the linear part of the function will simply reduce to a constant in the derivative ($f(x) = ax^2 + bx + c \rightarrow f'(x) = 2ax + b$). Thus, the linear relationship between the baseline diameter derivative and behavior likely reflected the quadratic relationship between baseline diameter and behavior that occurred independent of time-on-task.

A bio-feedback system could thus incorporate the derivative of pupil diameter and a receiver operating characteristic analysis could be performed to examine how reliably the signal preceding a behavioral response discriminates between lapse and non-lapse trials. Future studies could also incorporate purely momentary fluctuations in the derivative of the pupil (cf. (Reimer et al., 2014; McGinley et al., 2015a)) as opposed to changes during a longer window. These instantaneous fluctuations are, however, beyond the scope of the current study, as we were primarily interested in tonic fluctuations that evolve over longer time periods and how they relate to global fluctuations in attentional performance. Such global fluctuations are more akin to real-world fluctuations in behavior in settings that require prolonged sustained attention, as when an air-traffic controller must monitor a display for long periods of time.

In conclusion, our results demonstrate that time-on-task, a factor that is often ignored in studies on the relationship between pupil diameter and attentional state, can obscure non-linear pupil-behavior relationships. The non-linear (inverted U-shaped) relationship between baseline pupil diameter and attentional performance that we observed after partialling out time-on-task effects is consistent with the adaptive gain theory of LC-NE function (Aston-Jones and Cohen, 2005). Finally, our results indicate that the derivative of pupil diameter is a potential marker of attentional performance that could be used for the on-line prediction and prevention of attentional lapses.

6. Task-Free Spectral EEG Dynamics Track and Predict Patient Recovery From Severe Acquired Brain Injury

Abstract

For some patients, coma is followed by a state of unresponsiveness, while other patients develop signs of awareness. In practice, detecting signs of awareness may be hindered by possible impairments in the patient's motoric, sensory, or cognitive abilities, resulting in a substantial proportion of misdiagnosed disorders of consciousness. Task-free paradigms that are independent of the patient's sensorimotor and neurocognitive abilities may offer a solution to this challenge. A limitation of previous research is that the large majority of studies on the pathophysiological processes underlying disorders of consciousness have been conducted using cross-sectional designs. Here, we present a study in which we acquired a total of 74 longitudinal task-free EEG measurements from 16 patients (aged 6-22 years, 12 male) suffering from severe acquired brain injury, and an additional 16 age- and education-matched control participants. We examined changes in amplitude and connectivity metrics of oscillatory brain activity within patients across their recovery. Moreover, we applied multi-class linear discriminant analysis to assess the potential diagnostic and prognostic utility of amplitude and connectivity metrics at the individual-patient level. We found that over the course of their recovery, patients exhibited nonlinear frequency band-specific changes in spectral amplitude and connectivity metrics, changes that aligned well with the metrics' frequency band-specific diagnostic value. Strikingly, connectivity during a single task-free EEG measurement predicted the level of patient recovery approximately 3 months later with 75% accuracy. Our findings show that spectral amplitude and connectivity track patient recovery in a longitudinal fashion, and these metrics are robust pathophysiological markers that can be used for the automated diagnosis and prognosis of disorders of consciousness. These metrics can be acquired inexpensively at bedside, and are fully independent of the patient's neurocognitive abilities. Lastly, our findings tentatively suggest that the relative preservation of thalamo-cortico-thalamic interactions may predict the later reemergence of awareness, and could thus shed new light on the pathophysiological processes that underlie disorders of consciousness.

Chapter 6

6.1 Introduction

After awakening from coma, some patients remain unresponsive while others show behavioral features that are taken as signs of awareness (Jennett and Plum, 1972; Laureys et al., 2004). The reliance on behavioral criteria for the diagnosis of such disorders of consciousness (DOC) may be suboptimal, because impairments in the patients' motor system can obscure signs of consciousness (Giacino et al., 2014). These considerations have sparked the development of 'active paradigms' that rely on electroencephalography (EEG) or neuroimaging tools to detect signs of patient awareness during mental tasks (Kotchoubey et al., 2005; Owen et al., 2006; Wijnen et al., 2007; Fischer et al., 2010; Monti et al., 2010; Boly et al., 2011; Höller et al., 2011; Sitt et al., 2014). Though promising, some active paradigms rely on higher-order cognitive abilities such as language comprehension or attention. In addition, putative electrophysiological markers of awareness such as the mismatch negativity may be absent in patients that do show behavioral signs of consciousness (Kotchoubey et al., 2005; Wijnen et al., 2007; Fischer et al., 2010; Höller et al., 2011). Moreover, a necessity for active paradigms is that the patients' sensory pathways are intact, which may not always be the case. Thus, diagnostic tools that are independent of the patients' neurocognitive abilities and integrity of sensorimotor pathways may offer a substantial improvement on existing tools.

Accordingly, task-free paradigms, in which the patient is not required to follow instructions or process stimuli, have recently gained traction (Rosanova et al., 2012; Casali et al., 2013; Demertzi et al., 2015; Schurger et al., 2015; Estraneo et al., 2016; Schorr et al., 2016; Stender et al., 2016). For instance, using positron emission tomography, Stender et al. (2016) were able to predict the presence and later emergence of consciousness in patients with DOC. Similarly, the cortical spread of EEG activity following transcranial magnetic stimulation dissociates patients with unresponsive wakefulness syndrome (UWS) from those in the minimally conscious state (MCS) (Rosanova et al., 2012). However, these paradigms necessitate the use of costly or impractical equipment, and may therefore not offer the most convenient diagnostic procedures. Task-free EEG spectral amplitude and variance metrics have shown promise as diagnostic and prognostic markers (Schurger et al., 2015; Schorr et al., 2016), but thus far have been limited in their ability to dissociate UWS from MCS patients (Schurger et al., 2015), and provide only dichotomous prognoses without specifying the expected level of recovery (Schorr et al., 2016). In contrast to amplitude and variance metrics, the potential diagnostic and prognostic value of spectral EEG connectivity metrics during task-free measurements have yet to be explored.

Several findings suggest that spectral EEG characteristics may be indicative of the level of consciousness (LoC) in patients with DOC. Compared to fully conscious control participants, patients with DOC consistently show a reduction in the amplitude of oscillations in the α and β bands, and often show a concurrent increase in θ and δ amplitude (Lehembre et al., 2012; Lechinger et al., 2013; Chennu et al., 2014; Varotto

et al., 2014). Furthermore, during auditory processing, entropy metrics of cortical information exchange vary monotonically across LoC (King et al., 2013; Sitt et al., 2014). These and other (Laureys et al., 2000; Schiff et al., 2007; Giacino et al., 2014) findings have been proposed to reflect discontinuities in the thalamo-cortico-thalamic circuit that disrupt large-scale functional interactions, and thereby enable local cortical properties to shape the spectral dynamics (Schiff, 2010; Giacino et al., 2014; Schiff et al., 2014). However, it is unclear to what extent such accounts capture longitudinal spectral changes across patients' recovery, because comparisons between LoC have almost exclusively been conducted using cross-sectional (between-group) designs.

Here, we report a longitudinal study in which we acquired a total of 74 task-free EEG measurements over the course of patient recovery from severe acquired brain injury. We assessed the feasibility of diagnosis and prognosis of DOC within individual patients based on the amplitude and connectivity of neural oscillations, using state-of-the-art analysis methods. We found that nonlinear frequency band-specific changes in these metrics occur over the course of patients' recovery, and that these changes align well with the metrics' frequency band-specific diagnostic value. Strikingly, we found that connectivity during a single task-free EEG measurement predicted the level of patient recovery approximately 3 months later with a high level of accuracy. These results identify task-free EEG amplitude and connectivity as reliable diagnostic and prognostic markers of DOC, which can be inexpensively acquired at bedside and are completely independent of the patients' neurocognitive abilities. Furthermore, our results suggest that the preservation of reverberant thalamo-cortical interactions predicts later reemergence of consciousness, and thus yield new insights into the neural mechanisms underlying recovery following brain injury.

6.2 Materials and Methods

Participants. Sixteen patients (12 male) with severe brain injury, who participated in an 'Early Intensive Neurorehabilitation Programme' (Eilander et al., 2005) between November 2002 and January 2004, were included in the study. Age at the time of injury ranged from 5.5 to 25.2 years (M = 16.7 years; SD = 4.8). Time since injury at admission ranged from 44 to 136 days (M = 71.3 days; SD = 22.6). All but six patients suffered from traumatic brain injury caused by traffic accidents. Patients participated in the programme for 45 to 197 days (M = 103.8 days; SD = 37.6). See Table 1 for a detailed description of the patients' characteristics.

A healthy control group consisted of 16 individuals (8 male), aged from 5.8 to 25.2 years (M = 16.9; SD = 5.8). Patients and controls did not differ in age (t(15) = 0.71, p = 0.5). All patients and the healthy control group participated in this study following informed consent given by one of the parents, a legal representative or partner (patients and controls younger than 16 years), or by themselves (controls of 16 years or older). The study was approved by a medical ethics committee (METTOP).

Table 1. Patient characteristics.

P	Ms	M/F	Age	Cause	Initial CT scan (s)*	GCS	T1	T2	T3	LoC1	LoC-discharge
1	14	M	19,9	explosion	Skull fracture, diffuse swelling, high intracranial pressure, intracerebral and contusion haemorrhages, right (sub)cortical, left cortical, and brainstem lesion.	4	25	76	197	3	8
2	6	F	22,1	medulloblastoma	Encephalopathy	?	0	57	70	6	8
3	5	M	5,5	near-drowning	Hypoxia/anoxia, oedema, ischemia, atrophy, diffuse axonal injury	?	6	56	63	3	2
4	2	M	17,6	traffic accident	Epidural haematoma (right).	2t	72	80	139	3	2
5	7	M	6	near-drowning	Hypoxia/anoxia, diffuse axonal injury	3t	8	56	83	3	6
6	8	M	20,8	traffic accident	Punctual haemorrhages, intracerebral haemorages, contusion haemorrhages, atrophy, diffuse axonal injury	4	35	60	105	1	2
7	4	M	15,4	traffic accident	Skull fractures, arachnoid haemorrhages, contusion and punctual haemorrhages (right frontal, temporal, parietal), diffuse swelling.	4	33	136	112	2	5
8	7	M	20,4	traffic accident	N.A.	5	28	69	82	5	7
9	3	M	25,2	traffic accident	Skull fracture, oedema and punctual haemorrhages (cortical), diffuse swelling, and diffuse white matter lesions.	2t	65	64	77	2	7
10	4	M	8,4	cerebral haemorrhages	Intraventricular and intracerebral haemorrhages, left cortical.	2t	33	81	119	3	7
11	8	F	18,8	traffic accident	Oedema, ischemia, high intracranial pressure, diffuse swelling.	3	29	49	115	4	6
12	3	M	17,5	traffic accident	Oedema, intraventricular and intracerebral haemorrhages, focal lesions (subcortical, brainstem), diffuse white matter lesions.	4	13	44	92	3	8
13	7	M	21,8	traffic accident	Puntual haemorrhages, intraventricular haemorrhage (left), diffuse swelling, diffuse axonal injury.	5	26	71	105	2	3
14	5	F	15,7	traffic accident	Subarachnoid haemorrhage (right), high intracranial pressure, oedema (right subcortical and brainstem).	4	30	60	99	2	6
15	9	M	17,2	traffic accident	Intraventricual haemorrhages (bilateral), multiple punctual haemorrhages, Large haemorrhage in basal ganglia, and right frontal, oedema (mainly left perventricular white matter).	3	62	80	157	2	5
16	4	F	15,2	pneumonia + sepsis shock	Hypodensity in basal ganglia and cortical temporoparietal, anoxia, cortical and cerebellar atrophy, diffuse white matter lesion.	3	57	102	45	2	4

P = patient; Ms = participated measurements; F = female; M = male; Age = age at injury; * = diagnoses based on the medical reports of the acute phase; GCS = Glasgow coma scale at admission hospital; t = endotracheal tube; T1 = time at intensive care unit in days; T2 = time before admission to Rehabilitation Centre Leijpark in days; T3 = programme duration Rehabilitation Centre Leijpark in days; LoC1 = level of consciousness during the first EEG measurement; LoC–discharge= level of consciousness at discharge.

Observation scale. We categorized the patients' LoC based on the definitions described by 'the International Working Party on the Management of the Vegetative State' (Andrews, 1996), and the Aspen Neurobehavioural Conference (Giacino, 1997; Giacino et al., 2002). The categorization describes a comatose state, three vegetative sub-states, three nonvegetative sub-states, and a conscious state (see Supplementary Table 1 for a detailed classification scheme). This classification scale, now named the Post-Acute Level of Consciousness scale (PALOC-s), has a high reliability and validity (Eilander et al., 2009). In a second step this classification was reduced to four levels: levels 1, 2, and 3 were defined as UWS, levels 4, 5, and 6 as MCS, levels 7 and 8 as exit from MCS (eMCS) or conscious state.

Procedure. Nine days after a patient was admitted to the treatment programme the first measurements took place. Patients were examined while they were lying in a bed in a quiet room with a constant temperature (23 ± 1 °C). Every two weeks the EEG measurement of 3 minutes took place at the same time of the day (between 10:30 a.m. and 11:30 a.m).

Every two weeks the rehabilitation physician determined the LoC based on the categories described in Supplementary Table 1. These assessments were performed until the patient was discharged from the programme. The programme was completed when 1) a patient qualified for regular rehabilitation because of recovery of consciousness and cognitive abilities, or 2) a patient did not show any recovery for a period of at least six weeks. These different recovery courses led to a variation in time span of the patients' participation in the study and in the number of measurements.

EEG collection and preprocessing. Brain activity was recorded using actively shielded pin-electrodes, by means of the ActiveTwo System (BioSemi, The Netherlands) at a sampling rate of 2 kHz. The electrodes were placed by using a head cap and electrode gel (Parker Signa) according to the 10/20 system, at F3, Fz, F4, C3, Cz, C4, Pz, and Oz. Horizontal EOG was recorded from two electrodes placed at the outer canthi of both eyes. Vertical EOG was recorded from electrodes above and below the two eyes.

We used functions from the EEGLAB toolbox (Delorme and Makeig, 2004b) and custom MATLAB code to preprocess the EEG data. First, EEG data were down-sampled the data to 1 kHz to speed up computation and rereferenced off-line to the average of the mastoid electrodes. Next, we removed line noise by applying a notch filter (50 Hz), and removed any additional high-frequency noise (e.g., harmonics of line noise) by applying a low-pass filter at 100 Hz. Additionally, we removed slow drifts related to changes in galvanic skin properties using a high-pass filter with a 0.5-Hz cut-off. All filters were two-way, least-squares, finite impulse response filters, and designed using the 'fir1' function in MATLAB 2012a. This type of filter does not introduce spurious phase consistency of oscillatory activity (Cohen, 2014a; van den Brink et al., 2014), which can sometimes occur with infinite impulse response filters, and so will not bias connectivity estimates. After filtering, we rereferenced the pairs of vertical and horizontal EOG channels to each other, rereferenced all scalp electrodes

to the common average, and segmented the data into non-overlapping epochs of 2 s duration.

Next, segments containing artifacts were automatically detected using three criteria: the joint probability (3.5 SD cut-off), electrode kurtosis (3.5 SD cut-off), and a voltage threshold (±100 µV). In addition, data segments containing transient muscular activity or eye-movement-related artifacts were manually selected. On average, 14% (SD 11) of the data of the patient group and 10% (SD 2) of the data of the control group was marked as artefactual. After the rejection of artefactual segments, on average 170 s (SD 30) of clean data remained for the patient group, and 163 s (SD 8) for the control group. The amount of clean data did not differ between the patients and controls ($t(90) = 0.99$, $p = 0.33$).

Frequency band-specific amplitude. For all segments of clean EEG data we computed the fast Fourier transform (FFT). To enable the comparison of values across participants, we expressed the amplitude at each frequency as a percentage of the total spectrum, separately per electrode. We produced a metric of global frequency band-specific power by averaging FFT amplitude across electrodes and across frequencies within 4 canonical frequency bands: δ (1-3 Hz); θ (4-7 Hz); α (8-15 Hz); β (16-31 Hz). In addition, we computed the ratio in amplitude between the α and δ bands, as used in earlier studies (Fellinger et al., 2011; Cheadle et al., 2014). We did not include the γ band because of controversy over the ability of surface EEG to reliably detect it (Yuval-Greenberg et al., 2008).

EEG connectivity. We used correlation of orthogonalized amplitude envelopes as our measure of EEG connectivity (Hipp et al., 2012; Siems et al., 2016). The continuous (unsegmented) data were passed through a series of band-pass filters to isolate activity within the 4 canonical frequency bands (δ, θ, α, and β, see above). We filtered the continuous data rather than segmented data to prevent the introduction of edge artifacts that would otherwise occur around the segments' outer bounds. We again used two-way, least-squares, finite impulse response filters to ensure that no phase shifts would occur. For each EEG electrode and frequency band (f), excluding the segments that were previously identified as containing artifacts, we computed the complex analytic signal (X) over time (t) using the Hilbert transform, using the following equation:

$$X(t,f) = s(t,f) + i\left(\frac{1}{\pi}\text{P.V.}\int_{-\infty}^{+\infty}\frac{s(t,f)}{t-\tau}d\tau\right)$$

where P.V. denotes the Cauchy principal value, and s denotes the band-limited EEG signal. The analytic signal was calculated using the 'hilbert' function in MATLAB 2012a.

Given their heterogeneity in aetiology, the patients most likely differed from each other as well as from the control group in terms of volume conduction. That is, the patients' cerebral architecture is compromised, and in a way that varies across patients. Thus, the point spread of brain activity across the scalp most likely varies

across patients as well. To accurately estimate connectivity across scalp electrodes, we therefore needed to account for the influence of volume conduction and differences between groups / patients therein. To do so, we orthogonalized the complex analytic signal of each electrode to that of each other electrode (Hipp et al., 2012) using the following equation:

$$Y_{\perp X}(t,f) = \text{imag}\left(Y(t,f)\frac{X(t,f)^*}{|X(t,f)|} \right), \text{ where } X,Y \in S$$

and *S* denotes the set of analytic signals of all electrodes, and * denotes the complex conjugate. $Y_{\perp X}(t,f)$ represents the signal *Y* orthogonalized to signal *X*, at time point *t* and frequency band *f.* For each frequency band and electrode pair we then computed the Pearson correlation coefficient between $\ln(|Y_{\perp X}|)$ and $\ln(|X|)$. This can be interpreted as computing the correlation between the log-transformed orthogonalized amplitude envelopes. We performed the orthogonalization and correlation in both directions, from signal *X* to *Y* and from signal *Y* to *X*, yielding two correlation coefficients per electrode pair. These correlation coefficients were subsequently averaged. In all cases where correlation coefficients were averaged, we applied Fisher's r-to-z transform prior to averaging, and subsequently applied the z-to-r transform.

For each participant, this procedure resulted in a frequency band by electrode by electrode (4 by 8 by 8) matrix of correlation coefficients that indicated the strength of connectivity between pairs of electrodes, corrected for the effect of volume conduction. Next, we computed a frequency band-specific metric of global brain connectivity by averaging across the lower triangular part of the connectivity matrices (excluding the diagonal). This indicated, for each frequency band, the average connectivity across all unique electrode pairs. We focus on global connectivity for three reasons. First, the number of statistical tests is greatly reduced by collapsing across electrode pairs, which alleviates the need for a stringent correction for multiple comparisons. Second, as noted above, there was substantial heterogeneity across patients in aetiology. By considering only global dynamics, our results are less likely to be dominated by idiosyncratically located focal disturbances in brain processing. Instead, the metric putatively reflects (pathological) connectivity that is shared by the entire cortex and thus captures processes that are pervasive in nature. Third, such shared cortical dynamics arguably reflect processes that have more profound consequences for patient recovery than localized effects (Schiff et al., 2014). All *t* tests that involved connectivity were performed on Fisher's r-to-z transformed correlation coefficients.

To confirm that the orthogonalization procedure effectively reduced spurious correlations in the amplitude envelope across EEG electrodes, we compared the orthogonalized amplitude envelope correlation with the amplitude envelope correlation that was computed on non-orthogonalized signals, separately for each group and each frequency band, using paired sample *t*-tests. For both patients and

control participants, the orthogonalization reduced the strength of connectivity significantly for all frequency bands (all p's < 0.0001). Thus, the orthogonalization was effective in reducing spurious correlations.

Linear discriminant analysis. We used linear discriminant analysis (LDA) to explore whether frequency band-specific EEG amplitude and connectivity can be used to reliably dissociate patients with DOC from healthy control participants and from each other. That is, LDA was used to establish to what extent amplitude and connectivity metrics contain diagnostic information. In addition, we used receiver operated characteristic (ROC) analysis to examine to what extent the amplitude and connectivity of individual frequency bands contributed to the classifier. Second, we explored whether EEG amplitude and connectivity also contain prognostic information by using LDA to predict each patient's chances of recovery.

We implemented the LDA with a naïve Bayes classifier, using the 'classify' function in MATLAB 2012a. The classifier fitted a multivariate normal density to each group with diagonal covariance matrix estimates ('diaglinear' selected as 'type'), and then used likelihood ratios to assign observations to groups. 'Groups' here refers to either patient / control, patient groups (UWS / MCS), or outcome measures (UWS / MCS / eMCS). 'Observations' refer to the features that the classifier relied on: FFT amplitude, connectivity, or a combination of both. For each classification, unless mentioned otherwise, we report the combination of features that presented the highest degree of classification accuracy, quantified as the percentage of participants that were correctly assigned to their respective group by the classifier. In all cases, classification was performed using a leave-one-out procedure. Specifically, we first trained the classifier on the whole group of participants minus one, and we then used this trained classifier to predict to which group the left-out participant belonged. We did this for each participant separately so that eventually we obtained a prediction for each participant based on the rest of the participants.

The statistical significance of classification accuracy was assessed using non-parametric permutation testing. For 10,000 iterations we shuffled the assignment of observations to groups, and repeated the leave-one-out procedure. In cases where we tested multiple combinations of features, we computed all possible combinations of features in each iteration of the permutation test. This resulted in an aggregate distribution of 'accuracies' under the null hypothesis, corrected for the selection of a subset of features from the total possible feature set. We then calculated a p value (corrected for multiple comparisons across features) for the observed classification accuracy as the proportion of (aggregated) null accuracies that were more extreme than the true accuracy. Similarly, we tested the significance of the ROC analyses by comparing the area under the ROC curves to null distributions generated with permutation testing.

Longitudinal analyses. We used linear mixed models (McLean et al., 1991) with maximum likelihood estimation to assess changes in spectral amplitude and connectivity over the course of patient recovery. Mixed models are ideally suited for

repeated-measures designs with a varying number of samples per participant. We tested linear, exponential, and quadratic models with random slopes and intercepts across the 3 LoCs (UWS, MCS, and eMCS), with both the participants and LoC as random factors, and amplitude / connectivity as dependent variables. In each instance of the statistical test, we selected the covariance model that minimized the Akaike information criterion (Akaike, 1974) and Bayesian information criterion (Schwarz, 1978), and therefore provided the best fit. All mixed-model analyses were conducted using SPSS Statistics 23.

6.3 Results

Global and broad-band EEG activity distinguishes patients with DOC from controls, and from each other. We collected a total of 74 task-free EEG measurements at bedside from 16 patients diagnosed with DOC, using the Post-Acute Level of Consciousness scale (PALOC-s) (Eilander et al., 2009), and an additional 16 measurements from healthy age- and education-matched control participants. Our first objective was to characterize differences in spectral activity between patients and controls. To do so, we compared global spectral amplitude and connectivity during each patient's first measurement after entering the study to healthy controls, using independent-sample t tests. The patients showed an increased amplitude of oscillations in the δ and θ bands, but reduced amplitude in the α and β bands (Figure 1A; δ: $t(30) = 2.83$, $p = 0.004$; θ: $t(30) = 2.20$, $p = 0.018$; α: $t(30) = -3.17$, $p = 0.002$; β: $t(30) = -6.14$, $p < 0.001$). The full amplitude spectrum is shown in Supplementary
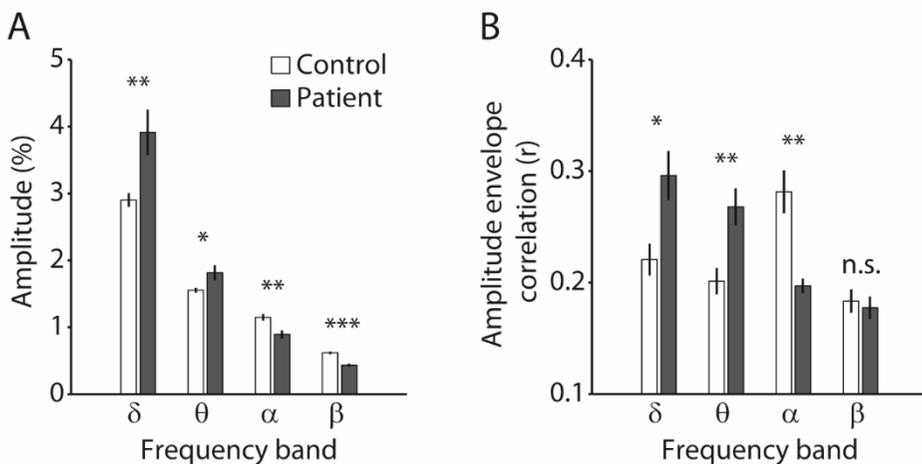
Figure 1. Global spectral amplitude and connectivity. **A)** Amplitude per frequency band for each group. **B)** Connectivity per frequency band for each group. Error bars denote the SEM. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; n.s. nonsignificant.

Figure 1. Similarly, the patients showed hypersynchronous activity in the δ and θ bands, and hyposynchronous activity in the α band (Figure 1B; δ: $t(15) = 2.51$, $p = 0.02$; θ: $t(15) = 2.95$, $p = 0.01$; α: $t(15) = -3.94$, $p = 0.001$; β: $t(15) = -0.38$, $p = 0.71$). Thus, compared to controls, the patients showed pronounced differences in both amplitude and connectivity that spanned a wide spectral range. Such global spectral disturbances in patients with DOC are consistent with earlier reports (Lehembre et al., 2012; Lechinger et al., 2013; Chennu et al., 2014; Varotto et al., 2014), and are indicative of widespread pathophysiological cortical activity.

Our next objective was to determine to what extent spectral amplitude and connectivity can aid the diagnosis of DOC at the level of individual patients. To do this, we used a naïve Bayes classifier. The classifier relied on frequency band-specific amplitude, connectivity, or a combination thereof, to predict the group of each
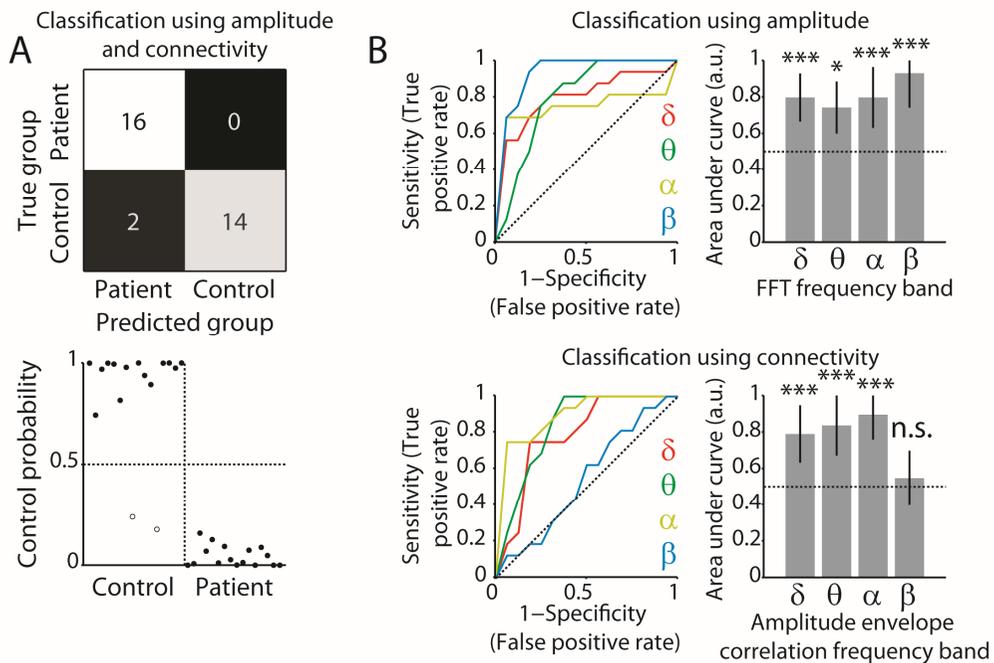
Figure 2. Classification of patients and controls. **A)** Top row, confusion matrix for classification distinguishing patients from controls, based on both amplitude (β band) and connectivity (δ, θ, α bands). Colors indicate the relative number of cases in each cell. Bottom row, associated classifier weights. Filled and open dots show correctly and incorrectly classified individuals, respectively. **B)** ROC curves and corresponding areas under the curve, indicating the extent to which each frequency band contributed to the classifier. Top row, for spectral amplitude. Bottom row, for amplitude envelope correlations. The area under the curve can be interpreted as the accuracy with which the individual participant/patient's group can be predicted based on the metric in that frequency band. The horizontal dotted line indicates chance performance. Error bars denote the 95% confidence interval of the permuted null distribution. *$p < 0.05$; ***$p < 0.001$; n.s. non-significant.

individual (patient or control). The statistical significance of classifier accuracy was assessed with permutation testing. When using amplitude in all frequency bands to distinguish the patients from participants in the control group, the classifier performed with an accuracy of 81% (26 out of 32 individuals assigned to the correct group, $p < 0.001$). Second, classification based on connectivity in the δ, θ, and α bands was also highly accurate (88%, 28 out of 32 participants correctly assigned, $p < 0.0001$). When the classifier relied on connectivity in the δ, θ, and α bands, and was additionally informed by amplitude in the β band, accuracy was highest (94%, 30 out of 32 participants correctly assigned, $p < 0.0001$). Figure 2A shows the confusion matrix for classification based on both amplitude and connectivity features. These results indicate that both spectral amplitude and connectivity can readily be used as metrics to distinguish patients from controls, but the combination of the two types of metrics yields additional information that cannot be inferred from either type of metric in isolation. ROC analysis indicated that the amplitude of all individual frequency bands contributed to the classifier, with the β band showing the highest accuracy (Figure 2B, top row). Moreover, connectivity in all but the β band contributed to the classifier (Figure 2B, bottom row).

A useful clinical diagnostic tool for the diagnosis of DOC does not only distinguish patients from controls, but also provides a reliable indication of the type of DOC within individual patients. Thus, we next set out to investigate to what extent a classifier could distinguish patients that were diagnosed with UWS from those that showed minimal signs of consciousness (MCS). These two types of DOC are most difficult to dissociate based on EEG metrics alone (Schurger et al., 2015), so classification of these two types of DOC provides a good benchmark to test the diagnostic value of amplitude and connectivity. Moreover, the longitudinal study design enabled us to sample an adequate number of measurements at these two LoC from within the patient group to be used for classification.

A classifier that relied on connectivity in the δ, θ, and α bands and amplitude in the β band, similar as used above, showed modest but above-chance-level performance in distinguishing the two groups (75% accurate, $p = 0.018$). However, accuracy improved when the classifier only used connectivity in the δ, θ and β bands as features (85% accurate, $p = 0.001$, Figure 3A). Thus, connectivity alone was most informative when distinguishing UWS patients from those patients that displayed minimal signs of consciousness. In agreement with this notion, ROC analysis showed that only connectivity in the δ, θ and β bands contributed to the classifier (Figure 3B). Control analyses ruled out patient age as a confound (see Supplementary Results). As noted above, for distinguishing patients from fully conscious control participants, the combination of amplitude and connectivity proved to be most informative. Together, these findings raise the hypothesis that changes in amplitude occur when patients transitioned from unconsciousness to consciousness, but changes in connectivity occur at the transition from UWS to MCS. To address this hypothesis, in
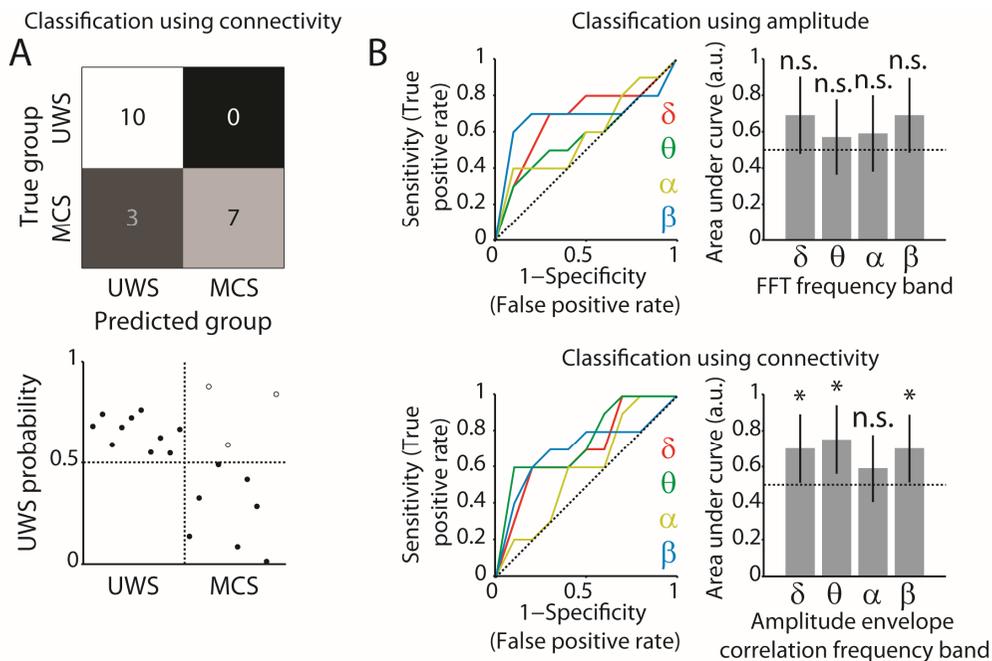
Figure 3. Classification between patient groups. **A)** Top row, confusion matrix for classification distinguishing UWS from MCS patients, based on connectivity (δ, θ, β bands). Colors indicate the relative number of cases in each cell. Bottom row, associated classifier weights. Filled and open dots show correctly and incorrectly classified patients, respectively. **B)** ROC curves and corresponding areas under the curve, indicating the extent to which each frequency band contributed to the classifier. Top row, for spectral amplitude. Bottom row, for amplitude envelope correlations. The area under the curve can be interpreted as the accuracy with which the individual participant/patient's group can be predicted based on the metric in that frequency band. The horizontal dotted line indicates chance performance. Error bars denote the 95% confidence interval of the permuted null distribution. *$p < 0.05$; ***$p < 0.001$; n.s. non-significant.

the next section we explore longitudinal changes in oscillatory amplitude and connectivity metrics across the patients' course of recovery.

Frequency band-specific amplitude and connectivity track longitudinal changes in patients' level of consciousness. Having established that spectral amplitude and connectivity can be used as reliable markers for the diagnosis of DOC, we next set out to investigate whether spectral amplitude and connectivity track the LoC over the course of patients' recovery. In the following set of analyses, we used linear mixed models to test if the individual metrics changed across the LoC. We explored linear, exponential, and quadratic changes in all frequency bands. Furthermore, we examined changes in the ratio between α and δ amplitude, as used in prior research (Fellinger et al., 2011; Cheadle et al., 2014). The amplitude and connectivity of all

frequency bands for each LoC are shown in Supplementary Figure 2 and Supplementary Figure 3.

The ratio between α and δ amplitude increased significantly across LoC ($F_{(2,23)}$ = 4.63, p = 0.021). As shown in Figure 4A, however, the data suggested that this increase was not linear over time, but instead was relatively stable for lower LoC and then exponentially increased, resulting in an overshoot compared to the control group. Consistent with this notion, including an exponential predictor in the model resulted in a significant exponential effect of LoC on α/δ amplitude ratio ($F_{(1,32)}$ = 6.31, p = 0.017), and rendered the linear effect nonsignificant ($F_{(1,40)}$ = 3.40, p = 0.073). Thus, the change in the α/δ amplitude ratio across LoC was best captured by an exponential increase instead of by a linear increase. Similarly, β amplitude increased linearly with LoC ($F_{(2,21)}$ = 3.75, p = 0.040), but an exponential model best explained the change across LoC ($F_{(1,14)}$ = 11.42, p = 0.005; Figure 4B). As opposed to a progressive increase across LoC, connectivity in the θ band showed a quadratic relationship with LoC ($F_{(1,45)}$ = 9.05, p = 0.024). Figure 4C shows that θ connectivity was low for UWS scores, increased for MCS scores, and recovered to normative levels for eMCS scores. In the following, we explore whether amplitude and connectivity also provide prognostic information.
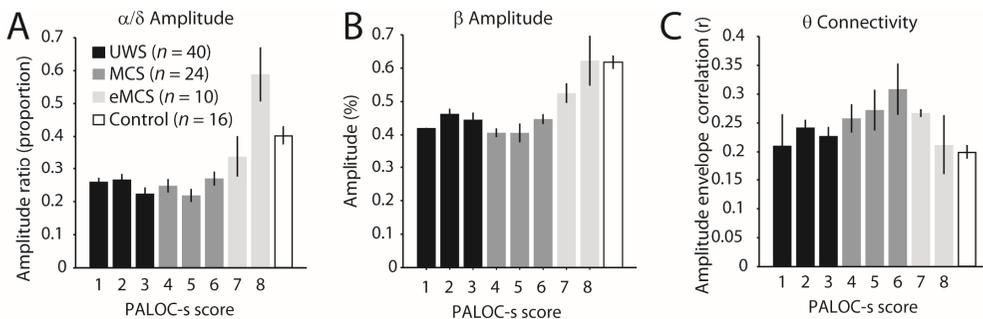
Figure 4. Longitudinal changes in EEG metrics. **A)** The ratio between α and θ amplitude increases with level of consciousness, and shows an overshoot for the patients with higher levels of consciousness. **B)** β amplitude increases with level of consciousness. **C)** θ connectivity shows an inverted-U relationship with level of consciousness. Controls are shown for visual comparison. Error bars denote the SEM. PALOC-s: Post-Acute Level of Consciousness scale.

Brain-wide connectivity predicts patient recovery. Thus far we have shown that global amplitude and connectivity can be used as markers for the diagnosis of DOC. Furthermore, frequency band-specific changes in these metrics occur across the course of patient recovery. We next asked if amplitude and connectivity can also be used as reliable prognostic markers. That is, can amplitude and connectivity during a single task-free EEG measurement, conducted upon the patients' admission to the

study, be used to predict the patients' level of recovery? To do this, we used a classifier to predict each patient's outcome diagnosis at the point of discharge from the rehabilitation center. The outcome diagnosis was either UWS, MCS, or eMCS, and thus chance-level classification accuracy was 33%.

When amplitude was used as features, the α band alone yielded the highest classification accuracy (62% accurate, p = 0.014). In line with the observations made above, amplitude dissociated relatively well between MCS and eMCS, but performed poorly at dissociating the lower LoC outcome measures (Figure 5A). However, connectivity in the θ, α and β bands proved to be more reliable features, resulting in an accuracy of 75% (p < 0.001, Figure 5B). Connectivity in isolation also out-performed a classifier that relied on both amplitude and connectivity (69% accurate, p = 0.003). Thus, based on connectivity during a single task-free EEG measurement, conducted upon the patents' admission to the study, it was possible to make a prognosis for patient recovery ~3 months later with 75% accuracy. The possibility that variation across patients in LoC at the time of measurement was driving the classifier cannot account for our findings, because LoC during the first measurement and outcome score were not significantly correlated (r = 0.33, p = 0.21). Additional control analyses ruled out patient age as a confound (see Supplementary Results). These
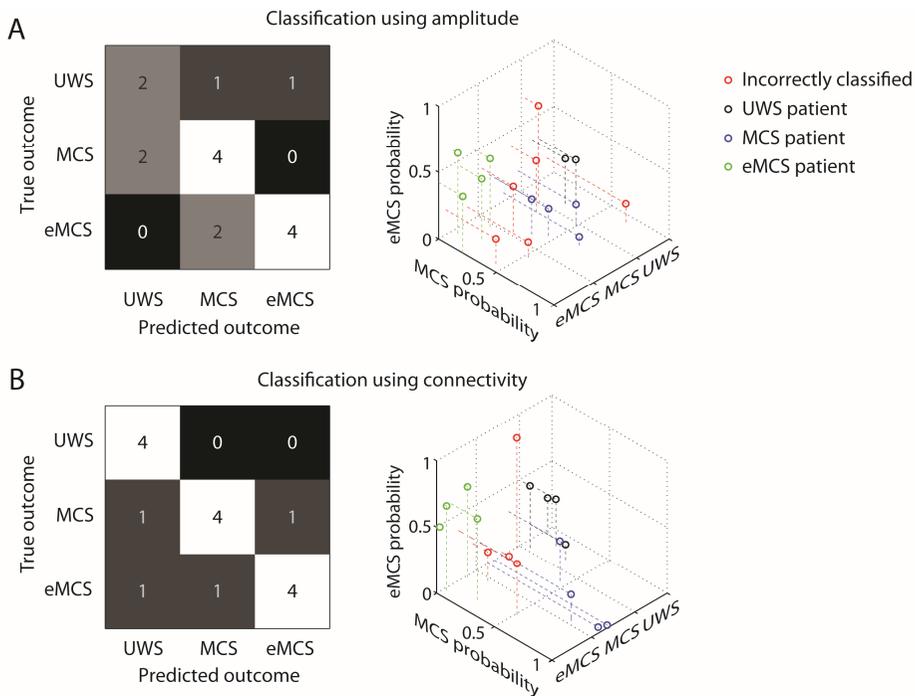
Figure 5. Classification of outcome measures. **A)** Confusion matrix for classification using α amplitude (left), and associated classifier weights (right). **B)** Confusion matrix for classification using θ, α and β connectivity (left), and associated classifier weights (right). Shades of grey and numbers in the confusion matrices indicate the relative number of cases in each cell.

Chapter 6

results identify EEG connectivity as a reliable marker of recovery from DOC. As we discuss below, these results also tentatively point to neural mechanisms that may underlie recovery from DOC.

6.4 Discussion

In the present study, we examined if and how task-free spectral EEG amplitude and connectivity metrics change over the course of patient recovery, following severe brain injury. Moreover, we examined if these metrics can be used to predict the current (diagnosis) and future (prognosis) LoC of individual patients. Our first key finding is that amplitude and connectivity can reliably be used as diagnostic markers of DOC (Figures 2 and 3). Dissociating patients from healthy control participants worked best when relying on δ, θ, and α band connectivity, and amplitude in the β band. Dissociating UWS from MCS patients was most successful based on δ, θ and β band connectivity. Our second key finding is that task-free spectral amplitude and connectivity do not vary monotonically across LoC, but instead show nonlinear dynamics (Figure 4). Specifically, we found that amplitude in the β band, and α/δ amplitude ratio, increased exponentially across LoC, while θ band connectivity showed an inverted-U relationship with LoC. Finally, our findings show that connectivity metrics (θ, α and β bands) are highly robust markers of patient prognosis (Figure 5B).

The exponential increase in amplitude (ratio) across LoC is broadly consistent with an account that posits that consciousness recovers only after neural function crosses a critical threshold level (Bagnato et al., 2013). Moreover, the inverted-U shaped relationship between LoC and θ connectivity may explain why amplitude and connectivity provide complementary diagnostic information when dissociating patients from controls, whereas connectivity alone is most informative when dissociating UWS from MCS patients. Whereas amplitude (ratio) is stable for UWS / MCS and then increases, θ connectivity deviates most strongly in MCS, but appears normative for UWS and eMCS. Accordingly, the three LoC are each marked by a unique spectral fingerprint (Siegel et al., 2012) that is apparent only when both amplitude and connectivity are considered. This suggests that a successful distinction between the three LoC requires multivariate classification, as we have used here.

A recent account has highlighted the central role of the thalamus in the regulation of arousal through its excitatory connections to the cortex and striatum (Schiff, 2010; Schiff et al., 2014). According to this account, pathologically elevated slow-wave *amplitude* indicates damage in the thalamo-cortico-thalamic loop. Such damage in the thalamo-cortical system causes a loss of excitatory drive to the cortex and consequently results in a general 'slowing down' of cortical rhythms (Giacino et al., 2014; Schiff et al., 2014), consistent with our findings (Figure 1a) and animal models of cortical deafferentation (Timofeev et al., 2000; Lemieux et al., 2014). Importantly,

combined with the finding of absent pathologically increased *connectivity* in UWS (Figure 4C), the pattern of results suggests that a lack of central thalamic coordination of oscillatory activity across the cortex in UWS, potentially due to the loss of excitatory drive from the thalamus to the cortex, or vice versa. Relatedly, the inverted-U shaped pattern of θ connectivity across LoC may explain why connectivity metrics in particular were most informative about later patient recovery. Elevated θ connectivity in UWS patients might be indicative of the relative sparing of projections within the thalamo-cortico-thalamic circuit, and hence the potential for recovery of reverberant excitatory drive and associated high-frequency activity. This notion is consistent with findings that the recovery of consciousness is paralleled by a restoration in thalamo-cortical interactions (Laureys et al., 2000), the spread of cortical activity following transcranial magnetic stimulation dissociates UWS from MCS patients (Rosanova et al., 2012; Casali et al., 2013), and thalamic stimulation can facilitate behavioral responsiveness (Schiff et al., 2007).

Notwithstanding the potential diagnostic and prognostic utility of amplitude and connectivity metrics, some limitations of the present study should be acknowledged. First, the classifier's false negative rate for the purpose of diagnosis as well as for prognosis was higher than the false positive rate (Figure 2, Figure 3, Figure 5), indicating that the classifier was somewhat pessimistic. Ideally, the false positive and false negative rates would be balanced. False negatives in diagnosis based on behavioral criteria have been attributed in part to temporal fluctuations in the patient's arousal state (Piarulli et al., 2016). This may also be the case for the neural markers employed here. This potential problem could be resolved by close monitoring of ultradian fluctuations in the patients' arousal state (Piarulli et al., 2016). Second, the nonlinear variations in amplitude and connectivity observed here (Figure 4) appear to be at odds with earlier reports of monotonic changes across LoC in entropy metrics of cortical interactions (King et al., 2013; Sitt et al., 2014). This discrepancy may be explained by the fact that in these studies patients were presented with auditory stimuli, which could evoke synchronous cortical states. Alternatively, the here employed measure of connectivity (orthogonalized amplitude envelope correlations) may reflect qualitatively different network interactions than the metrics used in previous studies. Third, it should be noted that our findings may not generalize to DOC with different aetiology (e.g., due to neurodegenerative disease). Finally, our findings strongly call for independent replication, preferably with a larger sample size, to determine the specific combination of features that yields the most accurate diagnosis and prognosis.

In conclusion, diagnosis and prognosis based on amplitude and connectivity from task-free EEG measurements is feasible. These measures can be acquired inexpensively, with low electrode density, at bedside, and are fully independent of the patients' neurocognitive abilities. Our longitudinal findings in the amplitude domain are consistent with an existing account that proposes that neural function crosses a threshold level prior to the reemergence of consciousness following DOC (Bagnato et

al., 2013). Furthermore, our findings in the connectivity domain lend support to a recent account that posits that dysfunction in the thalamo-cortical system underlies DOC (Schiff, 2010; Schiff et al., 2014), and further suggest that neural signatures of thalamo-cortical interactions are predictive of patient recovery. A rigorous assessment of the pathophysiological mechanisms underlying DOC may open the door to diagnostic taxonomies that are independent of behavioral criteria, and facilitate early targeted interventions that are tailored to the individual patient's needs.

6.5 Supplementary Materials

Ruling out participant age as a confound in classification analyses. Because of the heterogenous age distribution of our patient sample, and given that spectral activity has been reported to change across lifespan (van Albada et al., 2010), it is conceivable that patient age was driving the classification analyses. To rule out this possibility, we correlated participant age at the time of measurement with the classifier's graded output (i.e. the weighted posterior probability that the classifier assigned to each patient, indicating the classifier's estimated likelihood that the patient belonged to a particular group).

If patient age was driving the classifier's performance when distinguishing UWS from MCS participants (Figure 3A in the main text), then participant age should significantly correlate with the classifier's graded output. However, age and classifier output did not correlate significantly across the full patient sample ($r = 0.007$, $p = 0.98$), across UWS patients alone ($r = 0.22$, $p = 0.54$), or across MCS patients ($r = -0.12$, $p = 0.75$). Thus, patient age did not drive the performance of the classifier that distinguished UWS from MCS patients.

Similarly, to rule out that patient age was driving the classifier that most accurately predicted the patient outcome scores (Figure 5B in the main text), we correlated participant age with the classifier's assigned likelihood of each group, and found no significant effects (UWS: $r = -0.09$, $p = 0.74$; MCS: $r = -0.15$, $p = 0.58$; eMCS: $r = 0.30$, $p = 0.27$), thus ruling out age as a confound in the classification analyses.

Supplementary Table 1. Post-Acute Level of Consciousness Scale (PALOC-s)

Global Level	PALOC Score	Description of the levels
Coma		Eyes are closed all the time. No sleep-wake cycles present.
	1	All mayor body functions such as breathing, temperature control, or blood pressure can be disturbed. Generally, no reactions are noticed after stimulation. Sometimes reflexes (stretching or flexing) can be observed as a reaction when strong pain stimuli have been applied. No other reactions present.
LOC 1 ***Uresponsive wakefulness syndrome (UWS) / Vegetative state***		Patient has some sleep-wake cycles, but no proper day-night rhythm. Most of the body functions are normal. No further ventilation is required for respiration.
	2	**Very little response (hyporesponsive)** Generally no response after stimulation. Sometimes delayed presentation of reflexes is observed.
	3	**Reflexive state** Often stimuli result in massive stretching or startle reactions, without proper habituation. Sometimes these reactions evaluate into massive flexing responses. Roving eye movements can be seen, without tracking. Sometimes grimacing occurs after stimulation.
LOC 2 ***Minimally conscious state (MCS)***	**4**	**High active level and/or reactions in stimulated body parts** Generally spontaneous undirected movements. Retracting of a limb following stimulation. Orienting towards a stimulus, without fixating. Following moving persons or objects, without fixating.
	5	**Transitional state** Following and fixating of persons and objects. Generally more directed reactions to stimuli. Behaviour is automatic, i.e. opening of the mouth when food is presented, or reaching towards persons or objects. Sometimes emotional reactions are seen such as crying or smiling towards family or to specific (known) stimuli.
	6	**Inconsistent reactions** Sometimes, but not always, obeying simple commands. Totally dependent. Patient has profound cognitive limitations; neuropsychological testing is impossible. Level of alertness is fluctuating, but in general low.
LOC 3 ***Exit from MCS (eMCS) / Conscious state***	**7**	**Consistent reactions** Patient obeys simple commands. The level of alertness is high and stable. Many cognitive disturbances remain. Patient is totally dependent.
	8	Patient is alert and reacts to his/her environment spontaneously. Functional understandable mutual communication is possible, sometimes with technical support. As yet, cognitive and behavioural disturbances can be present.
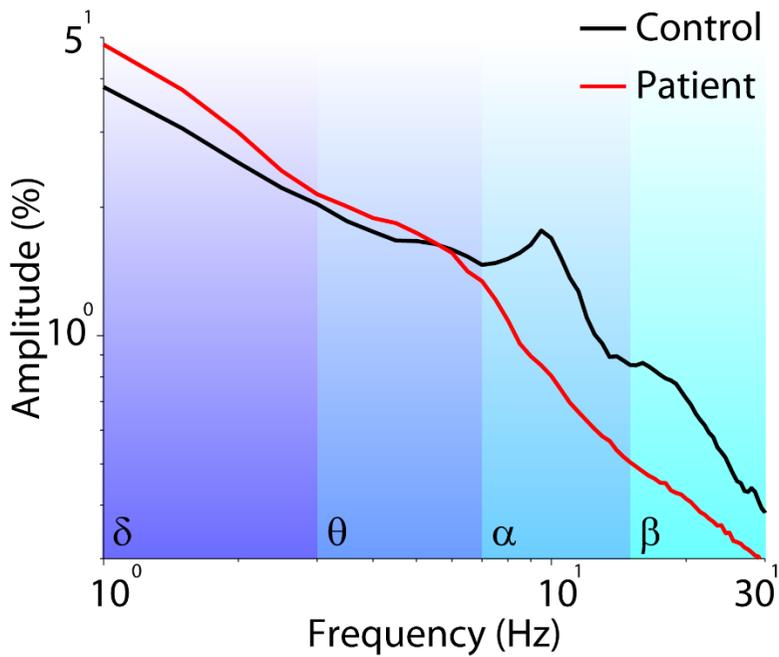
Supplementary Fig. 1. Log-plot of the global spectral amplitude for patients and control participants. Shaded areas indicate the different frequency bands.
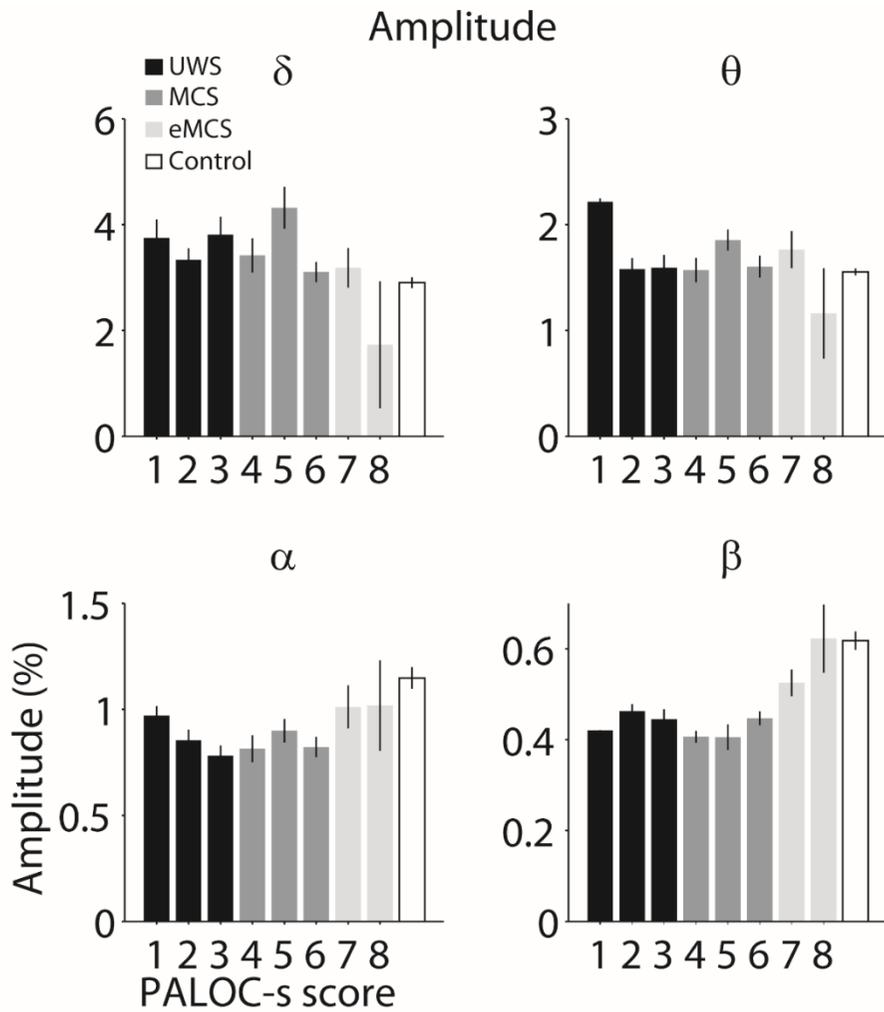
Supplementary Fig. 2. Amplitude per frequency band and level of consciousness. Error bars denote the SEM. PALOC-s: Post-Acute Level of Consciousness scale.
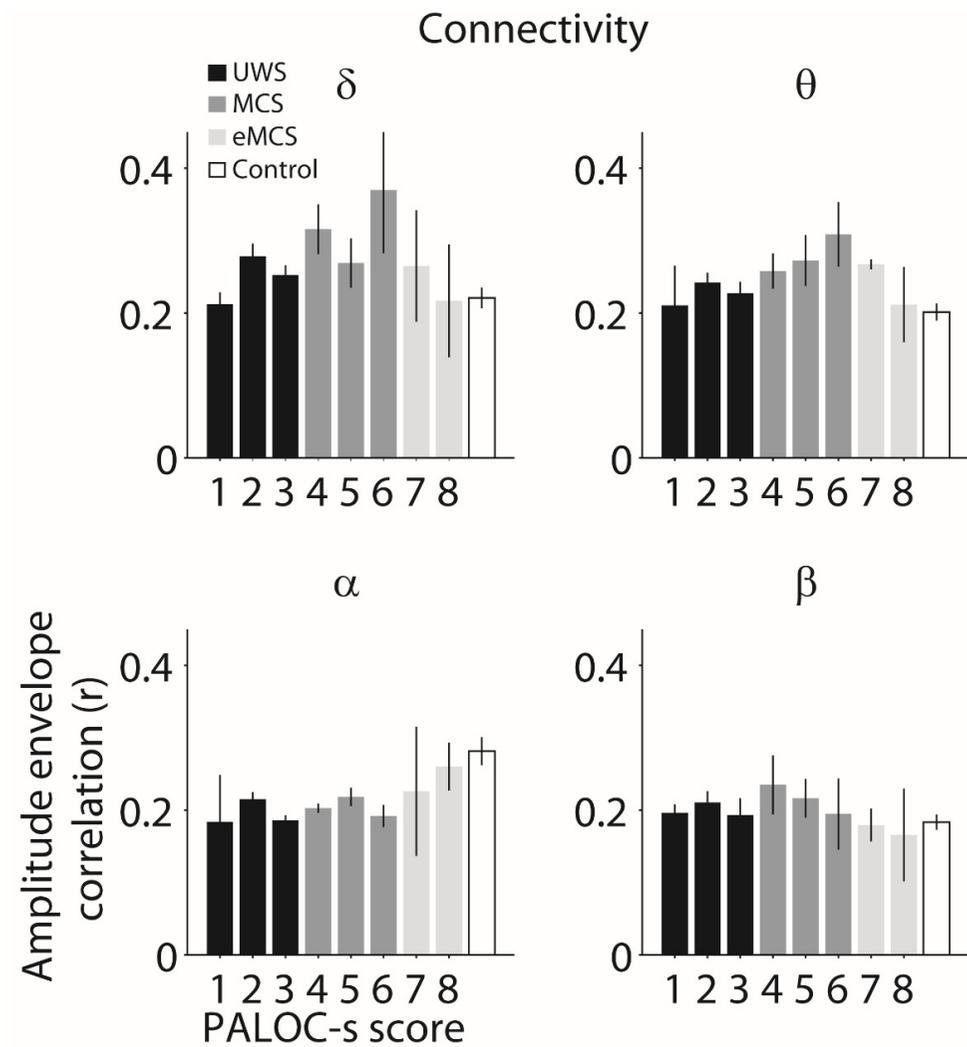
Supplementary Fig. 3. Connectivity per frequency band and level of consciousness. Error bars denote the SEM. PALOC-s: Post-Acute Level of Consciousness scale.

7. Dutch Summary

7.   Nederlandse samenvatting

In dit proefschrift zijn diverse thema's aan bod gekomen, zoals '*brain state*', neuromodulatie, en aandacht. Wat volgt is een beknopt overzicht van deze thema's, gevolgd door een samenvatting van dit proefschrift in de context van deze thema's.

7.1   Het locus coeruleus-noradrenerge systeem

De locus coeruleus (afgekort 'LC') is een kleine kern diep in de hersenstam, die zijn naam (letterlijk vertaald: 'blauwe kern') ontleent aan zijn kleur. In het volwassen brein bestaat de LC uit ongeveer 35.000 neuronen (hersencellen) in elke hersenhelft, en is ongeveer zo groot als een korrel rijst. Ondanks zijn grootte projecteert de LC naar vrijwel elk ander deel van het brein, waar het de stof noradrenaline (afgekort 'NE', afgeleid van het Engelse *norepinephrine*) afgeeft. Als gevolg van zijn wijdverbreide projecties is de LC de primaire bron van NE in het centrale zenuwstelsel, en speelt het LC-NE systeem een fundamentele rol bij diverse cognitieve functies.

Onderzoek met primaten heeft aangetoond dat de LC twee *modi operandi* heeft: in de 'fasische' modus vuren de neuronen van de LC gezamenlijk, kort, en op een manier die een sterke temporele correspondentie vertoont met gedragsmatige responsen. In de 'tonische' modus vuren de neuronen van de LC geleidelijk en langdurig, zonder directe temporele relatie met gedrag. Echter, over langere tijdsperiodes is er wel degelijk een relatie tussen tonische LC activiteit en gedrag: periodes van weinig tot geen tonische LC activiteit gaan samen met onoplettendheid of slaperigheid; periodes van gemiddelde tonische LC activiteit gaan samen met goede taakuitvoering; en periodes van hoge tonische LC activiteit gaan samen met afleidbaar gedrag of stress.

In tegenstelling tot andere neurotransmitters heeft NE geen eenduidig effect op de neuronen die het beïnvloedt, maar verandert NE het effect van andere neurotransmitters op deze neuronen, een proces genaamd neuromodulatie. Een eigenschap van neuromodulatie door NE is dat de vuurkarakteristieken van neuronen veranderen op een manier die er toe leidt dat dominante vuurpatronen nog dominanter worden, en minder dominante vuurpatronen zwakker. Dit proces staat bekend als *gain* modulatie, en vormt een belangrijk ingrediënt voor de *adaptive gain* theorie van LC-NE functie.

Deze theorie stelt dat het LC-NE systeem, door zijn effect op *gain*, de balans kan reguleren tussen uitbuiting van beloning van de huidige taak en exploratie van andere, mogelijk meer belonende, taken. Omdat fasische LC activiteit alleen plaatsvindt na stimuli die relevant zijn voor het uitvoeren van de huidige taak, zorgt fasische LC activiteit er voor (door middel van het effect van NE op *gain*) dat deze stimuli beter verwerkt worden, en ten koste van doel-irrelevante stimuli. Omgekeerd zorgt verhoogde tonische LC activiteit voor een onspecifieke verwerking (in de tijd) van stimuli, resulterend in een minder strenge selectie van sensorische informatie, en

daarmee mogelijke exploratie van andere taken. Dus door de balans tussen fasische en tonische LC activiteit te reguleren, kan de LC veranderingen in gedrag teweegbrengen in reactie op veranderingen in de omgeving.

Naast de modulatie van *gain* heeft onderzoek, voornamelijk in kreeftachtigen, aangetoond dat NE de totale 'activiteitsstaat' van een neuraal netwerk op diverse manieren kan hervormen. Afhankelijk van de concentratie van NE, de aanwezigheid van andere neuromodulatoren, en de huidige staat van het neurale netwerk, kan NE leiden tot het stilvallen of juist verhogen van de activiteit binnen het netwerk. Deze eigenschap vormt de basis van een andere belangrijke theorie over de functie van het LC-NE systeem: *network reset*. In tegenstelling tot adaptive gain stelt deze theorie dat de afleidbaarheid tijdens periodes van verhoogde LC activiteit wordt veroorzaakt door herhaalde '*network resets'*: abrupte veranderingen in de staat van het brein en bijkomende cognitieve *shifts*.

Breed genomen convergeren de bestaande theorieën over de functie van het LC-NE systeem (zoals *adaptive gain* en *network reset,* maar ook andere theorieën die hier niet zullen worden besproken) op de conclusie dat de LC veranderingen in gedrag reguleert, op basis van veranderingen in de omgeving. Echter, waar deze theorieën voornamelijk van elkaar verschillen, is de manier waarop het LC-NE veranderingen in gedrag implementeert op een neuraal niveau. Zoals hieronder besproken: sensorische informatie die een persoon informeert over de staat van de omgeving wordt niet eenduidig verwerkt door het brein, maar wordt sterk beïnvloed door spontane, intrinsieke, hersenactiviteit. Recent onderzoek suggereert dat NE mogelijk een belangrijke rol speelt bij het vormen van deze spontane hersenactiviteit en de interactie daarvan met de verwerking van sensorische informatie. Dit biedt mogelijk nieuwe inzichten in de manier waarop het LC-NE systeem gedrag reguleert.

7.2   De samenkomst van *brain state*, neuromodulatie, en cognitie

Hersenactiviteit volgt niet simpelweg uit de verwerking van externe (sensorische) informatie, maar ontstaat door een complexe en non-lineaire interactie tussen sensorische informatie en spontane – intern gegenereerde – hersenactiviteit. De staat van deze spontane hersenactiviteit, en de interactie met sensorische informatie, is niet statisch, maar fluctueert over de tijd.

Een bekend voorbeeld van zulke fluctuaties in de staat van het brein is het circadiaan ritme. Tijdens de diepe stadia van slaap schommelt hersenactiviteit ritmisch tussen periodes waarin neuronen gezamenlijk vuren, en periodes van bijna complete absentie van neuronaal vuren. Deze synchrone staat gaat samen met een verminderde responsiviteit van het brein op sensorische informatie, en vormt een sterk contrast met de waaktoestand, waarin neuronen voornamelijk asynchroon vuren en waarin men sensorische informatie over het algemeen normaal verwerkt. Recentelijk is aangetoond dat binnen de waaktoestand ook afwisselingen plaatsvinden tussen de

synchrone en asynchrone staat, zij het op een minder prominente manier dan over de gehele circadiane cyclus.

Onze capaciteit om doelrelevante sensorische informatie te selecteren en irrelevante informatie te negeren, genaamd *top-down* 'aandacht', vertoont soortgelijke neurale karakteristieken als de asynchrone toestand van het brein. Schommelingen tussen synchrone en asynchrone hersenactiviteit gaan samen met schommelingen in de responsiviteit van neuronen die sensorische informatie verwerken, en met schommelingen in onder andere het reactievermogen. Dit heeft geleid tot het voorstel dat aandacht mogelijk wordt gereguleerd door dezelfde neurale mechanismes als veranderingen in de staat van het brein.

De activiteit van de LC co-fluctueert ook met de hierboven beschreven schommelingen in de staat van het brein. Tijdens de transitie van de waaktoestand naar slaap neemt de activiteit van de LC geleidelijk af, en valt deze nagenoeg stil tijdens de diepere stadia van slaap. Daarnaast gaan overgangen van de synchrone staat naar asynchrone staat van het brein tijdens de waaktoestand samen met de afgifte van NE in de hersenschors, en met een toename in de diameter van de pupil. De grootte van de pupil co-fluctueert ook met de activiteit van de LC. Daarnaast gaan al deze schommelingen in *brain state*, pupil, en LC activiteit, samen met schommelingen in eigenschappen van gedrag zoals het reactievermogen, accuratesse, en afleidbaarheid.

Het bovenstaande geeft aan dat de staat van het brein, neuromodulatie, en cognitieve capaciteiten zoals aandacht, sterk met elkaar verbonden zijn. Deze samenkomst vormt het centrale thema van dit proefschrift. Wat volgt is een korte samenvatting van de empirische hoofdstukken van dit proefschrift waarin elk hoofdstuk wordt besproken in de context van *brain state*, neuromodulatie, cognitie, of een combinatie van deze sub-thema's.

7.3   Samenvatting van dit proefschrift

7.3.1   Hoofdstuk 2: Post-Error Slowing as a Consequence of Disturbed Low-Frequency Oscillatory Phase Entrainment

Een van de meest voorkomende bevindingen in reactietijd (RT) taken is dat RTs langer worden na het maken van fouten. Dit fenomeen heet *post-error slowing*, en treedt op onder verschillende taakcondities en in diverse respons-modaliteiten. Er zijn verschillende voorstellen gedaan over de aard van *post-error slowing*, waaronder de strategische aanpassing van gedrag met als doel het voorkomen van verdere fouten, als wel interferentie in stimulusverwerking als gevolg van de fout.

Zoals besproken in de voorgaande sectie, beïnvloedt de staat van het brein ons vermogen om relevante sensorische informatie te selecteren en er op te reageren. Voorgaand onderzoek suggereert dat onder omstandigheden waarin stimuli ritmisch

worden aangeboden de staat van ons brein dynamisch wordt aangepast aan dit ritme (zogenaamde *entrainment*), zodat stimuli geanticipeerd worden en optimaal worden verwerkt. Op deze manier faciliteert entrainment snelle en accurate gedragsmatige responsen op stimuli.

In hoofdstuk 2 van dit proefschrift testen we de nieuwe hypothese dat *post-error slowing* mogelijk veroorzaakt wordt door een verstoring van *entrainment.* Deze hypothese hebben we getest door proefpersonen een moeilijke RT taak te laten uitvoeren onder tijdsdruk, terwijl we ritmische EEG activiteit maten. Onze resultaten laten zien dat de staat van het brein zich inderdaad aanpast aan het ritme van de stimuluspresentatie, en dat de fase van deze ritmische hersenactiviteit de RT voorspelt. Ook tonen onze resultaten aan dat de fase van ritmische hersenactiviteit tijdelijk verstoord is na het maken van een fout, en dat de mate van deze verstoring de mate van *post-error slowing* voorspelt.

Onze resultaten zijn consistent met de *orienting account* van *post-error slowing*, die stelt dat fouten, als gevolg van het feit dat ze verrassende gebeurtenissen zijn, resulteren in de tijdelijke heroriëntatie van aandacht weg van de taak, en als gevolg van deze heroriëntatie, langere RTs op volgende trials. Een interessante observatie is dat de LC fasisch vuurt in reactie op verrassende stimuli, en er bestaan in de literatuur theoretische voorstellen die *post-error slowing* linken aan activiteit in het LC-NE systeem. Daarnaast stelt de *network reset* theorie dat activatie van het LC-NE systeem (in de huidige context, als gevolg van fouten) leidt tot een aanpassing van gedrag. We speculeren dat het mogelijk is dat *entrainment*, en de daaruit volgende responsstaat, verstoord wordt door een *orienting* respons in het LC-NE systeem. Deze verstoring hindert de prestatie op de taak op de korte termijn, maar is mogelijk adaptief op de langere termijn.

7.3.2   Hoofdstuk 3: Catecholaminergic Neuromodulation Shapes Intrinsic MRI Functional Connectivity in the Human Brain

Spontane hersenactiviteit correleert tussen hersengebieden. De ruimtelijke structuur van deze correlaties verandert dynamisch met veranderingen in de staat van bewustzijn en taakcondities. In hoofdstuk 3 van dit proefschrift testen we de hypothese dat fluctuaties in de sterkte van deze correlaties geïnduceerd worden door het LC-NE systeem. Proefpersonen kregen in deze dubbelblind uitgevoerde studie op twee aparte sessies placebo en atomoxetine. Atomoxetine remt de NE transporter, en verhoogt daarmee de beschikbaarheid van NE in het brein. Vervolgens onderzochten we wat het effect van deze verhoogde beschikbaarheid van NE was op de sterkte van spontane correlaties in hersenactiviteit, gemeten met functionele MRI.

 Zoals hierboven besproken heeft NE invloed op *gain*. Als gevolg daarvan vindt er een facilitatie plaats van het doorgeven van neurale signalen over het gehele brein. Computermodellen van neurale netwerken hebben gesuggereerd dat deze toename in het doorgeven van neurale signalen moet leiden tot een toename in de sterkte van

zowel positieve als negatieve correlaties in hersenactiviteit, alsmede de mate van clustering van deze correlaties. Op basis van deze voorspellingen, en op basis van pupil-gerelateerde bevindingen, verwachtten we dat atomoxetine de sterkte en de mate van clustering van spontane correlaties zou moeten verhogen. Daarnaast, gegeven de veronderstelde homogeniteit van de projecties van de LC over het brein heen, verwachtten we dat een atomoxetine-geïnduceerde toename in sterkte van correlaties ook homogeen zou moeten zijn over het brein.

Echter, tegen de verwachting in vonden we dat atomoxetine de sterkte van correlaties *verminderde*, en op een ruimtelijk gestructureerde manier. De afname in de sterkte van correlaties vertoonde een gradiënt over het brein heen, was afhankelijk van de *baseline* sterkte van correlaties, en was het sterkste voor connecties tussen verschillende netwerken in het brein. Onze bevindingen zijn de eersten die aantonen dat neuromodulatie de sterkte en ruimtelijke structuur van intrinsieke correlaties vormgeeft. De onverwachte afname in de sterkte van deze correlaties is een eerste indicatie dat het effect van neuromodulatie op spontane hersenactiviteit afhangt van de staat van het brein. Deze bevinding sluit goed aan bij theoretische voorstellen, maar is moeilijk te verklaren enkel door een toename in *gain*. Het ruimtelijk gestructureerde effect van atomoxetine op de sterkte van correlaties is mogelijk te verklaren door recente bevindingen die heterogeniteit in het LC-NE systeem aantonen, alsmede door heterogeniteit in de verdeling van NE receptoren over het brein heen.

7.3.3   Hoofdstuk 4: Catecholamines Modulate Intrinsic Long-range Correlations in the Human Brain

In hoofdstuk 4 testen we de voorspelling van de *network reset* theorie dat een toename in de beschikbaarheid van NE zou moeten leiden tot een reorganisatie van de functionele netwerken in het brein. Hiervoor hebben we de data gebruikt uit hoofdstuk 3, waarin proefpersonen atomoxetine toegediend kregen (waardoor de beschikbaarheid van NE verhoogd werd). We hebben twee complementaire analysemethodes toegepast om het effect van NE op de fijnmazige patronen van MRI correlaties te onderzoeken: '*dual regression*' en '*spatial mode decomposition*'. In tegenstelling tot hoofdstuk 3, onderzoeken we in hoofdstuk 4 of atomoxetine leidt tot veranderingen in de spatiële structuur (topologie) van spontane correlaties in hersenactiviteit, gemeten met fMRI, in plaats van alleen de sterkte van deze correlaties.

Beide analysemethodes leverden aansluitend bewijs voor een afname in de sterkte van correlaties tussen verschillende hersengebieden, met name tussen sensorische en motor-gerelateerde hersengebieden. Daarnaast leverde *spatial mode decomposition* bewijs voor een shift in dominantie van links- naar rechts-gelateraliseerde frontopariëtale netwerkactiviteit. Het belangrijkste effect van atomoxetine was het schalen van de sterkte van correlaties binnen bestaande

netwerken, in plaats van een globale reorganisatie van de topologie van deze netwerken.

Onze bevindingen zijn consistent met eerdere bevindingen in primaten: atomoxetine veroorzaakte verminderingen in de sterkte van correlaties binnen sensorische en motor-gerelateerde netwerken. Echter, wij tonen aan dat deze afnames in sterkte ook kwantitatief van aard kunnen zijn, in plaats van noodzakelijk het resultaat van een topologische verandering, zoals voorspeld door de *network reset* theorie. We concluderen dat NE dynamisch de sterkte van spontane correlaties aanpast, mogelijk met als doel om gedrag flexibel aan te passen aan de eisen omgeving.

7.3.4 Hoofdstuk 5: Pupil Diameter Tracks Lapses of Attention

Ons vermogen om voor langere tijd doel-relevante informatie te monitoren en er op te reageren, langdurige aandacht, is niet onbeperkt. Voorgaande studies over de relatie tussen verslappingen van de aandacht en psychofysiologische markers van langdurige aandacht, zoals pupildiameter, hebben tegenstrijdige resultaten opgeleverd. De *adaptive gain* theorie voorspelt dat *baseline* pupildiameter een omgekeerde-U relatie zou moeten vertonen met prestatie op taken waarbij aandacht besteed moet worden aan doel-relevante stimuli. Met andere woorden, de meeste verslappingen van de aandacht zouden onder *adaptive gain* moeten optreden op momenten dat de pupil zowel relatief groot is, als relatief klein. In hoofdstuk 5 testen we deze voorspelling. Daarnaast verkennen we additionele markers van de staat van aandacht, deels gebaseerd op voorgaand onderzoek dat een relatie heeft aangetoond tussen de afgeleide van pupildiameter en de staat van het brein.

We onderzoeken de relatie tussen tonische fluctuaties in de diameter van de pupil, en de prestatie op een uitdagende taak waarbij langdurige aandacht van belang is. Onze resultaten laten zien dat pupildiameter een robuuste lineaire relatie vertoont met verschillende maten van taakprestatie, waarbij de meeste verslappingen van de aandacht optreden op momenten dat de pupil relatief klein is. Echter, deze bevindingen werden voornamelijk gedreven door een gedeeld effect van *time-on-task* op zowel pupildiameter als prestatie op te taak. De lineaire relatie tussen pupildiameter en prestatie op de taak verdween nadat we statistisch corrigeerden voor het effect van *time-on-task.* In plaats daarvan vonden we een consistente omgekeerde-U relatie tussen pupildiameter en taakprestatie.

Daarnaast vonden we een sterke lineaire relatie tussen de afgeleide van pupildiameter en prestatie op de taak, die grotendeels onafhankelijk was van het effect van *time-on-task.* Onze resultaten helpen om tegenstrijdige bevindingen in de literatuur te verklaren, en zijn consistent met de *adaptive gain* theorie van LC-NE functie. Verder suggereren onze resultaten dat de afgeleide van pupildiameter mogelijk een nuttige psychofysiologische maat is die gebruikt zou kunnen worden voor het online voorspellen en voorkomen van verslappingen in de aandacht.

7.3.5 Hoofdstuk 6: Task-free Spectral EEG Dynamics Track and Predict Patient Recovery From Severe Acquired Brain Injury

Zoals besproken in de voorgaande sectie, hangt ons vermogen om sensorische informatie te verwerken en er op te reageren af van de staat van ons brein. Prominente voorbeelden waarbij dit vermogen niet goed functioneert, zijn de bewustzijnsstoornissen die kunnen optreden na hersenletsel. Sommige van deze patiënten ontwikkelen later weer tekenen van bewustzijn, terwijl andere patiënten non-responsief blijven. Neuraal gezien doen de pathofysiologische tekenen van bewustzijnsstoornissen denken aan de hypoactivatie van het LC-NE systeem die optreedt tijdens slaap, en de bijgaande synchronisatie van corticale activiteit. Zoals aangetoond in hoofdstuk 3 en 4 van dit proefschrift zijn fluctuaties in gesynchroniseerde herenactiviteit vatbaar voor neuromodulatie door NE. In hoofdstuk 6 van dit proefschrift verkennen we of de staat (gekwantificeerd als amplitude en connectiviteit) van zulke gesynchroniseerde herenactiviteit het niveau en het herstel van bewustzijn kan voorspellen bij patiënten met bewustzijnsstoornissen. Hievoor analyseren we een bestaande dataset van patiënten die deelnamen aan het '*Early Intensive Neurorehabilitation Programme*'.

Onze resultaten tonen aan dat, vergeleken met gezonde controle proefpersonen, de patiënten een algemene 'vertraging' laten zien van hersenritmes, waarbij laagfrequente (synchrone) schommelingen in hersenactiviteit dominant zijn. Daarnaast veranderde over de duur van het herstel van de proefpersonen de frequentieband specifieke eigenschappen van corticale ritmes op een non-ineaire manier. Deze veranderingen sloten goed aan bij de diagnostieke en prognostieke waarde van deze frequentiebanden. Opmerkelijk genoeg voorspelden correlaties in hersenactiviteit tussen hersengebieden het niveau van herstel van de patiënten ongeveer 3 maanden later 75% correct.

Onze bevindingen tonen aan dat de amplitude van, en correlaties in, de staat van spectrale hersenactiviteit het verloop en niveau van het herstel van patiënten kunnen voorspellen op een automatische manier, die los staat van de sensorische en cognitieve capaciteiten van de patiënt. Tot slot: onze bevindingen suggereren dat de mate waarin verbindingen tussen diepe hersenkernen (zoals bijvoorbeeld de LC of de thalamus) en de hersenschors gespaard zijn gebleven, de mate van het herstel van bewustzijn na hersenletsel kunnen voorspellen. Onze bevindingen geven mogelijk nieuwe inzichten in de pathofysiologische *brain state*-gerelateerde processen die ten grondslag liggen aan bewustzijnsstoornissen.

8. References

Achard S, Bullmore E (2007) Efficiency and cost of economical brain functional networks. PLoS Comput Biol 3:e17.

Akaike H (1974) New look at statistical-model identification. . IEEE Trans Automat Contr Ac 19:716 –723.

Allen EA, Damaraju E, Plis SM, Erhardt EB, Eichele T, Calhoun VD (2014) Tracking whole-brain connectivity dynamics in the resting state. Cereb Cortex 24:663-676.

Andrews K (1996) International Working Party on the Management of the Vegetative State: summary report. Brain Injury 10:797-806.

Aston-Jones G, Bloom FE (1981) Activity of the norepinephrine-containing locus ceuruleus neurons in behaving rats anticipates fluctuations in the sleep-saking cycle. J Neurosci 1:876-886.

Aston-Jones G, Cohen JD (2005) An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. Annu Rev Neurosci 28:403-450.

Aston-Jones G, Foote SL, Bloom FE (1984) Anatomy and physiology of locus coeruleus neurons: Functional implications. In: Norepinephrine (Ziegler MG, Lake CR, eds), pp 92–116. Baltimore: Williams & Wilkins.

Bagnato S, Boccagni C, Sant'angelo A, Fingelkurts AA, Fingelkurts AA, Galardi G (2013) Emerging from an unresponsive wakefulness syndrome: brain plasticity has to cross a threshold level. Neuroscience and biobehavioral reviews 37:2721-2736.

Bargmann CI, Marder E (2013) From the connectome to brain function. Nature Methods 10:483-490.

Barttfeld P, Uhrig L, Sitt JD, Sigman M, Jarraya B, Dehaene S (2015) Signature of consciousness in the dynamics of resting-state brain activity. Proc Natl Acad Sci USA 112:887-892.

Beatty J (1982) Task-evoked pupillary responses, processing load, and the structure of processing resources. Psychological Bulletin 91:276-292.

Beckmann CF (2009) Group comparison of resting-state FMRI data using multi-subject ICA and dual regression. OHBM.

Berens P (2009) CircStat: A MATLAB toolbox for circurlar statistics. Journal of statistics software 31:1-21.

Berridge CW, Waterhouse BD (2003) The locus coeruleus–noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. Brain Research Reviews 42:33-84.

Besle J, Schevon CA, Mehta AD, Lakatos P, Goodman RR, McKhann GM, Emerson RG, Schroeder CE (2011a) Tuning of the human neocortex to the temporal dynamics of attended events. J Neurosci 31:3176-3185.

Besle J, Schevon CA, Mehta AD, Lakatos P, Goodman RR, McKhann GM, Emerson RG, Schroeder CE (2011b) Tuning of the human neocortex to the temporal dynamics of attended events. The Journal of Neuroscience 31:3176-3185.

Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med 34:537-541.

Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Stat Mech: Theory and Experiment.

Boly M, Garrido MI, Gosseries O, Bruno MA, Boveroux P, Schnakers C, Massimini M, Litvak V, Laureys S, Friston K (2011) Preserved feedforward but impaired top-down processes in the vegetative state. Science 332:858-862.

Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict monitoring and cognitive control. Psychological Review 108:624-652.

Bouret S, Sara SJ (2005) Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. Trends in neurosciences 28:574-582.

Brown SB, Tona KD, van Noorden MS, Giltay EJ, van der Wee NJ, Nieuwenhuis S (2015) Noradrenergic and cholinergic effects on speed and sensitivity measures of phasic alerting. Behavioral neuroscience 129:42-49.

Bullmore E, Sporns O (2009) Complex brain networks: Graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10:186-198.

Bymaster FP, Katner JS, Nelson DL, Hemrick-luecke SK, Threlkeld PG, Heiligenstein JH, Mroin SM, Gehlert DR, Perry KW (2002) Atomoxetine increases extracellular levels of norepinephrine and dopamine in prefrontal cortex of rat: A potential mechanism for efficacy in attention deficit/hyperactivity disorder. Neuropsychopharmacol 27.

Casali AG, Gosseries O, Rosanova M, Boly M, Simone Sarasso, Casali KR, Casarotto S, Bruno M-A, Laureys S, Tononi G, Massimini1 M (2013) A Theoretically Based Index of Consciousness Independent of Sensory Processing and Behavior. Science Translational Medicine 5:198ra105.

Cavanagh JF, Cohen MX, Allen JJ (2009a) Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. J Neurosci 29:98-105.

Cavanagh JF, Cohen MX, Allen JJ (2009b) Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. The Journal of Neuroscience 29:98-105.

Chamberlain SR, Muller U, Blackwell AD, Robbins TW, Sahakian BJ (2006a) Noradrenergic modulation of working memory and emotional memory in humans. Psychopharmacology 188:397-407.

Chamberlain SR, Müller U, Blackwell AD, Clark L, Robbins TW, Sahakian BJ (2006b) Neurochemical Modulation of Response Inhibition and Probabilistic Learning in Humans. Science 311:861-863.

Chamberlain SR, Del Campo N, Dowson J, Muller U, Clark L, Robbins TW, Sahakian BJ (2007) Atomoxetine improved response inhibition in adults with attention deficit/hyperactivity disorder. Biol Psychiatry 62:977-984.

Chandler DJ, Gao W-J, Waterhouse BD (2014) Heterogeneous organization of the locus coeruleus projections to prefrontal and motor cortices. Proc Natl Acad Sci U S A 111:6816–6821.

Cheadle S, Wyart V, Tsetsos K, Myers N, de Gardelle V, Herce Castanon S, Summerfield C (2014) Adaptive gain control during human perceptual choice. Neuron 81:1429-1441.

Chennu S, Finoia P, Kamau E, Allanson J, Williams GB, Monti MM, Noreika V, Arnatkeviciute A, Canales-Johnson A, Olivares F, Cabezas-Soto D, Menon DK, Pickard JD, Owen AM, Bekinschtein TA (2014) Spectral signatures of reorganised brain networks in disorders of consciousness. PLoS Comput Biol 10:e1003887.

Cohen JD, Botvinick MM, Carter CS (2000) Anterior cingulate and prefrontal cortes: who's in control? Nat Neurosci 3:421-423.

Cohen M (2014a) Analyzing neural time series data: theory and practice. Cambridge, MA: MIT.

Cohen MX (2014b) Analyzing neural time series data: theory and practice: MIT Press.

Cohen MX, van Gaal S, Ridderinkhof KR, Lamme VA (2009) Unconscious errors enhance prefrontal-occipital oscillatory synchrony. Frontiers in human neuroscience 3:54.

Cohen MX, Bour L, Mantione M, Figee M, Vink M, Tijssen MA, van Rootselaar AF, van den Munckhof P, Schuurman PR, Denys D (2012) Top-down-directed synchrony from medial frontal cortex to nucleus accumbens during reward anticipation. Hum Brain Mapp 33:246-252.

Cole DM, Beckmann CF, Oei NY, Both S, van Gerven JM, Rombouts SA (2013) Differential and distributed effects of dopamine neuromodulations on resting-state network connectivity. Neuroimage 78:59-67.

Cole MW, Ito T, Bassett DS, Schultz DH (2016) Activity flow over resting-state networks shapes cognitive task activations. Nat Neurosci 19:1718-1726.

Colzato LS, de Rover M, van den Wildenberg WP, Nieuwenhuis S (2013) Genetic marker of norepinephrine synthesis predicts individual differences in post-error slowing: A pilot study. Neuropsychologia.

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3:201-215.

Corbetta M, Patel G, Shulman GL (2008) The reorienting system of the human brain: from environment to theory of mind. Neuron 58:306-324.

Coull JT, Middleton HC, Robbins TW, Sahakian BJ (1995) Clonidine and diazepam have differential effects on tests of attention and learning. Psychopharmacology 120:322-331.

Coull JT, Büchel C, Friston KJ, Firth CD (1999) Noradrenergically mediated plasticity in a human attentional neuronal network. Neuroimage 10:705-715.

Craddock RC, James GA, Holtzheimer PE, 3rd, Hu XP, Mayberg HS (2012) A whole brain fMRI atlas generated via spatially constrained spectral clustering. Hum Brain Mapp 33:1914-1928.

Cravo AM, Rohenkohl G, Wyart V, Nobre AC (2013) Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. The Journal of Neuroscience 33:4002-4010.

Crochet S, Petersen CC (2006) Correlating whisker behavior with membrane potential in barrel cortex of awake mice. Nat Neurosci 9:608-610.

Dang-Vu TT, Schabus M, Desseilles M, Albouy G, Boly M, Darsaud A, Gais S, Rauchs G, Sterpenich V, Vandewalle G, Carrier J, Moonen G, Balteau E, Degueldre C, Luxen A, Phillips C, Maquet P (2008) Spontaneous neural activity during human slow wave sleep. Proc Natl Acad Sci U S A 105:15160-15165.

Danielmeier C, Ullsperger M (2011) Post-error adjustments. Frontiers in psychology 2:233.

Danielmeier C, Eichele T, Forstmann BU, Tittgemeyer M, Ullsperger M (2011) Posterior medial frontal cortex activity predicts post-error adaptations in task-related visual and motor areas. The Journal of Neuroscience 31:1780-1789.

Dayan P, Yu AJ (2006) Phasic norepinephrine: a neural interrupt signal for unexpected events. Network: Computation in Neural Systems 17:335-350.

de Gee JW, Colizoli O, Kloosterman NA, Knapen T, Nieuwenhuis S, Donner T (2017) Dynamic modulation of decision biases by brainstem arousal systems. eLife 6:e23232.

De Luca M, Smith S, De Stefano N, Federico A, Matthews PM (2005) Blood oxygenation level dependent contrast resting state networks are relevant to functional activity in the neocortical sensorimotor system. Exp Brain Res 167:587-594.

deBettencourt MT, Cohen JD, Lee RF, Norman KA, Turk-Browne NB (2015) Closed-loop training of attention with real-time brain imaging. Nat Neurosci 18:470-475.

Deco G, Jirsa VK, McIntosh AR (2011) Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat Rev Neurosci 12:43-56.

Deco G, Ponce-Alvarez A, Mantini D, Romani GL, Hagmann P, Corbetta M (2013) Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. J Neurosci 33:11239-11252.

Delorme A, Makeig S (2004a) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. The Journal of Neuroscience Methods 134:9-21.

Delorme A, Makeig S (2004b) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9-21.

Demertzi A, Antonopoulos G, Heine L, Voss HU, Crone JS, de Los Angeles C, Bahri MA, Di Perri C, Vanhaudenhuyse A, Charland-Verville V, Kronbichler M, Trinka E, Phillips C, Gomez F, Tshibanda L, Soddu A, Schiff ND, Whitfield-Gabrieli S, Laureys S (2015) Intrinsic functional connectivity differentiates minimally conscious from unresponsive patients. Brain : a journal of neurology 138:2619-2631.

Devoto P, Flore G (2006) On the origin of cortical dopamine: Is it a co-transmitter in noradrenergic neurons? Curr Neuropharmacol 4:115-125.

Devoto P, Flore G, Pira L, Longu G, Gessa GL (2004) Alpha2-adrenoceptor mediated co-release of dopamine and noradrenaline from noradrenergic neurons in the cerebral cortex. J Neurochem 88:1003-1009.

Di Perri C, Bahri MA, Amico E, Thibaut A, Heine L, Antonopoulos G, Charland-Verville V, Wannez S, Gomez F, Hustinx R, Tshibanda L, Demertzi A, Soddu A, Laureys S (2016) Neural correlates of consciousness in patients who have emerged from a minimally conscious state: a cross-sectional multimodal imaging study. The Lancet Neurology 15:830-842.

Dimigen O, Sommer W, Hohlfeld A, Jacobs AM, Kliegl R (2011) Coregistration of eye movements and EEG in natural reading: analyses and review. Journal of experimental psychology General 140:552-572.

Donner TH, Siegel M (2011) A framework for local cortical oscillation patterns. Trends in cognitive sciences 15:191-199.

Donner TH, Nieuwenhuis S (2013) Brain-wide gain modulation: the rich get richer. Nat Neurosci 16:989-990.

Donner TH, Sagi D, Bonneh YS, Heeger DJ (2013) Retinotopic patterns of correlated fluctuations in visual cortex reflect the dynamics of spontaneous perceptual suppression. J Neurosci 33:2188-2198.

Drew PJ, Duyn JH, Golanov E, Kleinfeld D (2008) Finding coherence in spontaneous oscillations. Nat Neurosci 11:991.

Dudschig C, Jentzsch I (2009) Speeding before and slowing after errors: is it all just strategy? Brain Research 1296:56-62.

Dutilh G, Vandekerckhove J, Forstmann BU, Keuleers E, Brysbaert M, Wagenmakers EJ (2012a) Testing theories of post-error slowing. Attention, perception & psychophysics 74:454-465.

Dutilh G, van Ravenzwaaij D, Nieuwenhuis S, van der Maas HLJ, Forstmann BU, Wagenmakers E-J (2012b) How to measure post-error slowing: A confound and a simple solution. Journal of Mathematical Psychology 56:208-216.

135

Eichele H, Juvodden HT, Ullsperger M, Eichele T (2010) Mal-adaptation of event-related EEG responses preceding performance errors. Frontiers in human neuroscience 4.

Eilander HJ, Wijnen VJ, Scheirs JG, de Kort PL, Prevo AJ (2005) Children and young adults in a prolonged unconscious state due to severe brain injury: outcome after an early intensive neurorehabilitation programme. Brain Injury 19:425-436.

Eilander HJ, van de Wiel M, Wijers M, van Heugten CM, Buljevac D, Lavrijsen JC, Hoenderdaal PL, de Letter-van der Heide L, Wijnen VJ, Scheirs JG, de Kort PL, Prevo AJ (2009) The reliability and validity of the PALOC-s: a post-acute level of consciousness scale for assessment of young patients with prolonged disturbed consciousness after brain injury. Neuropsychological rehabilitation 19:1-27.

Eldar E, Cohen JD, Niv Y (2013) The effects of neural gain on attention and learning. Nat Neurosci 16:1146-1153.

Endrass T, Franke C, Kathmann N (2005) Error Awareness in a Saccade Countermanding Task. Journal of Psychophysiology 19:275-280.

Eriksen BA, Eriksen CW (1974) Effects of noise letters upon the identification of a target letter in a nonsearch task. Perception & Psychophysics 16:143-149.

Eschenko O, Magri C, Panzeri S, Sara SJ (2012) Noradrenergic neurons of the locus coeruleus are phase locked to cortical up-down states during sleep. Cereb Cortex 22:426-435.

Esterman M, Noonan SK, Rosenberg M, DeGutis J (2013) In the Zone or Zoning Out? Tracking Behavioral and Neural Fluctuations During Sustained Attention. Cereb Cortex 23:2712-2723.

Estraneo A, Loreto V, Guarino I, Boemia V, Paone G, Moretta P, Trojano L (2016) Standard EEG in diagnostic process of prolonged disorders of consciousness. Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology 127:2379-2385.

Fellinger R, Klimesch W, Schnakers C, Perrin F, Freunberger R, Gruber W, Laureys S, Schabus M (2011) Cognitive processes in disorders of consciousness as revealed by EEG time-frequency analyses. Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology 122:2177-2184.

Filippini N, MacIntosh BJ, Hough MG, Goodwin GM, Frisoni GB, Smith SM, Matthews PM, Beckmann CF, Mackay CE (2009) Distinct patterns of brain activity in young carriers of the APOE-epsilon4 allele. Proc Natl Acad Sci U S A 106:7209-7214.

Fischer C, Luaute J, Morlet D (2010) Event-related potentials (MMN and novelty P3) in permanent vegetative or minimally conscious states. Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology 121:1032-1042.

Fisher N (1993) Statistical analysis of circular data. Cambridge, UK Cambridge UP.

Forstmann BU, Ratcliff R, Wagenmakers EJ (2016) Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions. Annual review of psychology 67:641-666.

Fox MD, Raichle ME (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat Rev Neurosci 8:700-711.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci U S A 102:9673-9678.

Franklin MS, Broadway JM, Mrazek MD, Smallwood J, Schooler JW (2013) Window to the wandering mind: pupillometry of spontaneous thought while reading. Quarterly journal of experimental psychology 66:2289-2294.

Freeman J, Donner TH, Heeger DJ (2011) Inter-area correlations in the ventral visual pathway reflect feature integration. Journal of vision 11.

Friston KJ, Büchel C (2004) Functional connectivity. In: Human brain function, Ed 2 (Frackowiak RJS, ed). San Diego: Academic.

Gehring WJ, Fenscik DE (2001a) Functions of the medial frontal cortex in the processing of conflict and errors. The Journal of Neuroscience 21:9430-3437.

Gehring WJ, Fenscik DE (2001b) Functions of the medial frontal cortex in the processing of conflict and errors. Journal of Neuroscience 21:9430-3437.

Giacino JT (1997) Disorders of consciousness: differential diagnosis and neuropathologic features. Seminars in Neurology 17:105-111.

Giacino JT, Fins JJ, Laureys S, Schiff ND (2014) Disorders of consciousness after acquired brain injury: the state of the science. Nature reviews Neurology 10:99-114.

Giacino JT, Ashwal S, Childs N, Cranford R, Jennett B, Katz DI, Kelly JP, Rosenberg JH, Whyte J, Zafonte RD, Zasler ND (2002) The minimally conscious state: definition and diagnostic criteria. Neurology 58:349-353.

Gilzenrat MS, Nieuwenhuis S, Jepma M, Cohen JD (2010) Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. Cognitive, affective & behavioral neuroscience 10:252-269.

Glover GH, Li T-Q, Ress D (2000) Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magn Reson Med 44:162-167.

Grandchamp R, Braboszcz C, Delorme A (2014) Oculometric variations during mind wandering. Frontiers in psychology 5:31.

Greenberg DS, Houweling AR, Kerr JN (2008) Population imaging of ongoing neuronal activity in the visual cortex of awake rats. Nat Neurosci 11:749-751.

Greicius MD, Srivastava G, Reiss AL, Menon V (2004) Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proc Natl Acad Sci USA 101:4637-4642.

Griffanti L, Salimi-Khorshidi G, Beckmann CF, Auerbach EJ, Douaud G, Sexton CE, Zsoldos E, Ebmeier KP, Filippini N, Mackay CE, Moeller S, Xu J, Yacoub E, Baselli G, Ugurbil K, Miller KL, Smith SM (2014) ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. Neuroimage 95:232-247.

Guedj C, Monfardini E, Reynaud AJ, Farne A, Meunier M, Hadj-Bouziane F (2016) Boosting Norepinephrine Transmission Triggers Flexible Reconfiguration of Brain Networks at Rest. Cereb Cortex.

Hajcak G, McDonald N, Simons RF (2003) To err is autonomic: Error-related brain potentials, ANS activity, and post-error compensatory behavior. Psychophysiol 40:895-903.

Harris KD, Thiele A (2011) Cortical state and attention. Nat Rev Neurosci 12:509-523.

Haynes JD, Lotto RB, Rees G (2004) Responses of human visual cortex to uniform surfaces. Proc Natl Acad Sci USA 101:4286-4291.

Haynes JD, Driver J, Rees G (2005) Visibility reflects dynamic changes of effective connectivity between V1 and fusiform cortex. Neuron 46:811-821.

Henry MJ, Obleser J (2012) Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. Proc Natl Acad Sci U S A 109:20095-20100.

Hiltunen T, Kantola J, Abou Elseoud A, Lepola P, Suominen K, Starck T, Nikkinen J, Remes J, Tervonen O, Palva S, Kiviniemi V, Palva JM (2014) Infra-slow EEG fluctuations are correlated with resting-state network dynamics in fMRI. J Neurosci 34:356-362.

Hipp JF, Hawellek DJ, Corbetta M, Siegel M, Engel AK (2012) Large-scale cortical correlation structure of spontaneous oscillatory activity. Nat Neurosci 15:884-890.

Hoeks B, Levelt WJ (1993) Pupillary dilation as a measure of attention: A quantitative system analysis. Behavior Research Methods, Instruments & Computers 25:16-26.

Höller Y, Bergmann J, Kronbichler M, Crone JS, Schmid EV, Golaszewski S, Ladurner G (2011) Preserved oscillatory response but lack of mismatch negativity in patients with disorders of consciousness. Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology 122:1744-1754.

Hopstaken JF, van der Linden D, Bakker AB, Kompier MA (2015a) The window of my eyes: Task disengagement and mental fatigue covary with pupil dynamics. Biol Psychol 110:100-106.

Hopstaken JF, van der Linden D, Bakker AB, Kompier MA (2015b) A multifaceted investigation of the link between mental fatigue and task disengagement. Psychophysiol 52:305-315.

Hughes G, Yeung N (2011) Dissociable correlates of response conflict and error awareness in error-related brain activity. Neuropsychologia 49:405-415.

Invernizzi RW, Garattini S (2004) Role of presynaptic alpha2-adrenoceptors in antidepressant action: recent findings from microdialysis studies. Prog Neuropsychopharmacol Biol Psychiatry 28:819-827.

Ishimatsu M, Williams JT (1996) Synchronous activity in locus coeruleus results from dendritic interactions in pericoerulear regions. J Neurosci 16:5196-5204.

Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM (2012) Fsl. Neuroimage 62:782-790.

Jennett B, Plum F (1972) Persistent Vegetative State After Brain Damage: A Syndrome in Search of Name. The Lancet 1:734-737.

Jentzsch I, Dudschig C (2009) Why do we slow down after an error? Mechanisms underlying the effects of posterror slowing. Quarterly journal of experimental psychology 62:209-218.

Jepma M, Nieuwenhuis S (2011) Pupil Diameter Predicts Changes in the Exploration–Exploitation Trade-off: Evidence for the Adaptive Gain Theory. Journal of Cognitive Neuroscience 23:1587-1596.

Joshi S, Li Y, Kalwani Rishi M, Gold Joshua I (2016) Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. Neuron 89:221-234.

Kahneman D, Beatty J (1966) Pupil diameter and load on memory. Science 154:1583-1585.

Kahneman D, Beatty J (1967) Pupillary responses in a pitch-discrimination task. Perception and Psychophysics 2:101-105.

Kang OE, Huffer KE, Wheatley TP (2014) Pupil dilation dynamics track attention to high-level information. PLoS One 9:e102463.

Kempadoo KA, Mosharov EV, Choi SJ, Sulzer D, Kandel ER (2016) Dopamine release from the locus coeruleus to the dorsal hippocampus promotes spatial learning and memory. Proc Natl Acad Sci U S A 113:14835-14840.

King JA, Korb FM, von Cramon DY, Ullsperger M (2010) Post-error behavioral adjustments are facilitated by activation and suppression of task-relevant and task-irrelevant information processing. The Journal of Neuroscience 30:12759-12769.

King JR, Sitt JD, Faugeras F, Rohaut B, El Karoui I, Cohen L, Naccache L, Dehaene S (2013) Information sharing in the brain indexes consciousness in noncommunicative patients. Curr Biol 23:1914-1919.

Klaassens BL, Rombouts SA, Winkler AM, van Gorsel HC, van der Grond J, van Gerven JM (2017) Time related effects on functional brain connectivity after serotonergic and cholinergic neuromodulation. Hum Brain Mapp 38:308-325.

Klaassens BL, van Gorsel HC, Khalili-Mahani N, van der Grond J, Wyman BT, Whitcher B, Rombouts SA, van Gerven JM (2015) Single-dose serotonergic stimulation shows widespread effects on functional brain connectivity. Neuroimage 122:440-450.

Klumpers LE, Cole DM, Khalili-Mahani N, Soeter RP, Te Beek ET, Rombouts SA, van Gerven JM (2012) Manipulating brain connectivity with delta(9)-tetrahydrocannabinol: a pharmacological resting state FMRI study. Neuroimage 63:1701-1711.

Koda K, Ago Y, Cong Y, Kita Y, Takuma K, Matsuda T (2010) Effects of acute and chronic administration of atomoxetine and methylphenidate on extracellular levels of noradrenaline, dopamine and serotonin in the prefrontal cortex and striatum of mice. J Neurochem 114:259-270.

Kotchoubey B, Lang S, Mezger G, Schmalohr D, Schneck M, Semmler A, Bostanov V, Birbaumer N (2005) Information processing in severe disorders of consciousness: vegetative state and minimally conscious state. Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology 116:2441-2453.

Kristjansson SD, Stern JA, Brown TB, Rohrbaugh JW (2009) Detecting phasic lapses in alertness using pupillometric measures. Applied ergonomics 40:978-986.

Lakatos P, Schroeder CE, Leitman DI, Javitt DC (2013) Predictive suppression of cortical excitability and its deficit in schizophrenia. The Journal of Neuroscience 33:1692–11702.

Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. Science 320:110-113.

Laming D (1979) Choice reaction performance following an error. Acta Psychologica 43:199-224.

Large EW, Jones MR (1999) The dynamics of attending: How people track time-varying events. Psychological Review 106:119-159.

Laureys S, Owen AM, Schiff ND (2004) Brain function in coma, vegetative state, and related disorders. The Lancet Neurology 3:537-546.

Laureys S, Faymonville ME, Luxen A, Lamy M, Franck G, Maquet P (2000) Restoration of thalamocortical connectivity after recovery from persistent vegetative state. The Lancet 355:1790-1791.

Le Martelot E, Hankin C (2013) Multi-scale community detection using stability optimisation. Int J of Web Based Communities 9:323-348.

139

Lechinger J, Bothe K, Pichler G, Michitsch G, Donis J, Klimesch W, Schabus M (2013) CRS-R score in disorders of consciousness is strongly related to spectral EEG at rest. Journal of neurology 260:2348-2356.

Lehembre R, Bruno M-A, Vanhaudenhuyse A, Chatelle C, Cologan V, Leclercq Y, Soddu A, Marcq B, Laureys S, Noirhomme Q (2012) Resting-state EEG study of comatose patients: a connectivity and frequency analysis to find differences between vegetative and minimally conscious states. Functional Neurology 27:41-47.

Lemieux M, Chen JY, Lonjers P, Bazhenov M, Timofeev I (2014) The impact of cortical deafferentation on the neocortical slow oscillation. J Neurosci 34:5689-5703.

Leopold DA, Murayama Y, Logothetis NK (2003) Very slow activity fluctuations in monkey visual cortex: Implications for functional brain imaging. Cereb Cortex 13:422-433.

Li N, Ma N, Liu Y, He XS, Sun DL, Fu XM, Zhang X, Han S, Zhang DR (2013) Resting-State Functional Connectivity Predicts Impulsivity in Economic Decision-Making. Journal of Neuroscience 33:4886-4895.

Logothetis NK (2008) What we can do and what we cannot do with fMRI. Nature 453:869-878.

Luczak A, Bartho P, Harris KD (2009) Spontaneous events outline the realm of possible sensory responses in neocortical populations. Neuron 62:413-425.

Mackworth NH (1948) The breakdown of vigilance durning prolonged visual search. Quarterly journal of experimental psychology 1:6-21.

Mackworth NH (1968) Vigilance, arousal and habituation. Psychological Review 75:308-322.

Maltseva I, Giessler H-G, Başar E (2000) Alpha oscillations as an indicator of dynamic memory operations - anticipation of omitted stimuli. International Journal of Psychophysiology 36:185-197.

Marco-Pallarés J, Camara E, Münte TF, Rodríguez-Fornells A (2008) Neural mechanisms underlying adaptive actions after slips. Journal of Cognitive Neuroscience 20:1595-1610.

Marder E (2012) Neuromodulation of neuronal circuits: back to the future. Neuron 76:1-11.

Marder E, O'Leary T, Shruti S (2014) Neuromodulation of circuits with variable parameters: single neurons and small circuits reveal principles of state-dependent and robust neuromodulation. Annual review of neuroscience 37:329-346.

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. The Journal of Neuroscience Methods 164:177-190.

Maslov S, Sneppen K (2002) Speciþcity and stability in topology of protein networks. Science 296:910-913.

Mather M, Clewett D, Sakaki M, Harley CW (2015) Norepinephrine ignites local hot spots of neuronal excitation: How arousal amplifies selectivity in perception and memory. Behavioral and Brain Sciences:e200.

Mathôt S, Siebold A, Donk M, Vitu F (2015) Large Pupils Predict Goal-Driven Eye Movements. Journal of Experimental Psychology: General 144:513-521.

McCabe C, Mishor Z (2011) Antidepressant medications reduce subcortical-cortical resting-state functional connectivity in healthy volunteers. Neuroimage 57:1317-1323.

McGinley Matthew J, David Stephen V, McCormick David A (2015a) Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. Neuron 87:179-192.

McGinley MJ, Vinck M, Reimer J, Batista-Brito R, Zagha E, Cadwell CR, Tolias AS, Cardin JA, McCormick DA (2015b) Waking State: Rapid Variations Modulate Neural and Behavioral Responses. Neuron 87:1143-1161.

McLean RA, Sanders WL, Stroup WW (1991) A unified approach to mixed linear models. The American Statician 45:54-64.

Metzger CD, Wiegers M, Walter M, Abler B, Graf H (2015) Local and global resting state activity in the noradrenergic and dopaminergic pathway modulated by reboxetine and amisulpride in healthy subjects. Int J Neuropsychopharmacol:1-9.

Mitra PP, Pesaran B (1999) Analysis of dynamic brain imaging data. Biophysical Journal 76:691-708.

Mittner M, Boekel W, Tucker AM, Turner BM, Heathcote A, Forstmann BU (2014) When the brain takes a break: a model-based analysis of mind wandering. J Neurosci 34:16286-16295.

Moises HC, Woodward BJ, Hoffer BJ, Freedman R (1979) Interactions of norepinephrine with Purkinje cell responses to putative amino acid neurotransmitters applied by microiontophoresis. Experimental Neurology 64:493-515.

Monti MM, Vanhaudenhuyse A, Coleman MR, Boly M, Pickard JD, Tshibanda L, Owen AM, Laureys S (2010) Willful Modulation of Brain Activity in Disorders of Consciousness. The New England Journal of Medicine 362:579-589.

Mouton PR, Pakkenberg B, Gundersen HJ, Price DL (1994) Absolute number and size of pigmented locus coeruleus neurons in young and aged individuals. Journal of Chemical Neuroanatomy 7:185-190.

Murphy PR, Vandekerckhove J, Nieuwenhuis S (2014a) Pupil-Linked Arousal Determines Variability in Perceptual Decision Making. PLoS Comput Biol 10:e1003854.

Murphy PR, van Moort ML, Nieuwenhuis S (2016) The pupillary orienting response predicts adaptive behavioral adjustment after errors. PloS One 11:e0151763.

Murphy PR, Robertson IH, Balsters JH, O'Connell R G (2011) Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. Psychophysiol 48:1532-1543.

Murphy PR, Robertson IH, Allen D, Hester R, O'Connell RG (2012) An electrophysiological signal that precisely tracks the emergence of error awareness. Frontiers in human neuroscience 6:65.

Murphy PR, O'Connell RG, O'Sullivan M, Robertson IH, Balsters JH (2014b) Pupil diameter covaries with BOLD activity in human locus coeruleus. Hum Brain Mapp 35:4140-4154.

Nahimi A, Jakobsen S, Munk OL, Vang K, Phan JA, Rodell A, Gjedde A (2015) Mapping alpha2 adrenoceptors of the human brain with 11C-yohimbine. J Nucl Med 56:392-398.

Narayanan NS, Cavanagh JF, Frank MJ, Laubach M (2013) Common medial frontal mechanisms of adaptive control in humans and rodents. Nat Neurosci 16:1888-1895.

Ng BS, Schroeder T, Kayser C (2012) A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. The Journal of Neuroscience 32:12268-12276.

Nichols TE, Holmes AP (2001) Nonparametric permutation tests for functional neuroimaging: A primer with examples. Hum Brain Mapp 15:1-25.

Nieuwenhuis S, De Geus EJ, Aston-Jones G (2010) The anatomical and functional relationship between the P3 and autonomic components of the orienting response. Psychophysiol.

Nir Y, Hasson U, Levy I, Yeshurun Y, Malach R (2006) Widespread functional connectivity and fMRI fluctuations in human visual cortex in the absence of visual stimulation. Neuroimage 30:1313-1324.

Notebaert W, Houtman F, Opstal FV, Gevers W, Fias W, Verguts T (2009) Post-error slowing: an orienting account. Cognition 111:275-279.

Nunez Castellar E, Kuhn S, Fias W, Notebaert W (2010) Outcome expectancy and not accuracy determines posterror slowing: ERP support. Cognitive, affective & behavioral neuroscience 10:270-278.

Nunez MD, Srinivasan R, Vandekerckhove J (2015) Individual differences in attention influence perceptual decision making. Frontiers in psychology 8:18.

Okun M, Lampl I (2008) Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. Nat Neurosci 11:535-537.

Owen AM, Coleman MR, Boly M, Davis MH, Laureys S, Pickard JD (2006) Detecting awareness in the vegetative state. Science 313:1402.

Pace-Schott EF, Hobson JA (2002) The neurobiology of sleep: genetics, cellular physiology and subcortical networks. Nat Rev Neurosci 3:591-605.

Piarulli A, Bergamasco M, Thibaut A, Cologan V, Gosseries O, Laureys S (2016) EEG ultradian rhythmicity differences in disorders of consciousness during wakefulness. Journal of neurology 263:1746-1760.

Poulet JF, Petersen CC (2008) Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. Nature 454:881-885.

Power JD, Plitt M, Laumann TO, Martin A (2017) Sources and implications of whole-brain fMRI signals in humans. Neuroimage 146:609-625.

Rabbitt P (1966) Errors and error correction in choice-response tasks. Journal of Experimental Psychology 71:264-272.

Reimer J, Froudarakis E, Cadwell CR, Yatsenko D, Denfield GH, Tolias AS (2014) Pupil fluctuations track fast switching of cortical states during quiet wakefulness. Neuron 84:355-362.

Reimer J, McGinley MJ, Liu Y, Rodenkirch C, Wang Q, McCormick DA, Tolias AS (2016) Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nature communications 7:13289.

Ridderinkhof KR (2002) Micro- and macro-adjustments of task set: activation and suppression in conflict tasks. Psychological research 66:312-323.

Robertson IH, Manly T, Andrade J, Braddeley BT, Yiend J (1997) 'Oops!': Perfromance correlates of everyday attentional failures in traumatic brain injured and normal subjects. Neuropsychologia 35:747-758.

Robinson ES, Eagle DM, Mar AC, Bari A, Banerjee G, Jiang X, Dalley JW, Robbins TW (2008) Similar effects of the selective noradrenaline reuptake inhibitor atomoxetine on three distinct forms of impulsivity in the rat. Neuropsychopharmacol 33:1028-1037.

Rodgers EW, Krenz WD, Baro DJ (2011a) Tonic dopamine induces persistent changes in the transient potassium current through translational regulation. J Neurosci 31:13046-13056.

Rodgers EW, Fu JJ, Krenz WD, Baro DJ (2011b) Tonic nanomolar dopamine enables an activity-dependent phase recovery mechanism that persistently alters the maximal conductance of the hyperpolarization-activated current in a rhythmically active neuron. J Neurosci 31:16387-16397.

Rogawksi MA, Aghajanian GK (1980) Modulation of lateral geniculate neurone excitability by noradrenaline microiontophoresis or locus ceuruleus stimulation. Nature 287:731-734.

Rombouts SA, Barkhof F, Goekoop R, Stam CJ, Scheltens P (2005) Altered resting state networks in mild cognitive impairment and mild Alzheimer's disease: an fMRI study. Hum Brain Mapp 26:231-239.

Rosanova M, Gosseries O, Casarotto S, Boly M, Casali AG, Bruno MA, Mariotti M, Boveroux P, Tononi G, Laureys S, Massimini M (2012) Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. Brain : a journal of neurology 135:1308-1320.

Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: Uses and interpretations. Neuroimage 52:1059-1069.

Safaai H, Neves R, Eschenko O, Logothetis NK, Panzeri S (2015) Modeling the effect of locus coeruleus firing on cortical state dynamics and single-trial sensory processing. Proc Natl Acad Sci U S A 112:12834-12839.

Saleh M, Reimer J, Penn R, Ojakangas CL, Hatsopoulos NG (2010) Fast and slow oscillations in human primary motor cortex predict oncoming behaviorally relevant cues. Neuron 65:461-471.

Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM (2014) Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. Neuroimage 90:449-468.

Sara SJ, Bouret S (2012) Orienting and reorienting: the locus coeruleus mediates cognition through arousal. Neuron 76:130-141.

Schiff ND (2010) Recovery of consciousness after brain injury: a mesocircuit hypothesis. Trends in neurosciences 33:1-9.

Schiff ND, Nauvel T, Victor JD (2014) Large-scale brain dynamics in disorders of consciousness. Current opinion in neurobiology 25:7-14.

Schiff ND, Giacino JT, Kalmar K, Victor JD, Baker K, Gerber M, Fritz B, Eisenberg B, Biondi T, O'Connor J, Kobylarz EJ, Farris S, Machado A, McCagg C, Plum F, Fins JJ, Rezai AR (2007) Behavioural improvements with thalamic stimulation after severe traumatic brain injury. Nature 448:600-603.

Schölvinck ML, Maier A, Ye FQ, Duyn JH, Leopold DA (2010) Neural basis of global resting-state fMRI activity. Proc Natl Acad Sci USA 107:10238–10243.

Schorr B, Schlee W, Arndt M, Bender A (2016) Coherence in resting-state EEG as a predictor for the recovery from unresponsive wakefulness syndrome. Journal of neurology 263:937-953.

Schrantee A, Ferguson B, Stoffers D, Booij J, Rombouts S, Reneman L (2016) Effects of dexamphetamine-induced dopamine release on resting-state network connectivity in recreational amphetamine users and healthy controls. Brain imaging and behavior 10:548-558.

Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. Trends in neurosciences 32:9-18.

Schurger A, Sarigiannidis I, Naccache L, Sitt JD, Dehaene S (2015) Cortical activity is more stable when sensory stimuli are consciously perceived. Proc Natl Acad Sci U S A 112:E2083-2092.

Schwarz G (1978) Estimating dimension of a model. Ann Stat 6:461-464.

Schwarz LA, Luo L (2015) Organization of the Locus Coeruleus-Norepinephrine System. Curr Biol 25:R1051-1056.

Schwarz LA, Miyamichi K, Gao XJ, Beier KT, Weissbourd B, DeLoach KE, Ren J, Ibanes S, Malenka RC, Kremer EJ, Luo L (2015) Viral-genetic tracing of the input-output organization of a central noradrenaline circuit. Nature 524:88-92.

Seamans JK, Gorelova N, Durstewitz D, Yang CR (2001a) Bidirectional dopamine modulation of GABAergic inhibition in prefrontal cortical pyramidal neurons. J Neurosci 21:3628–3638.

Seamans JK, Durstewitz D, Christie BR, Stevens CF, Sejnowski TJ (2001b) Dopamine D1/D5 receptor modulation of excitatory synaptic inputs to layer V prefrontal cortex neurons. Proc Natl Acad Sci U S A 98:301-306.

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD (2007) Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci 27:2349-2356.

Sepulcre J, Liu H, Talukdar T, Martincorena I, Yeo BT, Buckner RL (2010) The organization of local and distant functional connectivity in the human brain. PLoS Comput Biol 6:e1000808.

Servan-Schreiber D, Printz H, Cohen JD (1990) A network model of catecholamiine effects: gain, signal-to-noise ratio, and behavior. Science 249:892-895.

Shine JM, Bissett PG, Bell PT, Koyejo O, Balsters JH, Gorgolewski KJ, Moodie CA, Poldrack RA (2016) The Dynamics of Functional Brain Networks: Integrated Network States during Cognitive Task Performance. Neuron 92:544-554.

Siegel M, Donner TH, Engel AK (2012) Spectral fingerprints of large-scale neuronal interactions. Nat Rev Neurosci 13:121-134.

Siems M, Pape AA, Hipp JF, Siegel M (2016) Measuring the cortical correlation structure of spontaneous oscillatory activity with EEG and MEG. Neuroimage 129:345-355.

Sitt JD, King JR, El Karoui I, Rohaut B, Faugeras F, Gramfort A, Cohen L, Sigman M, Dehaene S, Naccache L (2014) Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. Brain : a journal of neurology 137:2258-2270.

Smallwood J, Brown KS, Baird B, Mrazek MD, Franklin MS, Schooler JW (2012) Insulation for daydreams: a role for tonic norepinephrine in the facilitation of internally guided thought. PLoS One 7:e33706.

Smallwood J, Brown KS, Tipper C, Giesbrecht B, Franklin MS, Mrazek MD, Carlson JM, Schooler JW (2011) Pupillometric evidence for the decoupling of attention from perceptual input during offline thought. PLoS One 6:e18298.

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR, Beckmann CF (2009) Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci U S A 106:13040-13045.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23 Suppl 1:S208-219.

Stefanics G, Hangya B, Hernadi I, Winkler I, Lakatos P, Ulbert I (2010a) Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. The Journal of Neuroscience 30:13578-13585.

Stefanics G, Hangya B, Hernadi I, Winkler I, Lakatos P, Ulbert I (2010b) Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. Journal of neuroscience 30:13578-13585.

Steinborn MB, Flehmig HC, Bratzke D, Schroter H (2012) Error reactivity in self-paced performance: Highly-accurate individuals exhibit largest post-error slowing. Quarterly journal of experimental psychology 65:624-631.

Steinhauser M, Yeung N (2012) Error awareness as evidence accumulation: effects of speed-accuracy trade-off on error signaling. Frontiers in human neuroscience 6:240.

Stender J, Mortensen KN, Thibaut A, Darkner S, Laureys S, Gjedde A, Kupers R (2016) The Minimal Energetic Requirement of Sustained Awareness after Brain Injury. Curr Biol 26:1494-1499.

Sulzer D, Cragg SJ, Rice ME (2016) Striatal dopamine neurotransmission: regulation of release and uptake. Basal ganglia 6:123-148.

Swanson CJ, Perry KW, Koch-Krueger S, Katner J, Svensson KA, Bymaster FP (2006) Effect of the attention deficit/hyperactivity disorder drug atomoxetine on extracellular concentrations of norepinephrine and dopamine in several brain regions of the rat. Neuropharmacology 50:755-760.

Swanson LW, Hartman BK (1975) The central adrenergic system. An immunofluorescence study of the location of cell bodies and their efferent connections in the rat utilizing dopamine-beta-hydroxylase as a marker. J Comp Neurol 163:467-505.

Takeuchi T, Duszkiewicz AJ, Sonneborn A, Spooner PA, Yamasaki M, Watanabe M, Smith CC, Fernandez G, Deisseroth K, Greene RW, Morris RG (2016) Locus coeruleus and dopaminergic consolidation of everyday memory. Nature 537:357-362.

Tallon-Baudry C, Bertrand O (1999) Oscillatory gamma activity in humans and its role in object representation. Trends in cognitive sciences 3:151-162.

Tavor I, Parker Jones O, Mars RB, Smith SM, Behrens TE, Jbabdi S (2016) Task-free MRI predicts individual differences in brain activity during task performance. Science 352:216-220.

Thiebaut de Schotten M, Dell'Acqua F, Forkel SJ, Simmons A, Vergani F, Murphy DG, Catani M (2011) A lateralized brain network for visuospatial attention. Nat Neurosci 14:1245-1246.

Timofeev I, Grenier F, Bazhenov M, Sejnowski TJ, Steriade M (2000) Origin of slow cortical oscillations in deafferenated cortical slabs. Cereb Cortex 10:1185-1199.

Totah NK, Neves RM, Panzeri S, Logothetis NK, Eschenko O (2017) Monitoring large populations of locus coeruleus neurons reveals the non-global nature of the norepinephrine neuromodulatory system. bioRxiv.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15:273-289.

Ullsperger M, Harsay HA, Wessel JR, Ridderinkhof KR (2010) Conscious perception of errors and its relation to the anterior insula. Brain structure & function 214:629-643.

Unsworth N, Robison MK (2016) Pupillary correlates of lapses of sustained attention. Cognitive, affective & behavioral neuroscience.

van den Brink RL, Wynn SC, Nieuwenhuis S (2014) Post-error slowing as a consequence of disturbed low-frequency oscillatory phase entrainment. J Neurosci 34:11096-11105.

van den Brink RL, Pfeffer T, Warren CM, Murphy PR, Tona KD, van der Wee NJ, Giltay E, van Noorden MS, Rombouts SA, Donner TH, Nieuwenhuis S (2016)

Catecholaminergic Neuromodulation Shapes Intrinsic MRI Functional Connectivity in the Human Brain. J Neurosci 36:7865-7876.

Van Dijk KR, Sabuncu MR, Buckner RL (2012) The influence of head motion on intrinsic functional connectivity MRI. Neuroimage 59:431-438.

Van Orden KF, Jung T-P, Makeig S (2000) Combined eye activity measures accurately estimate changes in sustained visual task performance. Biological Psychology 52:221-240.

van Wijk BCM, Stam CJ, Daffershofer A (2010) Comparing brain networks of different size and connectivity density using graph theory. PLoS One 5:e13701.

Varazzani C, San-Galli A, Gilardeau S, Bouret S (2015) Noradrenaline and dopamine neurons in the reward/effort trade-off: a direct electrophysiological comparison in behaving monkeys. J Neurosci 35:7866-7877.

Varotto G, Fazio P, Rossi Sebastiano D, Duran D, D'Incerti L, Parati E, Sattin D, Leonardi M, Franceschetti S, Panzica F (2014) Altered resting state effective connectivity in long-standing vegetative state patients: an EEG study. Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology 125:63-68.

Wang J, O'Donnell P (2001) D1 dopamine receptors potentiate NMDA-mediated excitability increase in layer V prefrontal cortical pyramidal neurons. Cereb Cortex 11.

Wang X-J (2010) Neurophysiological and computational principles of cortical rhythms in cognition. Physiological Reviews 90:1195-1268.

Warren CM, van den Brink RL, Nieuwenhuis S, Bosch JA (2017) Norepinephrine transporter blocker atomoxetine increases salivary alpha amylase. Psychoneuroendocrinology 78:233-236.

Warren CM, van den Brink RL, Nieuwenhuis S, Bosch JA (in preparation) Testing the controversial claim that salivary alpha amylase is a biomarker of central noradrenergic activity.

Warren CM, Eldar E, van den Brink RL, Tona KD, van der Wee NJ, Giltay EJ, van Noorden MS, Bosch JA, Wilson RC, Cohen JD, Nieuwenhuis S (2016) Catecholamine-Mediated Increases in Gain Enhance the Precision of Cortical Representations. J Neurosci 36:5699-5708.

Waterhouse BD, Moises HC, Woodward BJ (1998) Phasic activation of the locus coeruleus enhances responses of primary sensory cortical neurons to peripheral receptive field stimulation. Brain Research 790:33-44.

Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-workd' networks. Nature 393:440-442.

Weissman DH, Roberts KC, Visscher KM, Woldorff MG (2006) The neural bases of momentary lapses in attention. Nat Neurosci 9:971-978.

White CN, Ratcliff R, Vasey MW, McKoon G (2010) Using diffusion models to understand clinical disorders. J Math Psychol 54:39-52.

Wijnen VJ, van Boxtel GJ, Eilander HJ, de Gelder B (2007) Mismatch negativity predicts recovery from the vegetative state. Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology 118:597-605.

Winterer G, Weinberger DR (2004) Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. Trends in neurosciences 27:683-690.

Wyart V, de Gardelle V, Scholl J, Summerfield C (2012) Rhythmic fluctuations in evidence accumulation during decision making in the human brain. Neuron 76:847-858.

Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zollei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106:1125-1165.

Yerkes RM, Dodson JD (1908) The relation of strength of stimulus to rapidity of habit-formation. Journal of Comparitive Neurology and Psychology 18:459-482.

Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. Neuron 46:681-692.

Yuval-Greenberg S, Tomer O, Keren AS, Nelken I, Deouell LY (2008) Transient induced gamma-band response in EEG as a manifestation of miniature saccades. Neuron 58:429-441.

Zagha E, Casale AE, Sachdev RN, McGinley MJ, McCormick DA (2013) Motor cortex feedback influences sensory processing by modulating network state. Neuron 79:567-578.

Zalesky A, Fornito A, Cocchi L, Gollo LL, Breakspear M (2014) Time-resolved resting-state brain networks. Proc Natl Acad Sci USA 111:10341-10346.

Zar JH (1999) Biostatistical Analysis, 5th Edition: Prentice Hall.

Zilles K, Amunts K (2009) Receptor mapping: architecture of the human cerebral cortex. Current opinion in neurology 22:331-339.

Acknowledgements

Curriculum Vitae

Ruud van den Brink was born in Delft, on the 22nd of September 1987. In 2007, he completed secondary school at the Dominicus College, Nijmegen. He obtained a Bachelor of Science, from the University of Amsterdam in 2010, where he studied Psychobiology. In 2012 he graduated with distinction (*cum laude*) from the research master's program Brain and Cognitive Sciences, also at the University of Amsterdam, earning the degree of Master of Science. In October 2012, he started to work as a PhD student at Leiden University, under the supervision of professor Sander Nieuwenhuis. The work focused on noradrenergic mechanisms for the control of brain state and cognition, and resulted in this thesis. Over the course of the PhD project, Ruud was involved in various extracurricular activities, such as an editorship at the quarterly journal of the Dutch psychnomics society, *de Psychonoom*. In the summer of 2017, he will start working as a postdoc in the lab of professor Tobias Donner, funded by an Alexander von Humboldt fellowship for postdoctoral researchers. The project is targeted at uncovering fast-acting neural mechanisms that facilitate transitions between cognitive states