Proceedings of the Master's Programme Cognitive Neuroscience of the Radboud University

Table of Contents

From the Editors in Chief

Dear Reader,

This year, the Proceedings of the Masters Programme of Cognitive Neuroscience turns 15 years old! Thankfully, journals do not become angsty hormonal teenagers. Bad jokes aside, it is an honour to present to you the 15.1 Issue, our first issue as Editors in Chief.

We have received numerous high-caliber theses, and the ones published here excelled in quality, creativity of the topic and societal relevance. We are proud to be publishing theses from all four tracks of the masters and being able to provide a wide range of topics to read. Not only are several topics covered, but also diverse methods: from the more classical fMRI, to tACS, to previously unexplored metrics of functional connectivity.

We want to congratulate the authors for their exceptional quality work and their willingness to cooperate with the journal during the review process.

Moreover, we would like to thank all junior and senior reviewers for their valuable input to the journal. While their work often occurs behind-the-scenes, it is vital for the success of our journal.

Finally, the last thanks have to be directed to our amazing team. Your enthusiasm and hard work has made bringing this journal to life possible, we feel very lucky to be working with you.

We hope you enjoy reading the theses in this issue as much as we did, and that you join us in celebrating the 15th Volume of our beloved journal.

Nijmegen, February 2020

Anne Leenders and Ilaria Lisi

Editors in Chief

From Prof. Dr. Ivan Toni

We are familiar with the "tree of life" sketched by Charles Darwin in his notebook in 1837 (Darwin, 2019). Darwin's drawing symbolizes the profound change he brought into human culture. Biological diversity is explained by the underlying universality of the selection mechanism, bringing conceptual unity to the phenomenological variety of life. There is another, less known tree of life, embodied in the ebony carvings of the Makonde craftsmen in Tanzania (Humphrey, 2007). The carved tree of life elaborates on Darwin's tree of life by emphasizing how humans are unified by their universal capacity, despite their differences, to work together as a community, for the community.

It is painfully easy to forget that writing a scientific paper is about carving a scholarly tree of life, working as a community, for the community of scholars sharing similar intellectual challenges. The mercantile attitude of several academic institutions and scholarly journals towards scientific publications makes it particularly hard for young scientists to think of those publications as little more than a badge of honour to pin on their academic uniform. I believe that this journal is an important tool for inoculating budding cognitive neuroscientists against those mercantile and individualistic attitudes towards science. By building the editorial process around the scientific contributors, i.e. the CNS MSc students, this journal structurally focuses their intellectual priorities towards the scientific dialogue embodied by a learned journal, through values of transparency and constructive criticism towards peers. These are exactly the values guiding community-based journals in neuroscience.

Peering towards the horizon, it is less clear whether and how the contributors to this journal are getting ready for a novel and more courageous way of doing science. In addition to the traditional reductionist approach of approximating complex systems into tractable toy-scenario, cognitive neuroscience has started to take seriously the fact that its objects of enquiry span multiple levels of organization, with multiple factors interacting within and between levels. Dealing with those complex scenarios goes well beyond what can be achieved by extending the range of our measurements, as it is done in big-data projects. Those complex scenarios requires the integration of different conceptual views of a scientific problem, pooling together expertise from different domains and methodologies. That is team science. Team science initiatives are highly labor-intensive, conflict-prone, and require substantial preparation, practice, and trust among team members (Stokols et al., 2008). Team science requires time and courage, e.g. time for its practitioners to educate themself about other's questions and tools, and courage to depend on goals and responsabilities shared with others. The topics and teams of the publications of this journal suggest that its young contributors might have already experienced the confusing thrill of working on a particular element of a larger research question. I hope that among those contributors there will be some resistant to short-term calculations of return on investment, and willing to wait for the exhilarating thrill experienced when individual elements of a larger question are brought together, carving out a new scientific tree of life.

Ivan Toni

About the cover

The magical underwater world consists of a vast variety of fascinating organisms. Not only can one find a large diversity of aquatic plants there, it is also inhabited by the most enchanting sea creatures. Depicted on the cover of this volume is one of such miraculous sea animals, scientifically named *Diploria Labyrinthiformis*, but more commonly known as brain coral. As you can probably imagine, brain coral was named after the beautiful object that connects us all due to its physical appearance. Similar to real human brains it has a spheroid shape, and a grooved surface resembling the brain's numerous gyri and sulci. Even though they do not have brains, the brown, yellow or gray brain corals can live up to 900 years and grow to a size of almost 2 meters in diameter.

Because of a phenomenon called meandroid tissue integration, which means that the brain corals are highly interconnected, it is hard to differentiate the separate animals. These separate coral creatures, also known as polyps, have a soft body and a hard outer skeleton made of limestone. Their surface is covered with thin tentacles that can be extruded to capture plankton and bacteria on which it survives. Most corals have a mutually beneficial relationship with algae. They reside within the coral's tissues where they are protected. The algae also use the coral's waste products for photosynthesis. In turn, the brain coral uses the products of the algae's photosynthesis to survive and grow. When brain corals are stressed due to changes in temperature, light or nutrients, they expel the algae which causes them to turn white. This does not kill them, but it makes them more vulnerable. This is of concern nowadays due to climate change.

Brain coral can be found offcoast in the Atlantic Ocean, the Gulf of Mexico and the Carribean Sea at 1 to 30 meters of depth, where they are of large ecological importance. Because they build strong structures during their growth and are highly interconnected, they form a great foundation to reefs. These reefs are very valuable since they can absorb up to 97 percent of the wave energy during tempestuous wheather.

The picture on the cover of this journal was taken by a remotely operated vehicle. The National Ocean Service used this research tool to discover the type of sea floor.

Suzanne de Jong

Influencing Social-Emotional Actions: The Effect of Theta-Gamma Coupled tACS on Social-Emotional Actions and aPFC-M1 Functional Connectivity

Amy Abelmann[1]
Supervisors: Bob Bramson[1], Karin Roelofs[1,2]

[1]*Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands*
[2]*Behavioural Science Institute, Radboud University, Nijmegen, Netherlands*

In our everyday life we are used to regulating our social-emotional behaviour, for example when you want to approach a fellow researcher at a conference for a possible collaboration. An important brain region involved in the implementation of social-emotional control is the anterior prefrontal cortex (aPFC). Previous research found that control over social-emotional behaviour is implemented by theta-band oscillations in the aPFC that are phase-locked to the power increase of gamma-band oscillations in the motor cortex. In this project we studied the effects of non-invasive brain stimulation on social-emotional behaviour using transcranial alternating current stimulation (tACS). We aimed to investigate (1) whether we can influence social-emotional actions by applying different types of theta-gamma coupled tACS and (2) how the functional connectivity patterns of the brain change as a consequence of the different stimulation conditions. In our tACS/functional magnetic resonance imaging (fMRI) study, 44 healthy participants performed the approach-avoidance task while receiving tACS stimulation (in-phase, anti-phase, or sham) in the magnetic resonance (MR) scanner. During the task, the participant used a joystick to either approach happy faces and avoid angry faces (affect-congruent condition) or avoid happy faces and approach angry faces (affect-incongruent condition). Error rates and reaction times were measured during the task, together with ongoing brain activity. We found reduced congruency effects for error rates but not for reaction times for in-phase stimulation as compared to anti-phase stimulation. Moreover, a stronger functional connectivity was found between aPFC (or lateral frontal pole) and primary motor cortex (M1) for in-phase compared to anti-phase stimulation. These findings show improvement in social-emotional control when theta-band oscillations of the aPFC are phase-coupled to the power increases of the gamma-band oscillations of the M1. This research has high clinical relevance as the findings suggest that theta-gamma coupled tACS could ultimately be used to develop possible new treatments for social, emotional, and behavioural disorders, such as social anxiety.

Keywords: social-emotional actions, control, anterior prefrontal cortex, lateral frontal pole, tACS, MRI

Corresponding author: Amy Abelmann**;** E-mail: a.abelmann@donders.ru.nl

Control over social-emotional behaviour plays an important role in our daily lives as is illustrated by our social interactions. Imagine a situation in which you are wandering around a conference and want to approach an interesting researcher to talk about their project or a possible collaboration. You probably have to exert some cognitive control to overcome your automatic "fear" of talking to this person, after which you approach them in order to achieve your goal. For some people this might come naturally. Others, for example, people who are more socially anxious, experience difficulties in overriding their automatic emotional responses in order to approach a person in this situation. Research has found that cortical oscillatory mechanisms support the control of human social–emotional actions through the coupling of slow theta- and fast gamma-band oscillations (Bramson, Jensen, Toni, & Roelofs, 2018). The most anterior section of the frontal cortex (aPFC) plays an important role in regulating social-emotional behaviour by orchestrating this theta-gamma band coupling. While research has been done on the role of the aPFC in overcoming automatic action tendencies and the neural mechanisms behind social-emotional control. However, limited research has been done on whether we can influence the oscillatory mechanisms underlying social-emotional control. Communication between neural networks and brain regions is influenced by coherence in oscillations between sender and receiver networks (Buzsáki & Wang, 2012; Womelsdorf et al., 2007). This type of neuronal coherence can also be observed between frequency bands and is known as cross-frequency coupling (Goodman et al., 2018; Varela, Lachaux, Rodriguez, & Martinerie, 2001). Previous research has found that the cross-frequency coupling between the phase of slow theta-band oscillations (4 – 8 Hz) and the amplitude of high gamma-band oscillations (30 – 100 Hz) underlies cognitive processes such as working memory (Engel, Fries, & Singer, 2001) and cognitive control (Cavanagh & Frank, 2014). This phenomenon is called theta-gamma coupling, which means that high-frequency gamma oscillations are modulated by low-frequency theta oscillations (Lisman & Jensen, 2013).

How does the function of theta- and gamma-band oscillations relate to theta-gamma coupling? Animal research has shown that theta oscillations play an important role in the communication between largely separated neural networks and that this is organized by a high level of coherence (Lisman & Jensen, 2013). Moreover, human research shows that frontal theta-band oscillations (4 – 8 Hz)

are involved in control over motivational behaviour (Cavanagh & Frank, 2014; Cooper et al., 2015) while the synchronization of gamma-band oscillations (30 – 100 Hz) in the motor system is related to action preparation and movement selection (Donner, Siegel, Fries, & Engel, 2009; Schoffelen, Poort, Oostenveld, & Fries, 2011). These theta- and gamma-band oscillations are coupled during information encoding between different brain regions and suggest that complex cognitive functions, such as cognitive control, can be influenced by a mechanism based on inter-regional phase encoding (Voytek et al., 2015).

Social-emotional control defines our ability to override automatic action tendencies in order to control approach and avoidance behaviour. The lateral anterior prefrontal cortex (aPFC)—also known as lateral frontal pole (FPl; Neubert et al., 2014)—is involved in regulating social-emotional control implemented by theta-gamma band coupled oscillations (Bramson et al., 2019; Bramson et al., 2018). Strong inhibition of this region leads to an impairment in control over social-emotional actions (Volman, Roelofs, Koch, Verhagen, & Toni, 2011). The aPFC controls these social-emotional behaviours by upregulating regions that are involved in rule selection, such as the parietal cortex and premotor area, and downregulating regions involved in automatic emotion evaluation, such as the amygdala (Volman et al., 2011; 2013). Moreover, the aPFC shares strong anatomical and functional connections to the limbic system (Tamnes et al., 2009; Tyborowska, Volman, Smeekens, Toni, & Roelofs, 2016). Other regions that share functional connectivity with the aPFC, even though not anatomically linked, are the posterior parietal cortex (Neubert, Mars, Thomas, Sallet, & Rushworth, 2014), primary motor cortex (M1) (Bramson et al., 2018), and thalamus/pulvinar (Tyborowska et al., 2016). Linking this to theta-gamma coupling: theta-band oscillations in the prefrontal cortex and gamma-band oscillations in the primary motor system may be coupled via a mechanism in which aPFC neurons phase lock their spiking and gamma-band firing to theta-band rhythms (Ardid et al., 2015; Voytek et al., 2015). The theta-band rhythms create overlapping periods of excitability across the cerebral network involved in social-emotional regulation (Colgin, 2013; Lisman & Jensen, 2013), especially in the aPFC-M1 network. Therefore, the aPFC and the aPFC-M1 network probably play a crucial role in the control over social-emotional actions.

Previous research has found that transcranial alternating current stimulation (tACS) can externally

influence phase-amplitude coupling of neural oscillations underlying complex cognitive processes. tACS is a non-invasive neuro-stimulation technique that uses a sinusoidal electrical current at a chosen frequency to directly interact with cortical rhythms (Antal & Paulus, 2013; Herrmann, Rach, Neuling, & Strüber, 2013). Adaptation of underlying brain processes can only occur when a person is already engaged in specific cognitive processes, such as memory retrieval or cognitive control, to synchronise the spike timing of neurons. As a consequence, tACS based on theta-gamma band coupled oscillations enhances communication between distant areas and consequently improves spatial working memory (Alekseichuk, Turi, de Lara, Antal, & Paulus, 2016) and value-based decision making (Polanía, Moisa, Opitz, Grueschow, & Ruff, 2015). Moreover, this technique can be used to revive working memory by synchronizing these rhythmic brain circuits (Reinhart & Nguyen, 2019). These findings show that theta-gamma band coupled tACS is able to improve cognitive processes and could possibly lead to long-term effects on cognition.

However, the possibilities of externally influencing existing neural rhythms underlying social-emotional control have not yet been investigated. Given that theta-gamma band coupling between the prefrontal cortex and primary motor area play an important role in regulating social-emotional behaviour, we think that theta-gamma coupled tACS can also be applied to influence the theta-gamma band coupled oscillatory mechanisms underlying this type of control. Non-invasive stimulation could possibly influence these mechanisms by synchronising the naturally occurring theta-gamma oscillations in these regions during social-emotional control.

In this study, we investigate whether we can improve social-emotional actions using theta-gamma band coupled tACS. Our first interest is the effect of theta-gamma coupled tACS stimulation on the control over social-emotional actions. This stimulation either synchronises (in-phase) or desynchronises (anti-phase) existing theta- and gamma-oscillations of the brain. We used the approach-avoidance task (AAT) in order to measure social-emotional control (Heuer, Rinck, & Becker, 2007). In the AAT, pictures of emotional faces are presented on a screen and the subject is told to either approach or avoid these faces by moving a joystick. Normally, people show the automatic response tendency to approach happy and avoid angry faces (affect-congruent) and have to exert control to override these automatic tendencies in order to approach angry and avoid happy faces (affect-incongruent). The AAT has extensively been proven to be a reliable task to investigate social emotional control in both healthy as well as clinical populations (Bramson et al., 2018; Roelofs, Elzinga, & Rotteveel, 2005; Roelofs, Minelli, Mars, Van Peer, & Toni, 2008; Volman et al., 2011; Volman et al., 2013). We expect that in-phase stimulation would lead to improved communication between the aPFC and primary motor areas while anti-phase stimulation would disrupt this communication. Therefore, we hypothesise that performance on the AAT will improve, reflected in reduced congruency effects, when theta-gamma coupled tACS is applied during in-phase stimulation as compared to anti-phase and sham (control) stimulation.

Secondly, we are interested in the functional connectivity change when theta-gamma coupled tACS is applied to the aPFC and primary motor cortex (M1). One of the advantages of tACS is that you can use the technique in the magnetic resonance imaging (MRI) scanner, which allows you to look at haemodynamic responses and inter-regional functional connectivity while externally influencing underlying brain processes. We expect to find a stronger functional connectivity between aPFC and M1 during in-phase theta-gamma coupled stimulation as compared to anti-phase or sham stimulation. If our hypotheses prove to be correct, then this provides causal evidence for the contribution of theta-gamma coupling during social-emotional control. Moreover, the use of tACS in adapting social-emotional control could be highly relevant for clinical practice. Theta-gamma coupled tACS could provide possible treatments for people who have difficulties in controlling their automatic emotional action tendencies, for example, patients with social anxiety (Radke, Roelofs, & De Bruijn, 2013). The technique could be used as an addition to extinction therapy, where patients are exposed repeatedly to stressful cues and situations, thereby improving the effectiveness of the therapy and increasing the speed of recovery.

Method

Participants

Forty-four male subjects participated in the study after providing written informed consent for MR and transcranial current stimulation (tCS) according to the guidelines of the local ethics committee. All participants were right-handed, had corrected-to-normal vision and had no history of neurological or psychiatric illness. They received payment or course

credits for their participation. Three participants were excluded due to technical difficulties. The participants took part in three sessions: one magnetic resonance spectroscopy (MRS) session (not further discussed in this manuscript) and two stimulation sessions (in which the subject performed the AAT).

Experimental task and procedure

The approach-avoidance task (AAT) was used as a measure to investigate social-emotional control in this study (Fig. 1) as used in several previous studies (Bramson et al., 2018; Volman et al., 2011). During the AAT, the participants were visually presented with happy and angry faces. The participants received written task instructions on the screen before each block. They either had to approach happy faces and avoid angry faces (affect-congruent) or avoid happy faces and approach angry faces (affect-incongruent). Participants had to respond to the congruent and incongruent conditions by either pulling the MRI-compatible joystick towards themselves (approach action) or pushing it away from themselves (avoidance action). The task rules changed each block. At the start of each trial a white fixation cross was presented on the screen for 1000 ms. After this, a face appeared on the screen for 100 ms and the participant had to respond by moving the joystick within 2000 ms after presentation of the stimulus. Responses were valid if they reached a 30% tilt of the joystick and if the participant responded within 2000 ms after the presentation of the stimulus. After responding, the participants had to bring back the joystick to the centre position. A warning message on the screen appeared if they did not bring the joystick back to the centre position or when they responded too late/missed a response. The inter-trial interval (ITI) was 2000 to 4000 ms. Reaction times (in ms) and error rates (percentage of correct trials) were measured. The participants performed the task in the MR scanner during two separate sessions. During the first session they did a practice run before the real task to get familiar with the task (this was left out during the second session). Each session consisted of 24 blocks with 12 trials. The block type of each first block was counterbalanced for each participant. The stimulation condition alternated at each block so that there was never a direct repetition of the same stimulation type in a row and the experiment included equal numbers of each block type (eight blocks per stimulation condition). The entire experiment included the total number of 576 trials (12 trials x 24 blocks x 2 sessions).

tACS procedure and conditions

Simultaneously with performing the AAT in the scanner, the participant received theta-band stimulation (6 Hz) over the aPFC and gamma-band stimulation (75 Hz) over the M1. An electrical current of 1 mA peak-to-peak was used during stimulation. The alternating current stimulation was administered using one transcranial electrical current stimulator for each site. Normally, tACS is applied by means of two single electrode patches that direct the electrical current between electrodes (Fig. 1). However, applying tACS to frontal areas in the classic montage could cause neurosensory effects (such as phosphenes), and does not allow for stimulation with two different frequencies on two different cortical locations, and thus cannot be used for localised theta-gamma band coupled stimulation (Bergmann, Karabanov, Hartwigsen, Thielscher, & Siebner, 2016).

Therefore, we used two sets of ring electrodes that consisted of a smaller inner circle and a larger outside ring, or a centre-ring montage. Each of the sets consisted of a small, circular, inner electrode (diameter 2 cm) and large, outer ring electrode (diameter 13 cm). According to Bergmann et al. (2016), this centre-ring montage set-up allowed for an effective focal stimulation of the area of interest and simultaneously reduced neurosensory side effects (such as phosphenes). For each block, the stimulation was either in-phase, anti-phase or sham (Fig. 2). For the in-phase condition, the power increases of the gamma-band oscillations

Figure 1. Schematic representation of the approach-avoidance task. In each block participants were presented with either affect-congruent conditions (approach-happy and avoid-angry faces) or affect-incongruent conditions (avoid-happy and approach-angry). Participants were instructed to respond accordingly by either pulling the joystick towards themselves or pushing it away from themselves. Reaction times and error rates were measured.

were phase-locked to the peaks of the theta-band oscillations. For the anti-phase condition, the power increases of the gamma-band oscillations were phase-locked to the troughs of the theta-band stimulation. These stimulation patterns were chosen to reflect the findings on increased theta-band power in aPFC and gamma-band power in premotor areas during control in the research of Bramson et al. (2018). For the sham condition, the stimulation was initiated in the first 10 seconds and then cancelled. Participants reported feeling no physical difference between the stimulation conditions. The stimulation blocks were randomly alternated.

Experimental procedure

The participants performed the approach-avoidance task (AAT) in the MR scanner while simultaneously receiving theta-gamma band tACS. Preparation included subject registration, preparation of the scalp sites and the application of the tACS electrodes. First, the Localite transcranial magnetic stimulation (TMS) Navigator was used for subject registration. Registration was done using a T1-weighted image of the subject that was acquired during a previous session and masks of the lateral frontal pole (FPl) and M1 that were specified for each individual subject. The FPl mask was obtained from Neubert et al. (2014) [MNI coordinates: 31 90 37] and the M1 mask from the MEG study by Bramson et al. (2018) [MNI coordinates: -28 -32 64]. We marked the target locations FPl and M1 on the scalp to indicate where the electrodes needed to be attached. Since all participants were right-handed and moved the joystick with their dominant hand, we used the M1 target location on the left hemisphere

(contralateral to the right hand). Consequently, we stimulated the FPl on the right hemisphere due to the size of the electrodes and to prevent interference of the two electrode pairs. Regions under the electrodes were first thoroughly prepared with Nuprep skin preparation gel after which two sets of ring electrodes were attached to the scalp of the subject using conductive paste (Ten20, Weaver and Co., USA) to maximize contact with the scalp. The electrical impedance of the electrodes was measured after application to ensure that the impedance was lower than 10 kΩ.

The participants were instructed about the AAT and the stimulation was tested outside the scanner. The tACS electrodes were connected to the stimulation box when the participant was lying in the scanner. The participant performed the AAT while a multi-band six sequence (MB6) was used to obtain functional images during the task. When the participant finished the task, a MB6 inverse sequence was obtained. Saliva samples were collected before and after the scan to obtain current testosterone and cortisol levels. This was done to either control for or explain interpersonal variabilities in testosterone and cortisol levels on the participant's performance. A practice trial was included during the second session but not during the third session. After the third session, the participant filled in the State Trait Anxiety Inventory (STAI) questionnaire to measure anxiety in adults. The STAI questionnaire was used to look at interpersonal variabilities in trait anxiety symptoms, which could possibly influence the results.

Materials and apparatus

Structural and functional MR data was acquired
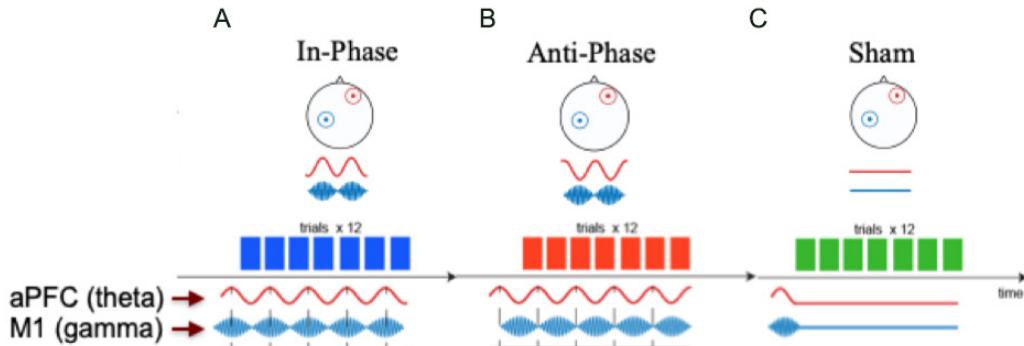
Figure 2. Theta-gamma band stimulation is applied to the primary motor cortex (blue circle) and anterior prefrontal cortex (red circle). A. In-phase stimulation in which the peaks of the slow theta (red) are phase locked to the peaks of the fast gamma (blue) oscillations (theta-gamma band coupling). B. Anti-phase condition in which peaks of the gamma waves are phase-amplitude locked to the troughs of the theta waves (no theta-gamma band coupling). C. Sham condition for which stimulation was only applied during the first 10 seconds and then cancelled.

on the institute 3 T MAGNETOM MRI scanner (Prisma and PrismaFit) (Siemens AG, Healthcare Sector, Erlangen, Germany) with a product 64-channel coil. The stimuli were projected to a screen located at the back of the scanner through an EIKI LC-XL100 beamer (resolution = 1024 x 768, refresh rate = 60 Hz). A mirror attached to the head coil was used to see the screen. The acquisition of functional scans was performed with a MB6 sequence (TR = 1000 ms, TE = 34 ms, flip angle = 60°, 104 x 104 x 68 FOV and 2 x 2 x 2 mm isotropic resolution). To correct for EPI artefacts, an inverse MB6 sequence was performed after the participant completed the main task (TR = 1000 ms, TE = 34 ms, flip angle = 60°, 104 x 104 x 68 FOV and 2 x 2 x 2 mm isotropic resolution). Structural T1 images were acquired in the sagittal orientation using a magnetisation prepared rapid gradient echo (MPRAGE) sequence (TR = 2300 ms, TI = 1100 ms, TE = 3 ms, flip angle = 8°). Parallel imaging (iPAT = 2) was used to accelerate the acquisition resulting in an acquisition time of 5 min and 21 sec.

Behavioural analysis

Reaction times (in ms) and error rates (percentage of correct trials) were used for the analysis of the behavioural data. Reaction times are defined as the time between stimulus presentation and displacement of the joystick. A displacement larger than 30% away from the centre in the sagittal plane was taken as a response. Trials with extremely short reaction times were removed (<150 ms). We also excluded trials for which the reaction times exceeded a threshold of 3 standard deviations above the mean reaction time of the subject per task condition (congruent and incongruent) separately. Reaction time analysis was only performed on correct responses.

Blocks with error rates above chance level were removed (>50% incorrect). Mean reaction times and error rates were calculated for each participant. Factors included stimulation condition (in-phase, anti-phase, sham), and congruency (approach-happy/ avoid-angry, avoid-happy/ approach-angry). We expected improved performance in participants on the AAT in the in-phase condition as compared to the anti-phase and sham condition. Therefore, we expected reduced congruency effects in error rates for the in-phase condition as compared to the anti-phase and sham condition. Previous research by Volman et al. (2011) only found increased congruency effects in error rate (or worse performance on the AAT) after inhibition of aPFC and no effects

on reaction times. Therefore, we expected to find no differences in congruency effects for reaction time between stimulation conditions.

fMRI analyses: Pre-processing steps

FMRI data processing was carried out using FEAT (FMRI Expert Analysis Tool) Version 6.00, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). The brain extraction tool (BET) of FSL was used in order to remove the neck and skull from the imaging data (Smith, 2002) and registration was performed using FNIRT and FLIRT (Andersson, Jenkinson, & Smith, 2007; Jenkinson, Bannister, Brady, & Smith, 2002). Pre-processing steps included motion correction (MCFLIRT), spatial filtering (spatial smoothing with 5mm FWHM Gaussian kernel) and global intensity normalization (Woolrich, Ripley, Brady, & Smith, 2001). The FSL topup tool was used to estimate and correct susceptibility induced distortions caused by distortions in opposite directions of the reversed phase-encoding sequences (Andersson, Skare, & Ashburner, 2003; Smith et al., 2004). Independent component analysis (ICA) in MELODIC was used to manually remove the components of the imaging data that do not reflect task effects, such as movement and cardiovascular responses (Beckmann & Smith, 2004).

fMRI analyses: Effective connectivity analysis (PPI)

The aim of the analysis was to investigate the effects of the different stimulation types on the functional connectivity between the FPl and the M1 during social-emotional actions. For this purpose, we used psychophysiological interactions (PPI) analysis. PPI analysis measures functional connectivity or, in other words, the statistical dependence between activity in between brain areas (O'Reilly, Woolrich, Behrens, Smith, & Johansen-Berg, 2012). This analysis investigates task-specific changes in the relationship between activity in different brain areas using fMRI data, and identifies voxels in which activity is more related to activity in a seed region of interest in a given psychological context (Friston et al., 1997; O'Reilly et al., 2012). More specifically, we tested for significant differences between the contrasts of parameter estimates (COPE) of FPl activity over M1 activity during the affect-incongruent versus the affect-congruent conditions of the AAT.

For this purpose, we created masks within the FPl as a seed region and a mask within M1 as volume of interest (VOI). The FPl, previously registered in

the study of Neubert et al. (2014), was the targeted area during the stimulation and is involved in human social-emotional control. We created separate masks for the FPl of the right hemisphere (rFPl, stimulation site: direct stimulation) and left hemisphere (lFPl, (contralateral to the stimulation site: no direct stimulation) (Fig. 3). First, we selected a voxel within the rFPl that showed a maximum intensity for the contrast between affect-congruent versus affect-incongruent conditions (incongruent – congruent) and created a 5 mm mask surrounding this peak voxel (MNI coordinates rFPl: 33 87 29, radius = 5 mm, intensity = 3.64). Afterwards, we extracted the fMRI time series from the mask (representative time course of activity) in order to create the physiological regressor. The mask and physiological regressor of the lFPl were created by the same procedure (MNI coordinates lFPl: 55 89 29, radius = 5 mm, z-value = 3.48). For the volume of interest, we created a mask surrounding the peak voxel in M1 of the left hemisphere (MNI coordinates M1: -28 -32 64, z-value = 3.70).

FPl-M1 functional connectivity for congruency effects

PPI analysis was run in FEAT according to the guidelines of FSL and the research of O'Reilly (O'Reilly et al., 2012; Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004; Woolrich et al., 2001), for the FPl of the right hemisphere and left hemisphere separately. For the first level analysis, we looked in voxels within the FPl with increased functional connectivity to the M1 during the in-phase, anti-phase and sham stimulation condition.

Explanatory variables included the congruency effects of the stimulation conditions convolved with a haemodynamic response function (HRF) (3 psychological/task regressors: in-phase, anti-phase and sham stimulation)[1] and the time course of the right FPl seed region (physiological regressor). In addition, regressors were included for the covariates of no interest, which included movement parameters (6 regressors: rotation and translation parameters in x/y/z direction), cerebrospinal fluid (CSF) noise, white matter noise and trials with missed responses. For the PPI, we tested the interaction between the seed region and the stimulation condition, or the interactions between the psychological and physiological regressors (in-phase congruency ~ rFPl, anti-phase congruency ~ rFPl, sham congruency ~ rFPl; z-threshold = 3.1). After the PPI, the average parameter estimates (congruency effect) were extracted from the M1 mask (MNI coordinates M1: 59 47 68) at the differences in functional connectivity between the right FPl and M1 during the in-phase, anti-phase and sham condition. Again, the same procedure was followed for the left FPl.

FPl-M1 functional connectivity for congruent versus incongruent trials

In order to investigate the PPI effects underlying
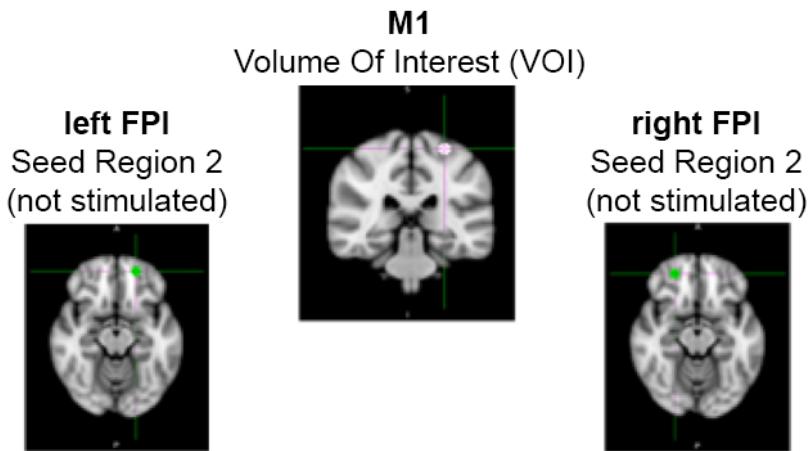
Figure 3. Seed regions and volume of interest (VOI) region of the PPI analysis. Interregional connectivity was investigated between the lateral frontal pole of the right hemisphere (green 5 mm mask at right FPl, seed region 1) and M1 (VOI, white 5 mm mask at primary motor cortex), and between the lateral frontal pole of the right hemisphere (green 5 mm mask at left FPl, seed region 2) and the M1. Right FPl and the M1 received tACS stimulation (stimulation side), whereas the left FPl did not receive local stimulation.

1 Design matrices for the congruency task regressors included the stimulus onset [tPict - start], the duration until response onset [RT] and the congruency condition [congruent = -1, incongruent = 1].

the findings of the VOI analysis above, we separated PPI analyses on the congruent and incongruent conditions for in-phase, anti-phase and sham stimulation. This was again done for the FPl in the left and right hemisphere separately. Task regressors included in-phase congruent, in-phase incongruent, anti-phase congruent, anti-phase incongruent, sham congruent, and sham incongruent[2], and physiological regressors in the left or right FPl. The average parameter estimates were again extracted from the M1 mask for these six factors.

Results

Behavioural outcomes

Average Reaction Times (RT) & Error Rates (ER)

Reaction times and error rates of individual subjects are presented in Figure 4. Mean group reaction times were longer for incongruent (M = 671 ms, SD = 133 ms) as compared to congruent conditions (M = 631 ms, SD = 108 ms) (t = -6.12, p < 0.001). In addition, participants had higher accuracy levels for congruent trials (M = 96.3%, SD = 2.9) as compared to the incongruent trials (M = 94.3%, SD = 4.2%) (t = 3.82, p < 0.001). These findings are in line with results of previous studies using the AAT (Bramson et al., 2018; Roelofs et al., 2008; Volman et al., 2011).

Error Rates

ER congruency effect

The congruency effect of error rates was calculated by subtracting the percentage of correct trials in the incongruent condition from the percentage of correct trials in the congruent condition[3].

We compared the congruency effect of error rates for the in-phase (M = 1.43%, SD = 3.6%), anti-phase (M = 2.51%, SD = 3.2%) and sham (M = 2.21%, SD = 3.5%) stimulation condition (one-sided t-test). A significantly smaller congruency effect in error rates was found for the in-phase as compared to the anti-phase condition (t(40) = -1.71, p = 0.047) (Fig. 5), which indicates an improvement in performance. A trend in reduced congruency effect was also found
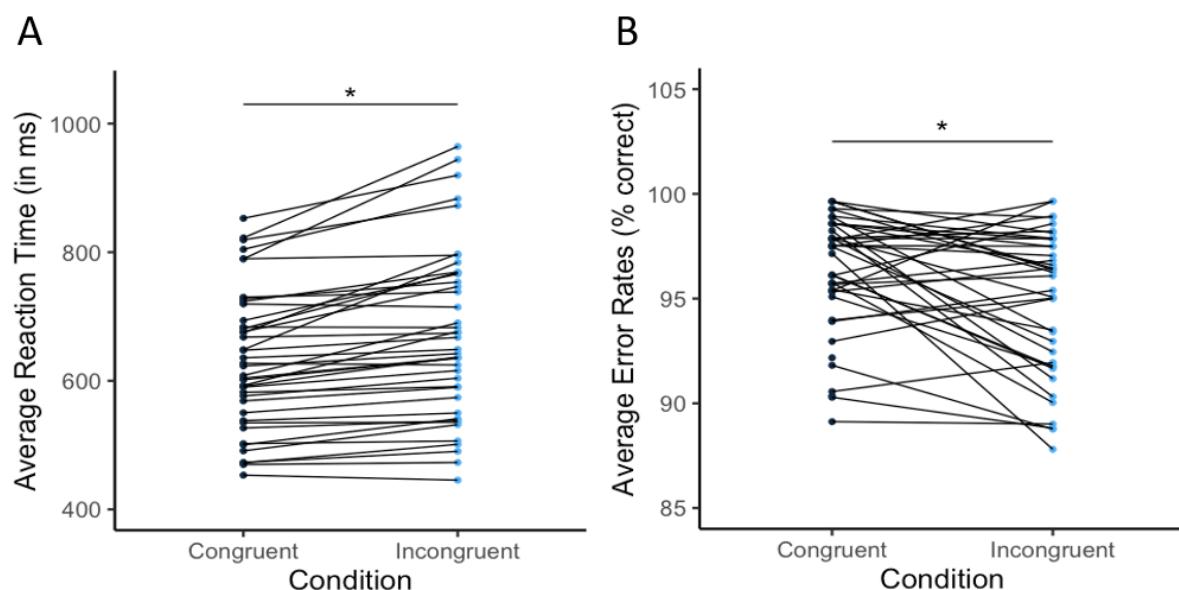
Figure 4. A. Average reaction times (in ms) per individual subject in congruent and incongruent conditions. Mean group reaction time is higher for incongruent conditions compared to congruent conditions, which indicates slower responses when a stimulus is affect-incongruent, and control needs to be applied by the participant. B. Average error rates (% correct) per individual subject in congruent and incongruent conditions. Mean error rates are higher for incongruent conditions as compared to congruent conditions, which indicates that subjects make more errors when control is needed to overcome automatic action tendencies.

2 Design matrices for the congruent task regressors included the stimulus onset [tPict - start], the duration until response onset [RT] and the congruency condition [congruent = 1, incongruent = 0].

Design matrices for the incongruent task regressors included the stimulus onset [tPict - start], the duration until response onset [RT] and the congruency condition [congruent = 1, incongruent = 0].

when comparing the in-phase and sham condition (t(40) = -1.05, p = 0.15). No significant differences in congruency effect were found between the anti-phase and sham condition (t(40) = 0.5, p = 0.31). In line with our hypothesis, we found that theta-gamma tACS reduces the congruency effect and improves performance in the AAT when subjects receive in-phase stimulation while having to control their automatic responses (incongruent condition).

ER congruent versus incongruent

In order to investigate what underlies these congruency effects in error rates, we took the group mean from the percentage of correct responses of the congruent and incongruent trials. Figure 6 shows the congruent and incongruent group means compared for the three stimulation conditions. As seen in the figure, the congruency effects in each stimulation condition are explained by a smaller percentage of correct trials for the incongruent trials (in-phase: M = 95.1%, SD = 3.8%; anti-phase: M = 94.5%, SD = 3.8%; sham: M = 94.8%, SD = 4.1%) as compared to the congruent trials (in-phase: M = 96.6%, SD = 3.6%; anti-phase: M = 97.2%, SD = 2.6%; sham: M = 97.3%, SD = 2.3%). We observed the trend that more errors were made during the in-phase as compared to the anti-phase trials (t(40) = 1.28, p = 0.10). However, we found no significant difference between the other conditions

(p > 0.20). We found no significant differences when comparing the congruent trials of the three stimulation conditions (for all comparisons, p > 0.40).

Reaction Times

RT congruency effect

The congruency effect of the reaction times is calculated by subtracting the mean reaction time (in ms) in the congruent conditions from the reaction time in the incongruent conditions[4] . We compared the congruency effect of reaction times for the in-phase (M = 43.8 ms, SD = 55.4 ms), anti-phase (M = 42.1 ms, SD = 41.1 ms), and sham (M = 48.4 ms, SD = 76.8 ms) stimulation condition (Fig. 5). No significant differences were found in the congruency effects of reaction times for the three stimulation conditions (in-phase vs anti-phase: t(40) = 0.21, p = 0.41; in-phase vs sham: t(40) = -0.50, p = 0.69; anti-phase vs sham: t(40) = -0.59, p = 0.28) (one-sided t-test).

RT congruent versus incongruent

When looking at the congruent and incongruent conditions separately, no significant differences in reaction times were found between stimulation conditions of the congruent trials (in-phase: M = 647 ms, SD = 116 ms, anti-phase: M = 641 ms, SD

A B
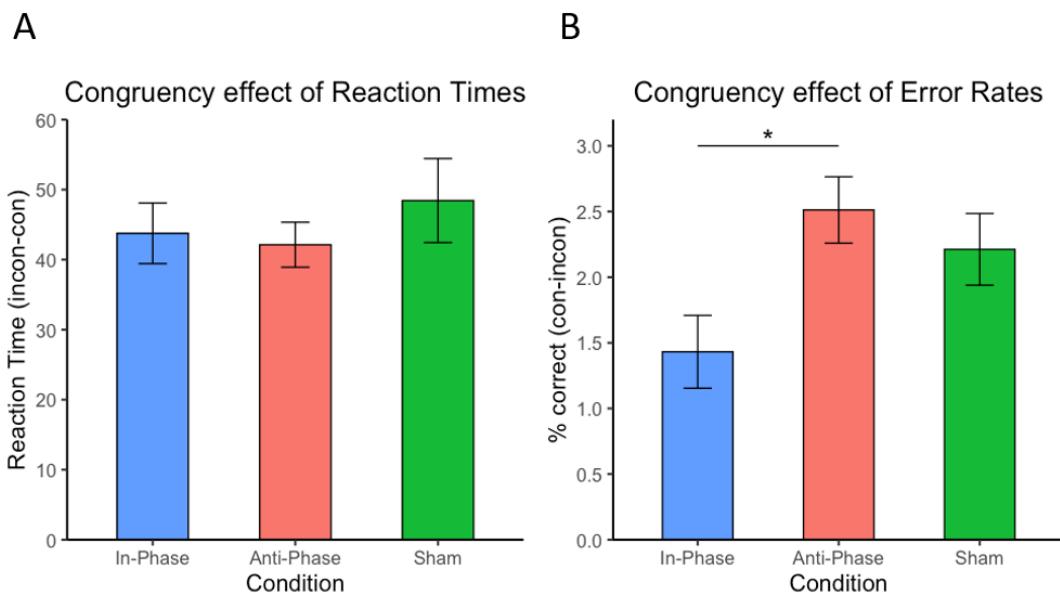
Figure 5. A. Congruency effect on reaction times. No significant differences were found when comparing the congruency effects of reaction times. **B.** Congruency effects on error rates. Reduced congruency effects (improved social-emotional control) was found during in-phase stimulation as compared to anti-phase and sham stimulation. A trend was seen in increased congruency effects for anti-phase compared to sham stimulation.

3 Congruency effect Error Rates = % correct congruent trials - % correct incongruent trials

= 105 ms, sham: M = 638 ms, SD = 106 ms) (p > 0.10 for all comparisons) or incongruent trials (in-phase: M = 690 ms, SD = 143 ms, anti-phase: M = 684 ms, SD = 127 ms, sham: M = 688 ms, SD = 139 ms) (p > 0.10 for all comparisons) (Fig. 6). Within the stimulation conditions, no significant differences were found between the congruent and incongruent conditions of the in-phase, anti-phase or sham condition (p> 0.10 for all comparisons).

These findings show that theta-gamma band stimulation does not result in differences in congruency effect in reaction times for the in-phase condition as compared to the anti-phase and sham condition, thus does not affect the response time of subjects during the task.

Neuroimaging outcomes

Functional connectivity between M1 and rFPl

The goal of our PPI is to identify task-specific changes in the relationship between the FPl and M1 (or functional connectivity) during the different stimulation conditions (O'Reilly et al., 2012). The task effect underlying the PPI is the congruency effect (incongruent – congruent trials). We found a negative task-specific change in functional connectivity between the right FPl (seed region) and the M1 (VOI) for in-phase stimulation (M = -6.5, SD = 2.6) and sham stimulation (M = -2.1, SD = 2.6) (Fig. 7B). These findings indicate that for in-phase and sham stimulation there is a stronger relationship (or stronger functional connectivity) between right

FPl and M1 in congruency effect. This relation is in the negative direction, which means that there is a negative relation between the difference in FPl activation for congruent and incongruent trials on the one hand and the difference in M1 activation between the congruent and incongruent trials on the other hand (incongruent-congruent difference scores are negatively correlated). Or, in other words, more positive congruency effects in the FPl are linked to more negative congruency effects in the M1. In contrast, we found that for the anti-phase stimulation (M = 0.8, SD = 1.2) there is a positive change in functional connectivity between the right FPl and M1. When comparing stimulation conditions, we find significantly stronger functional connectivity change for in-phase stimulation compared to anti-phase stimulation (t(40) = -2.1, p = 0.04). No significant differences are found between in-phase and sham stimulation (t(40) = -1.2, p = 0.22), and between anti-phase and sham stimulation (t(40) = 0.7, p = 0.43). These results indicate that when a subject receives in-phase stimulation there is a stronger functional connectivity change between the FPl and M1 for the congruency effect as compared to anti-phase and sham stimulation.

Functional connectivity between M1 and lFPl

Likewise, for the left FPl, we found a negative task-specific change in functional connectivity between the left FPl and the M1 for in-phase stimulation (M = -6.0, SD = 1.9) and sham stimulation (M = -2.6, SD = 2.1) while positive effects were seen for anti-phase stimulation (M = 3.1, SD = 2.5)(Fig. 7A).
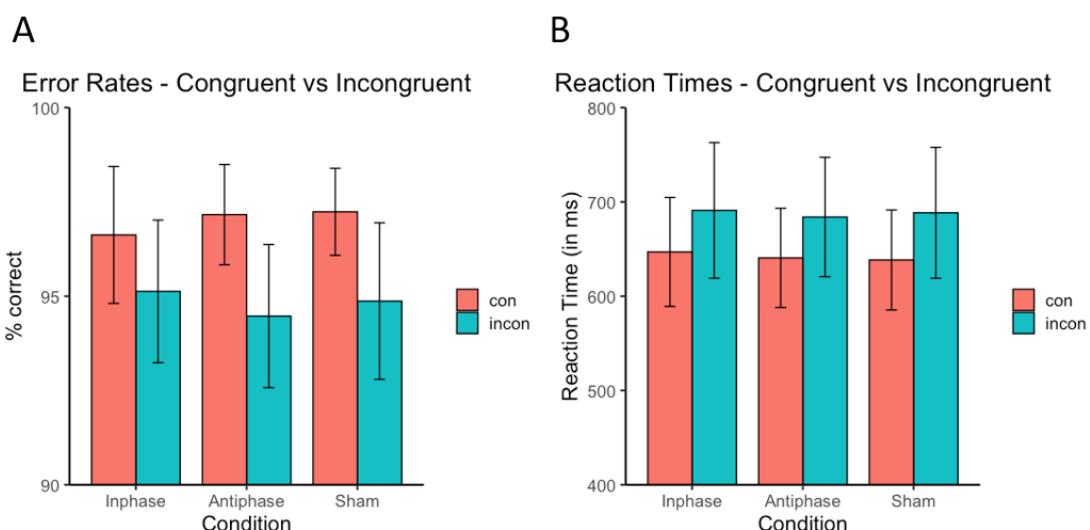
**Figure 6. A.** Error rates for congruent versus incongruent trials. **B.** Reaction times for in-phase, anti-phase and sham stimulation.

4 Congruency effect Reaction Time = mean reaction time incongruent trials - mean reaction time congruent trials

Significant differences were found for in-phase as compared to anti-phase stimulation ($t(40) = -2.7$, $p = 0.01$, SD = 22) and for anti-phase compared to sham stimulation ($t(40) = 2.2$, $p = 0.03$, SD = 16.9). The effect differences were not significant between in-phase and sham stimulation ($t(40) = -1.1$, $p = 0.26$, SD = 18.8). These results replicate the findings found in the right FPl, thereby showing that the effects are transferable to the contralateral hemisphere.

PPI congruent versus incongruent

These findings are supported by the PPI effects between FPl and M1 for congruent and incongruent conditions (Fig. 9). In-phase stimulation shows more positive effects for congruent (lFPl: M = 7.8, SD = 26 ; rFPl: M = 9.3, SD = 26) compared to incongruent trials (lFPl: M = -3.9, SD = 14, rFPl: M = -3.2, SD = 18). These findings are widely interpretable but could support the earlier finding that there is a larger task-related difference in functional connectivity (or stronger functional connectivity) between FPl and M1 for congruent trials as compared to incongruent trials.

Surprisingly, large negative PPI effects in the sham condition are found for the incongruent condition (lFPl: M = -5.8, SD = 15; rFPl: M = -7.2, SD = 21) as compared to the congruent condition (lFPl: M = -3.0, SD = 26 ; rFPl: M = -3.2, SD = 34). Smaller differences are found for the anti-phase condition between congruent trials (lFPl: M = -2.9, SD = 20; rFPl: M = 1.0, SD = 25) and incongruent trials (lFPl: M = 2.8, SD = 24 ; rFPl: M = 3.2, SD = 28). Similar PPI effects are found for both the right as well as the left FPL.

Whole-brain effects

A post hoc whole-brain PPI analysis was performed to investigate the functional connectivity between the aPFC and other regions in the brain comparing the in-phase and anti-phase stimulation conditions. For this purpose, we used the right FPl (stimulation site) as seed region and looked at whole-brain effects in the contrast between in-phase and anti-phase stimulation. Cluster-corrected results show clusters with increased activity in the right FPl (max intensity = 4.18, voxels = 1143, $p < 0.0001$, MNI coordinates = 36 79 41), cingulate gyrus (intensity = 3.71, voxels = 811, $p < 0.0001$, MNI coordinates = 45 40 46), right thalamus (max intensity = 3.7, voxels
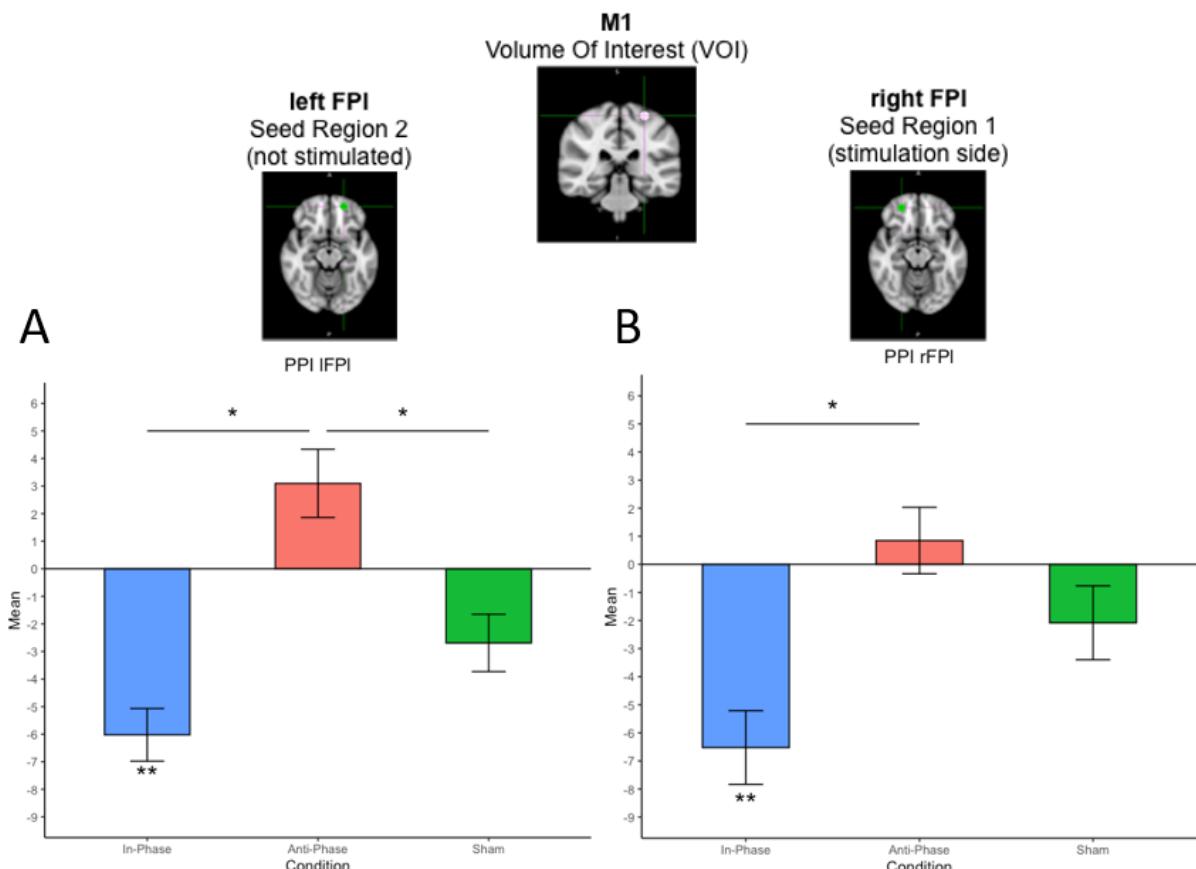
Figure 7. PPI effects. A. Left lateral frontal pole B. Right lateral frontal pole. Larger (negative) functional connectivity changes were found for in-phase as compared to anti-phase and sham stimulation.
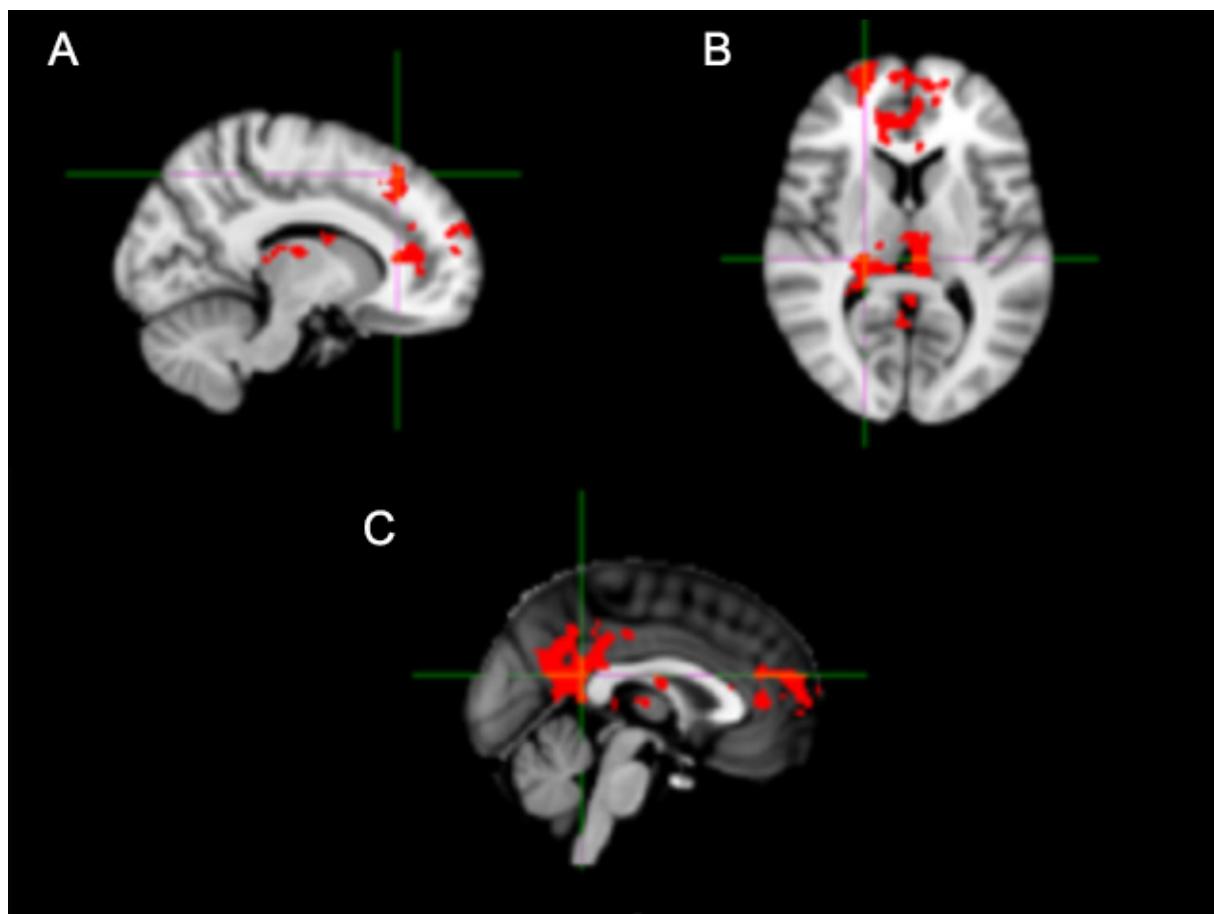
Figure 8. PPI cluster-corrected results for in-phase versus anti-phase stimulation. These individual images show clusters **A.** superior frontal gyrus and FPl **B.** Thalamus and pulvinar **C.** cingulate gyrus.
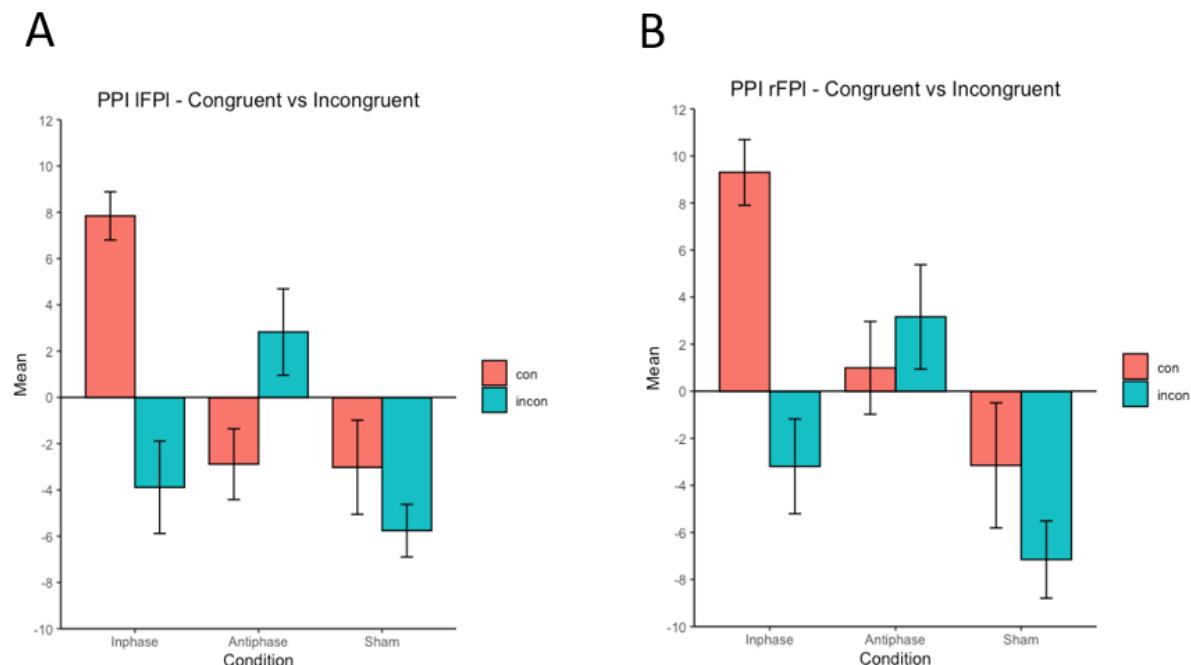
Figure 9. PPI effects for congruent and incongruent conditions separately. **A.** PPI effects for the left FPl **B.** PPI effects for the right FPl for congruent and incongruent trials. A larger functional connectivity was found the in-phase congruent as compared to incongruent trials. For the other stimulation conditions, larger effects were found for incongruent compared to congruent trials.

= 650, p < 0.001, MNI coordinates = 34 50 40) and superior frontal gyrus (max intensity = 4.3, voxels = 379, p < 0.01, MNI coordinates = 39 79 61) (Fig. 8). High intensity values were found in the right FPl (-3.67), left FPl (-2.91), pulvinar in left thalamus (-3.23), precuneus (part of the parietal lobule of the brain) (-3.0) and the cingulate gyrus (-3.5).

Correlation neuroimaging data and behavioural data

The relation between behavioural performance and functional connectivity for the three stimulation conditions is represented in Figure 10. For the right FPl, a significant positive correlation between congruency effect in reaction times and PPI score for the anti-phase condition was found (r = 0.3752, p = 0.0156). For the left FPl, a significant negative correlation between congruency effect in error rates and PPI score was found for the sham condition (r = -0.3371, p = 0.0312). No significant correlations were found for Reaction Time x PPI or Error Rate

x PPI in the other stimulation conditions for either the lFPl or the rFPl. These results show that there is limited to no interaction between performance and functional connectivity.

Discussion

Influence of theta-gamma coupled tACS on social-emotional actions

Our findings show that theta-gamma band coupled tACS influences social-emotional control and functional aPFC-M1 connectivity. Firstly, in-phase stimulation leads to improvement in overriding automatic emotional action tendencies when one is engaged in social-emotional control, compared to anti-phase stimulation. This confirms our hypothesis and shows that we can improve social-emotional behaviour when coupling theta- and gamma-oscillations between the aPFC and M1. This finding is supported by a stronger functional
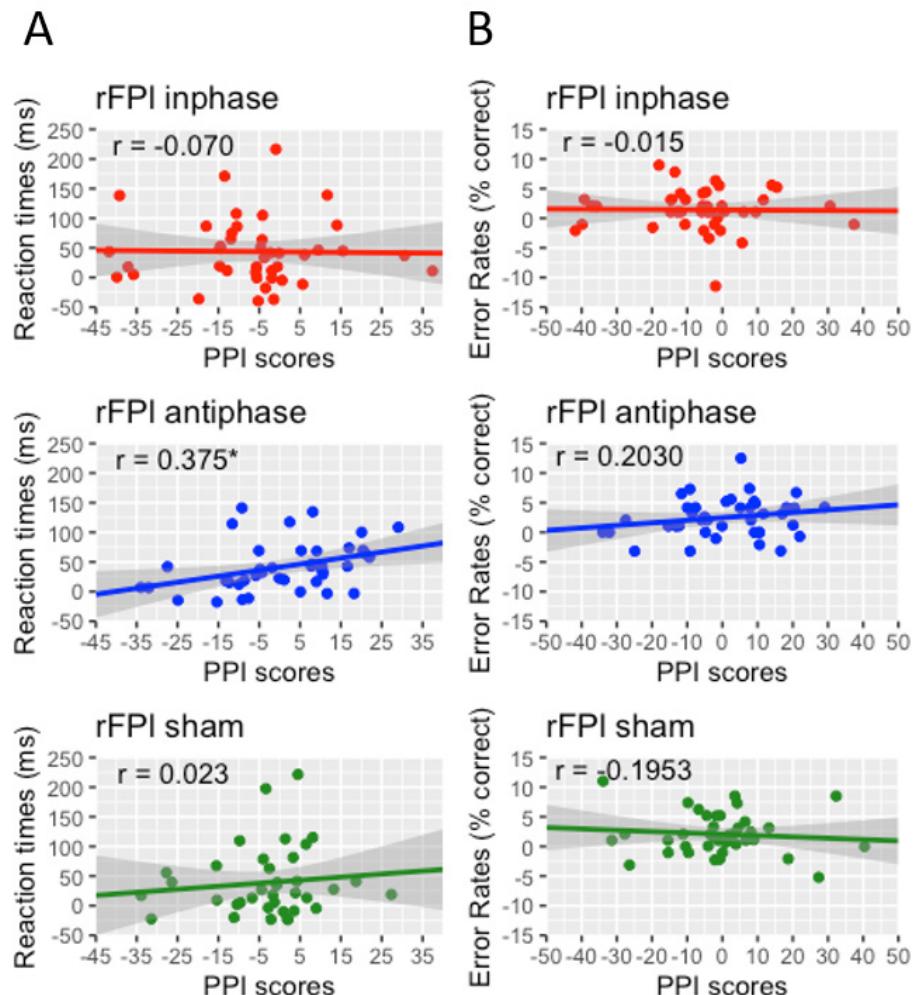

Figure 10. A. interaction between PPI scores of right FPl and congruency effect of reaction times **B.** PPI scores of right FPl and congruency effect of error rates.

connectivity between the lateral frontal pole and the primary motor cortex for in-phase as compared to the other stimulation conditions. Even though no interaction was found between behavioural performance and functional connectivity, these results may suggest that in-phase theta-gamma band stimulation improves social emotional control and strengthens the interregional connectivity in regions involved in social-emotional control.

Effect of theta-gamma band coupled tACS stimulation on approach/avoidance behaviour

In line with our hypothesis, subjects show improved control over social-emotional actions (reduced congruency effect) when stimulated in-phase as compared to anti-phase and sham stimulation. This improvement is specifically relevant for incongruent conditions, where control is needed to overcome automatic emotional action tendencies when avoiding happy faces and approaching angry faces. We applied theta-band tACS to the aPFC and gamma-band tACS to the M1 to influence theta-gamma coupling between the two brain regions. During the in-phase stimulation, the power increases of the M1 gamma-band stimulation are phase-locked to the peaks of the aPFC theta-band stimulation. In contrast, during anti-phase stimulation, these gamma-band power increases are phase-locked to the troughs of the aPFC theta-band stimulation. Our behavioural findings suggest that we can use in-phase theta-gamma coupled tACS to influence underlying cognitive processes related to social-emotional control. However, anti-phase tACS did not lead to significant impairments of social-emotional control, suggesting that desynchronising these oscillatory patterns does not lead to major disruptions in communication between the aPFC and M1.

Our finding that in-phase theta-gamma coupled stimulation enhances control supports the theory that social-emotional control exerted by the aPFC may be implemented by theta-gamma coupled oscillations. This is consistent with previous work by Bramson et al. (2018) which found stronger pre-movement gamma-band oscillations of premotor areas[5] during affect-incongruent trials compared to affect-congruent trials. Power increases of M1 gamma-band oscillations were phase locked to the peaks of aPFC theta-band oscillations, suggesting that these

gamma-band power increases during control could be guided by long-range communication between the aPFC and M1. Our research suggests that these oscillatory effects can be replicated by externally influencing the spike-timing of theta- and gamma-oscillations in these regions during social-emotional control. The spike-timing depends on the relative timing of pre- and postsynaptic action potentials (spike-timing dependent plasticity; (Song, Miller, & Abbott, 2000)). This means that the magnitude of change in synaptic strength between pre- and postsynaptic neurons is based on the timing of neuronal spikes. When the presynaptic spike reaches the postsynaptic neuron before the latter fires, then the synapse gets potentiated. When the presynaptic spike reached the postsynaptic neuron after it fires, the synapse gets depressed. When applying in-phase stimulation the presynaptic spike would be externally influenced by the tACS before the postsynaptic neuron fires after which the synapse would be potentiated. As a consequence, this potentiation would lead to an enhanced communication between aPFC and M1, thereby improving the ability of the aPFC to implement social-emotional control.

We expected that the participants would show an increase in congruency effect for the anti-phase condition as compared to the sham condition, which indicates that participants get worse in controlling their actions. While a trend was seen in this direction, no significant changes in congruency effect were found in the anti-phase and sham conditions for both reaction times as well as error rates. We speculate that during anti-phase coupling, the stimulation would influence the cortical oscillations such that the power increases of the gamma-band oscillations would phase-lock to the troughs of the theta-band oscillations of the aPFC. The inhibitory effects of theta-band oscillations would make it unlikely that presynaptic activity in aPFC drives postsynaptic spikes in the M1. Consequently, neural coherence and information transfer between the aPFC and the sensorimotor cortex would be impaired. This could lead to limited communication between the aPFC and distant regions as well as disrupted action preparation and selection by the premotor areas.

Looking at influencing social-emotional control specifically, our findings are in line with research by Volman et al. (2011), who found disruption of social-emotional control when the aPFC was inhibited using continuous theta-burst stimulation. They also only found differences in behavioural effects in error rates and not in reaction times. This

---

5 contralateral to the responding hand

implies that non-invasive stimulation affects one's ability to override automatic emotional tendencies but does not affect the time to respond to affect-incongruent trials.

In a broader perspective, our findings are in support of previous studies using a multi-electrode tACS set-up to influence complex cognitive processes. Previous research showed that the precision of value-based choice in humans can be manipulated by (de-)synchronising phase-coupling between parietal and frontal regions (Polanía et al., 2015). When tACS was used to disrupt fronto-parietal coherence, participants made less accurate choices between value-based food choices without affecting closely matched perceptual decisions. Other research found cognitive improvement on the revival of working memory (Reinhart & Nguyen, 2019). They applied customised electrical stimulation based on theta-gamma coupling to individual network dynamics in older adults. The revival effects lasted for 50 minutes, which may suggest that this technique can assist in long term improvements for cognitive processes. These studies applied tACS in a circle-ring set-up similar to our study. Our research gives further support to that hypothesis that tACS could influence coupling or decoupling of behaviourally relevant neural oscillations between distant cortical areas, and even has the possibility of long-term changes in cognitive processes.

Functional connectivity change during theta-gamma coupled tACS

Congruency effect

We investigated the effect of the different stimulation conditions by looking at the change in functional connectivity between the aPFC and the M1 during social-emotional actions. We found stronger aPFC-M1 connectivity after applying in-phase stimulation as compared to anti-phase and sham stimulation. These connectivity estimates are highly similar in contralateral aPFC-M1, indicating that this is not an artefact of potential distortion due to the electrode or electrical currents applied over the right aPFC. Therefore, the task-specific increase in the relationship between the aPFC and M1 (a PPI effect) is suggestive of an increased information flow between these regions. Together with the behavioural results, these findings suggest that the communication between these areas is improved by externally synchronising theta-gamma coupling resulting in improvement in social-emotional control.

More specifically, we found stronger functional connectivity effects for in-phase stimulation as compared to the anti-phase and sham conditions. This negative relation could suggest that more positive congruency effects in the FPl (larger difference in activity between congruent and incongruent trials) are linked to more negative congruency effects in the M1 (smaller difference in activity between congruent and incongruent trials). We suggest that this effect could be explained by the inhibitory nature of theta-band oscillations that would lead to stronger inhibitory effects implemented by the aPFC on other brain regions.

Congruent versus incongruent

We found surprising results in functional FPl-M1 connectivity when separating congruent and incongruent conditions. When applying sham stimulation, we found stronger task-related functional FPl-M1 connectivity changes for incongruent as compared to congruent conditions. This is what we expected, as we hypothesised that during affect-incongruent trials the aPFC/FPl has to exert more control in order to overcome automatic action tendencies. There would be a stronger functional connectivity to establish this and an increase in information flow between the FPl and M1. For anti-phase stimulation, larger PPI effects are also found for incongruent as compared to congruent trials but in the positive direction, which could indicate that more negative effects in the FPl are linked to more positive effects in the M1. This is surprising, as we would rather expect weaker functional connectivity between the aPFC and M1 that is related to impaired control. Both sham and stimulation show larger functional connectivity changes when control needs to be exerted during approach and avoidance actions (incongruent) compared to when less control is needed (congruent).

In line with the reasoning above, we therefore expected that during in-phase stimulation the FPl-M1 functional connectivity would be even stronger, matching the increased information transfer related to the improved behavioural performance we found. However, instead we found the opposite results, namely a stronger change in task-related activity for the congruent condition as compared to the incongruent condition. One speculation could be that, during congruent conditions, the aPFC does not need to be involved in that many control-regulating processes and therefore there would not be a lot of change in this low neural activity baseline. Stimulation would then increase neural activity even more from this low baseline as compared to when

the aPFC is already engaged in increased neural firing during control. However, further research into this matter is needed as we were not able to directly measure this baseline of neuronal activity. We can only speculate about the processes underlying these functional connectivity effects.

Interpretational issues

This study only focussed on the connectivity between aPFC and the motor cortex in social-emotional control. However, the connections between the aPFC and other regions should be taken into account when looking at influencing social-emotional control. We found no functional connectivity between the aPFC and the M1 during the whole-brain analysis but did find functional connectivity with the thalamus/pulvinar, cingulate gyrus/precuneus and superior frontal gyrus. Prefrontal-pulvinar functional connectivity was also found in the study by Tyborowska and colleagues (2016) as an alternative pathway for effective emotional control to the prefrontal-amygdalar connection. Functional connectivity between the aPFC and precuneus (part of the parietal lobule of the brain) can be explained by its role in emotion and social cognition. Moreover, this area shows connections with the dorsalmost nuclei of the thalamus (lateral pulvinar, amongst others; (Cavanna & Trimble, 2006), and lower task related deactivation is found for people with social anxiety as compared to healthy controls (Gentili et al., 2009). The mechanisms underlying social-emotional control are very complicated and are far more extensive than the interaction between prefrontal and sensorimotor areas.

Control over social-emotional actions by the aPFC or the FPl is likely mediated by the amygdala (Bramson et al., 2019). Impairment of prefrontal-amygdala connectivity, for example in social anxiety, is related to problems in social-emotional regulation (Gold et al., 2016; Radke et al., 2013). Previous research has shown that the aPFC inhibits the amygdala, which is involved in automatic emotion evaluation, and upregulates areas involved in action selection during social-emotional control (Volman et al., 2011; Volman et al., 2013). Therefore, we suggest that in-phase theta-gamma coupled stimulation could lead to a larger suppression of the aPFC over the amygdala and an enhanced communication between the aPFC and the M1. This mechanism could explain why participants become better at inhibiting the automatic emotional tendencies and improve in exerting social-emotional control.

Subsequently, we only found small improvements in behavioural performance when applying in-phase theta-gamma coupled stimulation as compared to the other stimulation types. Generally, healthy participants are already quite good at inhibiting their automatic emotional tendencies and exerting social-emotional control, so we did not expect to find very large improvements. Theta-gamma coupled tACS as an intervention strategy could therefore be more effective for people who experience problems in exerting social-emotional control. These people would benefit more from in-phase stimulation to improve the communication between prefrontal and motor areas, possibly mediated by a larger suppression of the amygdalar response, resulting in greater improvement in correctly inhibiting their automatic emotional tendencies.

Our hypothesis was based on the idea that theta-gamma band stimulation leads to the synchronisation of the neuronal activity of the targeted areas. However, another possibility is that the reported effects are caused by transcutaneous stimulation of peripheral nerves (Asamoah, Khatoun, & Mc Laughlin, 2019). It could be that tACS stimulation itself does not entrain neural activity but that the rhythmic activity from peripheral nerves entrains cortical neurons. We argue against this notion as our circle-ring set-up sends electrical currents between the inner and outer electrode instead of between two distant locations. This allows the current to flow over the superficial layer of the cortex. More importantly, the frontal and motor areas are stimulated in two different frequencies that are either in-phase or anti-phase. Behavioural and functional connectivity differences as a consequence of these stimulation types show that the effects are not just due to peripheral nerve stimulation but due to transcranial stimulation.

Conclusion

Taken together, our findings show that control over social-emotional actions can be influenced by theta-gamma band coupled tACS. When applying theta-band stimulation to the aPFC that is phase-locked to the power increase of gamma-band stimulation to the M1, participants showed improvements in overriding their automatic emotional action tendencies. Moreover, in-phase theta-gamma coupled stimulation was linked to stronger FPl-M1 functional connectivity. This suggests that in-phase stimulation can be used as a technique to facilitate communication between these areas. We speculate that the mechanism

underlying the improvement in social-emotional control would be an enhancement in theta-gamma coupling caused by externally synchronising theta- and gamma-band oscillations. This research has a high clinical relevance, as the technique could be used in treatment of patients who have difficulties regulating their social-emotional behaviour, such as in social anxiety. tACS influences naturally occurring oscillations and only leads to synchronisation of neuronal activity when one already engages in related cognitive processes, such as control. Therefore, the technique could be used as an addition to extinction therapy, where people are exposed repeatedly to stressful cues and situations, thereby improving the effect of the therapy and increase speed of recovery. This technique could provide promising possibilities for treatment of disorders that are related to social-emotional control.

References

Alekseichuk, I., Turi, Z., de Lara, G. A., Antal, A., & Paulus, W. (2016). Spatial working memory in humans depends on theta and high gamma synchronization in the prefrontal cortex. Current Biology, 26(12), 1513-1521.

Andersson, J. L., Jenkinson, M., & Smith, S. (2007). Non-linear registration aka Spatial normalisation FMRIB Technial Report TR07JA2. FMRIB Analysis Group of the University of Oxford.

Andersson, J. L., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. Neuroimage, 20(2), 870-888.

Antal, A., & Paulus, W. (2013). Transcranial alternating current stimulation (tACS). Current Biology.

Ardid, S., Vinck, M., Kaping, D., Marquez, S., Everling, S., & Womelsdorf, T. (2015). Mapping of functionally characterized cell classes onto canonical circuit operations in primate prefrontal cortex. Journal of Neuroscience, 35(7), 2975-2991.

Asamoah, B., Khatoun, A., & Mc Laughlin, M. (2019). tACS motor system effects can be caused by transcutaneous stimulation of peripheral nerves. Nature communications, 10(1), 266.

Beckmann, C. F., & Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE transactions on medical imaging, 23(2), 137-152.

Bergmann, T. O., Karabanov, A., Hartwigsen, G., Thielscher, A., & Siebner, H. R. (2016). Combining non-invasive transcranial brain stimulation with neuroimaging and electrophysiology: current approaches and future perspectives. Neuroimage, 140, 4-19.

Bramson, B., Folloni, D., Verhagen, L., Hartogsveld, B., Mars, R. B., Toni, I., & Roelofs, K. (2019). Human lateral Frontal Pole contributes to control over social-emotional behaviour. bioRxiv, 584896.

Bramson, B., Jensen, O., Toni, I., & Roelofs, K. (2018). Cortical oscillatory mechanisms supporting the control of human social-emotional actions. Journal of Neuroscience, 3382-3317.

Buzsáki, G., & Wang, X.-J. (2012). Mechanisms of gamma oscillations. Annual review of neuroscience, 35, 203-225.

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. Trends in cognitive sciences, 18(8), 414-421.

Cavanna, A. E., & Trimble, M. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. Brain, 129(3), 564-583.

Colgin, L. L. (2013). Mechanisms and functions of theta rhythms. Annual review of neuroscience, 36, 295-312.

Cooper, P. S., Wong, A. S., Fulham, W. R., Thienel, R., Mansfield, E., Michie, P. T., & Karayanidis, F. (2015). Theta frontoparietal connectivity associated with proactive and reactive cognitive control processes. Neuroimage, 108, 354-363.

Donner, T. H., Siegel, M., Fries, P., & Engel, A. K. (2009). Buildup of choice-predictive activity in human motor cortex during perceptual decision making. Current Biology, 19(18), 1581-1585.

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top–down processing. Nature Reviews Neuroscience, 2(10), 704.

Friston, K., Buechel, C., Fink, G., Morris, J., Rolls, E., & Dolan, R. (1997). Psychophysiological and modulatory interactions in neuroimaging. Neuroimage, 6(3), 218-229.

Gentili, C., Ricciardi, E., Gobbini, M. I., Santarelli, M. F., Haxby, J. V., Pietrini, P., & Guazzelli, M. (2009). Beyond amygdala: default mode network activity differs between patients with social phobia and healthy controls. Brain research bulletin, 79(6), 409-413.

Gold, A. L., Shechner, T., Farber, M. J., Spiro, C. N., Leibenluft, E., Pine, D. S., & Britton, J. (2016). Amygdala–cortical connectivity: associations with anxiety, development, and threat. Depression anxiety, 33(10), 917-926.

Goodman, M. S., Kumar, S., Zomorrodi, R., Ghazala, Z., Cheam, A. S., Barr, M. S., . . . Flint, A. (2018). Theta-gamma coupling and working memory in Alzheimer's dementia and mild cognitive impairment. Frontiers in Aging Neuroscience, 10, 101.

Herrmann, C. S., Rach, S., Neuling, T., & Strüber, D. (2013). Transcranial alternating current stimulation: a review of the underlying mechanisms and modulation of cognitive processes. Frontiers in human neuroscience, 7, 279.

Heuer, K., Rinck, M., & Becker, E. S. (2007). Avoidance of emotional facial expressions in social anxiety: The approach–avoidance task. Behaviour research therapy, 45(12), 2990-3001.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate

linear registration and motion correction of brain images. Neuroimage, 17(2), 825-841.

Lisman, J. E., & Jensen, O. (2013). The theta-gamma neural code. Neuron, 77(6), 1002-1016.

Neubert, F.-X., Mars, R. B., Thomas, A. G., Sallet, J., & Rushworth, M. F. (2014). Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. Neuron, 81(3), 700-713.

O'Reilly, J. X., Woolrich, M. W., Behrens, T. E., Smith, S. M., & Johansen-Berg, H. (2012). Tools of the trade: psychophysiological interactions and functional connectivity. Social cognitive affective neuroscience, 7(5), 604-609.

Polanía, R., Moisa, M., Opitz, A., Grueschow, M., & Ruff, C. (2015). The precision of value-based choices depends causally on fronto-parietal phase coupling. Nature communications, 6, 8090.

Radke, S., Roelofs, K., & De Bruijn, E. R. (2013). Acting on anger: social anxiety modulates approach-avoidance tendencies after oxytocin administration. Psychological Science, 24(8), 1573-1578.

Reinhart, R. M., & Nguyen, J. A. (2019). Working memory revived in older adults by synchronizing rhythmic brain circuits. Nature neuroscience, 22(5), 820.

Roelofs, K., Elzinga, B. M., & Rotteveel, M. (2005). The effects of stress-induced cortisol responses on approach–avoidance behavior. Psychoneuroendocrinology, 30(7), 665-677.

Roelofs, K., Minelli, A., Mars, R. B., Van Peer, J., & Toni, I. (2008). On the neural control of social emotional behavior. Social cognitive affective neuroscience, 4(1), 50-58.

Schoffelen, J.-M., Poort, J., Oostenveld, R., & Fries, P. (2011). Selective movement preparation is subserved by selective increases in corticomuscular gamma-band coherence. Journal of Neuroscience, 31(18), 6750-6758.

Smith, S. M. (2002). Fast robust automated brain extraction. Human brain mapping, 17(3), 143-155.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., . . . Flitney, D. E. (2004). Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage, 23, S208-S219.

Song, S., Miller, K., & Abbott, L. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nature neuroscience, 3(9), 919.

Tamnes, C. K., Østby, Y., Fjell, A. M., Westlye, L. T., Due-Tønnessen, P., & Walhovd, K. B. (2009). Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure. Cerebral cortex, 20(3), 534-548.

Tyborowska, A., Volman, I., Smeekens, S., Toni, I., & Roelofs, K. (2016). Testosterone during puberty shifts emotional control from pulvinar to anterior prefrontal cortex. Journal of Neuroscience, 36(23), 6156-6164.

Varela, F., Lachaux, J.-P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. Nature Reviews Neuroscience, 2(4), 229.

Volman, I., Roelofs, K., Koch, S., Verhagen, L., & Toni, I. (2011). Anterior prefrontal cortex inhibition impairs control over social emotional actions. Current Biology, 21(20), 1766-1770.

Volman, I., Verhagen, L., den Ouden, H. E., Fernández, G., Rijpkema, M., Franke, B., . . . Roelofs, K. (2013). Reduced serotonin transporter availability decreases prefrontal control of the amygdala. Journal of Neuroscience, 33(21), 8974-8979.

Voytek, B., Kayser, A. S., Badre, D., Fegen, D., Chang, E. F., Crone, N. E., . . . D'esposito, M. (2015). Oscillatory dynamics coordinating human frontal networks in support of goal maintenance. Nature neuroscience, 18(9), 1318.

Womelsdorf, T., Schoffelen, J.-M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., & Fries, P. (2007). Modulation of neuronal interactions through neuronal synchronization. Science, 316(5831), 1609-1612.

Woolrich, M. W., Behrens, T. E., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. Neuroimage, 21(4), 1732-1747.

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. Neuroimage, 14(6), 1370-1386.

Unravelling the Neurocognitive Mechanisms Underlying Counter-Conditioning

Jette de Vos[1,2]
Supervisors: Maxime Houtekamer[1,2], Dr. Marijn Kroes[1,2]

[1]*Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands*
[2]*Radboud University Medical Centre, Donders Institute for Brain, Cognition and Behaviour, The Netherlands*

Stress- and anxiety-related disorders are devastating for patients and a major burden on society. Although effective, a robust number of patients does not improve or relapse following exposure-based treatments. This highlights the need to find alternatives, like counter-conditioning. In counter-conditioning, the aversive outcome is replaced by a rewarding outcome instead of merely omitting the aversive outcome. Here we aim to unravel the effectiveness and neurocognitive mechanisms underlying counter-conditioning. We hypothesize that counter-conditioning either 1) enhances extinction learning, 2) integrates a fear and reward memory, or 3) overwrites the fear memory by a reward memory. We tested this in a two-day between-subjects study. On day 1, participants were fear conditioned to one of the two stimulus categories (animal vs. object pictures). Next, half of the participants underwent counter-conditioning ($N = 10$) during which novel trial-unique exemplars from the fear conditioned category were paired with a reward using monetary incentive delay. The other participants ($N = 10$) received standard extinction training. On day 2, return of fear responses and episodic memory for items from conditioning and counter-conditioning/extinction were tested. Results (pupil dilation, skin conductance (SCR), heart rate, reaction times, episodic memory, fMRI (functional magnetic resonance imaging) indicate feasibility of the design to address our hypothesis that one of the three above mentioned mechanisms supports the effectiveness of counter-conditioning. Moreover, category-specific conditioned SCRs seem to reduce faster during counter-conditioning than extinction, supporting the idea of counter-conditioning as a promising alternative to extinction. In the counter-conditioning group, episodic memory is enhanced for emotionally salient items from both the conditioning and counter-conditioning phase, which is in line with the hypothesis that counter-conditioning integrates a fear and reward memory.

Keywords: fear conditioning, extinction learning, counter-conditioning, fMRI

Corresponding author: Jette de Vos; E-mail: jette-de-vos@hotmail.com

Learning what to fear is a very valuable ability in everyday life. Imagine, if a dog bites you while walking through the park, a fear memory is formed of the association between the park and the possibility of getting bitten. This association will help you to be more cautious in the park and therefore to avoid getting bitten another time. If the dog were to move out of the neighbourhood, however, it would also be helpful to learn that the park is safe again and that you can have a relaxed walk in it in the future. Both learning aspects are important: the learning of the fear association and, if situations change as in the example above, the learning of the safety association. In practice, however, the first aspect, the fear association, turns out to be more persistent than the second aspect, the safety association. Clinical disorders such as specific phobias are examples of instances of difficult-to-eradicate fear associations (Vervliet, Craske, & Hermans, 2013). Disorders such as these are mostly treated with exposure-based therapies. However, a considerable number of patients who undergo treatment does not benefit. In addition, a large number of patients relapses after treatment and the fear association returns (Craske & Mystkowski, 2006), creating the need for finding alternative methods that are more effective in preventing the return of the fear association. Counter-conditioning may be such a promising method as supported by results from previous studies using counter-conditioning (Kerkhof, Vansteenwegen, Baeyens, & Hermans, 2011; Kaag et al., 2016; Karel et al., 2019). However, why it might be more effective remains unknown.

Conditioning and extinction

In laboratory settings, fear learning is mostly studied with Pavlovian classical conditioning paradigms (Pavlov, 1927). In classical conditioning, an intrinsically neutral stimulus (the conditioned stimulus, CS), for example a sound, is coupled with an aversive or appetitive stimulus (the unconditioned stimulus, US), for example an electrical shock. After (repeated) coupling of this CS and US, the CS gets to serve as a predictor of the aversive outcome (i.e., the shock). As a consequence, the CS, which initially did not evoke a response by itself, can evoke a response (the conditioned response, CR) now. The CR can be the same as the response evoked by the US (the unconditioned response, UR), although not necessarily. Conditioned associations can be both formed with aversive and appetitive unconditioned stimuli. Therefore, the conditioning paradigm can be used to study either memories with a positive or

negative associative value, depending on the valence of the US. As we will focus on fear conditioning in the current study, most examples used in the rest of this paper will address aversive rather than appetitive conditioning. The stimulus that is coupled with the US is referred to as the CS+; Stimuli in the same paradigm that are not coupled with the US are referred to as CS-.

Research investigating the process of learning that a CS no longer predicts a threat predominantly consists of studies implementing an extinction paradigm. Extinction paradigms aim to reduce conditioned fear responses. During extinction, the CS is repeatedly presented again, but now without the coupling with the US (Pavlov, 1927). The extinction paradigm has been found to be successful in reducing the CR (for a review, see Vervliet et al., 2013). In fear conditioning studies, repeated exposure to the CS without the aversive US does indeed initially reduce multiple aspects of the conditioned fear response (e.g., reduced startle-response and reduced skin conductance response to the CS). Unfortunately, the extinction procedure is less effective in preventing the return of the CR over longer periods (Vervliet et al., 2013). First, if the response to the CS is tested again after some time has passed, the CR is often found to recover. This phenomenon is described as spontaneous recovery (Quirk, 2002). Second, if the subject is confronted with presentations of the US, the response to the CS tends to return; The CR is reinstated with respect to the CS (Rescorla & Heth, 1975). Third, if the response to the CS is tested in a context other than the context in which the extinction procedure took place, again, the CR tends to recover. This phenomenon is known as renewal (Bouton, 2002). In sum, these three phenomena point out the shortcoming of the extinction procedure, namely, that fear may return even after initially successful extinction. This highlights the need for more effective alternatives to induce safety learning (Vervliet et al., 2013).

Several theories aim to explain what mechanisms underlie extinction and thereby aim to explain why extinction is unsuccessful in preventing the return of the CR. One, still widely accepted theory, is that of Bouton (1993) which states that extinction is not the unlearning of the conditioned CS-US association. Instead, according to this theory, an additional inhibitory (safety) association is formed during extinction. This safety memory is thought to have an inhibitory influence on the US memory. When confronted with the CS after extinction, the initial CS-US association and the safety memory will compete for expression. Bouton (1993) argues that

if the CS-US association is easier to retrieve than the safety memory, the CR will return in response to the CS. For example, if the CS-US association is triggered by confrontation with the US, this association might be easier to retrieve than the safety memory, causing reinstatement of the CR (Bouton, 2002; Craske, Liao, Brown, & Vervliet, 2012).

Also on a neural level, studies have been looking at the mechanisms of conditioning and extinction. In general, studies looking at the neural underpinnings of conditioning and extinction agree on the central role of the amygdala in forming the CS-US association (Hitchcock, & Davis, 1986; LeDoux, Sakaguchi, & Reis, 1984) during fear conditioning (Quirk & Milad, 2012). In turn, output of the amygdala is projected to the brainstem, which is thought to play a role in evoking the CR (Le Doux, Ruggiero, & Reis, 1985). The prefrontal cortex plays a role in both fear-expression after conditioning and inhibition of this fear expression after extinction. Activation in the dorsal anterior cingulate cortex (dACC) is positively correlated with fear expression, mediated by excitatory stimulation of the amygdala. In extinction, an important role of the ventromedial prefrontal cortex (vmPFC) is proposed to be in the formation of a safety memory and in inhibiting the output of the amygdala to the brainstem during extinction recall, thereby inhibiting the expression of conditioned fear responses (Phelps, Delgado, Nearing, & LeDoux, 2004). Altogether, neurocognitive mechanisms of conditioning and extinction have been studied quite extensively – interest in studying alternatives to extinction is rising, as discussed below.

Preventing the return of the conditioned response

As extinction is not as effective in preventing the return of the CR as one would hope for, efforts have been made to find better alternatives. One such alternative is counter-conditioning. Instead of merely omitting the US when presenting the CS, in counter-conditioning the CS is now coupled with a new unconditioned stimulus. This new stimulus is of opposite valence compared to the initial US. For example, if during conditioning a tone is coupled with a shock, this CS could be coupled with the delivery of food during counter-conditioning.

Correia, McGrath, Lee, Graybiel, and Goosens (2016) studied an aversive-to-appetitive counter-conditioning paradigm in rats. After initial auditory fear conditioning, a subgroup of the animals

received reward conditioning. The findings indicate that an extinction procedure followed by a counter-conditioning procedure is more effective in reducing spontaneous recovery of the fear response compared to extinction alone. Similarly, to the findings of Correia et al. (2016), with human subjects counter-conditioning seems to be more effective over the long term as well (Kang, Vervliet, Engelhard, van Dis, & Hagenaars, 2018). After a time interval, the return of threat expectancy is reduced after counter-conditioning: In other words, less spontaneous recovery takes place (Kang et al., 2018). Additionally, several studies have successfully prevented reinstatement of the CR in an appetitive-to-aversive counter-conditioning paradigm (in rodents: Tunstall, Verendeev, & Kearns, 2012; in humans: Kerkhof et al., 2011; Kaag et al., 2016). For example, Karel et al. (2018) showed that cue-conditioned cocaine administration in rats could be successfully counter-conditioned with a footshock, by reducing reinstatement of reward-seeking behaviour. Although recently promising results have been found, results of older studies do not support a difference in effectiveness between counter-conditioning and extinction with regard to preventing the return of fear responses (Brooks, Hale, Nelson, & Bouton, 1995), pointing out the need for more studies to assess the effectiveness of counter-conditioning compared to extinction.

Taken together, both studies looking at appetitive-to-aversive and aversive-to-appetitive show that counter-conditioning is a promising alternative to the common extinction procedures. In addition, some studies have addressed the neural correlates of counter-conditioning (in rodents: Correia et al., 2016; for appetitive-to-aversive counter-conditioning in humans: Bulganin, Bach, & Wittemann, 2014). However, to our knowledge the mechanisms of aversive-to-appetitive counter-conditioning have not yet been assessed in human subjects. Taken together, therefore the aim of the current study is to test the effectiveness of counter-conditioning compared to extinction in preventing the return of fear responses. Moreover, we aim to explore the neurocognitive mechanisms underlying counter-conditioning. In the current study, less return of the (physiological) fear response can be expected after counter-conditioning compared to regular extinction as found in previous studies, supporting the effectiveness of counter-conditioning (Kerkhof et al., 2011; Kaag et al., 2016; Karel et al., 2018). Additionally, based on previous work looking at counter-conditioning and other alternative procedures for extinguishing CRs, we hypothesize one of three neurocognitive

mechanisms to underlie counter-conditioning: enhanced extinction, competition or overwriting, which will be explained below.

Hypotheses

The first proposed mechanism is enhanced extinction. Similar processes could underlie both extinction and counter-conditioning. In this case, corresponding to the theory of Bouton (1993) that views extinction as the formation of a novel inhibitory safety memory, counter-conditioning may also induce an additional safety memory regarding the CS, yet a stronger safety memory than standard extinction. Dunsmoor et al. (2019) studied an alternative to extinction, namely: novelty-facilitated extinction (NFE). This paradigm is similar to counter-conditioning, as in that not only the CS is no longer paired with an US, but rather a new stimulus is paired with the CS, namely: a novel neutral stimulus (e.g., a tone). The difference between counter-conditioning and NFE is that the stimulus coupled with the CS is of greater appetitive valence in counter-conditioning than in NFE. Indeed, after NFE compared to extinction, Dunsmoor et al. (2019) found faster fear reduction and increased activation of the vmPFC which is involved in inhibiting the amygdala. Dunsmoor et al. (2019) argue that these results are induced by enhanced attention due to the level of surprise, prompting more learning during the NFE trials than during regular extinction. In the end, this results in a stronger safety memory that might be easier to retrieve and therefore better able to compete with the original CS-US association, leading to more inhibition of the latter.

The second proposed mechanism is competition. Correia et al. (2016) found increased activity of the basolateral amygdala (BLA)-ventral striatum (NAc) circuitry in rodents after counter-conditioning, which enhanced fear reduction. These findings suggest the involvement of reward processing areas (Schultz, Tremblay, & Hollerman, 2000) during counter-conditioning. Therefore, a second mechanism that we propose is that an additional reward memory is formed during counter-conditioning – instead of a safety memory – that competes with the fear memory for expression. Due to the emotional value of this reward memory, this memory might be of similar ease to retrieve, offering better competition with the fear association than a safety memory.

The third proposed mechanism is overwriting. We hypothesize that the initial CS-US association could be overwritten by a reward memory formed during counter-conditioning, due to reconsolidation-like processes. In reconsolidation procedures, a consolidated memory is reactivated and, thereby, brought to a labile state again. This labile state opens a window of opportunity for alternations to be made to this memory. If those changes are then reconsolidated, this results in a permanently altered memory (Alberini & LeDoux, 2013). Several studies suggest reconsolidation as an alternative to extinction for reducing CRs (Schiller, Monfils, Raio, Johnson, LeDoux, & Phelps, 2010; Kroes, Dunsmoor, Lin, Evans, & Phelps, 2017). In the current counter-conditioning paradigm, the initial fear association might still be in a labile state when the counter-conditioning takes place. Therefore, the memory might still be sensitive to alterations. This leads to our third hypothesis that the second reward association is not only formed coexisting with the original fear association, but rather interacts with this initial fear association during consolidation, overwriting this association. This results in only the newly acquired reward association getting consolidated.

The current study

In the current study, we aim to assess the neurocognitive mechanisms underlying counter-conditioning. Specifically, as mentioned above, we hypothesize that enhanced extinction, competition, or overwriting underlies counter-conditioning. To test our hypotheses, we ran a counter-conditioning paradigm in humans. In brief, in a between-subject functional magnetic resonance imaging (fMRI) study, all participants first acquired categorical fear-conditioned memories. Subsequently, half of the participants underwent a regular extinction and the other half a reward counter-conditioning paradigm. A day later, we tested for group differences in the return of threat CRs and episodic memory for items from the acquisition and extinction or counter-conditioning experience. An outline of the design can be found in Figure 1.

In our experiment, we used a categorical fear conditioning paradigm for a number of reasons I will explain. CRs tend to generalize to stimuli related to the CS (Dunsmoor & Murphy 2015). Humans generalize to stimuli that are both physically similar and conceptually related to the CS. This generalization creates the opportunity to use a categorical fear conditioning paradigm, shown to be feasible in other studies (Dunsmoor, Kragel, Martin, & LaBar, 2014; Kroes et al., 2017). The advantage of this paradigm is that unique pictures can be used over all trials and it allows testing episodic memory
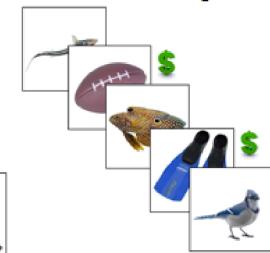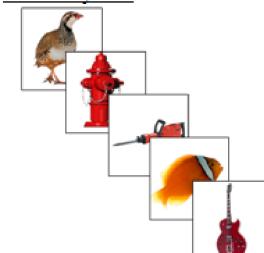
Figure 1. Experimental design. On day 1, after the category localizer, participants did a categorical fear conditioning task. The CS+ category stimuli (animals or objects) were followed by an electrical shock in 50% of the trials. This phase was segmented by a short break halfway through the trials. Next, half of the participants did an extinction procedure, the other half did a counter-conditioning task. In all tasks, trial-unique pictures were used, so on the second day the episodic memory for previously used items could be tested. In this task, participants were presented all pictures used in the conditioning and counter-conditioning or extinction task, mixed with the same number of new pictures. Participants had to indicate if they thought the pictures were new or old (previously seen in another task). In the mixed localizer, neutral abstract stimuli were used to asses brain areas involved in reward and fear learning, as well as general arousal related areas.

for items from the different parts of the experiment, effectively obtaining a retrospective index of memory for different points in time. In the current study, one category consists of pictures of animals and the other of pictures of objects, as animate (in this case the animals) and inanimate (in this case the objects) stimuli are represented by intrinsically different neural organizations (Dunsmoor et al., 2014). Animate stimuli are related to activity in lateral fusiform gyrus (FFG), posterior superior temporal sulcus (pSTS) and inferior occipital regions. Inanimate stimuli are related to activity in medial FFG, posterior middle temporal gyrus and middle occipital cortex (Martin, 2007). Taken together, using a categorical fear conditioning paradigm with the above mentioned categories gives the opportunity to study memory related mechanisms of counter-conditioning and asses changes in representations of stimuli, compared to well established representations of those stimuli under regular circumstances. Differences between extinction and counter-conditioning in those memory related mechanisms and changes in representations will help shed light on the mechanisms underlying counter-conditioning.

As a counter-conditioning procedure a modified version of a monetary incentive delay task was used. In animal studies looking at counter-conditioning, the animals are usually first food deprived in order to make food rewards a meaningful incentive.

However, for ethical and pragmatic constraints, food deprivation was not an option in the current study with human subjects. Alternatively, a monetary incentive delay was implemented (MID; Knutson, Westdorp, Kaiser, & Hommer, 2000). By getting participants actively involved in obtaining a reward during the counter-conditioning paradigm, the MID is found to provide enough incentive to instate CRs, as shown following pilot data collection within the current study design. Thus we use an instrumental conditioning procedure as a counter-conditioning procedure against a Pavlovian fear conditioning association. One may argue that mixing Pavlovian and instrumental fear memories is not a clean procedure. Yet, we argue that nearly every appetitive conditioning procedure has an instrumental component (e.g., an animal needs to approach food or lick to swallow sugar water) and this is simply unavoidable.

To induce event segmentation and allow distinct episodic memory traces to be consolidated for fear and safety, the phases (early and late conditioning and counter-conditioning or extinction) of the MID task are separated with a short break. This is in line with the study of Dunsmoor et al. (2018).

In general, we hypothesize that the counter-conditioning paradigm is more effective than the extinction paradigm. More specifically, we expect to find less return of conditioned fear responses in the

second session for the counter-conditioning group compared to the extinction group, as measured with both subjective and psychophysiological measures of the fear response. Moreover, we hypothesize that the three proposed mechanisms, that potentially underlie counter-conditioning, would be reflected in distinguishable results from the episodic memory task and from the neural outcome measure. First, if we would find increased vmPFC activity after counter-conditioning in the above mentioned design, suggesting increased reduction of the fear memory, our findings would support the hypothesis that counter-conditioning induces enhanced inhibition of the fear memory compared to extinction. If this is the case, furthermore, similar cognitive mechanisms are expected to underlie counter-conditioning and extinction. Episodic memory for stimuli coupled to the US (CS+ items) tends to be better than for CS-items, as observed in previous conditioning studies (Patil, Murty, Dunsmoor, Phelps, & Davachi, 2017; Dunsmoor, Martin, & LaBar, 2012; Dunsmoor & Murphy, 2015). In the case of enhanced extinction, we expect to find similar results as in those previous studies, in both the extinction and counter-conditioning group.

A second possibility is that a reward memory is formed during counter-conditioning, which competes with the original fear memory. As this reward memory has a meaningful emotional valence on its own, it may induce enhanced retroactive memory consolidation (Patil et al., 2017; Dunsmoor et al., 2012; Dunsmoor, Murty, Davachi, & Phelps, 2015), contributing to stronger competition with the initial CS-US association. If this is the case, we expect to find both involvement of fear and reward learning (neural) mechanisms. As found by Correia et al. (2016), involvement of the ventral striatum is expected after counter-conditioning, next to neural activity related to the conditioned fear memory. This would represent the presence of both a fear and the reward memory which compete after counter-conditioning. Furthermore, in the case of competition, both stimuli coupled with the initial US and stimuli coupled with reward during counter-conditioning are expected to be better remembered than CS- items, due to the fact that CS+ items are associated with the emotional valence of the initial US and the reward used in counter-conditioning.

In contrast to the two formerly proposed mechanisms, for our third hypothesis of overwriting, we expect no involvement of brain areas expressing or inhibiting the fear response. However, involvement of areas related to reward learning, like the ventral striatum, is expected.

Moreover, compared to the other two mechanisms, no memory benefits are expected for items related to the US of conditioning, due to the fact that this emotional association is overwritten by the new, opposite emotional association during counter-conditioning. More specifically, it is expected that CS+ items from the counter-conditioning phase will be better remembered than CS- items from this phase, however, no such difference is expected for items from the conditioning phase, in contrast to the two previously mentioned mechanisms.

In summary, the current study aims to asses counter-conditioning as a promising alternative to general extinction for reducing undesired CRs. So far, it has been shown in animal and human studies that counter-conditioning can be more effective compared to standard extinction in preventing the return of the CR. This study aims to extend the literature about counter-conditioning by unravelling the neurocognitive mechanisms underlying this increased effectiveness. This will be done by comparing physiological fear responses, neural activation and episodic memory in a categorical fear conditioning paradigm between groups that either undergo standard extinction or counter-conditioning. We hypothesize that counter-conditioning either results in increased inhibition of the initial fear memory, competition between the fear and reward memory, or overwriting of the fear memory by the reward memory. Shedding light on the effectiveness and mechanisms of counter-conditioning can, in the end, help to improve clinical treatments of undesired conditioned associations found in the clinical population.

Methods

All study proceedings, methods and analyses strategies were preregistered on the Open Science Framework (www.osf.io/).

Participants

Up to this point, 20 Dutch speaking participants took part in this study at the time of analyses (13 female, 7 male). All were healthy adults aged between 18 and 35 years, with normal or corrected-to-normal vision. All participants met the inclusion criteria for MRI (magnetic resonance imaging). People with a history in psychiatric and/or neurological disorders, and people with (by medication induced) altered endocrine and/or hormonal functioning (except by hormonal contraception) were excluded from

participation. Participants were recruited via an online recruitment system ('Radboud Research Participation System') or by personal approach. Participants received 10 € per hour in return for participation and could earn a certain extra bonus during part of the tasks. Participants were randomly assigned to either the counter-conditioning (7 females, 3 males, age: 22.0 ± 2.06) or extinction (6 females, 4 males, age: 22.3 ± 2.45) group. One participant from the extinction group did not come to the second session of the experiment. Participants with missing or incomplete data for a critical test were excluded from analyses for that test but will still be included in tests on other data, in accordance with our preregistration. The study was approved by the local ethical review board (CMO region Arnhem-Nijmegen) and in accordance with the Declaration of Helsinki.

Stimuli

A set of 220 animal and 220 object pictures was used for this study. All pictures were luminance equalized, including the grey background. The order of the pictures was pseudorandomized for each even and odd participant couple (participant 1 and 2, 3 and 4, 5 and 6, etc.), so that the participants of this couple had the same order of pictures but for each new couple, a new order was created. The category serving as the CS+ was as well randomly determined per even and odd couple. The first 32 pictures, from both categories, of this pseudorandom order were used for the categorical localizer task, the next 40 were used in the categorical fear conditioning task. In a similar fashion, 40 pictures were assigned to the counter-conditioning or extinction task and 20 pictures to the fear recovery test. In the episodic memory test, the pictures used in the acquisition and extinction phase were used together with 80 new pictures from each category. Phase scrambled versions of the pictures were created for the functional stimulus category localizer task (Reinders, Den Boer, & Büchel, 2005), for a control block in this task. For the mixed valence localizer task, three coloured squares were used. Those squares were also luminance equalized including their background.

Materials

All tasks described here were performed in the MRI scanner. Three of the tasks were performed in the first session, taking place on the first testing day. The other three tasks were completed in the second

session, which took place on the day after the first session (as illustrated in Fig. 1).

Functional stimulus category localizer

This task is based on the categorical localizer task used by Voogd, Fernandez, and Hermans (2016a). Note that we look at animals vs. objects instead of animals vs. fruits/vegetables, similar to Dunsmoor et al. (2014). fMRI data from this task is used to assess which areas are involved in processing pictures from the two categories in general. This task consists of three different block types. One block with pictures of animals, one with pictures of objects and one with phase-scrambled versions of pictures of animals and objects. Each block is repeated 12 times in a random order, with the random order of the blocks of the first half of the task being mirrored for the second half. Picture order within each block is randomized. Additionally, 12 rest blocks are included, consisting of a blank screen with a fixation cross. The blocks consist of 32 pictures being presented for 625 ms. The task for the participant is to watch the pictures and pay attention to a white circle which is presented superimposed on one of the pictures in half of the blocks. Once they see this circle, they have to react with a button press. With regard to the episodic memory task, pictures will only be used once in the tasks prior to the episodic memory test. Therefore, pictures used in this part of the experiment are not used again in other parts of the experiment.

Categorical fear conditioning task

This task is a combination of the categorical fear conditioning paradigm as described by Dunsmoor et al. (2012) and the MID as described by Knutson et al. (2000). The choice for the addition of aspects of the MID is based on reasons mentioned above with regard to the counter-conditioning task. For the purpose of keeping most tasks as similar as possible, these MID aspects are also incorporated into the acquisition phase, in the categorical fear conditioning task. In the practice version of this task, a green square is presented. After a variable interval, a cue (a white circle) appears superimposed on this square at a random location. The task of the participant is to press a button as soon as this cue appears. This is repeated ten times with variable inter-trial intervals (ITI).

In the actual version of this task, participants are first presented with a picture from one of the two categories (animals or objects). Figure 2 contains a representation of the setup of this task. After a variable interval (2.5 - 4 s) the cue appears on top of this picture, indicating that the participant has to
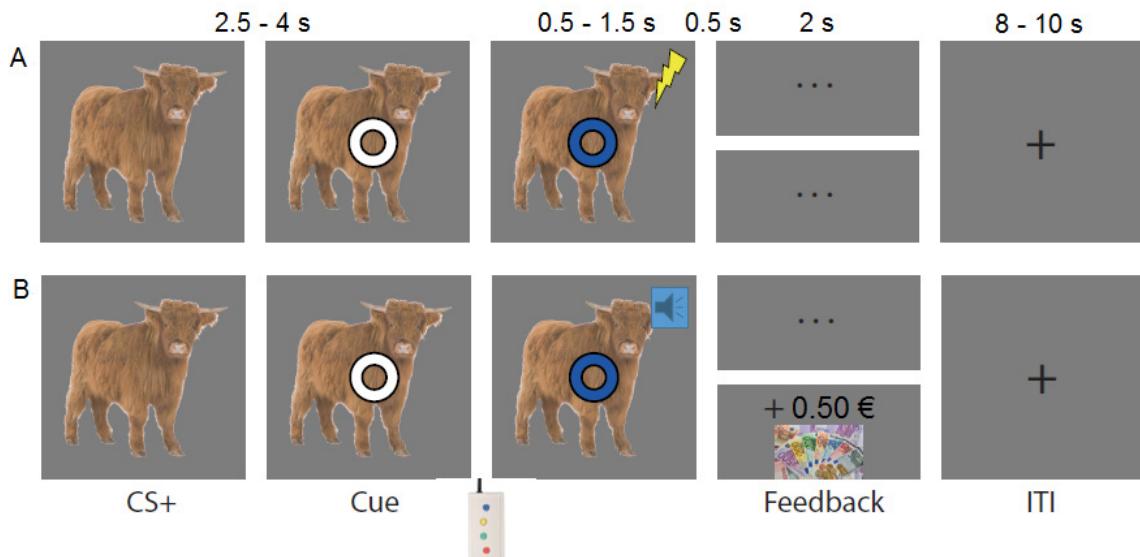
Figure 2. Categorical conditioning and counter-conditioning task including MID aspects. A. Representation of the MID aspect in the task used during the acquisition phase. In this phase 50% of the items from the CS+ category were coupled with a shock. B. Representation of the counter-conditioning task. The extinction group did the same task, however, they never received the monetary reward.

react by pressing the button as quickly as possible within the response window of 1 s. After pressing the button, the cue changes colour as feedback for the participant that their key press was effective. The cue remains on screen for 0.5 - 1.5 more seconds. In 50% of the trials from the CS+ category, a shock was administered during this time window. Subsequently, the neutral feedback of three dots (…) appears for 2 s after the picture and cue disappeared. This is followed by a variable inter-trial-interval (8 − 10 s) during which a fixation cross is presented. This is repeated 80 times with 40 trial-unique animal pictures and 40 trial-unique pictures of objects in a random order. In the middle of the conditioning phase, there is a short break of 10 s (as in Dunsmoor et al., 2018).

Counter-conditioning/extinction task

This task is used to compare two different strategies to extinguish a conditioned fear response. This task is largely the same as the categorical fear conditioning task: participants first see a picture and have to press a button as soon as the cue appears superimposed on the picture. The number of stimuli as well as the timings of the trials are the same as in the categorical fear conditioning task. In the extinction group, the pictures are no longer coupled with a shock. In the counter-conditioning group, none of the pictures are coupled with a shock either. However, now about 70% of the

pictures from the CS+ category are coupled with a 0.50 € reward and the message that they receive this reward because their response was fast enough. This percentage is set based on reinforcement rate pilot experiments that showed to induce effective conditioned responses. In order to create a smooth transition from the conditioning phase to the counter-conditioning phase, the first two CS+ trials are not rewarded. From the third trial onwards, the participants are able to receive rewards. The initial response time threshold for obtaining a reward is set to 10% faster than participants' responses during practice. During the remaining trials, this response threshold is altered dynamically in order to approach a 70% reinforcement rate. Note that this is a higher reinforcement rate than during fear acquisition (50%), but pilot studies indicated a higher reinforcement rate to be necessary to induce appetitive conditioned responses.

Fear recovery test

First, participants are presented 15 unique new stimuli from each of the two categories, again with the task to press the button as soon as possible after the cue appears on top of the picture. During this part of the task, the participants receive neither shocks nor rewards, therefore spontaneous recovery of the fear response can be assessed. Critically, pilot studies indicated that reward anticipation did not evoke anticipatory physiological responses.

Timings of the stimuli in this task are the same as in the previously described tasks. Next, we test for reinstatement of the fear response: participants receive three unannounced electrical shocks, followed by an additional five unique new stimuli per category, which are not followed by a reward or shock, to test for fear recovery in relation to those stimuli.

Episodic memory test

This task is used to get an indication of the episodic memory performance for CS+ and CS-items, which are used in different phases of the experiment. Participants are presented with all the pictures shown during the conditioning and counter-conditioning/extinction phase (80 pictures per category in total), one at a time. These are mixed in a random order with an equal amount of new pictures from the two categories. In total, this task consists of 320 items. The task of the participants is to indicate on a 6-point Likert scale how sure they are that a certain picture is an old (1) or new picture (6). They had to do this within 3 s after stimulus onset.

Mixed valence localizer

This task is used to localize brain activation evoked by arousal ([CS+shock & CS+reward] vs. [CS-]) and valence ([CS+shock] vs. [CS+reward]). This part consists of three novel neutral stimuli presented 20 times each in a random order. The stimuli used in this task are abstract figures (squares) in three different colours compared to the pictures of real life animals and objects as in the previously described tasks. One of the stimuli was followed by a shock 50% of the time. The second stimulus was followed by reward in 70% of the time, while the third stimulus was never followed by a reward or shock. The shock and reward rate as well as the timings of the events in this task are the same as in the rest of the experiment.

Valance, arousal and contingency awareness measures

Before conditioning, before extinction, after extinction, before the fear recovery test and after the fear recovery test, participants are asked to rate how negative or positive they experience pictures from the animal and object category at that moment in general (valence) and how arousing they experience the pictures to be in general (calm vs. excited) on a 9-point Likert scale. In addition, participants are asked to estimate for every different phase (conditioning, extinction, spontaneous recovery and reinstatement) what percentage of CS+ and CS- category images were paired with shocks. For participants in the counter-conditioning group, reward estimation and contingency awareness is also assessed for the different tasks.

Procedure

After receiving general information about the procedure and signing the informed consent form, participants were taken into the scanner. First, the shock intensity level was calibrated with an ascending staircase procedure starting with a low voltage (near a perceptible threshold) to reach a level deemed "maximally uncomfortable without being painful" by the participant. Participants were instructed that the intensity of the shocks would stay the same over the course of the experiment. Second, the peripheral psychophysiological and MRI measurements were prepared before the participant could start with the tasks.

On the first day, participants started with the functional stimulus category localizer task. The next task was the categorical fear conditioning task, preceded by practice trials of this task. In this acquisition phase, participants were fear conditioned by pairing trial-unique images of one category (animals or objects) with a mild electrical shock half of the time. Participants were assigned to either one of the fear conditioned categories by counterbalancing. Before starting the tasks, participants were told a contingency exists between the category of the stimuli and the outcome of the trials. Participants were instructed that their instrumental responses had no influence on the shocks they received. Moreover, they were instructed that for the rewards they could receive throughout the task, a relation exists between their reaction time in response to the cue and the possibility of receiving a reward. For half of the participants, the third task of the first session was an extinction phase and the other half received counter-conditioning.

At the start of the second session, participants were reminded that the shock intensity during this session was the same as set during the first session. Furthermore, they were reminded of the non-existing relation between reaction times and the possibility of receiving a shock, as well as of the existing relation between their responses and the possibility of receiving a reward. The first task of the second session was a test for the return of the conditioned fear response. Next, participants did an episodic memory test. The last task of the experiment was the mixed valence localizer paradigm.

Subjective valence and arousal ratings were sampled at the following time points: before and after the acquisition phase, after the extinction phase and before and after the fear recovery test. At the end of both sessions, participants estimated the contingency between the stimulus categories and the possibility of receiving a shock or reward.

Physiological measurements

Skin conductance responses

Electrodermal activity was measured using two Ag/AgCL electrodes attached to the middle- and forefinger of the left hand with a BrainAmp MR system and BrainVision Recorder software (Brain Products GmbH, Munich Germany). Before analysis, data were first cleaned from radio frequency (RF) artefacts and high frequencies (de Voogd, Fernández, & Hermans, 2016a, 2016b). Next, SCR responses were automatically scored and manually checked using Autonomate (Green, Kragel, Fecteau, & LaBar, 2014) implemented in Matlab 7.14 (MathWorks). This procedure is in line with previously published procedures (de Voogd et al., 2019). SCR amplitudes were calculated as the maximum through-to-peak deflection in the latency window of $0 - 12$ s after trial onset for each presentation of categorical stimuli. After pilot experiments with the current setup, it was decided to use this long latency window (instead of $0 - 4$ s which is used in most studies) because SCRs were found to arise relatively late. This might be due to the fact that in the current MID version of a classical conditioning task, the shock could only come relatively late after the trial onset (after the button press). Only trials in which no shock or reward happened were taken into account. SCRs were calculated for the following tasks: the categorical fear conditioning task, extinction/counter-conditioning, the fear recovery test and the mixed valence localizer.

Pupil dilation

The extent of pupil dilation in reaction to viewing the CS images was measured with a MR-compatible eye-tracker from SensoMotoric Instruments in the same tasks as the SCR measurements. This device was attached to the scanner bed. The pupil diameter was sampled at a rate of 50 Hz. Data were analysed using in-house software (Hermans, Henckens, Roelofs, & Fernandez, 2013) implemented in Matlab 7.14 (MathWorks), which was based on methods described previously by others (Siegle, Steinhauer, Stenger, Konecky, & Carter, 2003). Eyeblink

artefacts were identified and linearly interpolated 100 ms before and 100 ms after each identified blink. After interpolating missing values, time series were band-pass filtered at 0.05 to 5 Hz and z-scored (by subtracting the mean and dividing by the standard deviation) within each participant and run to account for between-subject variance in overall pupil size (including variance related to corrective lenses). Event-related pupil diameter responses were calculated by averaging pupil diameter during the 1 - 4 s period after stimulus onset and dividing by the averaged 1 s pre-stimulus pupil diameter (-1 - 0 s) (de Voogd et al., 2016a). The average of pupil dilation response (PDR) was computed per condition, phase and participant.

Respiration and heart rate

Respiration was measured using a respiration belt placed around the participant's diaphragm. Heart rate was measured using a pulse-oximeter. Raw pulse and respiratory data were processed offline using in-house software for visual artefact correction and peak detection, and were used for retrospective image-based correction (RETROICORplus) of physiological noise artefacts in BOLD (blood-oxygenation level dependent)-fMRI data (Glover, Li, & Ress, 2000). In line with procedures previously described by de Voogd et al. (2016a), processed pulse and respiratory data were used to specify fifth-order Fourier models of cardiac and respiratory phase-related modulation of the BOLD signal, yielding ten nuisance regressors per modality. An additional six regressors of no interest were calculated for heart rate frequency, heart rate variability, abdominal circumference, respiratory frequency, respiratory amplitude and respiration volume per unit time.

MRI data acquisition

During all tasks whole-brain functional imaging was conducted on a 3.0 Tesla PrismaFit MRI scanner equipped with 32-channel transmit-receiver head coil. The manufacturer's automatic 3D-shimming procedure was performed at the beginning of each experiment. Participants were placed in a light head restraint within the scanner to limit head movements during acquisition.

Functional images were acquired with multi-band multi-echo gradient echo-planar imaging (EPI) sensitive to the BOLD response using the following parameters: 51 oblique transverse slices, slice thickness = 2.5 mm, repetition time (TR) = 1.5 s, flip angle $\alpha = 75°$, echo times (TE) = 13.4, 34.8

and 56.2 ms, field-of-view (FOV) = 210 x 210 mm², matrix size 84 x 84 x 64, fat suppression.

A magnetic (B_0) field map was collected at the start of each session and used to unwarp the echo-planar images (Hutton, Bork, Josephs, Deichmann, Ashburner, & Turner, 2002). A structural image (1 mm isotropic) was acquired using a T1-weighted 3D magnetization prepared rapid gradient echo (MPRAGE) sequence with the following parameters: TR = 2300 ms, TE = 3.03 ms, flip angle α = 8°, 192 contiguous 1 mm slices, FOV = 256 x 256 mm², to be used for normalization procedures.

MRI data preprocessing

Data were preprocessed with *fMRIPrep* 1.2.6-1 (Esteban, Markiewicz, et al., 2018; RRID:SCR_016216). First, a reference volume and its skull stripped version were generated using a custom methodology of *fMRIPrep*. A deformation field to correct for susceptibility distortions was estimated based on a field map that was co-registered to the BOLD reference, using a custom workflow of *fMRIPrep* derived from D. Greve's *epidewarp.fsl* script and further improvements of HCP Pipelines (Glasser et al., 2013). Based on the estimated susceptibility distortion, an unwarped BOLD reference was calculated for a more accurate co-registration with the anatomical reference. Head motion parameters with respect to the BOLD reference (transformation matrices and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using *mcflirt* (FSL 5.0.9, Jenkinson, Bannister, Brady, & Smith, 2002). The BOLD time series were resampled onto their original, native space by applying a single composite transform to correct for head motion and susceptibility distortions. These resampled BOLD time series will be referred to as preprocessed BOLD in original space or just preprocessed BOLD. A T2* map was estimated from the preprocessed BOLD by fitting to a monoexponential signal decay model with loglinear regression. For each voxel, the maximal number of echoes with reliable signal in that voxel were used to fit the model. The calculated T2* map was then used to optimally combine preprocessed BOLD across echoes following the method described in Posse et al. (1999). The optimally combined time series was carried forward as the preprocessed BOLD and the T2* map was also retained as the BOLD reference. The BOLD reference was then co-registered to the T1 weighted (T1w) reference using *bbregister* (FreeSurfer) which implements boundary based registration (Greve &

Fischl, 2009). Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. The BOLD time series were resampled to MNI152NLin2009cAsym standard space, generating a preprocessed BOLD run in MNI152NLin2009cAsym space. For analyses in native space, weighted, realigned and unwarped functional images co-registered to each participants' structural scan are smoothed using a 6 mm FWHM (full width at half maximum) Gaussian.

Statistics

In general, dependent variables were submitted to repeated measures ANOVAs, using SPSS (IBM SPSS Statistics Inc.) to perform the analysis. Effects with a *p*-value lower than .05 were considered significant and were followed up with independent and paired samples t-tests. Partial eta-squares were reported as effect sizes of significant effects. Additionally, means and standard deviations were reported where relevant.

Neuroimaging data were analysed using SPM12 (Wellcome Trust Centre, www.fil.ion.ucl.ac.uk) implemented in MATLAB. Statistical thresholds of second-level whole-brain analyses were set at familywise error (FWE) corrected cluster $p < .05$. For the first level analysis of the categorical localizer, regressors were made for each block type (animals, objects and scrambled). Onsets and durations of the blocks were saved online during the execution of the task. Additionally, a nuisance regressor was used to account for events of no interest. For the first level analysis of the conditioning, counter-conditioning/extinction task and fear recovery test, regressors were made for the CS+ and CS- items. In order to be able to assess changes in brain activity within a task, the regressors for the conditioning and counter-conditioning/extinction task were split into regressors for CS+ items from the first and the second half of the task; The same was done for the CS- items. Again, a nuisance regressor was added to account for events of no interest. Contrast estimates from the first level analysis were taken into group level analysis.

Results

Valence and arousal ratings

Results of the subjective valence and arousal ratings of the two categories show that participants acquired conditioned responses during the conditioning task. Moreover, in both groups these
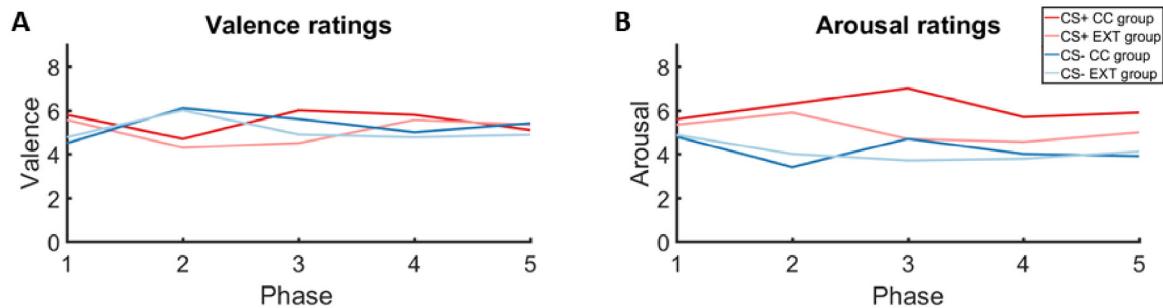
Figure 3. Mean subjective ratings of A. valence and B. arousal over the course of the experiment. Valence and arousal ratings were sampled before (phase 1) and after conditioning (phase 2), after counter-conditioning/extinction (phase 3), and before (phase 4) and after (phase 5) the fear recovery tests on the second day. Ratings were given on a 9-point Likert scale with 1 = very negative or absolutely not arousing, and 9 = very positive or extremely arousing.

conditioned associations were successfully reduced by the extinction and counter-conditioning task.

To test the effect of the conditioning and counter-conditioning (CC)/extinction (EXT) tasks on the evaluation of the two categories, valence and arousal rating were used in a group (EXT, CC) x time point (before conditioning, before extinction, after extinction, before fear recovery test, after fear recovery test) x CS-type (CS+, CS-) 2 x 4 x 2 repeated measures ANOVA. For the valence ratings, this revealed a significant time point*CS-type interaction ($F(4, 64) = 3.40$, $p = .014$, $\eta_p^2 = .18$). Right after the conditioning phase, there was a significant difference between the ratings for the CS+ ($M = 4.50$, $SD = 1.47$) and CS- items ($M = 6.10$, $SD = 2.00$). This was not the case for the other time points, as illustrated in Figure 3A and indicated by post-hoc paired-samples t-tests ($t(19) = -2.75$, $p = .013$). Thus, both groups show acquisition and later extinction of conditioned differential responses. Neither group shows return of these responses on the second day.

For the arousal ratings, both a main effect of CS-type ($F(1, 16) = 6.35$, $p = .023$, $\eta_p^2 = .28$) and a significant time point*CS-type interaction ($F(4, 64) = 5.46$, $p = .001$, $\eta_p^2 = .25$) were found. Overall, the arousal ratings were higher for the CS+ category than for the CS- category. However, when looking at the different time points in a post-hoc paired-samples t-test, this difference was only significant right after the conditioning phase (CS+ ratings: $M = 6.05$, $SD = 1.36$; CS- ratings: $M = 3.70$, $SD = 1.78$; $t(19) = 4.92$, $p < .001$, Fig. 3B). Neither measurement was influenced by the experimental group the participants were in. No main effects or interactions were found for the between-subjects factor group. If the extinction and counter-conditioning tasks would differ in effectiveness in preventing the return of the conditioned responses, differential CS+/CS- ratings would also be expected on the second day in

the extinction group. Which would suggest return of the fear association in the extinction group, while this return will then ideally be prevented in the counter-conditioning group.

Reaction time data

Reaction time data from the categorical fear conditioning task were subjected to a group (EXT, CC) x phase (early, late) x CS-type (CS+, CS-) 2 x 2 x 2 repeated measures ANONA. This analysis revealed a main effect of CS-type ($F(1, 18) = 4.65$, $p = .045$, $\eta_p^2 = .21$). Participants were faster in responding to the target in trials where they could receive a shock (respectively, $M = .40$, $SD = .05$; $M = .42$, $SD = .06$; $t(19) = -2.20$, $p = .040$). There was also a main effect of group ($F(1, 18) = 12.74$, $p = .002$, $\eta_p^2 = .41$): participants in the counter-conditioning group were in general slower in responding than the ones in the extinction group (respectively, $M = .44$, $SD = .05$; $M = .38$, $SD = .03$; $t(18) = 3.57$, $p = .002$). No interactions were found with the between-subjects factor group; There seems to be a baseline difference in response times between the groups, although, this difference was constant over time and between the CS-types.

To confirm the effectiveness of the MID aspect of the counter-conditioning task to learn the association between the CS+ items category and the possibility of receiving a reward after a fast enough button press, reaction time data of the counter-conditioning and extinction tasks were subjected to a group (EXT, CC) x phase (early, late) x CS-type (CS+, CS-) 2 x 2 x 2 repeated measures ANOVA. There was indeed a main effect of CS-type ($F(1, 18) = 20.75$, $p < .001$, $\eta_p^2 = .54$). In the counter-conditioning/extinction task, participants responded faster to targets in CS+ trials ($M = .38$, $SD = .01$)
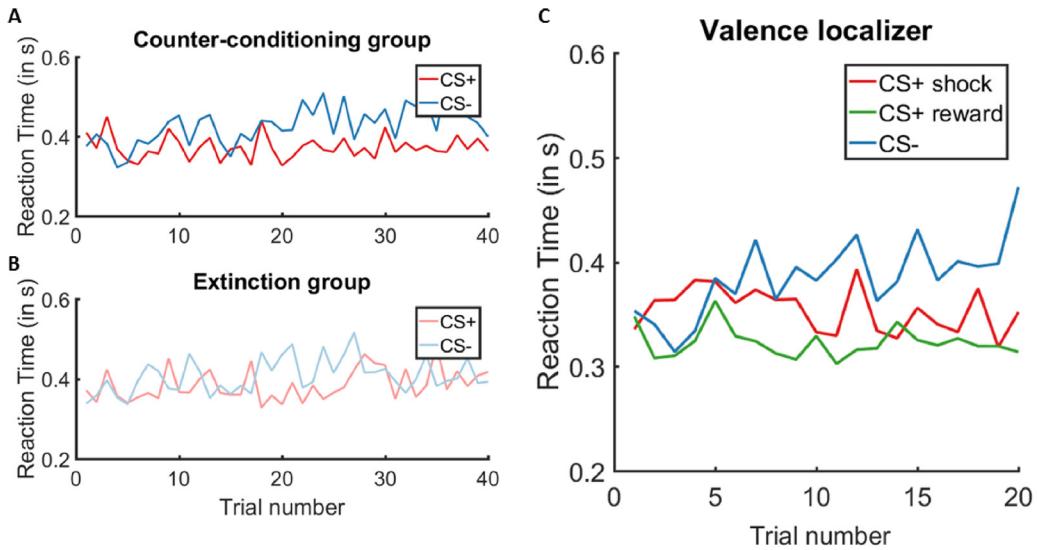
Figure 4. Reaction time data showing mean reaction times of participants in the two groups, illustrated per trial and per trial type (CS+ or CS-). A. & B. Reaction times from the counter-conditioning/extinction task illustrate participants keep responding fast if they can receive a reward in that trial (phase*CS-type*group ($F(1, 18) = 12.11$, $p = .003$, ηp2 = .402). C. Reaction times per trial and CS-type from the mixed valence localizer task, illustrating differential reaction times for the three different trial types (reward trials: M = .32, SD = .01; shock trials: M = .35, SD = .01; CS- trials: M = .39, SD = .01).

than they did in CS- trials ($M = .42$, $SD = .01$). Additionally, over the course of the task participants responded slower (first half: $M = .39$, $SD = .01$; last half: $M = .41$, $SD = .01$), shown by a main effect of phase ($F(1, 18) = 12.11$, $p = .003$, $\eta_p^2 = .40$). Most importantly, there were both a significant phase*CS-type*group ($F(1, 18) = 6.30$, $p = .022$, $\eta_p^2 = .26$) and a significant CS-type*group ($F(1, 18) = 5.00$, $p = .038$, $\eta_p^2 = .22$) interaction effect. As illustrated by Figure 4A (showing the data from the counter-conditioning group) and 4B (showing the data from the extinction group), participants in the counter-conditioning group kept responding relatively fast to the CS+ items, while showing a similar decrease in response times for the CS- items as the participants in the extinction group. Thus, participants in the counter-conditioning group clearly learned to differentiate between the CS+ and CS- items over time, as indicated by the reaction times.

To get an indication of how participants perceive the categories at the start of the second day compared to the end of the first day, reaction times from the spontaneous recovery task were analysed. Responses on the last four trials from the counter-conditioning/extinction task were compared with the first four trials of the spontaneous recovery task with a group (EXT, CC) x task (Ext/CC, spontaneous recovery) x CS-type (CS+, CS-) 2 x 2 x 2 repeated measures ANOVA. This analysis only revealed a main effect of phase ($F(1, 17) = 9.65$,

$p = .006$, $\eta_p^2 = .36$). Participants were faster at the beginning of the second session than at the end of the first (respectively, $M = .41$, $SD = .05$; $M = .36$, $SD = .05$; $t(18) = 3.19$, $p = .005$). There were no differences between the groups. An additional group (EXT, CC) x CS-type (CS+, CS-) 2 x 2 repeated measures ANOVA neither showed a significant influence of CS-type nor of group on the data. This might suggest that in neither group the fear or reward association with the CS+ category was influencing the response behaviour of the participants on the second day. If this would have been the case, in line with the first session, faster responses would have been expected for the CS+ than for the CS- items.

To confirm the effectiveness of the reward learning aspect in the mixed valence localizer task, reaction time data were used for a group (ECT, CC) x phase (first and second half of the trials) x CS-type (CS+ shock, CS+ reward, CS-) 2 x 2 x 3 repeated measures ANOVA. There was a significant main effect of CS-type ($F(2, 16) = 22.83$, $p < .001$, $\eta_p^2 = .74$). Participants were indeed faster in responding to the targets in CS+ reward trials ($M = .32$, $SD = .04$) than they were in CS+ shock ($M = .36$, $SD = .06$; $t(18) = 3.01$, $p = .008$) and CS- trials ($M = .39$, $SD = .05$; $t(18) = -6.96$, $p =< .001$). Additionally, participants also responded faster in CS+ shock trials than in CS- trials ($t(18) = -3.37$, $p = .003$).

In sum, participants respond faster in trials where a shock or a reward could be received. If neither a

reward nor a shock could be received, responses for CS+ and CS- trials are similar, as indicated in the extinction task. On the second day, responses in CS+ and CS- trials are comparable. The fear and reward associations formed on the first day do not seem to influence the response behaviour on the second day. This was the same for both groups, which suggests that similar mechanisms influence the responses on the second day for both groups.

Episodic memory test

Memory performances were analysed to assess whether they varied between tasks, groups and CS-types. In order to do so, per participant four memory scores were computed by subtracting the false alarm rates (pFA; false alarm rate was calculated by dividing the total number of 'very sure old', 'sure old', 'probably old' responses by the total number of responses to new CS+ or CS− items) from the hit rates (pHit; hit rate was calculated by dividing the total number of 'very sure old', 'sure old', 'probably old' responses divided by the total number of responses to old CS+ or CS− items). One memory score was computed for CS+ items previously seen in the conditioning phase and one for CS- items previously seen in the conditioning phase; The same was done for CS+ and CS- items previously used in the extinction/counter-conditioning task. Those memory scores were submitted to a group (EXT, CC) x CS-type (CS+, CS-) x task (conditioning task, extinction/counter-conditioning task) 2 x 2 x 2 repeated measures ANOVA. There was a significant main effect of task ($F(1, 17) = 10.14$, p = .005, η_p^2 = .37): items previously seen in the conditioning

phase were better remembered than items from the counter-conditioning/extinction task (respectively, $M = .38$, $SD = .03$; $M = .32$, $SD = .03$). Additionally, there was a main effect of group ($F(1, 17) = 6.00$, $p = .025$, $\eta_p^2 = .26$). Participants in the counter-conditioning group ($M = .42$, $SD = .12$) had overall a higher memory score than participants in the extinction group ($M = .29$, $SD = .10$; $t(17) = 2.45$, $p = .025$). As illustrated in Figure 5, there seems to be an enhanced memory performance for the CS+ items in the counter-conditioning group. However, these differences do not reach significance, as indicated by exploratory paired-samples t-tests on the data from this group (items from conditioning: CS+: $M = .49$, $SD = .18$; CS-: $M = .38$, $SD = .12$; $t(9) = 1.67$, $p = .129$; items from counter-conditioning: CS+: $M = .44$, $SD = .17$; CS-: $M = .36$, $SD = .10$; $t(9) = 1.93$, $p = .085$). In sum, data from the memory test tend to support the competition hypothesis, since CS+ items from both phases of the first session tend to be better remembered than CS- items in the counter-conditioning group.

In order to be able to analyse criterion scores of the memory test, hit and false alarm rates were z-transformed. Next, four criterion scores ($-0.5*(zHit + zFA)$) were computed for each participant and analysed with a group (EXT, CC) x CS-type (CS+, CS-) x task (conditioning task, extinction/counter-conditioning task) 2 x 2 x 2 repeated measures ANOVA. This analysis also yielded a main effect of task ($F(1, 17) = 8.65$, $p = .009$, $\eta_p^2 = .34$). A higher criterion was related to items from the counter-conditioning/extinction task ($M = .10$, $SD = .06$) compared to the conditioning
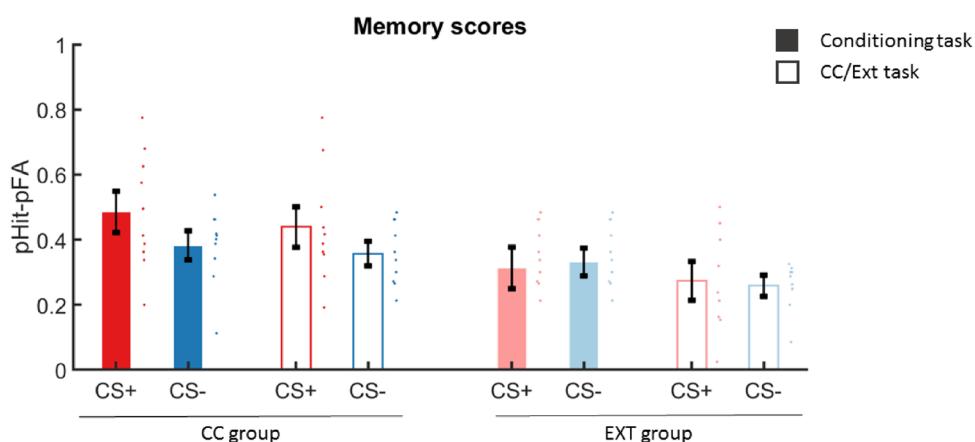

Figure 5. Mean memory scores (Hit rate - False alarm rate), standard errors of the mean (represented by the black bars) and individual memory scores (represented by the scatter plots). Per group memory, scores for CS+ and CS-items previously seen in the conditioning and in the counter-conditioning/extinction task are represented with different bars.
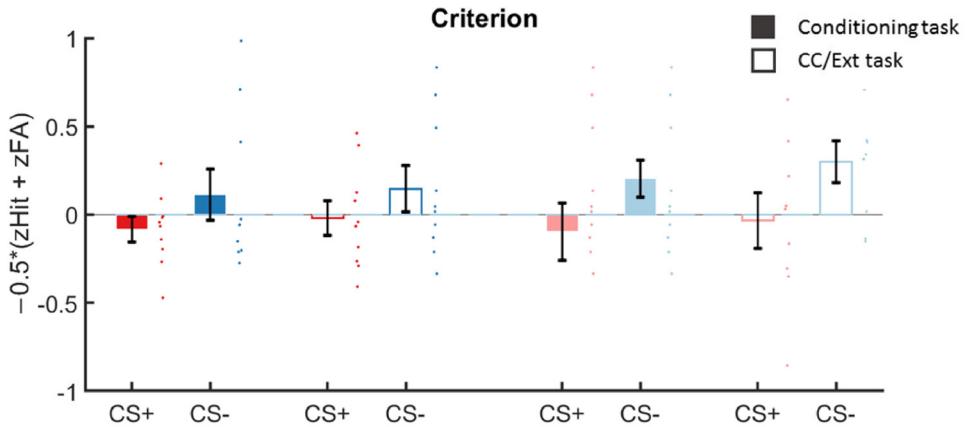
Figure 6. Criterion scores, standard errors of the mean and individual criterion scores from the episodic memory test, presented per group and per task in which the pictures were previously used. Higher criterion scores represent lower likeliness of the participants to state a certain item is old. As illustrated, criterion scores are significantly higher for items from counter-conditioning/extinction.

task (M = .03, SD = .06). The effect of CS-type was trend significant ($F(1, 17)$ = 4.23, p = .055, η_p^2 = .20) with a higher criterion for CS- items compared to CS+. In sum, participants are more conservative in stating that an item from the counter-conditioning/extinction task is old rather than new, as they are for items from the conditioning task. Furthermore, participants tend to be more conservative concerning CS- items.

Skin conductance responses (SCR)

Categorical fear conditioning task

To check the effectiveness of the conditioning paradigm, SCR data from this task were analysed using a group (EXT, CC) x phase (early conditioning [first half of the trials], late conditioning [second half of the trials]) x CS-type (CS+, CS-) 2 x 2 x 2 repeated measures ANOVA. Only the trials where no shock occurred were taken into account. There was a significant main effect of CS-type ($F(1, 18)$ = 17.72, p = 0.001, η_p^2 = .50). A post-hoc paired-samples t-test confirmed that SCRs to CS+ items were greater than to CS- items (respectively, M = .45, SD = .43; M = .22, SD = .25; $t(19)$ = 4.21, p < .001), confirming the effectivity of the paradigm in inducing anticipatory SCRs to the CS+. Additionally, there was a main effect of phase ($F(1, 18)$ = 8.16, p = .010, η_p^2 = .31): SCRs to both CSs declined at the end of the conditioning task (early: M = .43, SD = .45; late: M = .25, SD = .24; $t(19)$ = 2.90, p = .009) compared to the first half. There were no group effects: Both groups showed conditioned responses specific to the CS+.

Counter-conditioning phase

To compare the effectiveness of the extinction and counter-conditioning tasks in reducing threat responses, the SCR data were submitted to a group (EXT, CC) x phase (early, late) x CS-type (CS+, CS-) 2 x 2 x 2 repeated measures ANOVA. For CS-type, a significant main effect was found ($F(1, 18)$ = 8.87, p = 0.008, η_p^2 = .33). A post-hoc paired-samples t-test showed that in both groups and over the course of the whole task, the SCRs for CS+ items were greater than for CS- items (respectively, M = .22, SD = .27; M = .16, SD = .20; $t(19)$ = 2.80, p = 0.12). There was also a significant difference between the groups ($F(1, 18)$ = 4.48, p = .049, η_p^2 = .20) with participants in the counter-conditioning group having lower SCRs (respectively, M = 0.90, SD = .10; M = .29, SD = .28).

Fear recovery test

To test the effect of the counter-conditioning procedure on the return of threat responses during the spontaneous recovery test, SCRs on the last two trials of the extinction or counter-conditioning task were compared with the first two trials of the spontaneous recovery task in a group (EXT, CC) x task (Ext/CC, spontaneous recovery) x CS-type (CS+, CS-) 2 x 2 x 2 repeated measures ANOVA. Only a significant main effect of CS-type was found ($F(1, 17)$ = 5.48, p = 0.32, η_p^2 = .24). Similarly to previous analyses, SCRs for CS+ items were bigger than for CS- items (respectively, M = .45, SD = .44; M = .32, SD = .32). However, when separately analysing the first two trials of the spontaneous recovery task in a group (EXT, CC) x CS-type (CS+, CS-) 2 x 2 repeated measures ANOVA, no significant main
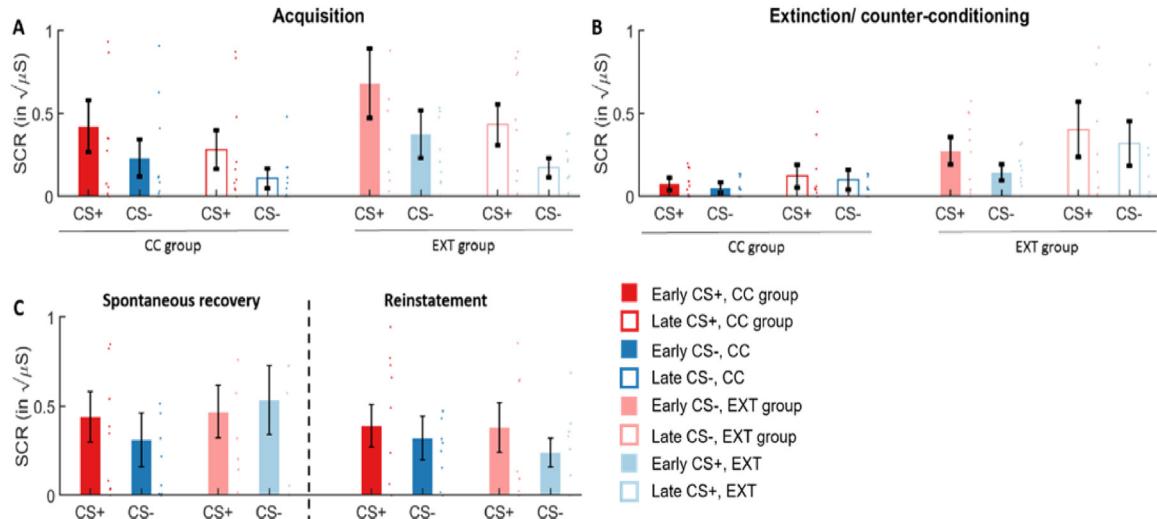
Figure 7. Skin conductance data (SCR) of A. the conditioning task, B. the counterconditioning and extinction tasks for the two groups. C. Shows the data from the first two trials of the spontaneous recovery test at the start of the second session, on the left. On the right, the first trial of the reinstatement test is shown. Figures represent the mean of the square roots of the SCRs, which is measured in micro Siemens. Black bars represent the standard errors of the mean, individual means are represented by the scatter plots.

effect of CS-type remained. As shown by additional paired samples t-tests, neither group shows a differential CS+/CS- SCR at the start of the second session (CC group: CS+: $M = .30$, $SD = .22$, CS-: $M = .21$, $SD = .18$, $t(9) = 1.78$, $p = .110$; EXT group: CS+: $M = .21$, $SD = .24$, CS-: $M = .20$, $SD = .14$, $t(8) = .11$, $p = .912$). Taken together, neither group shows spontaneous recovery of the fear response. The on the first day conditioned increased SCRs to CS+ items did not spontaneously recover after the time interval between the two sessions.

To test for reinstatement in the second part of the fear recovery test, a reinstatement index was calculated (first trial of reinstatement test - last trial of spontaneous recovery test; similar to Lucas, Luck, & Lipp, 2018) and submitted to a group (EXT, CC) x CS-type (CS+, CS-) 2 x 2 repeated measures ANOVA. Neither group shows a significant reinstatement of the fear response for neither the CS+ nor the CS- items. To explore the differences between CS+ and CS- items separately for the two groups, additional t-tests were conducted on the first two trials of the reinstatement phase. In the counter-conditioning, there was no difference between the CS+ and CS- trials (respectively, $M = .34$, $SD = .28$; $M = .25$, $SD = .21$; $t(9) = 1.32$, $p = .219$). The same holds for the extinction group (respectively, $M = .26$, $SD = .26$; $M = .28$, $SD = .20$; $t(8) = -.30$, $p = .773$). Similarly to the spontaneous recovery, no return of SCRs to the CS+ trials was found after triggering reinstatement of this response. Contrary to expectations, in both parts of the retention tests there tends to be more CS+/CS- differences

in the counter-conditioning group. Although not significant, the counter-conditioning phase therefore seems to be less effective in preventing the return of the fear responses.

Valence localizer

SCR data from the mixed valence functional localizer task were subjected to a group (ECT, CC) x phase (first and second half of the trials) x CS-type (CS+ shock, CS+ reward, CS-) 2 x 2 x 3 repeated measures ANOVA. There were both a main effect of CS-type ($F(2, 16) = 7.37$, $p = .005$, $\eta_p^2 = .48$) and of phase ($F(1, 17) = 4.65$, $p = .045$, $\eta_p^2 = .22$). The anticipatory SCR for CS+ shock items ($M = .83$, $SD = .84$) was bigger than both the SCR for the CS+ reward ($M = .33$, $SD = .39$; $t(18) = 3.92$, $p = .001$) and CS- ($M = .24$, $SD = .24$; $t(19) = 3.87$, $p = .001$). The SCR for CS+ reward and CS- did not differ from each other ($t(18) = 1.45$, $p = .164$); confirming the effectiveness of the fear learning aspect of this task. Moreover, these results support the efficacy of SCR data as a unique metric for anticipatory aversive arousal but not reward anticipation. Similar to the SCR data from the conditioning task, SCRs decreased towards the end of the experiment compared to the beginning (respectively, $M = .52$, $SD = .52$; $M = 40$, $SD = .43$; $t(18) = 2.21$, $p = .040$).

In sum, SCR data support the effectiveness of the categorical fear conditioning task in inducing greater anticipatory SCRs in CS+ trials than in CS- trials. In both groups, this differential SCR is successfully reduced at the end of the counter-conditioning/ extinction task. The counter-conditioning paradigm
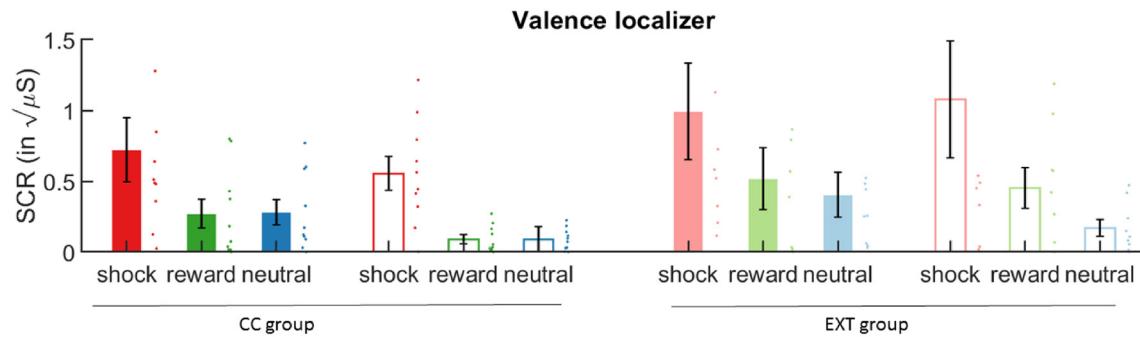
Figure 8. Skin conductance data from the mixed-valence localizer task. Red bars represent the SCR in anticipation of a shock. Green bars represent the anticipatory SCR in trials where a reward could be received. Blue bars show the SCR in CS- trials. As expected, SCRs are higher in anticipation of a shock than of a reward or in neutral trials (CS+ shock: M = .83, SD = .84; CS+ reward: M = .33, SD = .39; CS- M = .24, SD = .24).

seems to reduce fear responses faster. However, no differences were found in effectiveness of preventing the return of the fear responses. On the second day, contrary to expectations, this differential SCR returned in neither group when testing for spontaneous recovery and reinstatement of this fear response.

Pupil dilation data

The pupil dilation responses (PDR) were analysed to assess the effectivity of the categorical fear conditioning task in inducing fear related PDRs. Namely, by submitting the PD (pupil dilation) data to a group (EXT, CC) x phase (early conditioning [first half of the trials], late conditioning [second half of the trials]) x CS-type (CS+, CS-) 2 x 2 x 2 repeated measures ANOVA. This analysis did not reveal significant main effects. However, a significant CS-type*phase*group interaction ($F(1, 12) = 5.84$, $p = .033$, $\eta_p^2 = .33$) and a significant CS-type*group ($F(1, 12) = 7.82$, $p = .016$, $\eta_p^2 = .39$) were found. These might reflect different starting points of the groups with regard to the different categories. To explore whether the two groups were comparable at the end of the conditioning phase, a group (EXT, CC) x CS-type (CS+, CS-) 2 x 2 x 2 repeated measures ANOVA was conducted with the data from the second half of the experiment. This yielded a nearly significant main effect of CS-type ($F(1, 12) = 4.10$, $p = .066$, $\eta_p^2 = .25$). At the end of the task, anticipatory PDRs for CS+ trials seem to be bigger than for CS- trials (respectively, $M = .54$, $SD = 1.16$; $M = -.232$, $SD = .59$; $t(13) = 2.00$, $p = .067$). So the conditioning phase seems to successfully induce conditioned PDRs in both groups.

For the counter-conditioning/extinction task, the PD data were analysed with a group (EXT, CC) x phase (early EXT/CC, late EXT/CC) x

CS-type (CS+, CS-) 2 x 2 x 2 repeated measured ANOVA. This analysis did not reveal significant results, supporting the effectiveness of the counter-conditioning/extinction procedure in reducing conditioned PDRs. To explore whether the absence of an effect of phase, which would be expected if the reduction in conditioned PDRs takes place over the course of the tasks, might be explained by rapid extinction within the first few trials, additional ANOVAs for the separate phases were conducted. This analysis also did not reveal significant results. As described below, data from the valence localizer show a lower PDR for stimuli predicting rewards as for stimuli predicting shocks with PDRs in neutral trials lying in between those two. Since PDRs on day two did not differentiate between the CS-types in both groups, no shock or reward associations that are formed on the first day seem to influence the PDRs on the second day.

To test for spontaneous recovery of the fear response, the last two trials of the counter-conditioning/extinction task and the first two trials of the spontaneous recovery task were submitted to a group (EXT, CC) x CS-type (CS+, CS-) 2 x 2 repeated measures ANOVA. Neither this analysis nor an analysis (group (EXT, CC) x CS-type (CS+, CS-) 2 x 2 repeated measures ANOVA) computed on reinstatement indexes (first trial of reinstatement test - last trial of spontaneous recovery test) of the PD data revealed significant effects.

PD data from the mixed valence localizer were submitted to a group (EXT, CC) x phase (mean of first half, mean of the second half of trials) x CS-type (CS+ shock, CS+ reward, CS-) 2 x 2 x 3 repeated measures ANOVA. This revealed a significant main effect of CS-type ($F(2, 10) = 5.0$, $p = .031$, $\eta_p^2 = .50$). Post-hoc paired-samples t-tests show that pupil responses to CS+ shock ($M = .45$, $SD = .62$) were

bigger than those for CS+ reward trials (M = -.34, SD = .76; $t(12)$ = 3.19, p = .008). The responses to the CS- items (M = .39, SD = 2.08) did not differ from both CS+ items. Additionally, the above mentioned repeated measures ANOVA revealed a significant main effect of phase ($F(1, 11)$ = 5.52, p = .039, η_p^2 = .33). Responses increased over the course of the experiment (in the first half: M = -.11 SD = .85; in the second half: M = .43, SD = .89; $t(12)$ = -2.37, p = .036).

Taken together, at the end of the conditioning task, there seems to be a conditioned PDR specific to the CS+ category stimuli. Next, no differences between PDRs in the CS+ and CS- trials were left in the counter-conditioning and extinction tasks. No spontaneous recovery or reinstatement seems to take place in the second session, therefore no support was found for a higher effectiveness of the counter-conditioning procedure in preventing fear recovery.

Functional MRI data

Category representation localizer task

BOLD responses to the animals, objects and phase-scrambled blocks were modelled with separate regressors for each participant. Subsequent group level analysis of those responses revealed only a small cluster responsive to animals > objects, namely in the right fusiform gyrus (cluster size = 18 voxels, cluster p < .001, corrected). The opposite contrast of objects > animals revealed two small active clusters, namely in the left fusiform gyrus (cluster size = 9 voxels, cluster p < .001, corrected) and right fusiform gyrus (cluster size = 9 voxels, cluster p < .001, corrected).

Acquisition

Whole brain analysis of the BOLD responses during the categorical fear conditioning task revealed differential activity in several areas for the CS+ > CS- contrast. An overview of those areas can be found in Table 1. Amongst others, activity was found in the right (cluster size = 78 voxels, cluster p < .001, corrected) and left (cluster size = 86 voxels, cluster p < .001, corrected) insular cortex. The reversed contrast of CS- > CS+ did not reveal any significant cluster activity.

Counter-conditioning and extinction task

BOLD responses during the counter-conditioning task were subjected to a second level whole brain analysis. We looked for between-group differences and whether differential CS+ and CS- related activity changed from the first half to the second half of the experiment. Neither differential activity was found between CS+ and CS- items nor did this change over time or differ between the two groups. Exploratory analysis with uncorrected p-values also did not reveal significant effects.

Retention test

Whole brain analyses were conducted on the BOLD responses during the spontaneous recovery and reinstatement tests. No differences were found in activity specific to the CS+ or CS- items. Both categories seem to be retrieved in a similar way with involvement of similar brain areas. Moreover, when exploring group differences in CS+ and CS- related activity, no differences were found either. Neither were there differences when looking at significant activity with an uncorrected p-value.

Mixed valence localizer

To identify regions involved in fear and reward anticipation, BOLD activity in this task was modelled with separate regressors for CS+ shock, CS+ reward and CS- items. Activity was found in the left insular cortex (cluster size = 114 voxels, cluster p < .001, corrected) when looking at the CS+ shock > CS+ reward & CS- contrast, indicating involvement of this area in fear anticipation because it was only active in shock trials and not in neutral trials or trials in which a reward could follow. Activity was found in the superior temporal gyrus (cluster size = 33 voxels, cluster p < .001, corrected) for reward anticipation, represented by the CS+ reward > CS+ shock & CS- contrast.

Differential activity in relation to the CS- compared to the other regressors was analysed in order to indicate areas related to situations where no aversive or appetitive outcome was anticipated. Activity was found in frontal middle gyrus (cluster size = 96 voxels, cluster p < .001, corrected). Other involved areas can be found in Table 1.

Lastly, arousal related activity was assessed by a CS+ shock & CS+ reward > CS- contrast. This revealed activity in the bilateral inferior frontal gyrus (left: cluster size = 104 voxels, cluster p < .001, corrected; right: cluster size = 204 voxels, cluster p < .001, corrected).

Discussion

Extinction is a commonly used procedure to reduce undesirable conditioned associations. However, extinction procedures often do not succeed in preventing the return of the extinguished

conditioned response (CR). Therefore, the current study aimed to assess the effectiveness of a counter-conditioning paradigm compared to an extinction paradigm. Moreover, this study aimed to assess the neurocognitive mechanisms underlying counter-conditioning and to test the hypotheses that counter-conditioning either induces an enhanced extinction memory, induces a reward memory that competes with the initial fear memory or induces a reward memory that overwrites the fear memory via reconsolidation-like processes. Preliminary results highlight the effectiveness of the study design to

Table 1. Peak voxel coordinates and cluster statistics and size for the localizer paradigm.

Region	Side	x	y	z	Cluster p (FWE corrected)	z-value	Size (in mm^3)
		MNI Coordinates					
Functional localizer							
Animals > Objects							
Fusiform gyrus	R	30	-44	-26	1.6534e-06	6.01	144
Objects > Animals							
Fusiform gyrus	L	-31	-50	-6	8.6314e-05	5.70	72
Fusiform gyrus	R	32	-47	-10	8.6314e-05	5.56	72
Acquisition							
CS+ > CS-							
Posterior Insular cortex	R	36	8	12	1.2448e-11	6.49	624
Posterior Insular cortex	L	-36	3	14	2.8082e-12	6.13	688
Rolandic operculum	L	-41	-22	22	1.8874e-15	6.12	1032
Superior temporal gyrus	R	46	-34	20	7.0911e-09	6.01	376
Rolandic operculum	R	54	0	10	2.5200e-07	5.88	256
MCC	L	-8	10	40	1.0514e-10	5.68	536
Mixed valence localizer							
Fear > reward + neutral							
Posterior Insular cortex	L	-38	10	10	1.2448e-11	5.95	912
Reward > fear + neutral							
Superior temporal gyrus	R	64	-34	7	1.6248e-07	5.50	264
Neutral > fear + reward							
Angular gyrus	L	-51	-70	30	1.0190e-8	5.45	312
Angular gyrus	R	54	-67	27	1.4559e-10	5.39	448
Middle frontal gyrus	L	-28	28	42	2.7534e-14	5.30	768
Arousal (fear + reward > neutral)							
Inferior frontal gyrus	L	-31	26	12	5.7732e-15	5.68	832
Inferior frontal gyrus	R	44	18	7	<.001	5.60	1632

Notes: All coordinates are defined in MNI152 space. All reported statistics are significant at p < .05, cluster corrected.

induce conditioned skin conductance responses (SCRs) and pupil dilation responses (PDRs) and to increase subjective arousal and negative valence ratings with regard to the CS+ category. SCR data from the counter-conditioning and extinction task suggest both procedures are effective in reducing conditioned fear responses. Furthermore, analysis of reaction times recorded during the counter-conditioning and mixed valence localizer tasks indicate that the monetary incentive delay (MID) components of tasks were successful in producing sufficient reward incentive.

We found successful category conditioning with the current task design. After the conditioning phase, subjective ratings, SCRs and PDRs differentiated between CS+ and CS- items. These differences were in the expected directions, namely CS+ items were rated more negative and arousing than CS- items and CS+ trials were related to increased SCRs and PDRs. These results suggest the presence of a learned association between the CS+ category and the possibility of a negative outcome, in this case an electrical shock. The psychophysiological data from the mixed valence localizer, in which fear anticipation responses were measured alongside reward anticipation and neutral trials, support the directions of the findings from the conditioning task.

The reaction time data we collected support the effectiveness of the task design we used for counter-conditioning. In order to create an appetitive US with enough incentive to counteract the initial aversive US, we decided to make the participants actively involved in obtaining the reward rather than passively, which was the case for the aversive US. This was achieved by using an MID task (Knutson et al., 2000). As reaction time data from the counter-conditioning task show, participants' responses become slower over time, however, for the targets in the CS+ trials participants seem to stay motivated to respond fast. Additionally, in the mixed valence localizer the motivation to respond was greater in trials where a reward could be received. Interestingly, even though participants were instructed that their responses would have no influence on the possibility of receiving a shock, participants also show shorter reaction times in CS+ trials in the conditioning phase, as they do for the trials where they could receive a shock in the mixed valence localizer task. This could be explained by a state of increased action preparedness, which is proposed to be induced in the anticipation of threat when responses are available (Gladwin, Hashemi, Van Ast, & Roelofs, 2016). Hashemi et al. (2019) found a similar decrease in response time in relation to physiological threat anticipatory responses.

After the counter-conditioning in one and extinction procedure in the other group, the differential conditioned fear response, that was formed during the conditioning phase, disappeared again. This suggests an effective reduction of the earlier present fear association towards the CS+ category. Again, this was the case for the subjective ratings and the SC and PD data. When looking at the SC data, the CRs seem to fade away faster in the counter-conditioning group than in the extinction group, in line with the hypothesis that counter-conditioning is more effective in reducing fear responses compared to extinction. However, this group difference in effectiveness is not found with regard to the subjective ratings. Neither is this difference found in the PD data, notwithstanding, the PDRs were found to extinguish rapidly in both groups. This ceiling effect in both groups could prevent observing between group differences in effectiveness.

If counter-conditioning and extinction paradigms have a different efficacy in preventing the return of conditioned fear responses, then we hypothesize a reduction in return of subjective and psychophysiological fear responses in the participants undergoing counter-conditioning compared with extinction. However, the data does not support this hypothesis as no group differences were found. These results suggest that the two paradigms have comparable efficacy. An explanation for the absence of a group difference could be that the interval of one day between sessions is not long enough to observe differences in the rate of fear recovery between the groups. Quirk (2002) suggests that fear responses recover gradually with the passing of time after extinction, with an interval of minimum ten days until full recovery in rats and only minor recovery after one day.

Findings from the episodic memory task show that participants in the counter-conditioning group have a better memory for the items they have seen during the first session than the participants in the extinction group. In line with previous studies (Patil et al., 2017; Dunsmoor et al., 2012; Dunsmoor et al., 2015), participants in the counter-conditioning group tend to have higher memory scores for items from the CS+ category than from the CS- category. This difference does not seem to be present in the extinction group. These results from the counter-conditioning group support the competition hypothesis as for both the conditioning and counter-conditioning phase, memory for the emotionally

relevant stimulus category seems enhanced. This suggests both the existence of a fear and a reward association (Hamann, Ely, Grafton, & Kilts, 1999). Alternatively, the later formed reward association with the CS+ category may have caused a beneficial memory effect for items from both phases, inducing improved recollection of CS+ items in general (Hamann, 2001). This would explain why in the extinction group no CS+ related memory enhancement seems to be present.

Although data from the episodic memory test imply that both a fear and reward memory exist after counter-conditioning that compete for expression, results from the reaction time data point more in the direction of the enhanced extinction hypothesis to underlie counter-conditioning. As indicated by the conditioning, counter-conditioning and mixed valence localizer tasks, participants respond faster in trials where they could receive a shock or earn a reward, compared to neutral CS- trials. At the start of the second day, during the test for spontaneous recovery and reinstatement, participants respond similarly in CS+ and CS- trials. Moreover, this was the same in both groups, suggesting that similar mechanisms underlie counter-conditioning as well as extinction. In neither group, a fear or reward association with the CS+ category seemed to influence the response behaviour of the participants on the second day.

So far, the brain regions that have been found active in relation to the tasks are in line with the expected regions. In line with earlier studies using a task similar to the currently used category localizer (Dunsmoor et al., 2013; de Voogd et al., 2016a), we found activity in the fusiform gyrus. In the conditioning and mixed valence localizer task, the insular cortex was found to be involved in fear learning, as could be expected based on the role of the insula in monitoring emotions and anticipation of affective stimuli (Nicholson et al., 2016). More data might provide a more reliable indication of which subpart of the insula is involved in those tasks, since the subareas of the insular cortex are found to have distinguishable functionalities (Nelson et al., 2010). An area that was not found to be active during fear learning was the amygdala. This is unexpected, as based on studies naming the amygdala as crucial for acquiring conditioned fear associations (Quirk & Milad, 2012; Fragkaki, Thomaes, & Sijbrandij, 2016), it was expected to play a role in the current study as well. However, multiple other conditioning studies also do not report differential amygdala involvement during fear learning (Mechias, Etkin, & Kalisch, 2010; Bulganin et al., 2014; de Voogd et al., 2016a;

Fullana et al., 2016). The proposed distinction between situations with certain and uncertain threats by Fox and Shackman (2019) could shed light on the absence of amygdala involvement. They propose that in situations with a certain threat the amygdala and BNST are involved, while in situations with an uncertain threat more cortical regions, like the insular cortex, are involved. This would indeed be in line with the current paradigm. Since we used a partial reinforcement of 50% of the trials, there was uncertainty of whether or not a shock would follow in the CS+ trials. We indeed found involvement of the insular cortex rather than the amygdala and BNST.

Several analyses are yet to be done when data collection is finished and the total sample size of 48 participants is reached. First, a representational similarity analysis (Kriegeskorte, Mur, & Bandettini, 2008) will be conducted on the functional MRI data from the episodic memory test. As Dunsmoor et al. (2013) found, representations of CS+ category members encoded during conditioning are more similar than CS- category items. This analysis can help elucidate the mechanisms of counter-conditioning. If counter-conditioning is akin to extinction, a similar effect on the category representations is expected to be found as in the study of Dunsmoor et al. (2013). Another possibility is that due to the reward learning in the counter-conditioning phase, items previously seen in this phase also have increased similarity in representation. Yet another possibility is that this reward learning induces a reward association with the CS+ category that overwrites the fear association. In this case, a representational similarity analysis can support this hypothesis by showing that only items previously seen in the counter-conditioning phase have enhanced representational similarity. Second, alterations in functional connectivity will be assessed, especially to differentiate the hypothesis that extinction mechanisms are enhanced during counter-conditioning from the hypothesis that a reward association is formed that competes with the conditioned fear association. If the former were correct, similar to during extinction increased vmPFC-amygdala activity during extinction recall would be expected (Sierra-Mercado, Padilla-Coreano, & Quirk, 2011; Krabbe, Gründemann, & Lütthi, 2017; Dunsmoor et al., 2019) as a marker of the vmPFC playing a role in regulating the expression of the fear response. If the second hypothesis were correct, however, involvement of an amygdala-ventral striatum circuitry is expected, in line with the findings of Correia et al. (2016).

As mentioned above, a limitation of the used

design is the short interval between the extinction or counter-conditioning task and the measurement of spontaneous recovery. Having a longer interval might have allowed more spontaneous recovery to take place (Quirk, 2002) and thereby allowing a better comparison of the effectiveness of counter-conditioning and extinction in preventing the return of this fear response. Another limitation of the current results is that they are based on a subset of the final participant sample. In this paper, preliminary results based on 20 participants are discussed. Our study has a final intended sample size of 48, as indicated during preregistration. More participants are needed to successfully find group level effects of conditioning on all psychophysiological outcome measurements and pinpoint the neurocognitive mechanisms of counter-conditioning. Moreover, a limitation of the current study design is that extinction and counter-conditioning immediately follow-up on the conditioning phase. Maren (2014) suggests that immediate extinction is less effective than extinction that takes place at least a day after conditioning. Nonetheless, preliminary data of this design, as shown in this paper, indicate neither spontaneous recovery nor reinstatement of the fear response. This suggests that the current design is strong enough to induce extinction of the fear response. This could be explained by the fact that the paper of Maren (2014) is focused on rats, in which very high stress levels are induced, which is not the case in human studies. Furthermore, since the interval between conditioning and counter-conditioning and extinction is this short, findings are less translatable to clinical situations, where the extinction-based therapies are likely to take place long after the initial formation of fearful associations. A final limitation is that, for pragmatic reasons, the length of the interval between the two sessions varied between participants. Although the second session was always on the day after the first session, the exact amount of time between them varied. If the effectiveness of the counter-conditioning procedure relies on consolidation dependent mechanisms, this might complicate the results.

In conclusion, the current study looks at the effectiveness and neurocognitive mechanisms of counter-conditioning as an alternative to extinction procedures. Preliminary results indicate the effectiveness of the study design in inducing both fear and reward learning. Tentative effects of the counter-conditioning procedure on the episodic memory tend to support the hypothesis that during counter-conditioning a reward association is established, which competes with the conditioned fear association for expression. If results from the final sample also support the effectiveness of a counter-conditioning procedure, then this underlines counter-conditioning as a promising alternative to extinction. Understanding counter-conditioning will provide neurocognitive insight into fear learning and regulation and may help to improve treatments for stress- and anxiety-related disorders.

References

Alberini, C. M., & LeDoux, J. E. (2013). Memory reconsolidation. *Current Biology, 23*(17), R746-R750.

Bouton, M. E. (1993). Context, time and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114*(1), 80–99.

Bouton, M. E. (2002). Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biological Psychiatry, 52*(10), 976–986.

Bouton, M. E., & Bolles, R. C. (1979). Role of conditioned contextual stimuli in reinstatement of extinguished fear. *Journal of Experimental Psychology-Animal Behavior Processes, 5*(4), 368-378.

Brooks, D. C., Hale, B., Nelson, J. B., & Bouton, M. E. (1995). Reinstatement after counterconditioning. *Animal Learning & Behavior, 23*(4), 383-390.

Bulganin, L., Bach, D. R., & Wittemann, B. C. (2014). Prior fear conditioning and reward learning interact in fear and reward networks. *Frontiers in Behavioral Neuroscience, 8*:67.

Correia, S. S., McGrath, A. G., Lee, A., Graybiel, A. M., & Goosens, K. A. (2016). Amygdala ventral striatum circuit activation decreases long-term fear. *eLife,* 1–25.

Craske M. G., Liao B., Brown L., & Vervliet, B. (2012). Role of inhibition in exposure therapy. Journal of Experimental Psychopathology, 3(3), 322–345.

Craske, M. G., & Mystkowski, J. L. (2006). Exposure Therapy and Extinction: Clinical Studies. In M. G. Craske, D. Hermans, & D. Vansteenwegen (Eds.), *Fear and learning: From basic processes to clinical implications* (pp. 217-233). Washington DC, US: American Psychological Association.

de Voogd, L. D., Fernandez, G., & Hermans, E. J. (2016a). Awake reactivation of emotional memory traces through hippocampal–neocortical interactions. *Neuroimage, 134,* 563-572.

de Voogd, L. D., Fernandez, G., & Hermans, E. J. (2016b). Disentangling the roles of arousal and amygdala activation in emotional declarative memory. *Social Cognitive and Affective Neuroscience, 11*(9), 1471–1480.

de Voogd, L. D., Murray, Y. P., Barte, R. M., van der Heide, A., Fernandez, G., Doeller, C. F., & Hermans, E. J. (2019). The role of hippocampal spatial representations in contextualization and generalization of fear. *bioRxiv preprint.*

Dunsmoor, J. E., Kragel, P. A., Martin, A., & LaBar, K. S. (2014). Aversive Learning Modulates Cortical Representations of Object Categories. *Cerebral Cortex, 24*(1), 2859-2872.

Dunsmoor, J. E., Kroes, M. C. W., Li, J., Daw, N. D., Simpson, H. B., & Phelps, E. A. (2019). Role of human ventromedial prefrontal cortex in learning and recall of enhanced extinction. *Jorunal of Neuroscience,* 2713-2718.

Dunsmoor, J. E., Kroes, M. C. W., Moscatelli, C. M., Evans, M. D., Davachi, L., & Phelps, E. A. (2018). Event segmentation protects emotional memories from competing experiences encoded close in time. *Nature Human Behaviour, 2,* 291-299.

Dunsmoor, J. E., Martin, A., & LaBar, K. S. (2012). Role of conceptual knowledge in learning and retention of conditioned fear. *Biological Psychology, 89*(2), 300-305.

Dunsmoor, J. E., & Murphy, G. L. (2015). Categories, concepts, and conditioning: how humans generalize fear. *Trends in Cognitive Sciences, 19*(2), 73-77.

Dunsmoor, J. E., Murty, V. P., Davachi, L. & Phelps, E. A. (2015). Emotional learning selectively and retroactively strengthens memories for related events. *Nature, 520*(7547), 345-348.

Esteban, O., Markiewicz, C., Blair, R., Moodie C., Isik, A. I., Aliaga, … & Kent, J., et al. (2018). fMRIPrep: A Robust Preprocessing Pipeline for Functional MRI. *Nature Methods,* 16(1), 111-116.

Fox, A., & Shackman, A. J. (2019). The central extended amygdala in fear and anxiety: Closing the gap between mechanistic and neuroimaging research. *Neuroscience letters, 693*(S1), 58-67.

Fragkaki, I., Thomaes, K., & Sijbrandij, M. (2016). Posttraumatic stress disorder under ongoing threat: a review of neurobiological and neuroendocrine findings. *European Journal of Psychotraumatology,* 7:30915.

Fullana, M. A., Harrison, B. J., Soriano-Mas, C., Vervliet, B., Cardoner, N., Avila-Parcet, A., & Radua, J. (2016). Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. *Molecular Psychiatry, 21* (4), 500-508.

Gladwin, T. E., Hashemi, M. M., Van Ast, V., & Roelofs, K. (2016). Ready and waiting: Freezing as active action preparation under threat. *Neuroscience Letters, 619,* 182-188.

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., et al. (2013). The Minimal Preprocessing Pipelines for the Human Connectome Project. *NeuroImage, Mapping the connectome, 80,* 105–24.

Glover, G. H., Li, T. Q., & Ress, D. (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magnetic Resonance in Medicine, 44,* 162-167

Green, S. R., Kragel, P. A., Fecteau, M. E., & LaBar, K. S. (2014). Development and validation of an unsupervised scoring system (Autonomate) for skin conductance response analysis. *International Journal of Psychophysiology, 91*(3), 186–93.

Greve, D. N., & Fischl, B. (2009). Accurate and Robust Brain Image Alignment Using Boundary-Based Registration. *NeuroImage, 48* (1), 63–72.

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Science, 5*(9), 394-400.

Hamann, S., Ely, T. D., Grafton, S. T., & Kilts, C. D. (1999). Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience, 2*(3), 289-293.

Hashemi, M. M., Zhang, W., Kaldewaij, R., Koch, S. B. J., Jonker, R., Figner, … & Roelofs, K. (2019). Human defensive freezing is associated with acute threat coping, long term hair cortisol levels and trait anxiety. *BioRxiv.*

Hermans, E. J., Henckens, M. J. A. G., Roelofs, K., & Fernandez, G. (2013). Fear bradycardia and activation of the human periaqueductal grey. *NeuroImage, 66,* 278-287.

Hitchcock, J., & Davis, M. (1986). Lesions of the amygdala, but not of the cerebellum or red nucleus, block conditioned fear as measured with the potentiated startle paradigm. *Behavioral Neuroscience, 100*(1), 11–22.

Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J., & Turner, R. (2002). Image distortion correction in fMRI: A quantitative evaluation. *Neuroimage, 16*(1), 217-240

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage, 17*(2), 825–841.

Kaag, A., Schluter, R. S., Karel, P., Homberg, J., van den Brink, W., Reneman, L., & van Wingen, G. A. (2016). Aversive Counterconditioning Attenuates Reward Signaling in the Ventral Striatum. *Frontiers in Human Neuroscience, 10*:418.

Kang, S., Vervliet, B., Engelhard, I. M., van Dis, E. A. M., & Hagenaars, M. A. (2018). Reduced return of threat expectancy after counterconditioning versus extinction. *Behaviour Research and Therapy, 108,* 78-84.

Karel, P., Almacellas-Barbanoj, A., Prijn, J., Kaag, A., Reneman, L., Verheij, M., & Homberg, J. (2019). Appetitive to aversive counter-conditioning as intervention to reduce reinstatement of reward-seeking behavior: the role of the serotonin transporter *Addiction Biology.*

Kerkhof, I., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2011). Counterconditioning An Effective Technique for Changing Conditioned Preferences. *Experimental Psychology, 58*(1), 31-38.

Knutson, B., Westdorp, A., Kaiser, E., & Hommer, D. (2000). FMRI Visualization of Brain Activity during a Monetary Incentive Delay Task. *Neuroimage, 12*(1), 20-27.

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis connecting the branches of systems neuroscience. *Frontiers in systems neuroscience,* 2:4.

Krabbe, S., Gründemann, J., & Lüthi, A. (2017) Amygdala inhibitory circuits regulate associative fear. *Biological Psychiatry, 83*(10), 800-809.

Kroes, M. C. W., Dunsmoor, J. E., Lin, Q., Evans, M., & Phelps, E. A. (2017). A reminder before extinction strengthens episodic memory via reconsolidation but fails to disrupt generalized threat responses. *Scientific reports,* 10858.

LeDoux, J. E., Ruggiero, D. A., & Reis, D. J. (1985). Projections to the subcortical forebrain from anatomically defined regions of the medial geniculate body in the rat. *Journal of Comparative Neurology, 242*(2), 182–213.

LeDoux, J. E., Sakaguchi, A., & Reis, D. J. (1984). Subcortical efferent projections of the medial geniculate nucleus mediate emotional responses conditioned to acoustic stimuli. *Journal of Neuroscience, 4*(3), 683–698.

Lucas, K., Luck, C. C., & Lipp, O. V. (2018). Novelty-facilitated extinction and the reinstatement of conditional human fear. *Behaviour research and therapy, 109*, 68-74.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58,* 25-45.

Mechias, M. L., Etkin, A., & Kalisch, R. (2010). A meta-analysis of instructed fear studies: Implications for conscious appraisal of threat. *Neuroimage, 49*(2), 1760-1768.

Milad, M. R., & Quirk, G. J. (2012). Fear extinction as a model for translational neuroscience: ten years of progress. *Annual Review of Psychology, 63*, 129–151.

Nelson, S. M., Dosenbach, N. U. F., Cohen, A. L., Wheeler, M. E., Schlaggar, B. L., & Petersen, S. E. (2010). Role of the anterior insula in task-level control and focal attention. Brain Structure & Function, 214(5-6), 669-680.

Nicholson, A. A., Sapru, I., Densmore, M., Frewen, P. A., Neufeld, R. W. J., Theberge, J., McKinnon, M. C., & Lanius, R. A. (2016). Unique insula subregion resting-state functional connectivity with amygdala complexes in posttraumatic stress disorder and its dissociative subtype. *Psychiatry Research-Neuroimaging, 250*, 61-72.

Patil, A., Murty, V. P., Dunsmoor, J. E., Phelps, E. A., & Davachi, L. (2017). Reward retroactively enhances memory consolidation for related items. *Learning & Memory, 24*(1), 65-69.

Pavlov, I. P. (1927). *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex.* London: Oxford University Press.

Phelps, E. A., Delgado, M. R., Nearing, K. I., & LeDoux, J. E. (2004). Extinction learning in humans: role of the amygdala and vmPFC. *Neuron, 43*(6), 897–905.

Posse, S., Wiese, S., Gembris, D., Mathiak, K., Kessler, C., Grosse-Ruyken, M., … & Kiselev, V. G. (1999). Enhancement of BOLD Contrast Sensitivity by Single-Shot Multi-Echo Functional MR Imaging. *Magnetic Resonance in Medicine 42*(1), 87–97.

Reinders, A. A. T. S., Den Boer, J. A., & Büchel, C. (2005). The robustness of perception. *European Journal of Neuroscience, 22*(2), 524-530.

Rescorla, R. A., & Heth, D. (1975). Reinstatement of fear to an extinguished conditioned stimulus. *Journal of Experimental Psychology-Animal Behavioral Processes, 104*(1), 88–96.

Rescorla R. A., & Wagner A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black, W. F. Prokasy, Classical conditioning II (pp. 64-99). New York, NY: Appleton-Century-Crofts

Schiller, D., Monfils, M. H., Raio, C., Johnson, D. C., LeDoux, J. E., & Phelps, E. A. (2010). Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature, 463*(7277), 49-U51.

Schultz, W., Tremblay, L., & Hollerman, J. R. (2000). Reward Processing in Primate Orbitofrontal Cortex and Basal Ganglia. *Cerebral Cortex, 10*(3), 272-283.

Siegle, G. J., Steinhauer, S. R., Stenger, V. A., Konecky, R., & Carter, C. S. (2003). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. *NeuroImage, 20*, 114-124.

Sierra-Mercado, D., Padilla-Coreano, N., & Quirk, G. J. (2011). Dissociable Roles of Prelimbic and Infralimbic Cortices, Ventral Hippocampus, and Basolateral Amygdala in the Expression and Extinction of Conditioned Fear. *Neuropsychopharmacology, 36*(2), 529-538.

Tunstall, B. J., Verendeev, A., & Kearns, D. N. (2012). A comparison of therapies for the treatment of drug cues: Counterconditioning vs. extinction in male rats. *Experimental and Clinical Psychopharmacology, 20*(6), 447-453.

Vervliet, B., Craske, M. G., & Hermans, D. (2013). Fear extinction and relapse: state of art. *Annual Review of Clinical Psychology, 9,* 215-248.

Quirk, G. J. (2002). Memory for extinction of conditioned fear is long-lasting and persists following spontaneous recovery. *Learning & Memory, 9*(6), 402–407.

Conflict-Related Theta Activity in the Midfrontal Cortex:
An Electrophysiological Rodent Model for Response Conflict

Jordi ter Horst[1,2]

Supervisors: Nils Zuiderveen Borgesius[1,2], Michael X Cohen[1,2]

[1]*Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands*
[2]*Radboud University Medical Centre Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands*

During response conflict, when a stimulus is associated with multiple responses, there is need for a mechanism that enables selection of the goal-relevant response, while suppressing a habitual response. Due to the lack of spatial and temporal precision in previous studies that tried to understand the mechanisms underlying response conflict, four male wild-type Long-Evans-Tg(TH-Cre)3.1Deis rats were trained on an adapted Simon task. After training they were implanted with a custom-made probe to record local field potentials from the midfrontal cortex during execution of the Simon task. During conflict, when the stimulus side did not match with the response side, reaction times and errors increased. Additionally, intracranial electrophysiology in the midfrontal cortex revealed that theta power (5-8 Hz) increased during conflict. This newly established electrophysiological rodent model opens up new possibilities for studying the underlying neural circuits of response conflict.

Keywords: response conflict, electrophysiology, rodent model, midfrontal cortex

During everyday life people sometimes face the problem of response conflict, a phenomenon in which an external stimulus causes an internal conflict since the stimulus is associated with multiple potential responses. In most cases of response conflict there is a habitual response that has to be prevented in order to react with the goal-relevant response. Response conflict has been studied a lot in the past using the Simon task, in which participants are asked to manually respond to a stimulus characteristic (e.g. colour) while ignoring the stimulus location. It is typically found that participants respond faster when the stimulus side corresponds with the response hand side (congruent) as compared to when the stimulus side does not correspond with the response hand side (incongruent), and erroneous responses are made more often during incongruent trials (Simon & Rudell, 1967; Simon, Hinrichs, & Craft, 1970). The effect on reaction time, known as the Simon effect, is dependent on the congruency of the previous trial (Stürmer, Leuthold, Soetens, Schröter, & Sommer, 2002; Egner, 2007). When the previous trial is incongruent, this can result in a decrease or even a reversal of the Simon effect as compared to when the previous trial is congruent.

Besides the typical effects of stimulus-response congruency on reaction times and error rates, several electrophysiological manifestations have been shown to occur during response conflict. The first one is a conflict-related power increase in the theta band (4-8 Hz) around 400 ms after cue onset, located at electrodes covering the midfrontal cortex (MFC) (Cohen & Donner, 2013; Cohen & Ridderinkhof, 2013; Nigbur, Ivanova, & Stürmer, 2011). This conflict-related theta modulation appears to be predominantly non-phase-locked to stimulus and response, meaning that ongoing endogenous theta band oscillations change in amplitude as a function of conflict, rather than being elicited by the appearance of the stimulus or the button press. In addition, non-phase-locked theta power can predict congruency and correlates positively with reaction times (Cohen & Donner, 2013). Whereas theta modulations occur around the frontal part of the midline, there are other oscillatory features that arise after conflict. Using source reconstruction, Cohen and Ridderinkhof (2013) found evidence for network shifting driven by the congruency of a trial. Whereas midfrontal theta couples with lateral prefrontal theta during incongruent trials, during congruent trials there is stimulus-contralateral parietal gamma coupling with ventrolateral frontal theta, followed by coupling between midfrontal theta and stimulus-contralateral parietal alpha. It is suggested that

gamma decoupling during incongruent trials occurs to prevent the conflicting spatial location to hinder the correct response. Decreased coupling in the alpha range during incongruent trials might reflect the prevention from keeping the contralateral stimulus attended, but this seems to be unrelated to conflict resolution since this decoupling usually occurs after the response is made. Theta coupling between the MFC and lateral prefrontal cortex (LPFC) during incongruent trials probably reflects the inhibition of the incorrect response, as the amount of coupling negatively correlates with the size of the Simon effect (Cohen & Ridderinkhof, 2013).

Several cognitive models have tried to explain the mechanisms underlying response conflict, but they lack a biologically plausible mechanism for the effects that can be found (Cohen, 2014). To be able to understand how cognitive control is implemented during response conflict, and how it consecutively affects behaviour, it is necessary to monitor the underlying processes in the physical substrate. Although previous studies already took much effort in trying to capture the ongoing physiology during response conflict, they are still difficult to interpret in terms of how stimulus input step-by-step leads to response outcome.

One way to investigate the physiological processes in more detail is to measure electrophysiological signals intracranially, which gives higher spatial resolution as compared to the more common extracranial methods used in previous studies on response conflict. To measure intracranial electrophysiology, there is need for a rodent model of the Simon task, which also shows the typical effects reported in human studies. Courtière, Hardouin, Burle, Vidal, and Hasbroucq (2007) adapted the Simon task to a rat-compatible version and showed that the Simon effect is also evident in rats, as indicated by increased reaction times and more response errors when the stimulus side is not congruent with the response side. As a follow-up, Marx et al. (2012) used positron emission tomography (PET) to capture the brain areas involved in conflict processing during the adapted Simon task. Similar to human studies, they found involvement of midfrontal and premotor areas, given the increased metabolic activity after rats performed the task. Together with the work done by Courtière et al. (2007) this suggests that this rodent model is suitable for studying response conflict. However, the underlying biological processes are still unknown, since PET does not provide insight into ongoing neuronal processes while the rat is preparing a response. Especially time-frequency dynamics could use more elaboration to better

understand how stimulus input consecutively results in response outcome during response conflict.

To better understand how stimulus input step-by-step leads to response outcome during response conflict, there is need for a rodent model that enables investigation of ongoing electrophysiology. In this study, rats were trained on the Simon task after which they were implanted with a custom-made probe in the MFC. As intracranial electrophysiological signals during the Simon task have never been studied before, this thesis primarily focused on conflict-related power changes in the MFC. First, typical behavioural manifestations were verified to replicate validation of this rodent model for response conflict. Second, time-frequency analysis showed conflict-related theta modulations in the MFC.

Method

Animals

Four male wild-type Long-Evans-Tg(TH-Cre)3.1Deis rats took part in this study, weighing 275-310 g before the training procedure started (Rat Resource & Research Center, Columbia, Missouri, USA). They were housed pairwise in Makrolon type III cages (UNO B.V., Zevenaar, The Netherlands) with a reversed 12-hour light-dark cycle (lights dimming at 06:30 and being off at 07:00) in a temperature- and humidity-controlled room (21 ± 2°C, 60 ± 15%). When rats weighed more than 350 g they were housed in Makrolon type IVS cages. After probe implantation, the rats were housed individually and the flat conventional cage lids were substituted with high lids to prevent implant damage. The rats had *ad libitum* access to food pellets and were restricted on water consumption except for half an hour each training and testing day. Water restriction was used because rewards during the training procedure and testing phase were given in the form of water drops. During weekends, when the rats were not trained and tested, they had *ad libitum* access to food pellets and water. Water bottles were taken away on Sundays so rats would be motivated to obtain water on the consecutive training or testing day. Weight and health were monitored on a daily basis to ensure welfare of the rats. All animal procedures were approved by the Animal Welfare Body of the Radboud University Nijmegen and the Animal Experiment Committee, according to national and international laws, to protect appropriate welfare under experimental conditions.

Materials

Training and testing sessions took place in a custom-built Skinner box (25 x 27 x 25 cm), in which one side wall contained three nose poke holes (bottom-left, bottom-centre, bottom-right), four yellow-coloured light-emitting diodes (LED; top-left, top-right, bottom-left, bottom-right), and two small loudspeakers (left, right; see Fig. 1). Each nose poke hole contained an infrared LED transmitter and receiver to continuously monitor potential poking by the rat. The left and right nose poke hole also contained an entry for a small tube in the bottom to enable provision of 50 μL water rewards, driven by solenoid pumps (The Lee Company, Westbrook, Connecticut, USA). Rats were randomly assigned to either receive sound cues (10 vs. 15 kHz) or light cues (top vs. bottom LED), and to the stimulus-response association that would lead to a reward (e.g., one of the rats was assigned to sound cues and was eventually trained to respond in the left nose poke hole if a low pitched tone (10 kHz) was presented, and in the right nose poke hole if a high pitched tone (15 kHz) was presented). The entire Skinner box was controlled by custom MATLAB software.

To record local field potentials (LFPs) in the MFC, a custom-made probe was used. The probe consisted of 30 Tungsten wires (California Fine Wire Company, Grover Beach, CA, USA) of a diameter of 50 μm eachand all were wrapped in a styrene-isoprene-styrene polymer except for the tip. The layout of the probe was designed based on the anatomy of the MFC (see Fig. 2). Before each testing session, the probe was connected to a 32-channel headstage (Omnetics, Dublin, Ireland) of which the cable went through the roof of the Skinner box, suspending onto an elastic band. The headstage cable was connected to an acquisition board (Open Ephys, Lisbon, Portugal).

Training procedure

The behavioural task on which the rats were trained was adapted from Marx et al. (2012), with the only difference that they used food pellets as a reward instead of water drops. The training procedure was also very similar to the one used by Marx et al. (2012), consisting of 1) habituation, 2) central nose poke initiation, 3) bilateral cue discrimination, 4) unilateral cue discrimination, and 5) the Simon task. All rats were handled for two weeks before the training procedure to enable the rats to get used to the experimenter and the lab environment. After handling for two weeks, all rats were put on water restriction. During the habituation phase, rats were put in the Skinner box for 30

minutes for two consecutive days, in which they received a pseudorandom bilateral cue with a fixed time interval of 10 seconds. A reward was directly given at cue onset, independent of their response. This phase helped the rats to familiarize with the Skinner box, cues, nose poke holes and rewards. In the central nose poke initiation phase, rats had to learn to initiate a trial themselves by poking in the central nose poke hole for 1.5 seconds. At the first day of this phase, poking for just a very short time was already sufficient to initiate a trial, causing the occurrence of a pseudorandom bilateral cue and an immediate reward at the associated response side. The time necessary to poke in the central nose poke hole was increased with 0.1 seconds for the next training session if the rat could initiate at least 80 trials per hour in the previous training session. When rats were able to poke for 1.5 seconds and could initiate at least 80 trials, they proceeded to the bilateral cue discrimination phase. From this phase on, rats were trained for one hour each training day for a maximum of 152 trials, and only got a reward following accurate responding. The bilateral cue discrimination phase was meant to teach each rat its stimulus-response association leading to a reward. When rats responded correctly in at least 80% of the trials in three consecutive days, they could go to the unilateral cue discrimination phase. In this phase cues were all unilateral and congruent, meaning that they were only presented at the left or right, and only presented at the side where they had to respond. This phase was meant to familiarize the rats with unilateral cues. When they responded correctly in at least 80% of the trials in three consecutive days, they went to the Simon task (see Fig. 1). In this last phase, rats were presented with congruent and incongruent unilateral cues. Rats were ready for probe implantation when they obtained an accuracy score of at least 80% in three consecutive days.

Surgical procedure

Before rats underwent surgery for probe implantation, they got *ad libitum* water access for five days. Rats were anaesthetised with 5% isoflurane in a mixture of equal parts O_2 and air and a flow rate of 1 L/min for approximately five minutes. Then, isoflurane was set to 3% to keep the animal anaesthetised during the surgery. Breathing, body temperature, O_2 and heart rate levels were monitored every hour to assess whether the isoflurane level should be adjusted. After shaving the top surface of the head, it was fixated in a stereotactic frame and cleaned with ethanol and betadine. The isoflurane flow was mounted to the stereotactic frame to ensure proper anaesthesia. Eye ointment was applied several times during the surgery to protect the eyes from drying out. Then, carprofen (Rimadyl, 0.5 mg/kg) and lidocaine (Xylocaine, 0.4 mL) were injected subcutaneously at the incision site. A midline incision was made from behind the eyes ($\pm$ 3 cm in front of Bregma) to the attachment of the neck muscle ($\pm$ 1 cm behind Lambda). The skull surface was cleaned with sterile cloths and bleached with 5% H_2O_2 to rough up the surface, making it easier for dental cement to stick to it later on. Then, a 0.7 mm hole was drilled, contralateral to the side of probe implantation, to enable insertion of a skull screw for the ground wire of the probe. After measuring and marking the implantation location (see Fig. 2), a craniotomy was made with a stereotactic mounted drill. A bent 27G needle was used to perform durotomy above the implantation site, after which the opening was covered with saline-soaked gelatine sponges to prevent the brain from drying out. Subsequently, the probe was carefully implanted using the stereotactic frame and Vaseline was used to cover the exposed brain surface. The ground wire coming from the probe was attached to the skull screw and covered with conductive paint. To fix and reinforce the probe, dental cement (Parkell, Edgewood, NY, USA) was applied. Eventually the skin was stitched in such a way that the rat could easily blink the eyes. Following that, isoflurane flow was turned off and the rat was removed from the stereotactic frame. The rat was put in a new clean cage on a paper towel, with a heat mat put under the cage while ensuring an unheated location in the cage for the first night. Food was provided in the form of softened and sweetened pellets, as well as a daily subcutaneous injection of carprofen (Rimadyl, 5 mg/kg), both for three consecutive days. Regular food pellets were given for the rest of the recovering period, lasting two weeks. Water was accessible *ad libitum* and weight, wound healing, and general health were monitored daily.

After this recovering period, rats were put on water restriction again to make them motivated to respond correctly during the Simon task. From this moment on intracranial electrophysiology was recorded from the MFC on a daily basis while performing the Simon task. A recording session took approximately one hour, or 152 trials if the rat was faster.

Behavioural pre-processing

All analyses were performed in MATLAB

(R2016b, The MathWorks, Inc., Natick, Massachusetts, US). For each trial, the MATLAB environment saved the congruency, response side, response correctness, reaction time (RT) and movement time (MT). RTs were defined as the time between cue onset and central nose poke release, whereas MTs were defined as the time between central nose poke release and left or right nose poke response (see Fig. 1). Only sessions with an average performance of 80% or higher were included in analyses, as an average performance of 75% could already be obtained by responding correctly in all congruent trials and guessing in incongruent trials. All trials with an RT faster than 150 ms were eliminated, as it was very likely that the rat was not attending the stimulus, or was randomly poking in one of the lateral nose poke holes. All trials with RTs exceeding two standard deviations above the median were also discarded, because it was unsure whether the rat was really engaged in these trials. Trials with incorrect responses were not included for evaluating the difference in RTs and MTs between congruent and incongruent trials. The cut-off for significance was set at $p = .05$.

LFP pre-processing

LFP data were acquired at 30 kHz and saved in raw format. The data were then downsampled offline at 1000 Hz and epoched in time windows between -1 and 2 seconds surrounding cue onset. Although the actual time window of interest is between -0.5 and 1 seconds, slightly broader windows were epoched to prevent edge artifacts in the time-frequency representation. Trials were visually inspected in EEGLAB (Delorme & Makeig, 2004) and manually rejected if they contained movement artifacts or other artifacts that were clearly distinguishable from true electrophysiology from the brain (see Fig. 3 for typical raw data). Independent component analysis (ICA) was used to remove components in the data that were clearly driven by non-brain sources, but were not removed if they also contained a sufficient amount of true brain-generated electrophysiological activity. Trial rejection and component removal was done blind to conditions and responses. Similar to behavioural pre-processing, all trials with incorrect responses and reaction times faster than 150 ms or slower than two standard deviations above the median were eliminated from further analyses, as well as data from sessions that had a performance lower than 80%. These pre-processing steps eventually led to rejection of 27% of the trials. For this thesis, only a randomly chosen subset of sessions from one rat was used for LFP analysis due to the time-consuming pre-processing steps.
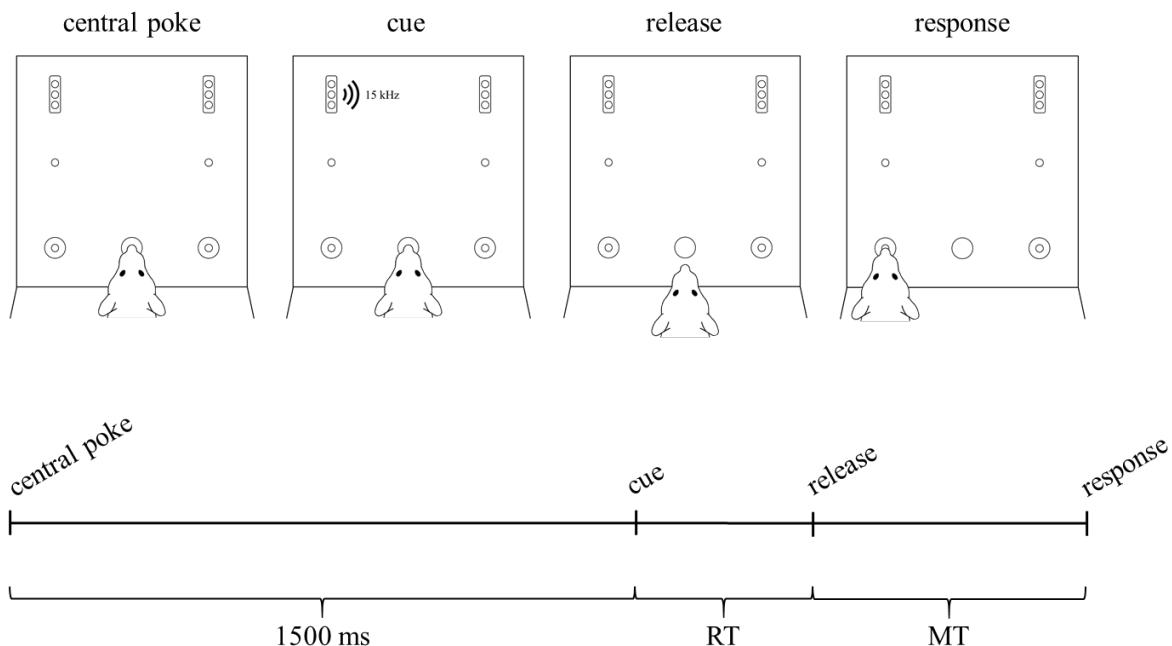

Figure 1. The Simon task adapted for rats. Rats had to initiate a trial themselves by poking in the central nose poke hole for 1500 ms. Then, a pseudorandom cue (congruent or incongruent) was presented for 300 ms, associated with a left or right reward. Reaction time (RT) was defined as the time between cue onset and central nose poke release, and movement time (MT) as the time between central nose poke release and response at either the left or right nose poke hole. Note that the cue and response side in this figure are examples.
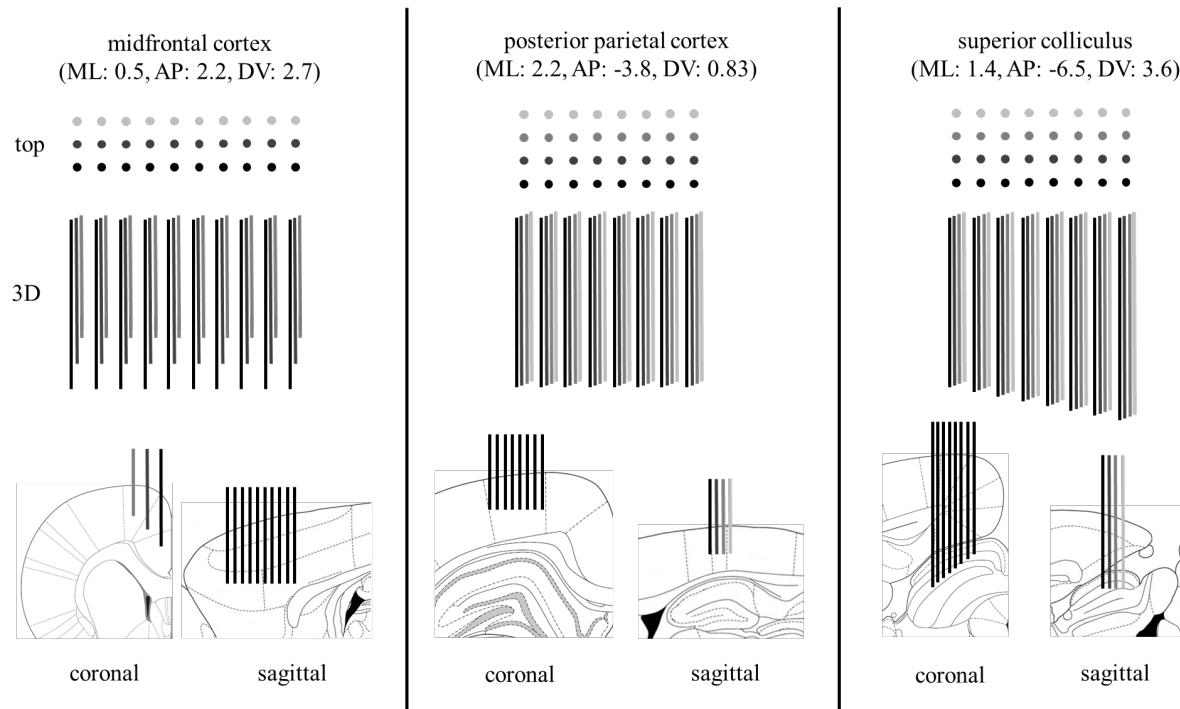
Figure 2. Probe configuration for the midfrontal cortex. The top view of the probe is depicted in the top left corner, showing the 3x10 (500x2250 μm) layout. The bottom left illustrates the probe configuration from a three-dimensional perspective, and the right half of the figure shows the intended positioning of the probe in the rat brain. The coronal slice is from the right hemisphere, and the sagittal slice is shown with the rostral side on the left. Coordinates mentioned in the top right of the figure describe the centre of the probe.
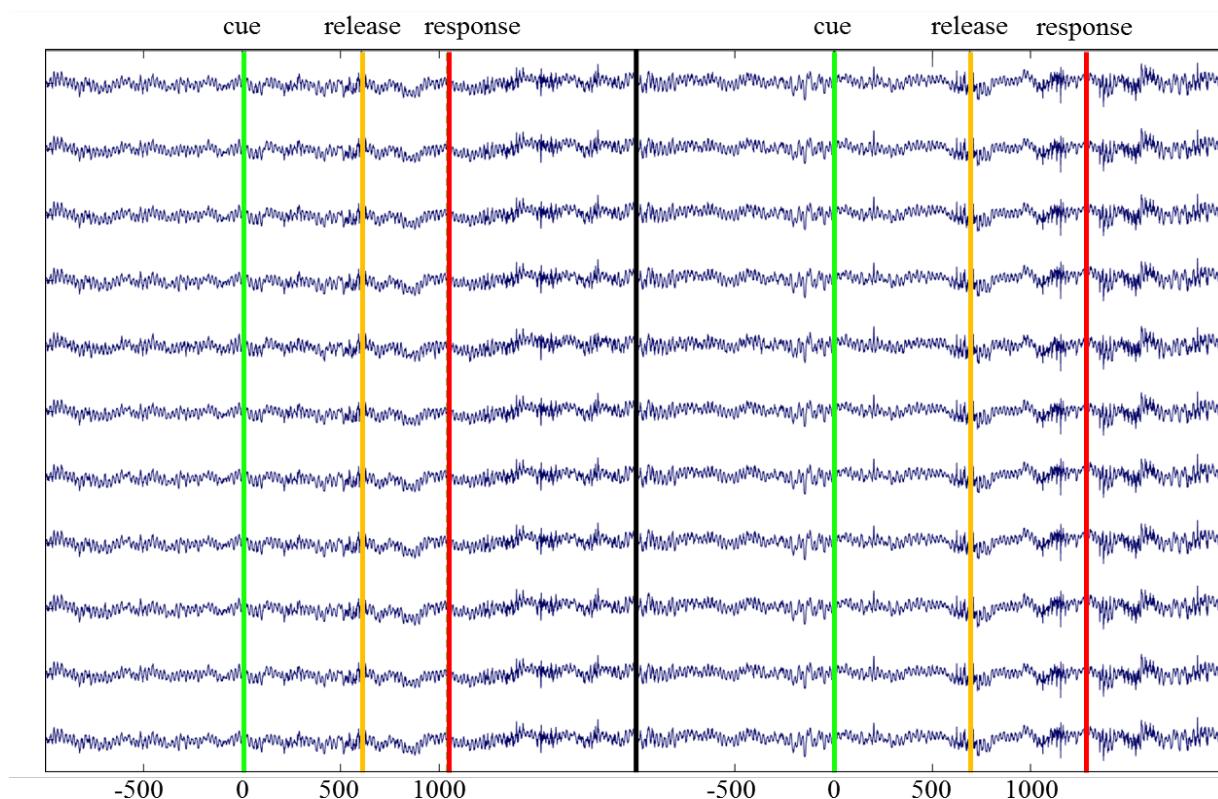

Figure 3. Epochs of typical raw data from a random selection of channels and trials, with markers indicating the cue (green), release (yellow) and response (red). Time (ms) is displayed on the x-axis and channels on the y-axis, scaled to 1500 μV for each channel.

LFP analysis

For each channel the trial data were decomposed into their time-frequency characteristics by pointwise multiplying the fast-Fourier-transformed data with the fast-Fourier-transformed complex Morlet wavelets ($e^{i2\pi tf}e^{-t^2/2s^2}$), where t is time, f is frequency (ranging from 1 to 40 in 40 logarithmically spaced steps), and s the width of the Gaussian for each frequency ($n/2\pi f$) with n reflecting the number of cycles (ranging from 3 to 8 in 40 logarithmically spaced steps). After this multiplication, the inverse fast Fourier transform was taken to obtain a complex signal from which power for each frequency at each time point could be extracted by squaring the complex signal. The result of this was then decibel power normalized ($10log10(power/baseline)$) for each frequency separately, where the baseline was picked as -500 to -200 ms before cue onset.

To test for statistically significant power changes in the MFC related to response conflict, single-trial power values for each channel, frequency and time point were extracted for permutation-based statistics. At each of the 1000 permutations, the condition label was randomly swapped on a trial-by-trial basis. The difference map between congruent and incongruent trials was computed, after which values in the difference map were transformed to z-scores. If z-scores did not exceed the critical z-value for significance, they were set to zero. After this, the biggest cluster size was stored based on which the null hypothesis distribution was eventually created. Then, the observed clusters were transformed to z-scores and set to zero if they did not exceed the cluster threshold that was based on the null hypothesis distribution of biggest cluster sizes with $p = .05$.

Results

Behavioural results

Three t-tests for independent samples were used for each rat with condition (congruent, incongruent) as independent variable to test whether the congruency affected reaction time, movement time, and error rate. Because electrophysiological analysis focusses on one rat, this paragraph will also focus on this rat (rat 1, red-coloured in Fig. 4 and Table 1). The statistics of the other three rats are reported in Table 1, showing effects in line with the rat reported here, or with even bigger effect sizes. Following behavioural data in human Simon tasks, RTs were slower on incongruent trials compared to congruent trials ($t(5462) = 3.34$, $p < .001$, $d = 0.09$). However, MTs did not differ depending on the congruency ($t(5347) = 1.27$, $p = .203$, $d = 0.03$). In line with the conflict-effect on RTs, the rat also made more errors on incongruent trials compared to congruent trials ($t(6041) = 2.60$, $p = .009$, $d = 0.07$).

As can be seen in Figure 4, there is a tendency for rats to have slower RTs and longer MTs if the cue is presented in the auditory domain as compared to the visual domain. However, claims cannot be made about this due to a low sample size, as it was not the primary goal to test for the effect of stimulation domain on RT. Another remarkable aspect is a low conflict effect for rat 1, which is most visible on the error rate. It is suspected that this is due to the larger amount of Simon task sessions this rat has done as compared to the other three rats, enabling this rat to be more trained to deal with response conflict. Overall, it can be concluded that the rats showed the Simon effect.

Task-related power changes

Evaluation of power dynamics averaged over conditions shows that two major task-related components were elicited, as can be seen in the upper plots in Figure 5. Around 300 ms after cue onset there was an increase in alpha power (~7-14 Hz), as well as an increase in high-delta power around 500 ms after cue onset (~2.5-4.5 Hz). Post-hoc data driven t-tests were used to identify whether these alpha and delta power increases were affected by conflict, by averaging the power values within each window for each trial, for each channel separately. The alpha window was defined between 250-900 ms and 7-14 Hz, whereas the delta window was defined between 350-700 ms and 2.5-4.5 Hz (see contours in upper plots of Fig. 5). None of the channels captured significant differences in the average alpha or delta power between congruent and incongruent trials (alpha: $t(175) = -0.05 - 1.06$, $p = .292 - .977$, $d = -0.01 - 0.16$; delta: $t(175) = 0.10 - 0.70$, $p = .482 - .922$, $d = 0.01 - 0.11$).

Conflict-related power changes

As illustrated with black-coloured contours in the bottom plots of Figure 5, cluster-based permutation testing revealed several windows that were statistically significant ($p < .05$, uncorrected for multiple comparisons) between congruent and incongruent trials. As expected, there was a conflict-

related increase in theta power (5-8 Hz) around 250 ms after cue onset, which is before the average reaction time (congruent trials: 597 ms, incongruent trials: 667 ms). An even stronger conflict-related power increase was observed in the beta-band (12-30 Hz) around 750 ms following cue onset, which is shortly after the average reaction time. Although all contoured windows did not survive correction for multiple comparisons, it is interesting to see how conflict-related midfrontal theta and beta are distributed over the channels in the MFC. Therefore, the average power within the conflict-related theta and beta contour was computed for each channel separately. As can be seen in Figure 6, visual inspection shows there was no smooth pattern transition in power across the channel layout for both conflict-related theta and beta, except that there was more theta power in frontal channels, and more beta power in posterior channels. A post-hoc t-test indeed showed that the 15 frontal channels displayed significantly more theta power as compared to the 15 posterior channels ($t(28) = 3.63$, $p = .001$, $d = 1.32$). Similarly, but in the opposite direction, there was more beta power in the 15 posterior channels as compared to the 15 frontal channels ($t(28) = -2.13$, $p = .042$, $d = -0.78$). A Spearman's correlation showed that conflict-related theta and beta power did not correlate with each other ($r = -.05$, $p = .784$).

Total power versus non-phase-locked power

Subtracting the phase-locked signal (ERP, event-related potential) from the time-series data for both conditions separately did not result in clearly visible changes in the time-frequency domain, as illustrated with the left versus right column of Figure 5. This indicated that the majority of task-related and conflict-related power changes were due to amplitude modulations of ongoing endogenous oscillations, not being time-locked to cue onset. As there was no clear difference in total power versus non-phase-locked power, all reported analyses were performed with the total power series data.

Relation between theta power and reaction time

To investigate the possible relationship between theta power and reaction time during congruent and incongruent trials, theta power between 5 and 8 Hz was extracted for each channel separately from 100 to 350 ms after cue onset. This window was chosen based on the contoured conflict-related theta window (see lower plots in Fig. 5). Power values within this time window were averaged for each trial and correlated with the RT in that same trial. Although 29 out of 30 MFC channels showed the expected positive correlation during incongruent trials, Spearman's correlation was not significant for any of the 30 channels in the probe ($r = -0.01 - 0.11$, $p = .288 - .967$). Contrary to incongruent trials, 28 out of 30 MFC channels showed negative correlations during congruent trials, but none of the Spearman's correlations were significant ($r = -0.14 - 0.01$, $p = .226 - .955$).

Discussion

Towards a rodent model for studying neural circuits of response conflict

Table 1.

Statistics on reaction time, movement time and error rate for each individual rat. Every rat colour in the left column corresponds to the colours in Figure 4. Note that the degrees of freedom reported here are much larger than the degrees of freedom reported in electrophysiological analyses, as only a randomly chosen subset of sessions from one rat was used for electrophysiological analyses due to the time-consuming pre-processing steps.

|	Reaction time	Movement time	Error rate
Rat 1 sound cues	$t(5462) = 3.34$, $p < .001$, $d = 0.09$	$t(5347) = 1.27$, $p = .203$, $d = 0.03$	$t(6041) = 2.60$, $p = .009$, $d = 0.07$
Rat 2 sound cues	$t(1144) = 3.48$, $p < .001$, $d = 0.21$	$t(1101) = 2.30$, $p = .022$, $d = 0.14$	$t(1334) = 6.80$, $p < .001$, $d = 0.37$
Rat 3 light cues	$t(2007) = 8.83$, $p < .001$, $d = 0.39$	$t(2014) = 2.35$, $p = .019$, $d = 0.11$	$t(2291) = 8.47$, $p < .001$, $d = 0.35$
Rat 4 light cues	$t(2556) = 16.52$, $p < .001$, $d = 0.65$	$t(2581) = 0.04$, $p = .970$, $d = 0.00$	$t(2943) = 6.84$, $p < .001$, $d = 0.25$
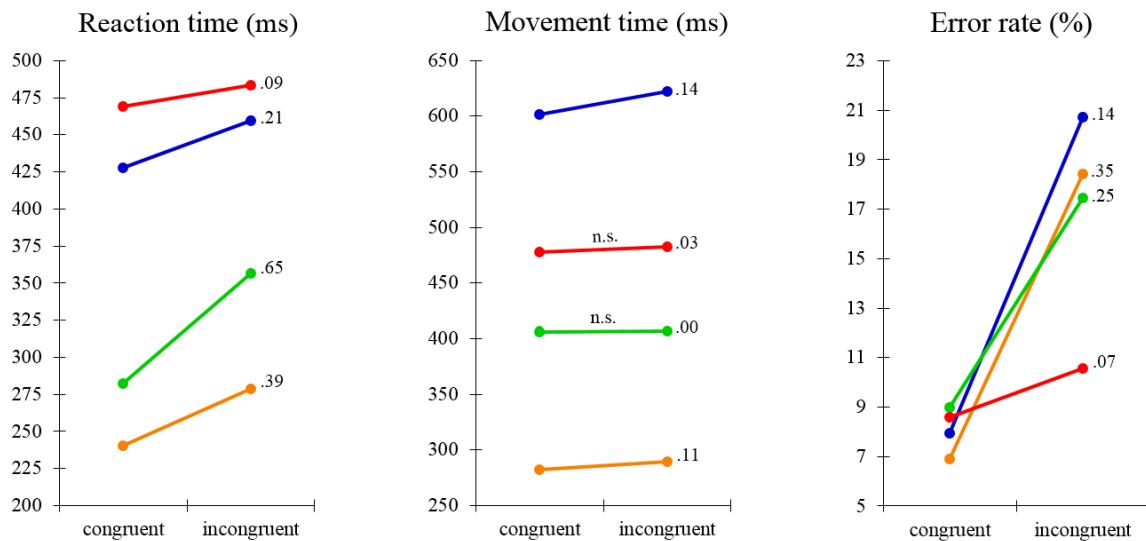
Figure 4. Mean reaction time, movement time and error rate depicted for congruent and incongruent trials, for each individual rat in a separate colour (rats coloured red and blue received auditory cues, rats coloured green and yellow received visual cues). The numbers next to each line represent effect sizes (Cohen's d). Non-significance is indicated by n.s. on top of the line. The cut-off for significance was set at p = 0.05.
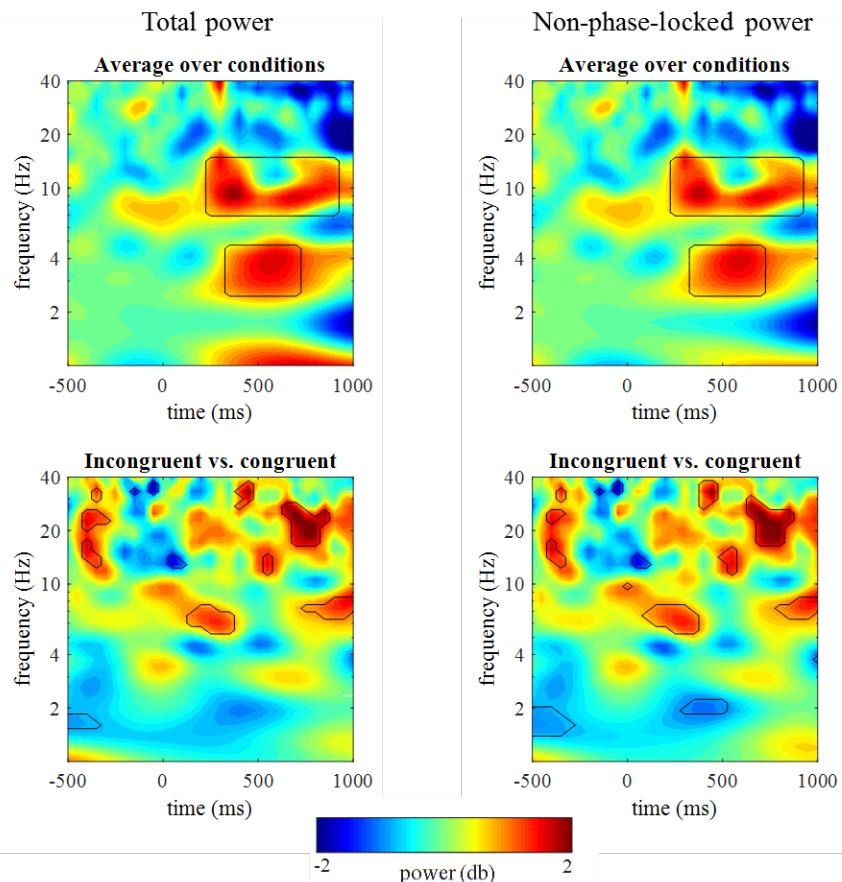
Figure 5. Time-frequency plots for the total power (left) and non-phase-locked power (right) relative to cue onset. The upper two plots show the average task-related power over both congruent and incongruent trials, whereas the bottom two plots show conflict-related power changes. Contours in the top two plots surround the windows of interest for post-hoc analysis. Contours in the bottom two plots indicate significant (p < .05) windows of conflict-related power changes. None of these contours survived correction for multiple comparisons. Note that y-axes are logarithmically scaled. Data is from a typical MFC channel.
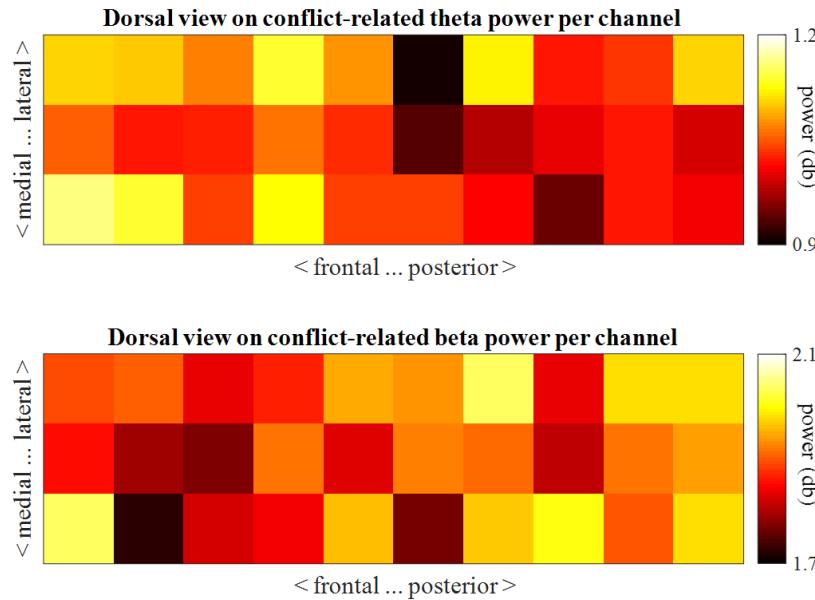
Figure 6. Dorsal view on the 3x10 channel layout of the MFC probe with conflict-related theta (top) and beta (bottom) power shown for each channel separately. Note the difference in scaling of power values between the top and bottom plot.

Previous studies in humans on response conflict lacked anatomical precision, due to the use of extracranial electrophysiology. On the other hand, the animal studies that were already executed by Courtière et al. (2007) and Marx et al. (2012) did not provide the temporal precision needed to study the neural circuits underlying response conflict. The current study aimed to validate an intracranial electrophysiological rodent model for response conflict, to enable researchers in the future to study the mechanisms in the physical substrate with both anatomical and temporal precision. Comparable with human response conflict, this study shows that the Simon effect in rats manifests in both the visual and auditory domain, and that there is an increase in theta power during conflict in the MFC. This rodent model opens up possibilities to study response conflict in more detail, and specifically, study the role of ongoing electrophysiological processes within and between different brain regions during response conflict and resolution.

Conflict-related theta activity in the midfrontal cortex

The contrast between congruent and incongruent trials revealed several changes in power, of which one frequency band was most prominent before the average reaction time. Comparable with human response conflict, conflict-related theta power increases after cue onset, is predominantly non-phase-locked, and occurs before the average reaction time. This theta power increase is relatively stronger in the anterior half of the MFC probe, suggesting that this power change has its origin in more frontal regions in the MFC. Although histology should give a definite answer about the exact coordinates of the channels, and thus the brain region in which the theta increase is most prominent, it might well be that the frontal channels of the MFC probe were located close to or in the prelimbic cortex (see left panel in Fig. 2).

The prelimbic cortex in rats is thought to be associated with response selection, response planning, reward anticipation, and working memory (Emmons, Ruggiero, Kelley, Parker, & Narayanan, 2016; Seamans, Floresco, & Phillips, 1995; Seamans, Lapish, & Durstewitz, 2008; Vertes, 2004). Lesions to the prelimbic cortex result in impaired performance on tasks in which multiple responses are optional (Delatour & Gisquet-Verrier, 1996). Given the projections that originate in the prelimbic cortex, this region is comparable with the dorsolateral prefrontal cortex (DLPFC) in humans (Seamans et al., 2008; Vertes, 2004). In human studies, the (D)LPFC is found to be associated with conflict resolution, potentially through attentional biasing or action inhibition, by theta-coupling between the MFC and LPFC (Cohen & Ridderinkhof, 2013; Cavanagh, Cohen, & Allen, 2009; Kim, Chung, & Kim, 2013; Swart et al., 2018). Because it is very plausible that theta in the current study originates

from the prelimbic cortex, these lines of evidence suggest that conflict-related theta increases in the current study are homologous to the conflict-related theta increases found in human studies.

Activity in other frequency bands

Beta power increased during conflict after the average RT. Contrary to conflict-related theta power, this beta power is relatively stronger in the posterior half of the MFC probe. These posterior channels were probably located near to or in the cingulate cortex (see left panel in Fig. 2). The cingulate cortex is associated with a multitude of cognitive functions, such as selective attention, response flexibility, and effort- and reward-guided decision making (Cowen, Davis, & Nitz, 2012; Kim, Wasserman, Castro, & Freeman, 2016; Porter, Hillman, & Bilkey, 2019; Seamans et al., 1995). Inactivation of the cingulate cortex with muscimol impairs selective attention during the presence of a distractor stimulus (Kim et al., 2016), and the majority of neurons in the anterior part of the cingulate cortex code for specific amounts of effort (Porter et al., 2019). However, there is a lack of studies describing beta power fluctuations in this region during cognitive tasks in rats, which makes it difficult to grasp which of these potential processes it might reflect. Another possibility is that this beta power reflects a response inhibition process through synchronisation in motor-related areas. The striatum is located very close to the posterior half of the MFC probe (< 2 mm), and it is well known that response inhibition processes are accompanied by striatal beta synchronisation, which eventually suppresses motor cortex output (Aron, Herz, Brown, Forstmann, & Zaghloul, 2016; Eagle & Robbins, 2003; Wessel et al., 2016; Zavala et al., 2018). Although local field potentials are thought to reflect the sum of current fluctuations from a very small area surrounding the probe, volume conduction still enables probes to pick up signals from neighbouring tissue up to 6 mm away (Kajikawa & Schroeder, 2011). It is not unlikely that the increase in conflict-related beta power reflects the inhibition of an initiated movement towards the incorrect nose poke hole, so the correct movement direction can become dominant in order to obtain a reward.

Across conditions increases in alpha and high-delta power were found, both appearing after cue onset. Increases in frontal alpha and delta power during cognitive tasks in humans and non-human primates are associated with a variety of cognitive functions such as attention, working memory, inhibition, and mental calculation (Bonnefond &

Jensen, 2012; Harmony, 2013; Jensen & Bonnefond, 2013). However, translating primate-based findings to rodent electrophysiology might be inaccurate to do. Concerning rat studies, some researchers argue for the role of the MFC in auditory detection. A study by Herzog, Salehi, Bohon, and Wiest (2014) showed increased alpha activity after auditory cue onset, and another study found spiking activity related to the detection of auditory targets in the MFC, but not related to movement direction (Bohon & Wiest, 2014). In addition to increased alpha power after auditory cues, it has been previously found that auditory cues are followed by increases in midfrontal delta power if the cue is followed by a reward after a delay period, suggesting that post-cue delta power reflects cue-driven reward anticipation (Emmons et al., 2016). It might very well be that the increase in high-delta in the current study reflects something similar, as the rat could always expect a reward after moving to the correct lateral nose poke hole. Indepth investigation of the potential relationship between post-cue alpha/high-delta power and behaviour is needed to elucidate whether it plays a role in response conflict.

Limitations

One of the benefits of using intracranial electrophysiology during task execution is to investigate the relationship between trial-by-trial variations in power dynamics and behavioural measures. However, the current study failed to find a relationship between these two measures, as theta power and RT did not correlate significantly in both congruent and incongruent trials. Interestingly, during incongruent trials almost all channels showed the expected direction, namely higher theta power during trials with slower RTs. On the other hand, during congruent trials almost all channels showed the opposite direction, namely higher theta power during trials with faster reaction times. In human response conflict, Cohen and Donner (2013) found a positive correlation in both congruent and incongruent trials, which raises the question why the direction of effect in the current study is not identical across congruent and incongruent trials. It might be the case that midfrontal theta during congruent trials does not reflect conflict, but something like an attentional or motivational process. However, caution should be taken about the cause of this heterogeneity to prevent over-interpretation of these non-significant correlations.

Another limitation of the current study is the amount of data included for electrophysiological

analysis. Only a total of 244 trials were included, leaving 177 trials after pre-processing, which might have caused underpowered statistical analyses. However, this study demonstrates the richness of data collected with intracranial electrophysiology during a behavioural task, as it enables a wide range of analyses to be performed to better understand the underpinnings of response conflict and cognitive control in general.

Conclusions and future directions

In conclusion, this study successfully shows that this rodent model is well suited for studying the electrophysiological mechanisms underlying response conflict. Typical behavioural manifestations of response conflict in rats are evident in both the auditory and visual domain, showing increased reaction times and more errors during conflict. In addition, intracranial electrophysiology shows that theta power increases during conflict, originates from anterior parts of the MFC, and occurs after cue onset.

To better understand the circuitry underlying response conflict, future analyses could focus on functional connectivity between the MFC and other brain regions during conflict processing. Besides this, investigating the anatomical connections between them will shed light on the potential flow of information from stimulus to response. When these circuits are better known, interference with optogenetics can provide causal evidence about the dynamics that drive conflict processing and resolution. A point of attention in this whole process is to study response conflict in multiple sensory domains, to elucidate which components are specific to visual or auditory conflict, and which components are shared and therefore more fundamental in response conflict.

References

Aron, A. R., Herz, D. M., Brown, P., Forstmann, B. U., & Zaghloul, K. (2016). Frontosubthalamic circuits for control of action and cognition. *The Journal of Neuroscience, 36,* 11489–11495.

Bohon, K. S., & Wiest, M. C. (2014). Role of medio-dorsal frontal and posterior parietal neurons during auditory detection performance in rats. *PloS ONE, 9,* 1–12.

Bonnefond, M., & Jensen, O. (2012). Alpha oscillations serve to protect working memory maintenance against anticipated distracters. *Current Biology, 22,* 1969–1974.

Cavanagh, J. F., Cohen, M. X., & Allen, J. J. B. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *The Journal of Neuroscience, 29,* 98–

105.

Cohen, M. X. (2014). A neural microcircuit for cognitive conflict detection and signaling. *Trends in Neurosciences, 37,* 480–490.

Cohen, M. X., & Donner, T. H. (2013). Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior. *Journal of Neurophysiology, 110,* 2752–2763.

Cohen, M. X., & Ridderinkhof, K. R. (2013). EEG source reconstruction reveals frontal-parietal dynamics of spatial conflict processing. *PloS ONE, 8,* 1–14.

Courtière, A., Hardouin, J., Burle, B., Vidal, F., & Hasbroucq, T. (2007). Simon effect in the rat: A new model for studying the neural bases of the dual-route architecture. *Behavioural Brain Research, 179,* 69–75.

Cowen, S. L., Davis, G. A., & Nitz, D. A. (2012). Anterior cingulate neurons in the rat map anticipated effort and reward to their associated action sequences. *Journal of Neurophysiology, 107,* 2393–2407.

Delatour, B., & Gisquet-Verrier, P. (1996). Prelimbic cortex specific lesions disrupt delayed-variable response tasks in the rat. *Behavioral Neuroscience, 110,* 1282–1298.

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

Eagle, D. M., & Robbins, T. W. (2003). Inhibitory control in rats performing a stop-signal reaction-time task: Effects of lesions of the medial striatum and d-amphetamine. *Behavioral Neuroscience, 117,* 1302–1317.

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective, & Behavioral Neuroscience, 7,* 380–390.

Emmons, E. B., Ruggiero, R. N., Kelley, R. M., Parker, K. L., & Narayanan, N. S. (2016). Corticostriatal field potentials are modulated at delta and theta frequencies during interval-timing task in rodents. *Frontiers in Psychology, 7,* 1–9.

Felsen, G., & Mainen, Z. F. (2008). Neural substrates of sensory-guided locomotor decisions in the rat superior colliculus. *Neuron, 60,* 137–148.

Harmony, T. (2013). The functional significance of delta oscillations in cognitive processing. *Frontiers in Integrative Neuroscience, 7,* 1–10.

Herzog, L., Salehi, K., Bohon, S., & Wiest, M. C. (2014). Prestimulus frontal-parietal coherence predicts auditory detection performance. *Journal of Neurophysiology, 111,* 1986–2000.

Ignashchenkova, A., Dicke, P. W., Haarmeier, T., & Their, P. (2004). Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. *Nature Neuroscience, 7,* 56–64.

Jensen, O., & Bonnefond, M. (2013). Prefrontal alpha- and beta-band oscillations are involved in rule selection. *Trends in Cognitive Sciences, 17,* 10–12.

Kajiwaka, Y., & Schroeder, C. E. (2011). How local is the local field potential? *Neuron, 72,* 847–858.

Kim, C., Chung, C., & Kim, J. (2013). Task-dependent response conflict monitoring and cognitive control in anterior cingulate and dorsolateral prefrontal cortices. *Brain Research, 1537,* 216–223.

Kim, J., Wasserman, E. A., Castro, L., & Freeman, J. H. (2016). Anterior cingulate cortex inactivation impairs rodent visual selective attention and prospective memory. *Behavioral Neuroscience, 130,* 75–90.

Krauzlis, R. J., Liston, D., & Carello, C. D. (2004). Target selection and the superior colliculus: Goals, choices and hypotheses. *Vision Research, 44,* 1445–1451.

Marx, C., Lex, B., Calaminus, C., Hauber, W., Backes, H., Neumaier, B., ... Endepols, H. (2012). Conflict processing in the rat brain: Behavioral analysis and functional μPET imaging using [^{18}F] fluorodeoxyglucose. *Frontiers in Behavioral Neuroscience, 6,* 1–12.

Nigbur, R., Ivanova, G., & Stürmer, B. (2011). Theta power as a marker for cognitive interference. *Clinical Neurophysiology, 122,* 2185–2194.

Porter, B. S., Hillman, K. L., & Bilkey, D. K. (2019). Anterior cingulate cortex encoding of effortful behavior. *Journal of Neurophysiology, 121,* 701–714.

Posner, M. I., Walker, J. A., Friedrich, F. J., & Rafal, R. D. (1984). Effects of parietal injury on covert orienting of attention. *The Journal of Neuroscience, 4,* 1863–1874.

Seamans, J. K., Floresco, S. B., & Philips, A. G. (1995). Functional differences between the prelimbic and anterior cingulate regions of the rat prefrontal cortex. *Behavioral Neuroscience, 109,* 1063–1073.

Seamans, J. K., Lapish, C. C., & Durstewitz, D. (2008). Comparing the prefrontal cortex of rats and primates: Insights from electrophysiology. *Neurotoxicity Research, 14,* 249–262.

Simon, J. R., Hinrichs, J. V., & Craft, J. L. (1970). Auditory S-R compatibility: Reaction time as a function of ear-hand correspondence and ear-response-location correspondence. *Journal of Experimental Psychology, 86,* 97–102.

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology, 51,* 300–304.

Stürmer, B., Leuthold, H., Soetens, E., Schröter, H., & Sommer, W. (2002). Control over location-based response activation in the Simon task: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1345–1363.

Swart, J. C., Frank, M. J., Määttä, J. I., Jensen, O., Cools, R., & den Ouden, H. E. M. (2018). Frontal network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in motivated action. *PloS Biology, 16,* 1–25.

Vertes, R. P. (2004). Differential projections of the infralimbic and prelimbic cortex in the rat. *Synapse, 51,* 32–58.

Wessel, J. R., Ghahremani, A., Udupa, K., Saha, U., Kalia, S. K., Hodaie, M., ... Chen, R. (2016). Stop-related subthalamic beta activity indexes global motor suppression in Parkinson's disease. *Movement Disorders, 31,* 1846–1853.

Whitlock, J. R. (2017). Posterior parietal cortex. *Current Biology, 27,* 681–701.

Yang, F. C., Jacobsen, T. K., & Burwell, R. D. (2017). Single neuron activity and theta modulation in the posterior parietal cortex in a visuospatial attention task. *Hippocampus, 27,* 263–273.

Zavala, B., Jang, A., Trotta, M., Lungu, C. I., Brown, P., & Zaghloul, K. A. (2018). Cognitive control involves theta power within trials and beta power across trials in the prefrontal-subthalamic network. *Brain, 141,* 3361–3376.

The Role of Ventral Fibre Pathway in Language Production in Health and Disease

Margot Mangnus[1]
Supervisors: Ardi Roelofs[1], Nikki Janssen[1]

[1]*Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands*

While neuroimaging research on language production has traditionally focused primarily on grey matter, several recent studies highlight the involvement of ventral and dorsal white matter pathways. The exact functional role of these pathways is a debated issue. The ventral pathway has been suggested to underlie top-down control in language production, but the functional roles of each specific white matter tract within this pathway, like the inferior fronto-occipital fasciculus (IFOF), the inferior longitudinal fasciculus (ILF), and the uncinate fasciculus (UF), have not yet been elucidated. To investigate the involvement of these tracts in top-down control, 16 patients with primary progressive aphasia (PPA), an acquired language deficit due to neurodegenerative disease, and 22 age-matched healthy controls performed a picture-word interference task. The task elicited an interference effect, which served as a behavioural measure of the participants' top-down interference control. Furthermore, the microstructural integrity of the IFOF, the ILF, and the UF was calculated as a neuroanatomical measure with diffusion tensor magnetic resonance imaging (MRI) and tractography, based on Fractional Anisotropy (FA) and Mean Diffusivity (MD) values. A linear mixed-effects model revealed that patients were more susceptible to the interference effect in reaction time compared to controls. Moreover, the integrity of all three tracts was altered in patients compared to controls. Importantly, the integrity of the IFOF was not associated with the interference effect in reaction times or accuracy. The integrity of the ILF was associated with the interference effect in reaction times, but not accuracy. The integrity of the UF was associated with the interference effect in both reaction times and accuracy. These results indicate that patients with PPA manifest impaired top-down control processes in language production and that these processes are mediated by the microstructural properties of the inferior longitudinal fasciculus and uncinate fasciculus in both patients with PPA and healthy individuals.

Keywords: language production, ventral pathway, top-down control, tractography, primary progressive aphasia

Corresponding author: Margot Mangnus, E-mail: mangnusm@gmail.com

In its early beginnings, research on the neuroanatomy of language was restricted to one specific region, Broca's area, which was thought to be the centre of language production (Broca, 1861). It was not long until another area involved in language processing, located in the posterior part of the superior temporal gyrus, was discovered and found to be necessary for speech comprehension (Wernicke, 1874). Wernicke suggested that these areas are directly connected and damage to this connection would lead to a specific speech deficit, which he called conduction aphasia. Post-mortem dissections by Dejerine and Dejerine-Klumpke (1895) confirmed the existence of this connection and the fibre pathway was named arcuate fasciculus. This fibre pathway starts in the superior and middle temporal gyrus and arches around the Sylvian fissure to terminate in the posterior part of the inferior frontal gyrus (Catani, Jones, & ffytche, 2005; Geschwind, 1970; Glasser & Rilling, 2008). It is this tract that lies at the centre of the dorsal pathway of language processing. Historically, less attention has been paid in research to a ventral pathway of processing, which consists of a number of fibre tracts that run under the Sylvian fissure from posterior temporal to anterior temporal and frontal areas (Hickok & Poeppel, 2007; Rauschecker & Scott, 2009). The work reported in this thesis aimed to contribute to a better understanding of the role of these ventral fibre tracts in language, in particular, language production.

The dorsal-ventral pathway distinction originates from the visual processing domain, indicating two information processing streams that differ in anatomy (occipitoparietal vs. occipitotemporal), and function ('how', spatial relation vs. 'what', object recognition) (Ungerleider & Haxby, 1994). This analogy has since been extended to auditory processing (Kaas & Hackett, 1999; Rauschecker & Tian, 2000), and language in particular (Hickok & Poeppel, 2004; Hickok & Peoppel, 2007; Wise, 2003). In these theories, it is generally agreed upon that the dorsal pathway is responsible for mapping sound representations (e.g., phonemes or words) to motor representations (e.g., articulatory movements) and the ventral pathway for mapping sound to meaning (e.g., concepts) representations. Evidence supporting this dual-stream processing of language comes from a study in which participants performed an overt repetition task and a listening task (Saur et al., 2008). Results of this study showed that performance in the repetition task was associated with the dorsal stream and performance in the listening task with the ventral stream. Therefore,

the authors concluded that the dorsal stream was necessary for sublexical repetition of speech and the ventral stream for comprehension. More research on the distinction between the two pathways confirms the involvement of the dorsal pathway in repetition (Fridriksson et al., 2010) and the involvement of the ventral pathway in comprehension (Harvey, Wei, Ellmore, Hamilton, & Schnur, 2013; Scott, Blank, Rosen, & Wise, 2000). These were all studies on repetition of meaningless speech or speech comprehension. However, no agreement exists on the involvement of these pathways in production of meaningful speech.

This anatomical pathway of language production (i.e., mapping meaning onto speech output representations) has been an issue of debate. Analysis of data from both healthy participants as well as patients with aphasia has yielded two different views on the issue, namely the dorsal production view and the ventral production view. The dorsal production view entails that mapping meaning onto speech output happens primarily through the arcuate fasciculus. This view has been implemented in the WEAVER++/ARC computational model (Roelofs, 2014). This model suggests the ventral pathway to have an indirect role in language production, namely a role in top-down executive control. Importantly, the model successfully accounts for research from stroke patients where damage to the arcuate fasciculus, but not to the uncinate fasciculus or extreme capsule (components of the ventral pathway), predicts impairment in language production (Marchina et al., 2011; Wang, Marchina, Norton, Wan, & Schlaug, 2013). However, this view has potentially been challenged by a study from Wilson et al. (2011), who found no effects of damage of the arcuate fasciculus on picture naming in patients of different subtypes of primary progressive aphasia (PPA). This is a neurodegenerative disease that primarily affects language processing and will be discussed in more detail below, as this patient group is of direct relevance to the current study. In the ventral production view, it is assumed that mapping meaning onto speech output primarily happens through the ventral pathway. This pathway includes the white matter tracts of the inferior fronto-occipital fasciculus (IFOF), inferior longitudinal fasciculus (ILF), uncinate fasciculus (UF), extreme capsule (EmC) and the recently described middle longitudinal fasciculus (MdLF). The ventral production view has also been implemented computationally in the Lichtheim 2 model (Ueno, Saito, Rogers, & Lambon Ralph, 2011). This model successfully simulates several findings in the aphasia literature but seems

to be challenged by the finding that damage to the arcuate fasciculus, but not the uncinate fasciculus or extreme capsule, predicts impairment in language production (Marchina et al., 2011; Wang et al., 2013).

In order to measure if white matter tracts are involved in certain tasks or processes, diffusion tensor imaging (DTI) and tractography are often used. DTI is a magnetic resonance imaging (MRI) sequence that maps the diffusion of water molecules in the brain and tractography is used to visually represent 3D images of white matter tracts with a high spatial resolution (Alexander, Lee, Lazar, & Field, 2007). These techniques allow us to measure microstructural integrity of specific white matter tracts in vivo as a proxy for anatomical connectivity. The two most used indicators of microstructural integrity are fractional anisotropy (FA), the strength of diffusivity along a particular direction, and mean diffusivity (MD), the total amount of diffusion in a volume of brain tissue. Typically, higher FA and lower MD values indicate stronger connectivity. Consequently, lower FA and higher MD values could indicate a neurodegenerative disorder. Multiple patient studies using these techniques provide support for the dorsal/phonological and ventral/lexical-semantic distinction. Studies on white matter damage in the different subtypes of PPA found lower FA and higher MD in the UF and ILF in patients of the semantic variant. Interestingly, these patients showed very poor performance on a naming task compared to other subtypes and controls (Agosta et al., 2013; Galantucci et al., 2011). Two studies with stroke patients showed that the occurrence of semantic paraphasias in a picture naming task was strongly associated with loss of integrity in the left IFOF, UF and ILF (Han et al., 2013; McKinnon et al., 2018). A semantic paraphasia occurs when a participant names a word that is semantically related to the target (e.g., tiger instead of lion).

In intraoperative studies on awake patients undergoing brain surgery, direct electrical stimulation to a specific site can disrupt a particular cognitive function. This methodology has been used to investigate the regions involved in language production and has provided much knowledge on the involvement of the ventral pathway. Studies by Duffau and colleagues show that stimulation of the IFOF induced semantic paraphasias in picture naming. Importantly, stimulation of the UF, ILF and MdLF did not induce these naming disturbances (De Witt Hamer, Moritz-Gasser, Gatignol, & Duffau, 2011; Duffau et al., 2005; Duffau, Gatignol, Moritz-Gasser, & Mandonnet, 2009; Duffau, Herbet, &

Moritz-Gasser, 2013; Mandonnet, Nouet, Gatignol, Capelle, & Duffau, 2007). The authors propose that there is a direct route of semantic processing that consists of the IFOF and an indirect route that consists of the UF, ILF and possibly the MdLF.

In most of the previously mentioned research, a picture naming task was used to assess the language production performance of participants. However, naming a picture constitutes of many linguistic as well as non-linguistic processes. Therefore, it is difficult to determine whether the ventral pathway is involved in mapping meaning to motor representations, which relate to processes in speech, or in more domain-general control processing. From language comprehension studies, it is known that the ventral pathway is involved in semantic control (Harvey et al., 2013; Whitney, Kirk, O'Sullivan, Lambon Ralph, & Jefferies, 2011). However, a different task than simple picture naming is necessary to investigate this for language production. The picture-word interference task proves to be a valid way to study interference and control processing in language production (Glaser & Düngelhoff, 1984).

In the picture-word interference paradigm, participants are instructed to name a picture with superimposed letter strings that could induce different effects. For example, for a picture of a dog, the superimposed text could be neutral (e.g., a string of letters, XXX), semantically related (e.g., cat), unrelated (e.g., banana) or the word of the picture itself (dog). Experiments with this paradigm show that unrelated words vs. a string of letters (lexical interference) and related vs. unrelated words (semantic interference) induce slower reaction times. These mean reaction times in turn predict the inhibitory control of the participant (Shao, Roelofs, & Meyer, 2012). Multiple picture-word interference studies have shown the ventral pathway to be associated with top-down control. For instance, in an intraoperative picture-word interference experiment, stimulation of the ventral pathway induced semantic paraphasias, causing the participants to name semantically related words instead of the target (Ries et al., 2019). In a picture-word interference experiment with patients with lesions in the left prefrontal cortex, a lexical interference effect was found (Piai, Ries, & Swick, 2015). Furthermore, a semantic interference effect was found in the prefrontal cortex in an fMRI experiment (de Zubicaray, Wilson, McMahon, & Muthiah, 2001). Importantly, the prefrontal cortex is a termination site for the ventral pathway (Catani & Thiebaut de Schotten, 2008; Forkel et al., 2014; Sarubbo, De Benedictis, Maldonado, Basso, & Duffau, 2013). This could indicate that the ventral

pathway is involved in the top-down resolution of these interference effects.

As previously mentioned, the dorsal and ventral production views describe different views on which pathway is involved language production. Hence these views differ in the function of the ventral pathway in language production. In the dorsal view, it is only involved in top-down cognitive control and in the ventral view, it is involved in mapping meaning to speech output. It makes sense to hypothesize that top-down control processes are involved in language production, because word planning in language production requires attention (Roelofs & Piai, 2011). The fact that the lateral prefrontal cortex is associated with executive function and control (Petrides, 2005) makes a good case to consider that IFOF is involved in control processes in language, supporting the dorsal production view.

More research on the ventral pathway and its role in top-down control comes from non-linguistic pathologies that show a clear deficit in inhibition and interference control. For obsessive-compulsive disorder (OCD), lower FA values were consistently found in bilateral IFOF and also correlated with symptom severity and neuropsychological performance (Garibotto et al., 2010; Li et al., 2014; Peng et al., 2012). Interestingly, OCD patients show decreased response inhibition and Stroop task performance (Abramovitch, McCormack, Brunner, Johnson, & Wofford, 2019; Gruner & Pittenger, 2017; Shin, Lee, Kim, & Kwon, 2014). Moreover, schizophrenia patients show consistently lower FA in left IFOF and ILF, with severity of disorder being related to IFOF and disorganised factor to ILF integrity (Cheung et al., 2008; Epstein et al., 2014; Koch et al., 2010; Rigucci et al., 2013; Yao et al., 2013). In this disorder, a deficit in selective attention and inhibitory processes and an increased interference effect in the Stroop task are found as well (Boucart, Mobarek, Cuervo, & Danion, 1999; Henik & Salo, 2004; Westerhausen, Kompus, & Hugdahl, 2011). Also, for attention hyperactivity deficit disorder (ADHD), one study found lower FA in left ILF and higher MD in left IFOF, with attentional performance negatively related to MD in left ILF (Konrad et al., 2012). For adults with ADHD that were more inattentive, lower FA values were found in bilateral IFOF and left UF (Shaw et al., 2015). In a meta-analysis, ADHD patients of all age groups showed lower interference control than healthy controls (Lansbergen, Kenemans, & Van Engeland, 2007). This multidisciplinary research shows that the ventral pathway, and the IFOF in particular, is involved in top-down control. The question remains, however, if this is also the case in language production.

Comparing a patient group with language deficits and altered tract integrity to a healthy population enables the investigation of white matter tracts and their involvement in language production in health and disease. This comparison is possible with patients of primary progressive aphasia (PPA), as it is a neurodegenerative disease characterised by speech and language deficits and relatively spared other cognitive domains (Mesulam, 1982). This makes the comparison between healthy controls and patients with PPA of different subtypes suitable to investigate certain questions about language production. Moreover, the picture-word interference effect has not been investigated in patients with PPA. Therefore, the current study provides a novel approach by investigating top-down control in language production in patients with PPA. The three subtypes of PPA can be described by the differences in their main deficits. Logopenic PPA is characterised by deficits in word retrieval and sentence repetition, nonfluent (or agrammatic) PPA by agrammatism and effortful, slow speech, and semantic PPA (or semantic dementia) by deficits in naming and single-word comprehension (Gorno-Tempini et al., 2011). Interestingly, significant differences in white matter integrity have been found in patients with PPA and identified even in the different subtypes. Damage of the UF and ILF has been found to be different in patients with PPA and controls (D'Anna et al., 2016). Moreover, a study on the white matter damage in the subtypes of PPA shows that the behavioural symptoms converge with the areas in which tracts are damaged (Galantucci et al., 2011). In the nonfluent variant, changes in DTI metrics were mostly found in dorsal tracts. In the semantic variant, these changes were found in the UF and ILF. Lastly, for the logopenic variant, no difference in metrics of entire tracts was observed, but FA and diffusivity changes were found in the left temporoparietal region of the dorsal pathway. The possibility to measure differences in white matter structure in PPA and its subtypes combined with isolated language deficits make PPA a relevant disorder to measure language production and white matter changes.

Aim and approach

From the existing literature and research on language production and white matter pathways, several questions remain unresolved, which we aim to answer. First, the current study investigates

whether the ventral white matter pathway is involved in meaning-to-motor mapping or top-down control in language production. This will provide evidence for either the dorsal or the ventral production view. Secondly, a picture-word interference task performed by patients with PPA and age-matched healthy controls with a regular naming and interference condition will allow us to assess the amount of top-down control. With this task, we will be able to investigate if patients with PPA have damaged top-down interference control in language processing behaviourally, which is not known from the literature. Because of their language processing deficits, we expect patients with PPA to perform worse than controls. Lastly, research is inconclusive about which specific tracts in the ventral pathway could underlie language production. Intraoperative studies are effective in showing a causal relationship, but the methods have a low spatial sampling. Therefore, the present study will use DTI and probabilistic tractography to measure the microstructural differences in the three main tracts in the ventral pathway, namely the IFOF, ILF and UF. Finally, with this approach, we will be able to determine which specific tracts underlie top-down interference control in language processing in healthy individuals and patients with PPA. We expect to see a correlation between tract integrity and top-down control in the IFOF and UF but not ILF in both groups because of the literature on top-down control and the fact that these tracts connect frontal (control) and temporal (semantic) areas. This study will deepen the understanding of the neurobiology of language and control. The findings on patients with PPA, their interference control, and the involvement of the ventral pathway could also aid in the diagnosis of PPA.

Method

Participants

Sixteen patients with PPA (mean age = 70.0, eight females) were recruited from multiple hospitals in the Netherlands, including Radboud University Medical Centre (Nijmegen), Erasmus Medical Centre (Rotterdam) and Jeroen Bosch Hospital ('s-Hertogenbosch). Clinical diagnoses were established based on extensive assessment, including neuropsychological testing and neurological tests. These tests consisted of neuroimaging and, when available, liquor diagnostics. Among the patients with PPA were six of the logopenic variant, two of the

non-fluent variant, and eight of the semantic variant (semantic dementia). Twenty-two controls (mean age = 67.1, seven females) participated. Their age was not significantly different from that of the patients. All participants were right-handed and native Dutch speakers. All patients were tested with the approval of the local ethics committee (CMO Arnhem-Nijmegen, CMO 2016-2340, NL56842.091.16) and healthy controls were tested with the approval of the local ethics committee (CMO Arnhem-Nijmegen) under the general ethics approval ("Imaging Human Cognition", CMO 2014/288). All participants gave their written consent before participation.

Task and materials

A picture-word interference task was used to measure the participants' top-down control processes recruited to resolve the interference in the task. Twelve pictures divided over two semantic categories (animals and fruit) were used as stimuli. On each picture, words were superimposed as distractors. In the neutral condition, a number of X's was superimposed on the picture. In the incongruent condition, the distractor words were picture names from the same semantic category as the picture (e.g., a picture of a banana with the word kiwi printed on top). In Figure 1, an example trial is shown. All the distractor words were part of the response set of the experiment. Both conditions consisted of 60 items, with the incongruent condition including each picture with all other picture names of the same category as distractor words. Stimuli were randomised with the following constraints: stimulus cannot be shown twice in a row, distractor cannot be shown twice in a row, a maximum of four items of the same category can be shown in a row. Each participant received a different order of stimuli.

Procedure

Stimuli were presented in six blocks of twenty items using Presentation software (http://nbs.neurobs.com). The pictures appeared in the centre of a screen with a white background for 4000 ms. In-between trials, a blank screen was presented for 1000 ms. Responses were recorded for the purpose of determining accuracy and response time.

First, participants completed twelve practice trials, after which they were allowed to ask clarification questions about the experiment. The instruction for the experiment was to name the picture as fast as possible and ignore the superimposed word.
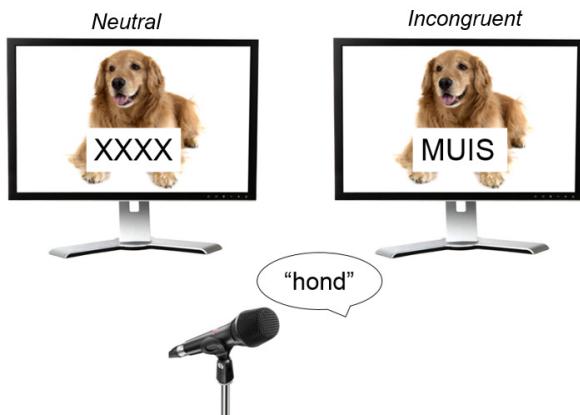
Figure 1. Example of a trial in the two conditions of the picture-word interference task.

Image acquisition

Structural and diffusion-weighted images were acquired in a single session from all participants using a Siemens Prisma Fit 3T scanner and a 32-channel head coil at the Donders Centre for Cognitive Neuroimaging, Nijmegen.

Diffusion weighted images were acquired with a simultaneous-multislice diffusion-weighted Echo Planar Imaging (EPI) sequence. Acquisition parameters were the following: multiband factor = 3; TR (repetition time) = 2282 ms; TE (echo time) = 71.2 ms; in-plane acceleration factor = 2; voxel size = 2×2×2 mm3; 9 unweighted scans; b-values = 1250 and 2500 s/mm2; 100 diffusion-encoding gradient directions in multiple shells for both b-values; Taq (total acquisition time) = 8 min 29 s.

Anatomical scans were acquired for accurate localization of anatomical structures in the tractography procedure, using the MP2Rage sequence (Marques et al., 2010). The following parameters were used: 176 slices; voxel size = 1×1×1 mm³; TR = 6000 ms; TE = 2.34 ms; Taq = 7 min 32 s.

DTI analysis

Preprocessing

Diffusion MR images were realigned and corrected for eddy-current induced distortion with SPM12 (Penny, Friston, Ashburner, Kiebel, & Nichols, 2006). Furthermore, brain extraction and correction for artefacts resulting from head/cardiac motion was performed with the PATCH algorithm (Zwiers, 2010). The diffusion tensor model was fitted at each voxel with weighted least-squares regression in DTIFIT from the FMRIB Diffusion Toolbox (FDT) to produce maps of DTI measures such as FA and MD, as well as the eigenvectors and eigenvalues.

Tractography

To model for crossing fibres, the distribution of diffusion parameters was estimated with a multi-shell generalised ball-and-sticks model at each voxel using BedpostX (Jbabdi, Sotiropoulos, Savio, Grana, & Behrens, 2012). We only included the tracts in the left hemisphere because of the strong lateralisation of the language system in the brain. For reconstructing the UF and ILF tracts, TRActs Constrained by UnderLying Anatomy (TRACULA) from Freesurfer was used, which is an automated global probabilistic tractography method (Yendiki et al., 2011). It uses prior anatomical information about the tracts derived from training subjects of which 18 white matter tracts were manually labelled. The output represents a probabilistic distribution of the reconstruction of each tract, of which we used the reconstruction of the UF and ILF. The tractography algorithm of TRACULA with its inclusion of prior anatomical information fits the research question at hand very well since we are looking for known tracts. The IFOF was not included in the tracts that are defined in TRACULA. Therefore, probabilistic tractography was performed with Probtrackx for this tract (Behrens et al., 2003; Behrens, Johansen-Berg, Jbabdi, Rushworth, &Woolrich, 2007). In this algorithm, the probability distribution of streamlines is estimated, starting from the seed mask and ending in a termination mask, not having been rejected by exclusion masks. All seeds were drawn in native diffusion space to maximise anatomical accuracy for each participant. For the IFOF, two seeds were drawn, one localised in the frontal and one in the occipital lobe. The frontal seed was defined as a rectangle over the entire left hemisphere on the coronal slice a few slices anteriorly to the corpus callosum. The occipital seed was defined as a rectangle over the left cerebrum on the coronal slice approximately five slices posteriorly to the curve of the arcuate fasciculus. Exclusion masks were added to constrain the fibre tracking to the relevant tract and to prevent the algorithm from including neighbouring tracts. After creating the masks, fibre tracking was performed with 10000 streamline samples initiated from all voxels within the seed mask, a curvature threshold of 0.2 and a step length of 0.5 mm. The output of fibre tracking, a probability distribution of the reconstruction of the tract, was then normalised by log transforming and dividing by the maximum value. This was done to control for the size of the seed mask. Tracking was

done in two directions, from seed to waypoint mask and the other way around. After that, the average of the two connectivity maps was calculated. This was done because the tracking procedure favours voxels close to the seed mask. The distributions were then thresholded to exclude spurious connections and noise. The connectivity maps can be seen in Figure 2.

For all three tracts, a mask was created from the connectivity map and used to derive the FA and MD from these images.

Statistical analysis

Analyses were performed on both reaction times (RT) and accuracy of the recorded responses. RTs were manually assessed using Praat software (Boersma & Weenink, 2019). For the RT analyses, the median RT of responses was used because very long RTs are expected for the patients with severe language impairment, and the median is more robust to outliers and a skewed distribution than the mean. Responses that were defined as errors included hesitations, semantic and phonological paraphasias, naming the distractor word and no response. Accurate responses where the target word was preceded by an article or when the RT was otherwise not reliable were not included in the RT analysis.

Statistical analyses were performed with R version 3.5.0 (R Core Team, 2016), using the lme4 package for its linear mixed-effects models (Bates et al., 2019).

To test for behavioural differences in RT and
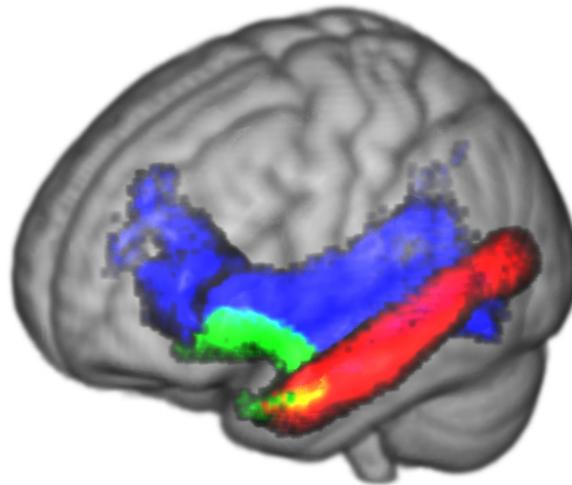
Figure 2. Tractography results of all three ventral tracts. The tracts represent the mean tracts of all participants. Green indicates the UF. Blue indicates the IFOF. Red indicates the ILF. The results are displayed on a 3D 2 mm3 template in the Montreal Neurological Institute standard space.

accuracy between patients and controls as well as between conditions, two linear mixed-effects models were used. For both models, group (controls, patients), condition (incongruent, neutral), and interaction between group and condition were fixed effects. A main effect of condition served as a measure of the interference effect. Random effects of condition by participant and a random intercept for item were also included.

Furthermore, the differences in FA and MD of all three tracts were analysed between controls and patients. FA is the most widely used measure in DTI research and is shown to be altered in neurodegenerative diseases (Galantucci et al., 2011). We felt it was important to include MD as well, as it reflects additional information about white matter properties over the information that merely FA can show (Alexander et al., 2011; Alexander, Lee, Lazar, & Field, 2007). These analyses were performed with Kruskal-Wallis tests, which do not assume normality. This test was chosen because of our small sample size and the fact that the dependent variables (FA and MD) are not normally distributed.

To assess whether the differences in RT and accuracy between conditions are influenced by the integrity of the three ventral tracts, linear mixed-effects models were used. However, FA and MD of each white matter tract turned out to be very collinear (Spearman's $\varrho < -0.6$ for all tracts). In many cases, the best way to control for high collinearity would be to simply remove one of the collinear variables from the analysis. Yet, we believe that the two measures provide important complementary information about the biological properties of the white matter. Therefore, we decided to perform a principal component analysis (PCA) using the psych package in R (Revelle, 2019). In a model with variables that may be correlated, a PCA is able to produce the underlying component that accounts for the largest amount of variability in the model. A PCA with two factors was performed on the FA and MD of each tract and the first component of the analysis, the one which explains the most variance, was used as a comprehensive measure of white matter integrity.

The PCAs showed that for all tracts two components explain 100% of the variance in FA and MD. Component 1 (C1) explained 83% in the IFOF, 78% in the UF, and 88% in the ILF and component 2 (C2) explained the other 17% in the IFOF, 22% in the UF, and 12% in the ILF. The loadings of the dependent variables for C1 were 0.91 for FA and -0.91 for MD in the IFOF, 0.88 for FA and -0.88 for MD in the UF, and 0.94 for FA and -0.94 for MD in

the ILF. For C2, the loadings were 0.41 for both FA and MD in the IFOF, 0.47 in the UF, and 0.35 in the ILF. Because of the negative relationship between FA and MD and the fact that C1 accounts for a very large part of the variance, this component was used for the following linear mixed-effects models for all tracts.

For the linear mixed-effects models, C1, condition and the interaction between C1 and condition were included as fixed effects. A random intercept for participant was also included.

To measure goodness of fit, we compared these models with an ANOVA to the same models without fixed effects. Lastly, additional linear-mixed effects models with the same variables and effects were used on a subset of the data, indexing the specific groups. This was done to investigate the effect of integrity of the tracts on the interference effect for the two groups separately.

Results

Interference effect

In the picture-word interference task, patients with PPA responded slower than controls ($\beta = 0.37$, SE = 7.4 x 10-2, t[36] = 4.97, p < .001, Fig. 3A). Moreover, all participants responded slower in the incongruent condition compared to the neutral condition (i.e., the interference effect) ($\beta = 0.14$, SE = 1.9 x 10-2, t[30] = 7.07, p < .001). There was also an interaction of group and condition ($\beta = 0.11$, SE

= 3.1 x 10-2, t[33] = 3.66, p < .001, Fig. 3B).

Regarding the accuracy rates, patients also made more errors than controls ($\beta = -3.58$, SE = 0.60, z = -5.94, p < .001, Fig. 3C). More errors were made in the incongruent condition than in the neutral condition for all participants ($\beta = -1.61$, SE = 0.53, z = -3.01, p = .003). However, there was no interaction of group and condition for accuracy rates ($\beta = 0.77$, SE = 0.54, z = 1.41, p = .16, Fig. 3D).

White matter integrity

Patients with PPA showed a highly decreased FA compared to controls for the UF ($\chi^2[1, N = 38] = 872.42$, p < .001), IFOF ($\chi^2[1, N = 38] = 371.10$, p < .001), and ILF ($\chi^2[1, N = 38] = 2327.40$, p < .001). Likewise, patients with PPA showed a highly increased MD compared to controls for the UF ($\chi^2[1, N = 38] = 872.80$, p < .001), IFOF ($\chi^2[1, N = 38] = 635.14$, p < .001), and ILF ($\chi^2[1, N = 38] = 1077.90$, p < .001). Figure 4 shows the differences between groups for all tracts.

White matter integrity and interference effect

IFOF

For the IFOF, the linear mixed-effects model did not show a main effect of tract integrity on RT ($\beta = -9.6$ x 10-2, SE = 5.0 x 10-2, t[32] = -1.90, p = .07). There was also no significant interaction between condition and tract integrity on RT ($\beta = -1.0$ x 10-2,
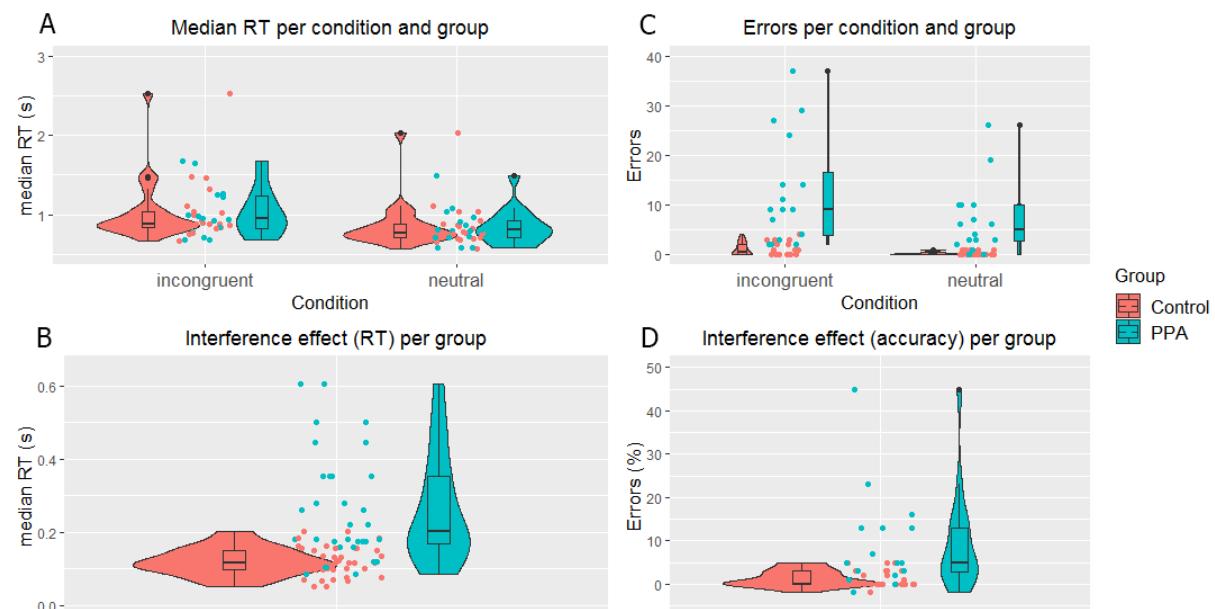

Figure 3. Behavioural results of the picture-word interference task. A. Median reaction times for each condition and group. B. Interference effect in reaction time for each group. C. Error rates for each condition and group. D. Interference effect in error percentage for each group.
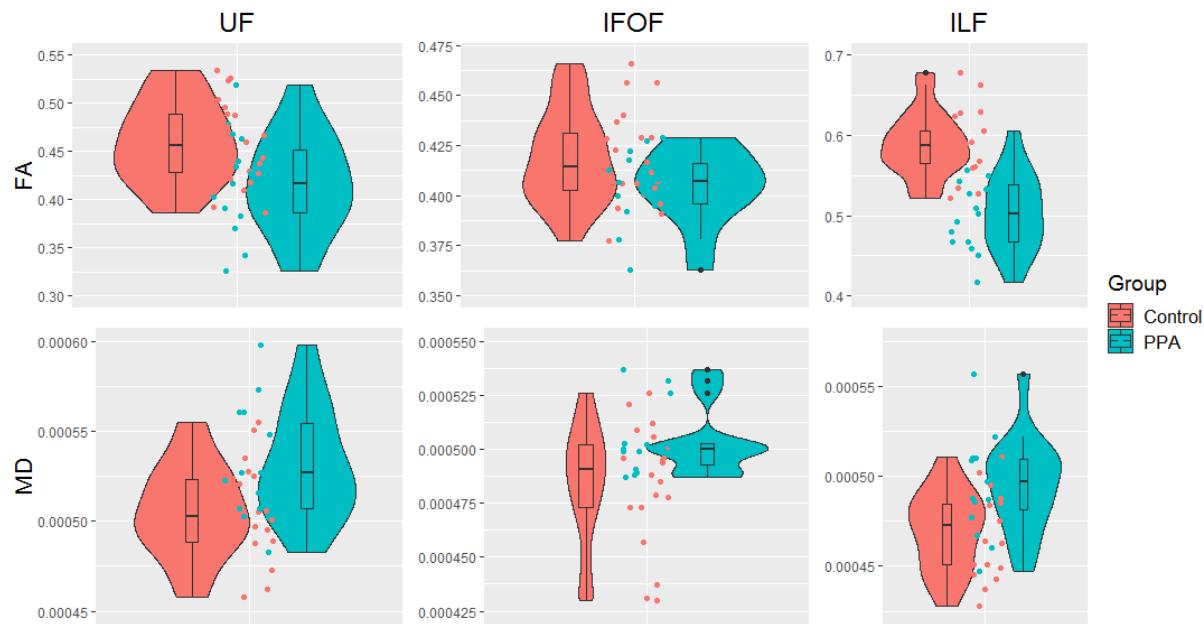
Figure 4. Differences in white matter integrity between patients of PPA and healthy controls. Top row: Differences in FA values. Bottom row: Differences in MD values. Left column: Differences in UF metrics. Middle column: Differences in IFOF metrics. Right column: Differences in ILF metrics.

SE = 8.0 x 10-3, t[3750] = -1.19, p = .23, Fig. 5AB). The generalised linear mixed-effects model also did not show a main effect of tract integrity on accuracy rates (β = 0.49, SE = 0.37, z = 1.35, p = .18). There was also no significant interaction between condition and tract integrity on accuracy rates (β = 0.11, SE = 0.19, z = 0.59, p = .56, Fig. 5CD). Figure 6 displays the relationships between tract integrity of all three tracts as measured by FA and MD and performance in the two conditions.

These results show that the integrity of the IFOF is not associated with performance in picture naming or the interference effect for both RT and accuracy.

UF

For the UF, the linear mixed-effects model did not show a main effect of tract integrity on RT (β = -9.6 x 10-2, SE = 5.3 x 10-2, t[33] = -1.81, p = .08). There was, however, a significant interaction between condition and tract integrity on RT (β = -2.0 x 10-2, SE = 9.0 x 10-3, t[3813] = -2.28, p = 2.3 x 10-2, Fig. 5AB). The generalised linear mixed-effects model showed a main effect of tract integrity on accuracy rates (β = 1.00, SE = 0.34, z = 2.91, p = .004). Moreover, there was a significant interaction between condition and tract integrity on accuracy rates (β = -0.31, SE = 0.15, z = -2.00, p = .046, Fig. 5CD).

So, the tract integrity of the UF is associated with performance in picture naming in accuracy rates, but not in RT. Furthermore, for both accuracy and

RT the integrity of the UF is associated with the interference effect.

ILF

For the ILF, the linear mixed-effects model showed a main effect of tract integrity on RT (β = -0.14, SE = 4.6 x 10-2, t[35] = -3.07, p = .004). There was also a significant interaction between condition and tract integrity on RT (β = -3.4 x 10-2, SE = 8.4 x 10-3, t[4039] = -4.09, p < .001, Fig. 5AB). The generalised linear mixed-effects model also showed a main effect of tract integrity on accuracy rates (β = 1.27, SE = 0.32, z = 3.94, p < .001). However, there was no significant interaction between condition and tract integrity on accuracy rates (β = -0.20, SE = 0.18, z = -1.10, p = .27, Fig. 5CD).

This means that for both RT and accuracy the integrity of the ILF was associated with performance in picture naming. Furthermore, for RT but not for accuracy the integrity of the ILF was associated with the interference effect.

Model fit

The ANOVAs for the linear mixed-effects models for RT (χ^2 [3, N = 2] = 407.74, p < .001) and accuracy (χ^2 [3, N = 2] = 56.93, p < .001) and IFOF integrity, RT (χ^2 [3, N = 2] = 395.31, p < .001) and accuracy (χ^2 [3, N = 2] = 56.05, p < .001) and UF integrity, and RT (χ^2 [3, N = 2] = 442.15, p < .001) and accuracy (χ^2 [3, N = 2] = 64.65, p < .001)
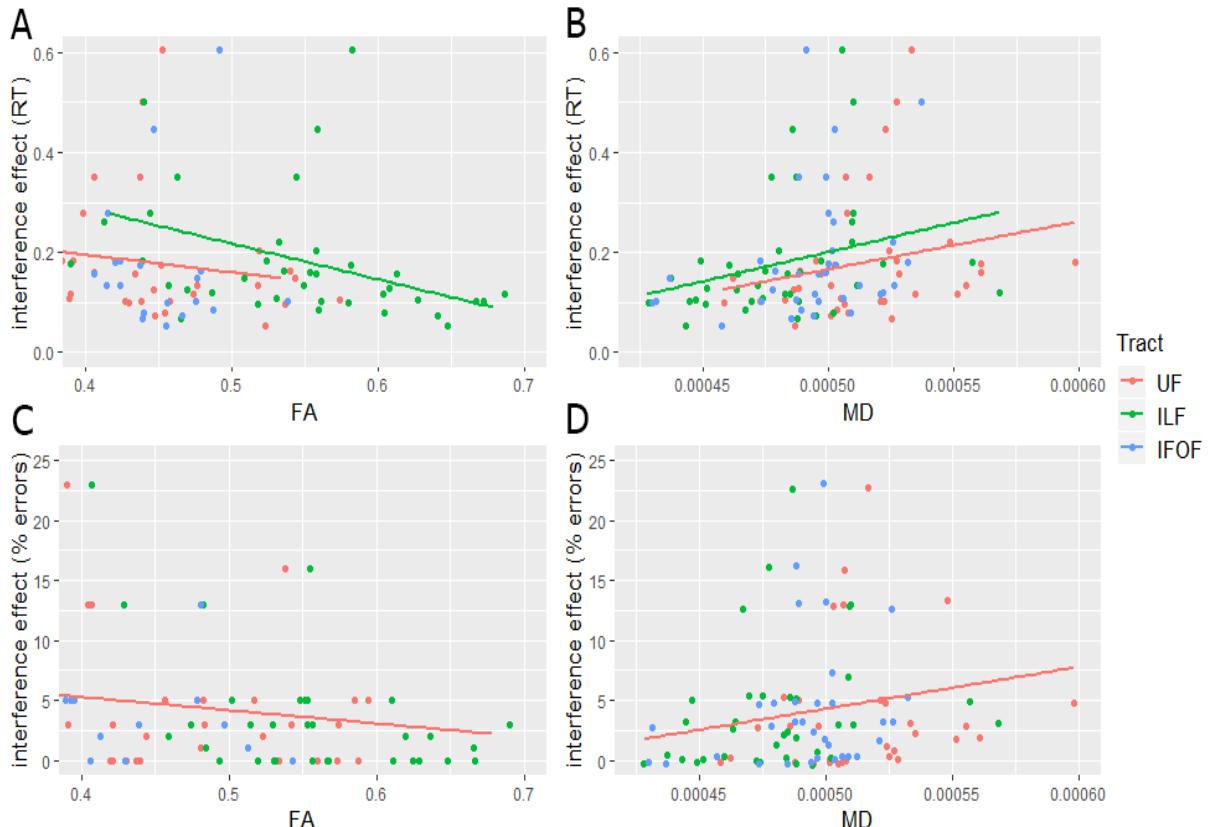
Figure 5. Relationship between tract integrity and the interference effect for all three ventral tracts. A. Relationship between fractional anisotropy (FA) and reaction time (RT). B. Relationship between mean diffusivity (MD) and reaction time (RT). C. Relationship between fractional anisotropy (FA) and error rates. D. Relationship between mean diffusivity (MD) and error rates. Regression lines show significant associations between tract integrity (as measured by the first component of the performed PCA) and the interference effect. Non-significant associations are not displayed as regression lines.
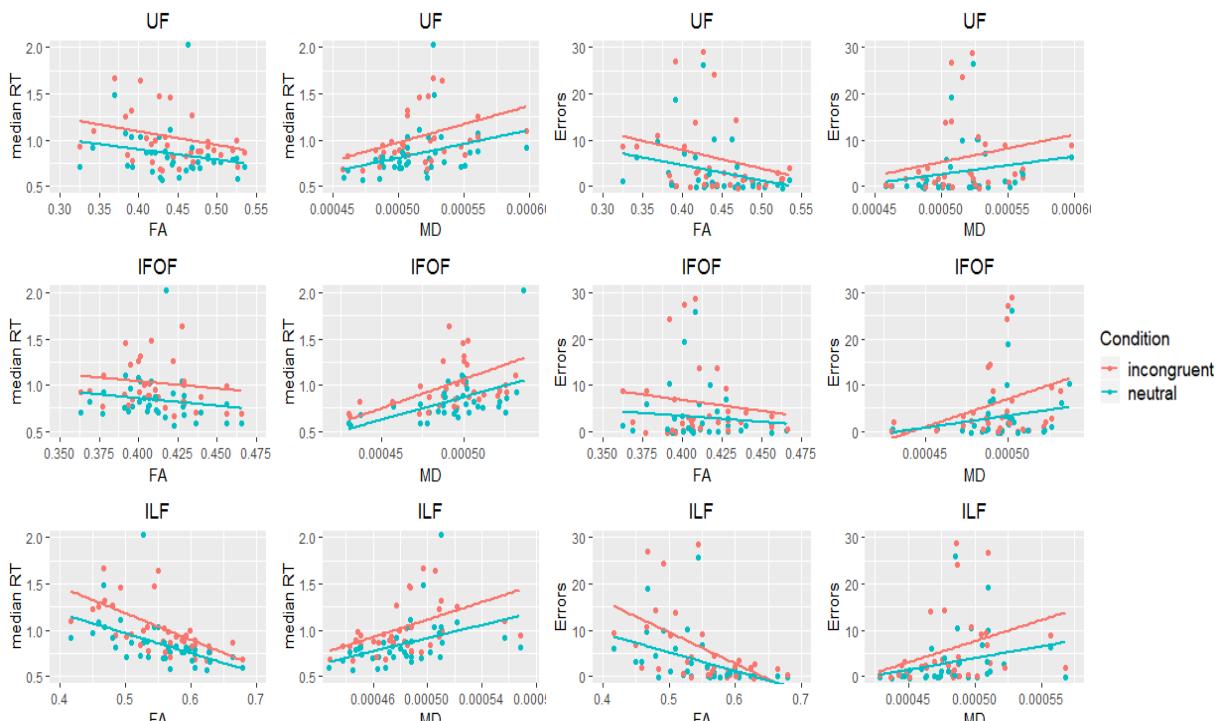
Figure 6. Relationships between tract integrity and performance in both conditions of the picture-word interference task. Left six panels: Results for median RT. Right six panels: Results for error rates.

and ILF integrity all showed that the fixed effects are justified and explain more of the variance in the data.

White matter integrity and interference effect within groups

From the separate analyses of the effect of tract integrity on the interference effect within the patient and control groups, almost all of them did not show an effect. The integrity of the IFOF in controls was related to picture naming in RT (β = -5.1 x 10-2, SE = 1.9 x 10-2, t[18] = -2.64, p = .02). Moreover, the integrity of the UF in controls was related to the interference effect in RT (β = 1.7 x 10-2, SE = 7.9 x 10-3, t[2580] = 2.21, p = .03). All other analyses did not show an effect.

Discussion

This study is the first to investigate the interference effect in a picture-word inference task in patients with PPA as a proxy for top-down control. Moreover, it is novel in linking this interference effect to the integrity of white matter tracts. The effect of tract integrity on top-down control was also assessed within the two groups, patients with PPA and healthy individuals.

Our results show that patients with PPA respond slower and less accurately than healthy controls in both conditions of the picture-word interference task. Moreover, the interference effect (i.e., difference in performance between the two conditions) is larger in patients with PPA than in healthy controls. In the comparison of integrity of the ventral white matter tracts between groups, we find that the integrity of all three tracts is lower in patients with PPA than in controls. Our results also show that the UF and ILF, but not the IFOF, are associated with picture naming and the interference effect in the whole sample. When these effects are investigated within groups, the relationship between integrity of the IFOF and picture naming in RT and the integrity of the UF and the interference effect in RT are observed only for controls.

Our finding that patients with PPA are slower and make more errors in picture naming is in line with the PPA pathology. This goes for the logopenic and semantic variant in particular, showing deficits in naming (Gorno-Tempini et al., 2011). The interference effect of the picture-word interference task is also corroborated (Glaser & Düngelhoff, 1984). All participants responded slower and made more errors in the incongruent condition than in the neutral condition. A novel finding from this study is the fact that the interference effect was

larger in patients with PPA than in healthy controls. Inhibitory control in patients with PPA has not been studied in an experiment with a specific task manipulation before. In the Stroop task, patients with semantic dementia have been found to perform worse in one study, while another study only found these effects for nonfluent PPA, and not semantic dementia (Desgranges et al., 2007; Heidler-Gary et al., 2007). Interestingly, lack of inhibitory control can be seen in the disinhibition in behaviour in the behavioural variant of frontotemporal dementia (Brun et al., 1994). The other subtypes of frontotemporal dementia are the three PPA variants (Chare et al., 2014). Symptoms of the behavioural variant can often be seen in patients with PPA when their disorder has progressed into a later stage, particularly in semantic dementia. This suggests that the degeneration of the frontotemporal network results in impairment in inhibitory control in both language and behaviour.

The comparison of the integrity of the ventral tracts between patients with PPA and healthy controls showed that for patients with PPA in all three tracts the FA was decreased and the MD increased. These results are in line with previous findings on the altered integrity of the ILF and the UF in subtypes of PPA (Agosta et al., 2013; Galantucci et al., 2011). A new finding of this study is the degeneration of the IFOF in patients with PPA, which is seen in both FA and MD values.

No association was found between the integrity of the IFOF and picture naming or the interference effect. This is surprising, as the IFOF has been implicated to play a role in picture naming (Han et al., 2013; Harvey & Schnur, 2015). Furthermore, several studies have found an association between semantic control and the IFOF (Harvey & Schnur, 2015; Nugiel, Alm, & Olson, 2016). Our findings are also different from the findings obtained by Duffau and colleagues using electrostimulation during intraoperative surgery, who found consistent involvement of the IFOF in picture naming (Duffau et al., 2005; Duffau et al., 2009; Moritz-Gasser, Herbet, & Duffau, 2013). A limitation of these awake studies, however, is that their anatomical precision is not optimal due to "brain shifting" after tumour removal (Bello et al., 2010). Additionally, the white matter fibres of IFOF and ILF run closely next to each other through the occipital and up until the anterior temporal lobe (Catani & Thiebaut de Schotten, 2008). Therefore, this method may not be ideal to delineate them accurately. To that end, tractography is a more reliable method. However, as discussed later, we had to use the Probtrackx

rather than the TRACULA algorithm for the IFOF, which makes use of different methods and types of information to reconstruct white matter tracts. This presumably reduced the quality of the tractography for the IFOF, which may explain the absence of a correlation with picture naming.

For the UF (using TRACULA), we find an association between the tract integrity and picture naming for accuracy but not for RT across conditions. This is consistent with some studies, linking the UF to processing semantics and specific semantic categories in picture naming (Han et al., 2013; Papagno et al., 2011). These effects could reflect the process of mapping semantic to motor representations, as held by the ventral production view and computationally implemented in the Lichtheim 2 model (Ueno et al., 2011). However, other studies that use a simple picture naming task, such as the Boston Naming Test, do not find an effect of the integrity of the UF on picture naming (Marchina et al., 2011; Wang et al., 2013). This is different from our findings, but our task includes superimposed letter strings in both conditions. In both of these conditions, top-down control is recruited, although more in the incongruent condition than in the neutral condition. No top-down control is necessary in simple picture naming. Therefore, this difference in naming paradigm (i.e., picture-word interference vs. plain picture naming) could explain why we obtained an effect for the UF (i.e., top-down control was required in our study) and others did not (where top-down control was required less or not at all). If the UF maps meaning onto articulation in picture naming (as the ventral pathway view maintains), then the difference between our results and those of Marchina et al. (2011) and Wang et al. (2013) remain unexplained. However, if the UF underpins top-down control in picture naming, then the difference in results is readily explained (i.e., the difference is due to differential requirement of top-down control).

Supporting the view that the UF mediates top-down control in picture naming is our finding that the integrity of the UF is associated with the magnitude of the interference effect for both accuracy and RT. This finding provides direct evidence that the UF is involved in top-down control rather than meaning-to-motor mapping. This is in line with the WEAVER++/ARC model of Roelofs (2014) and the literature on the ventral pathway connecting the left inferior frontal gyrus as a source of top-down control on semantic information within the anterior temporal lobe (Badre, Poldrack, Pare-Blagoev, Insler, & Wagner, 2005; Badre & Wagner, 2007;

Barredo, Verstynen, & Badre, 2016; Bedny, McGill, & Thompson-Schill, 2008; Novick, Trueswell, & Thompson-Schill, 2010; Roelofs, 2014; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). These findings are also in line with research showing involvement of the arcuate fasciculus but not of the ventral tracts in simple picture naming (Marchina et al., 2011; Wang et al., 2013). As indicated, the picture-word interference task requires top-down control and relies on the contribution of the ventral tracts, while simple picture naming, investigated by Marchina et al. and Wang et al., does not. Summarizing, our findings provide a strong case for the dorsal production view, where meaning-to-motor mapping happens through the dorsal pathway, while top-down influences are relayed by the ventral pathway, most likely the UF.

A strong association of the integrity of the ILF can be seen in picture naming for accuracy and RT. This tract has been described as being part of the ventral visual stream (Ungerleider & Haxby, 1994). Lesions in the tract are associated with impairments in visual object recognition (Shinoura et al., 2010). Anatomically, this makes sense, as the ILF connects visual (occipital) areas with conceptual (ATL) areas. Therefore, decreased integrity of the ILF can lead to impaired picture recognition, which could lead to slower or incorrect responses. Apart from object recognition, the ILF has also been proposed as a central tract in lexical retrieval in picture naming by Herbet et al. (2016). However, the ILF is strongly connected with the medial temporal gyrus (MTG) through U-shaped fibres along the tract and the MTG is a crucial area for lexical retrieval in picture naming (Catani, Jones, Donato, & Ffytche, 2003; Indefrey & Levelt, 2004). Therefore, the effects of the ILF ascribed by Herbet et al. (2016) to lexical retrieval processes actually originate from a later process in picture naming, namely lexical retrieval mediated by the MTG. Herbet et al. used phonological cueing in their picture naming task. If a patient could not name a picture (e.g., a poodle), a phonological cue was given (e.g., the initial phoneme /p/), which helped retrieving the picture name. According to Herbet et al., this cueing effect demonstrates that the ILF mediates lexical retrieval. However, if damage to the ILF leads to poor object recognition (e.g., it is unclear for the patient whether the object is a poodle or a horse), the retrieval of the picture name mediated by the MTG may be helped by a phonological cue (i.e., /p/). Thus, there is no need to assume that the ILF is directly underpinning word retrieval, as Herbet et al. do, but problems in object recognition due to low integrity of the ILF may propagate into the later stage of word retrieval,

mediated by the MTG.

A surprising finding of our study is that the interference effect is associated with the integrity of the ILF in RT, but not in accuracy. While the study by Nugiel et al. (2016) relates the ILF to semantic control with a verb generation task, previous research generally does not provide much support for the idea that the ILF is involved in top-down control in language production. Considering the areas the ILF connects to and the type of information it passes on, one would not assume the ILF to be involved in control processes. What might cause this effect is the possibility that the integrity of the ILF could impair object recognition and in turn affect lexical retrieval, which is more difficult in the incongruent than in the neutral condition of the picture-word interference task. It is important to note that the research on the functional role of the ILF is very limited. Thus, by putting these findings into perspective, future studies on lesions, neurodegeneration, and tractography of the ILF will be able to tell us more about the role of the ILF in top-down control.

Differences between healthy and degenerated white matter tracts in the ventral pathway and their influence on picture naming and interference effect were also assessed. Only for the healthy controls an association was found between the integrity of the IFOF and picture naming in RT, and integrity of the UF and the interference effect in RT. The effect of the integrity of the IFOF on picture naming was not seen in the analysis of the whole sample, but in fact corresponds more closely with previous research (Han et al., 2013; Harvey & Schnur, 2015). The effect of the integrity of the UF on the interference effect was also seen in the analysis on the whole sample and here, it is corroborated for healthy individuals, but not for patients with PPA. No other relationship between integrity of the ventral tracts and interference effect in RT or accuracy was found within groups. This may indicate that for the patients with PPA, the integrity of the tracts is degenerated to such a degree that on average the tract integrity is too low compared to healthy controls to show a linear relationship with interference effect for the whole sample.

The current study has its limitations. An evident limitation of this study is that different tractography programs and algorithms were used for the delineation of the tracts. For the delineation of the IFOF, Probtrackx was used, and for the UF and ILF, TRACULA was used. This was done because TRACULA does not include the IFOF in their atlas and therefore it cannot be reconstructed with this program. An inherent difference between the two

tractography algorithms is that the former is local and the latter is global. Local tractography fits the pathway from the seed to the waypoint mask one step at a time using the local diffusion orientation. Global tractography fits the entire pathway at once using the diffusion orientation of all of the voxels along the pathway length. Moreover, TRACULA uses prior anatomical knowledge from an atlas, while Probtrackx does not use prior information. For TRACULA, this results in tracts that are much more constrained to their probable location compared to the more error-prone algorithm of Probtrackx. While thresholding and multiple subject-specific exclusion masks were used to limit the differences in tract reconstruction, the connectivity maps of the IFOF look very different from those of the other tracts. Therefore, the results of the IFOF in this study have to be regarded more as exploratory than conclusive. A second limitation of the study is the fact that the sample size of the PPA subtypes withheld us from analysing differences between them. Such analyses have previously been done for the UF and ILF, where the integrity of the UF in semantic dementia is seen to be altered (Agosta et al., 2013; Galantucci et al., 2011). PPA subtypes show atrophy and microstructural degeneration in different areas of the brain (Gorno-Tempini et al., 2004). Thus, it can be expected that subtypes with atrophy in brain areas related to top-down control perform worse on the picture-word interference task and show decreased integrity in tracts located in these areas. Therefore, an analysis on the integrity of the three ventral tracts in all three PPA subtypes and how this relates to the interference effect would have been an informative addition to this study. Lastly, our conclusions of the influence of tract integrity on top-down control for the whole sample should be interpreted with some caution. The data of the whole sample are not normally distributed, because they contain data of both patients and controls and these two groups have a different distribution. As a result, the relationship between tract integrity and interference effect is not a straightforward linear relationship driven by a normal distribution, but rather by the differences between the patients and the healthy controls. The variance of the patients is much larger than the variance of the controls, which skews the regression (black line in Fig. 7) slightly towards the patient data (blue dots in Fig. 7). Thus, one could say that a lower integrity leads to a higher interference effect, but not in the form of a simple linear relationship. Rather, it shows that white matter degeneration is related to a high interference effect and healthy white matter is not.
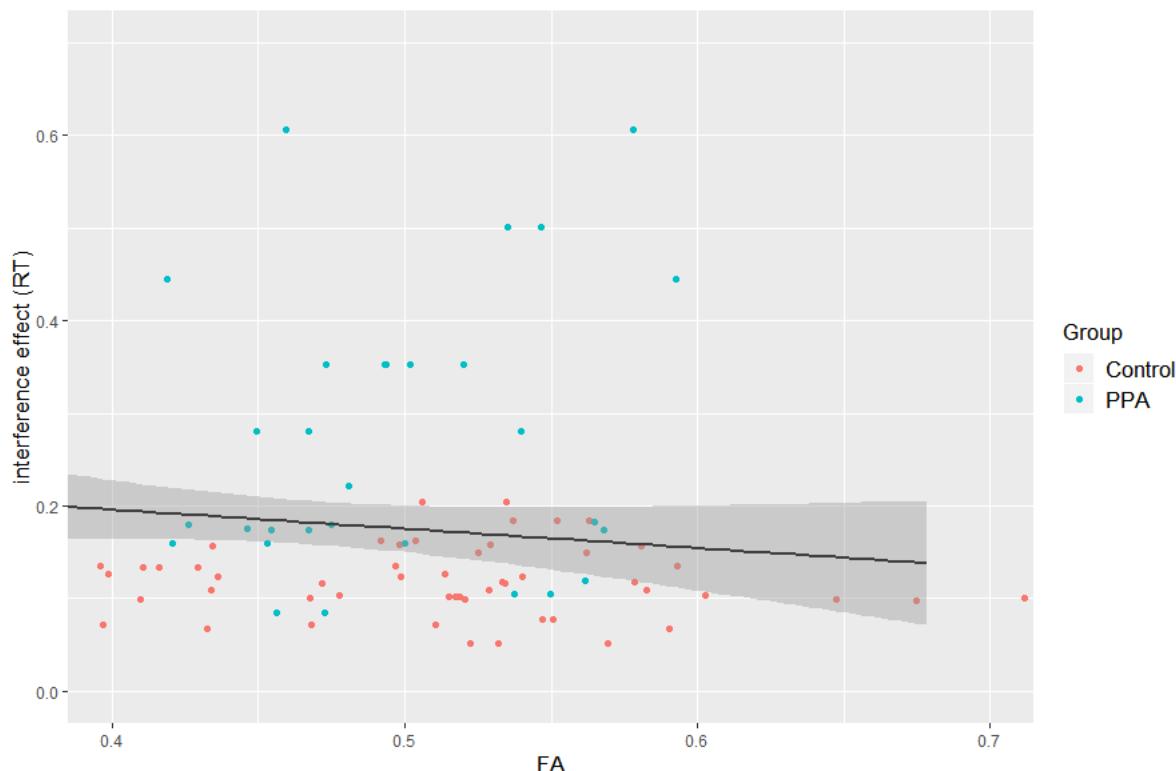
Figure 7. Variance of data between groups shown for the relationship between interference effect in RT and FA.

Conclusion

The impairment in top-down control of semantic information in patients with PPA is established in this study and could be integrated or added to diagnostic practices. Future work could investigate the top-down control of non-linguistic information in this patient group, as patients with PPA also show impairments in non-linguistic domains such as memory, despite language being their main deficit (Eikelboom et al., 2018). Top-down control in healthy language production is also a topic that calls for further investigation, since it has been linked to the ability to speak coherently in healthy aging (Hoffman, Loginova, & Russell, 2018). Finally, we have shown that damage to ventral white matter tracts, specifically the UF and ILF, affects top-down control in language production. These findings are in line with the dorsal production view. Yet, research on the function of white matter tracts in linguistic processes is in its infancy. Language research with MRI methods has historically focused on grey matter and fMRI. Only recently, with the development of tractography, white matter tracts have been linked to language function. More research on the involvement of white matter in linguistic processes in health and disease will provide meaningful insights on the neurobiology of language.

References

Abramovitch, A., McCormack, B., Brunner, D., Johnson, M., & Wofford, N. (2019). The impact of symptom severity on cognitive function in obsessive-compulsive disorder: A meta-analysis. *Clinical Psychology Review, 67*, 36-44.

Agosta, F., Galantucci, S., Canu, E., Cappa, S. F., Magnani, G., Franceschi, M., . . . & Filippi, M. (2013). Disruption of structural connectivity along the dorsal and ventral language pathways in patients with nonfluent and semantic variant primary progressive aphasia: a DT MRI study and a literature review. *Brain and Language, 127*(2), 157-166.

Alexander, A. L., Hurley, S. A., Samsonov, A. A., Adluru, N., Hosseinbor, A. P., Mossahebi, P., . . . & Field, A. S. (2011). Characterization of Cerebral White Matter Properties Using Quantitative Magnetic Resonance Imaging Stains. *Brain Connectivity, 1*(6), 423-446.

Alexander, A. L., Lee, J. E., Lazar, M., & Field, A. S. (2007). Diffusion tensor imaging of the brain. *Neurotherapeutics, 4*(3), 316-329.

Badre, D., Poldrack, R. A., Pare-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron, 47*(6), 907-918.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45*(13), 2883-2901.

Barredo, J., Verstynen, T. D., & Badre, D. (2016). Organization of cortico-cortical pathways

supporting memory retrieval across subregions of the left ventrolateral prefrontal cortex. *Journal of Neurophysiology, 116*(3), 920-937.

Bates, D., Maechler, M., Bokler, B., Walker, S., Bojesen, R. H., Singmann, H., . . . & Fox, J. (2019). Package 'lme4' (Version 1.1-21). https://cran.r-project.org/web/packages/lme4/lme4.pdf.

Bedny, M., McGill, M., & Thompson-Schill, S. L. (2008). Semantic adaptation and competition during word comprehension. *Cerebral Cortex, 18*(11), 2574-2585.

Behrens, T.E.J., Johansen-Berg, H., Jbabdi, S., Rushworth, M.F.S., & Woolrich, M.W. (2007). Probabilistic diffusion tractography with multiple fibre orientations. What can we gain? *NeuroImage, 34*(1), 144-155.

Behrens, T.E.J., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., …, & Smith, S.M. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine, 50*(5),1077-1088.

Bello, L., Castellano, A., Fava, E., Casaceli, G., Riva, M., Scotti, G., . . . & Falini, A. (2010). Intraoperative use of diffusion tensor imaging fiber tractography and subcortical mapping for resection of gliomas: technical considerations. *Neurosurgical Focus, 28*(2), E6.

Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program] (Version 6.0.52). http://www.praat.org/.

Boucart, M., Mobarek, N., Cuervo, C., & Danion, J. M. (1999). What is the nature of increased Stroop interference in schizophrenia? *Acta Psychologica (Amst), 101*(1), 3-25.

Brun, A., Englund, B., Gustafson, L., Passant, U., Mann, D. M. A., Neary, D., & Snowden, J. S. (1994). Clinical and neuropathological criteria for frontotemporal dementia. The Lund and Manchester Groups. *Journal of neurology, neurosurgery, and psychiatry, 57*(4), 416-418.

Catani, M., Jones, D. K., Donato, R., & Ffytche, D. H. (2003). Occipito-temporal connections in the human brain. *Brain, 126*(Pt 9), 2093-2107.

Catani, M., Jones, D. K., & ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology, 57*(1), 8-16.

Catani, M., & Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex, 44*(8), 1105-1132.

Chare, L., Hodges, J. R., Leyton, C. E., McGinley, C., Tan, R. H., Kril, J. J., & Halliday, G. M. (2014). New criteria for frontotemporal dementia syndromes: clinical and pathological diagnostic implications. *Journal of neurology, neurosurgery, and psychiatry, 85*(8), 865-870.

Cheung, V., Cheung, C., McAlonan, G. M., Deng, Y., Wong, J. G., Yip, L., . . . & Chua, S. E. (2008). A diffusion tensor imaging study of structural dysconnectivity in never-medicated, first-episode schizophrenia. *Psychological Medicine, 38*(6), 877-885.

D'Anna, L., Mesulam, M. M., Thiebaut de Schotten, M., Dell'Acqua, F., Murphy, D., Wieneke, C., . . . & Catani, M. (2016). Frontotemporal networks and behavioral symptoms in primary progressive aphasia. *Neurology, 86*(15), 1393-1399.

De Witt Hamer, P. C., Moritz-Gasser, S., Gatignol, P., & Duffau, H. (2011). Is the human left middle longitudinal fascicle essential for language? A brain electrostimulation study. *Human Brain Mapping, 32*(6), 962-973.

De Zubicaray, G. I., Wilson, S. J., McMahon, K. L., & Muthiah, S. (2001). The semantic interference effect in the picture-word paradigm: an event-related fMRI study employing overt responses. *Human Brain Mapping, 14*(4), 218-227.

Dejerine, J., & Dejerine-Klumpke, A. (1895). *Anatomie des centres nerveux: Méthodes générales d'étude-embryologie-histogénèse et histologie. Anatomie du cerveau* (Vol. 1): Rueff.

Desgranges, B., Matuszewski, V., Piolino, P., Chetelat, G., Mezenge, F., Landeau, B., . . . & Eustache, F. (2007). Anatomical and functional alterations in semantic dementia: a voxel-based MRI and PET study. *Neurobiology of Aging, 28*(12), 1904-1913.

Duffau, H., Gatignol, P., Mandonnet, E., Peruzzi, P., Tzourio-Mazoyer, N., & Capelle, L. (2005). New insights into the anatomo-functional connectivity of the semantic system: a study using cortico-subcortical electrostimulations. *Brain, 128*(Pt 4), 797-810.

Duffau, H., Gatignol, P., Moritz-Gasser, S., & Mandonnet, E. (2009). Is the left uncinate fasciculus essential for language? A cerebral stimulation study. *Journal of Neurology, 256*(3), 382-389.

Duffau, H., Herbet, G., & Moritz-Gasser, S. (2013). Toward a pluri-component, multimodal, and dynamic organization of the ventral semantic stream in humans: lessons from stimulation mapping in awake patients. *Frontiers in Systems Neuroscience, 7*, 44.

Eikelboom, W. S., Janssen, N., Jiskoot, L. C., van den Berg, E., Roelofs, A., & Kessels, R. P. C. (2018). Episodic and working memory function in Primary Progressive Aphasia: A meta-analysis. *Neuroscience & Biobehavioral Reviews, 92*, 243-254.

Epstein, K. A., Cullen, K. R., Mueller, B. A., Robinson, P., Lee, S., & Kumra, S. (2014). White matter abnormalities and cognitive impairment in early-onset schizophrenia-spectrum disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 53*(3), 362-372 e361-362.

Forkel, S. J., Thiebaut de Schotten, M., Kawadler, J. M., Dell'Acqua, F., Danek, A., & Catani, M. (2014). The anatomy of fronto-occipital connections from early blunt dissections to contemporary tractography. *Cortex, 56*, 73-84.

Fridriksson, J., Kjartansson, O., Morgan, P. S., Hjaltason, H., Magnusdottir, S., Bonilha, L., & Rorden, C. (2010). Impaired Speech Repetition and Left Parietal Lobe Damage. *Journal of Neuroscience, 30*(33), 11057-11061.

Galantucci, S., Tartaglia, M. C., Wilson, S. M., Henry, M. L., Filippi, M., Agosta, F., . . . & Gorno-Tempini, M. L. (2011). White matter damage in primary progressive aphasias: a diffusion tensor tractography study. *Brain, 134*(Pt 10), 3011-3029.

Garibotto, V., Scifo, P., Gorini, A., Alonso, C. R., Brambati,

S., Bellodi, L., & Perani, D. (2010). Disorganization of anatomical connectivity in obsessive compulsive disorder: a multi-parameter diffusion tensor imaging study in a subpopulation of patients. *Neurobiology of Disease, 37*(2), 468-476.

Geschwind, N. (1970). The organization of language and the brain. *Science, 170*(3961), 940-944.

Glaser, W. R., & Düngelhoff, F. J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 640-654.

Glasser, M. F., & Rilling, J. K. (2008). DTI tractography of the human brain's language pathways. *Cerebral Cortex, 18*(11), 2471-2482.

Gorno-Tempini, Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., . . . & Miller, B. L. (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Annals of Neurology, 55*(3), 335-346.

Gorno-Tempini, Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., . . . & Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology, 76*(11), 1006-1014.

Gruner, P., & Pittenger, C. (2017). Cognitive inflexibility in Obsessive-Compulsive Disorder. *Neuroscience, 345*, 243-255.

Han, Z., Ma, Y., Gong, G., He, Y., Caramazza, A., & Bi, Y. (2013). White matter structural connectivity underlying semantic processing: evidence from brain damaged patients. *Brain, 136*(Pt 10), 2952-2965.

Harvey, D. Y., & Schnur, T. T. (2015). Distinct loci of lexical and semantic access deficits in aphasia: Evidence from voxel-based lesion-symptom mapping and diffusion tensor imaging. *Cortex, 67*, 37-58.

Harvey, D. Y., Wei, T., Ellmore, T. M., Hamilton, A. C., & Schnur, T. T. (2013). Neuropsychological evidence for the functional role of the uncinate fasciculus in semantic control. *Neuropsychologia, 51*(5), 789-801.

Heidler-Gary, J., Gottesman, R., Newhart, M., Chang, S., Ken, L., & Hillis, A. E. (2007). Utility of behavioral versus cognitive measures in differentiating between subtypes of frontotemporal lobar degeneration and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders, 23*(3), 184-193.

Henik, A., & Salo, R. (2004). Schizophrenia and the stroop effect. *Behavioral and Cognitive Neuroscience Reviews, 3*(1), 42-59.

Herbet, G., Moritz-Gasser, S., Boiseau, M., Duvaux, S., Cochereau, J., & Duffau, H. (2016). Converging evidence for a cortico-subcortical network mediating lexical retrieval. *Brain, 139*(11), 3007-3021.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition, 92*(1-2), 67-99.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience, 8*(5), 393-402.

Hoffman, P., Loginova, E., & Russell, A. (2018). Poor coherence in older people's speech is explained by impaired semantic and executive processes. *eLife, 7*, e38907.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1-2), 101-144.

Jbabdi, S., Sotiropoulos, S.N., Savio, A., Grana, M., & Behrens, T.E.J. (2012). Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems. *Magnetic Resonance in Medicine, 68*, 1846-1855.

Kaas, J. H., & Hackett, T. A. (1999). 'What' and 'where' processing in auditory cortex. *Nature Neuroscience, 2*(12), 1045-1047.

Koch, K., Wagner, G., Dahnke, R., Schachtzabel, C., Schultz, C., Roebel, M., . . . & Schlosser, R. G. (2010). Disrupted white matter integrity of corticopontine-cerebellar circuitry in schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience, 260*(5), 419-426.

Konrad, A., Dielentheis, T. F., El Masri, D., Dellani, P. R., Stoeter, P., Vucurevic, G., & Winterer, G. (2012). White matter abnormalities and their impact on attentional performance in adult attention-deficit/hyperactivity disorder. *European Archives of Psychiatry and Clinical Neuroscience, 262*(4), 351-360.

Lansbergen, M. M., Kenemans, J. L., & Van Engeland, H. J. N. (2007). Stroop interference and attention-deficit/hyperactivity disorder: a review and meta-analysis. *Neuropsychology, 21*(2), 251.

Li, F., Huang, X., Tang, W., Yang, Y., Li, B., Kemp, G. J., . . . & Gong, Q. (2014). Multivariate pattern analysis of DTI reveals differential white matter in individuals with obsessive-compulsive disorder. *Human Brain Mapping, 35*(6), 2643-2651.

Mandonnet, E., Nouet, A., Gatignol, P., Capelle, L., & Duffau, H. (2007). Does the left inferior longitudinal fasciculus play a role in language? A brain stimulation study. *Brain, 130*(Pt 3), 623-629.

Marchina, S., Zhu, L. L., Norton, A., Zipse, L., Wan, C. Y., & Schlaug, G. (2011). Impairment of speech production predicted by lesion load of the left arcuate fasciculus. *Stroke, 42*(8), 2251-2256.

Marques, J. P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P. F., & Gruetter, R. (2010). MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *Neuroimage, 49*(2), 1271-1281.

McKinnon, E. T., Fridriksson, J., Basilakos, A., Hickok, G., Hillis, A. E., Spampinato, M. V., . . . & Bonilha, L. (2018). Types of naming errors in chronic post-stroke aphasia are dissociated by dual stream axonal loss. *Scientific Reports, 8*(1), 14352.

Mesulam, M. M. (1982). Slowly progressive aphasia without generalized dementia. *Annals of Neurology, 11*(6), 592-598.

Moritz-Gasser, S., Herbet, G., & Duffau, H. (2013). Mapping the connectivity underlying multimodal (verbal and non-verbal) semantic processing: a brain electrostimulation study. *Neuropsychologia, 51*(10),

1814-1822.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2010). Broca's Area and Language Processing: Evidence for the Cognitive Control Connection. *Language and Linguistics Compass, 4*(10), 906-924.

Nugiel, T., Alm, K. H., & Olson, I. R. (2016). Individual differences in white matter microstructure predict semantic control. *Cognitive, Affective, & Behavioral Neuroscience, 16*(6), 1003-1016.

Papagno, C., Miracapillo, C., Casarotti, A., Romero Lauro, L. J., Castellano, A., Falini, A., . . . & Bello, L. (2011). What is the role of the uncinate fasciculus? Surgical removal and proper name retrieval. *Brain, 134*(Pt 2), 405-414.

Peng, Z., Lui, S. S., Cheung, E. F., Jin, Z., Miao, G., Jing, J., & Chan, R. C. (2012). Brain structural abnormalities in obsessive-compulsive disorder: converging evidence from white matter and grey matter. *Asian Journal of Psychiatry, 5*(4), 290-296.

Penny, W., Friston, K., Ashburner, J., Kiebel, S., & Nichols, T. (2006). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Cambridge, Massachusetts, United States: Academic Press.

Petrides, M. (2005). Lateral prefrontal cortex: architectonic and functional organization. *Philosophical Transactions of the Royal Society B Biological Sciences, 360*(1456), 781-795.

Piai, V., Ries, S. K., & Swick, D. (2015). Lesions to Lateral Prefrontal Cortex Impair Lexical Interference Control in Word Production. *Frontiers in Human Neuroscience, 9*, 721.

R Core Team. (2016). R: A Language and Environment for Statistical Computing. https://www.R-project.org/.

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience, 12*(6), 718-724.

Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America, 97*(22), 11800-11806.

Revelle, W. (2019). Package 'psych' (Version 1.8.12). https://cran.r-project.org/web/packages/psych/psych.pdf.

Ries, S. K., Piai, V., Perry, D., Griffin, S., Jordan, K., Henry, R., . . . & Berger, M. S. (2019). Roles of ventral versus dorsal pathways in language production: An awake language mapping study. *Brain and Language, 191*, 17-27.

Rigucci, S., Rossi-Espagnet, C., Ferracuti, S., De Carolis, A., Corigliano, V., Carducci, F., . . . & Comparelli, A. (2013). Anatomical substrates of cognitive and clinical dimensions in first episode schizophrenia. *Acta Psychiatrica Scandinavica, 128*(4), 261-270.

Roelofs, A. (2014). A dorsal-pathway account of aphasic language production: the WEAVER++/ARC model. *Cortex, 59*, 33-48.

Roelofs, A., & Piai, V. (2011). Attention demands of spoken word planning: a review. *Frontiers in Psychology, 2*, 307.

Sarubbo, S., De Benedictis, A., Maldonado, I. L., Basso, G., & Duffau, H. (2013). Frontal terminations for the inferior fronto-occipital fascicle: anatomical dissection, DTI study and functional considerations on a multi-component bundle. *Brain Structure and Function, 218*(1), 21-37.

Saur, D., Kreher, B. W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M. S., . . . & Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the United States of America, 105*(46), 18035-18040.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain, 123 Pt 12*, 2400-2406.

Shao, Z., Roelofs, A., & Meyer, A. S. (2012). Sources of individual differences in the speed of naming objects and actions: the contribution of executive control. *Quarterly Journal of Experimental Psychology (Hove), 65*(10), 1927-1944.

Shaw, P., Sudre, G., Wharton, A., Weingart, D., Sharp, W., & Sarlls, J. (2015). White matter microstructure and the variable adult outcome of childhood attention deficit hyperactivity disorder. *Neuropsychopharmacology, 40*(3), 746-754.

Shin, N. Y., Lee, T. Y., Kim, E., & Kwon, J. S. (2014). Cognitive functioning in obsessive-compulsive disorder: a meta-analysis. *Psychological Medicine, 44*(6), 1121-1130.

Shinoura, N., Suzuki, Y., Tsukada, M., Yoshida, M., Yamada, R., Tabei, Y., . . . & Yagi, K. (2010). Deficits in the left inferior longitudinal fasciculus results in impairments in object naming. *Neurocase, 16*(2), 135-139.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proceedings of the National Academy of Sciences of the United States of America, 94*(26), 14792-14797.

Ueno, T., Saito, S., Rogers, T. T., & Lambon Ralph, M. A. (2011). Lichtheim 2: synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron, 72*(2), 385-396.

Ungerleider, L. G., & Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Current Opinion in Neurobiology, 4*(2), 157-165.

Wang, J., Marchina, S., Norton, A., Wan, C., & Schlaug, G. (2013). Predicting speech fluency and naming abilities in aphasic patients. *Frontiers in Human Neuroscience, 7*(831).

Wernicke, C. (1874). *Der aphasische Symptomencomplex: eine psychologische Studie auf anatomischer Basis*: Cohn.

Westerhausen, R., Kompus, K., & Hugdahl, K. (2011). Impaired cognitive inhibition in schizophrenia: a meta-analysis of the Stroop interference effect. *Schizophrenia Research, 133*(1-3), 172-181.

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A., & Jefferies, E. (2011). The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. *Cerebral Cortex, 21*(5), 1066-1075.

Wilson, S. M., Galantucci, S., Tartaglia, M. C., Rising, K., Patterson, D. K., Henry, M. L., . . . Gorno-Tempini, M. L. (2011). Syntactic processing depends on dorsal language tracts. *Neuron, 72*(2), 397-403.

Wise, R. J. (2003). Language systems in normal and aphasic human subjects: functional imaging studies and inferences from animal studies. *British Medical Bulletin, 65*, 95-119.

Yao, L., Lui, S., Liao, Y., Du, M. Y., Hu, N., Thomas, J. A., & Gong, Q. Y. (2013). White matter deficits in first episode schizophrenia: an activation likelihood estimation meta-analysis. *Progress in Neuro-psychopharmacol Biological Psychiatry, 45*, 100-106.

Yendiki, A., Panneck, P., Srinivasan, P., Stevens, A., Zöllei, L., Augustinack, J., . . . & Fischl, B. (2011). Automated Probabilistic Reconstruction of White-Matter Pathways in Health and Disease Using an Atlas of the Underlying Anatomy. *Frontiers in Neuroinformatics, 5*(23).

Zwiers, M. P. (2010). Patching cardiac and head motion artefacts in diffusion-weighted images. *Neuroimage, 53*(2), 565-575.

Quantified Sleep:
Single-Subject Study on Weekly Variation of Sleep Quality and Quantity in Regard to Well-Being Across one Year

Sebastian Idesis[1]
Supervisors: Frederik Weber[1], Martin Dresler[1]

[1]Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands

Having enough sleep is relevant to avoid cognitive and emotional problems. Longitudinal studies are necessary to track how such problems might emerge in relation to sleep. However, the majority of such studies rely on behavioural and self-reported measures of sleep. This study aims to investigate how sleep varies in quality and quantity across time. We took an exploratory approach, using a multitude of measurements, investigating whether these variations in sleep are associated with any other variables such as subjective well-being. Therefore, during a continuous longitudinal study, we followed a single participant across one year on a weekly basis and acquired polysomnography (PSG) sleep recordings, self-reported questionnaires and behavioural tasks. The results show that our participant falls within the normal ranges of his normative group regarding the length of sleep stages but deviated in sleep efficiency and duration of sleep. We found that higher reaction times in the psychomotor vigilance task (PVT) predicted increased feeling of sickness, while higher sleep efficiency predicted decreased feeling of sickness and a more positive mood. Moreover, a longer duration of sleep in minutes during the night predicted a lower level of tiredness. Furthermore, the frequency analysis revealed that certain ranges of frequencies (2-4 Hz and 12-14 Hz) fluctuated more across time compared to other frequencies. Nevertheless, these fluctuations were not associated with behavioural changes. Overall, these results could help developing personalized analysis, diagnostics and treatments. Through their high specificity, individualized interventions could thereby lead to an improvement in life quality.

Keywords: sleep, EEG, longitudinal

Corresponding author: Sebastian Idesis; E-mail: Sebastian.idesis@gmail.com

All humans need sleep. It is as important as food or water. We spend about a third of our lives sleeping. Sleep is defined as "a natural and reversible state of reduced responsiveness to external stimuli and relative inactivity, accompanied by a loss of consciousness" (Rasch & Born, 2013). Even though the amount of research on sleep is continuously increasing, the function of sleep is still not entirely clear. Different functions such as thermoregulation, metabolic regulation and adaptive immune functions have been proposed (Rasch & Born, 2013). In fact, altered sleep behaviour has been linked to various negative outcomes such as mood disorders, obesity, cardiovascular diseases, diabetes, and mortality (de Zambotti, Goldstone, Claudatos, Colrain, & Baker, 2018).

Sleep is classified into two major states: rapid-eye-movement (REM) and non-rapid-eye-movement (Non-REM) sleep. These sleep stages alternate in a cyclic way in humans as well as in other animals. Commonly, more Non-REM sleep is observed during the first half of the night, while REM sleep prevails during the second half of the night. The most widely used technique in human sleep research is electroencephalography (EEG) (De Gennaro, Ferrara, Vecchio, Curcio, & Bertini, 2005). Though other sleep assessment techniques are improving, the gold standard for sleep measurement is polysomnography (PSG) which involves recording electrooculography (EOG), electromyography (EMG) and EEG activity. To this end, multiple electrodes are applied to the head and other body parts. Depending on the placement of the electrodes, they will provide information about eye movement, muscle activity, and EEG rhythms. This information can then be used to visually score wake and sleep stages (N1, N2, N3, and REM sleep) (Berry et al, 2015).

A common problem in sleep research is the high costs of PSG. Furthermore, PSG is rather inconvenient, as it requires an expert technician and preparation time. Therefore, data are usually only acquired during a few nights instead of longitudinally. One way to solve this issue of not having enough data is to assess the phenomena that occurred during sleep behaviourally on the following day. A simple approach to assess the effect of sleep disturbances is by using reaction time (RT) tasks that measure performance during continuous work (Dinges & Powell, 1985). One of the most common RT tasks is the psychomotor vigilance task (PVT) (Graeber, Rosekind, Connell, & Dinges, 1990). Although motor tasks and mood questionnaires are easily assessable and relatively

quick, they do not take into account neural data and only focus on short timescales (seconds/minutes). In contrast, there are other measures from diverse fields that approach their research by using long timescales (years/decades) with samples spread over time (non-continuous measurement), to assess the effects after several years (Mischel et al., 2011). Thus far, very limited empirical research has been done on the intermediate level: continuous measurements over an extended period.

There are several caveats for conducting research with continuous, longitudinal data. The lack of motivation and a high rate participants' dropout across time may be the most common one, increasing the complexity of acquiring repeated and frequent measures. Testing for variability over different sessions within single individuals can be inconvenient and costly. Therefore, participants should ideally be aware of the value of the data and have a high intrinsic motivation for successful completion of the study. Therefore, researchers may subject themselves to their experiment. However, the use of self-experimentation may come with expectations regarding the results of the study, which may bias the analysis and interpretation of the data (Poldrack et al., 2015).

A prevalent problem of research, in general, is the generalisability of the results. Clearly, this problem increases with fewer participants. The longitudinal approach solves the inconvenience of statistical conclusions from group studies that cannot be generalised to the intra-individual level. This is only possible in cases in which the processes in question are ergodic (i.e., "if its statistical properties can be deduced from a single, sufficiently long, random sample of the process") (Fisher, Medaglia & Jeronimus, 2018). Cognitive neuroscience studies usually extrapolate the results of single time points as representative of an individual (Poldrack et al., 2015). This is a problem as measures of brain structure and function have been shown to vary as a consequence of, for instance, exercise, water consumption, and caffeine intake. These are just a few of many examples of potential sources of intra-individual variability. As imaging results expressing between-person differences cannot be generalised to within-person time series, intra-individual variability must be studied explicitly, and appropriate methods must be developed for this purpose (Filevich et al., 2017).

Recently, longitudinal studies have benefited from technological improvement and increased access to wearable devices. The use of wearables is increasing at an exponential rate. In the US, a large group of

people uses wearable technology, providing millions and millions of data points (de Zambotti et al., 2018). These immense amounts of data, however, have not been used in continuous longitudinal studies next to neuroimaging techniques. Sleep trackers are usually part of larger devices and only used for validity for a few nights. For example, in one study the authors compared a Fitbit (Fitbit Inc., San Francisco, United States), a consumer multi-sensory wristband wearable, which among other things measures sleep quality, to PSG while assessing sleep/wake state. Nevertheless, the authors assessed 44 healthy adults by measuring only a single night (de Zambotti et al., 2018).

Other studies used neuroimaging techniques to acquire within-participant sleep data. While looking at the power spectra, one study found what they called a trait-like consistency in non-rapid eye movement sleep in adults, which involves almost perfect within-subject stability across multiple nights (Ong, Lo, Patanaik, & Chee, 2019). These correlations were lower in the low-frequency range (usually observed in slow-wave sleep stages) and in spindle frequency bins when comparing a sleep altered group compared with a control group. The researchers claimed that these "trait-like" characteristics could be considered as an electrophysiological "fingerprint". Those characteristics stayed consistent even when external manipulations such as changes in sleep schedules and drug administration were performed.

A famous study on longitudinal and single-case research is the so-called "MyConnectome" study (Poldrack et al., 2015). The data from this study has been used for several projects (Filevich et al., 2017; Mirchi, Betzel, Bernhardt, Dagher, & Mišic, 2019). In the original study, the researchers acquired Magnetic Resonance Imaging (MRI) data next to genetic data of a healthy adult over the course of 18 months. In this experiment, the authors claimed that brain function dynamics are associated with psychological and biological variability. This was not the only attempt to study the brain dynamics across time. Another, less prominent, example is single-case study where they repeatedly measured over more than a year involving 51 MRI sessions (Laumann et al., 2015). Using MyConnectome data, another study reported that functional connectivity patterns reliably predicted daily fluctuations in mood that was assessed with the Positive and Negative Affect Schedule (PANAS questionnaire) (Mirchi et al., 2019). Taking notice of the few variables used in the MyConnectome study, other researchers tried to expand the protocol by measuring over 50 covariates next to each session. One project added

the measurement of, among others, the hours of sunshine, anxiety, blood pressure and physical activity (Filevich et al., 2017). The researchers also included sleep assessments based on self-report questionnaires, which included sleep quality, at what time the participant went to bed the previous night and how much time was spent in bed the previous night. In the same sense, the majority of sleep studies have focused on group measures that emphasize the differences between subjects rather than those within subjects. In order to solve that, de Gennaro and colleagues (2005) acquired several sleep EEG data from 10 subjects where they found a "fingerprint-like" topographic distribution of the EEG power during non-REM sleep. The authors argued that this stability could be due to genetically determined functional brain anatomy, instead of sleep-dependent mechanisms (de Gennaro et al., 2005). This result was also observed in other studies using a similar procedure (e.g., Finelli, Achermann, & Borbély, 2001).

In summary, until this date, sleep studies failed to generalize results from intra-individual data to a whole population and vice versa. This could be resolved by the implementation of continuous longitudinal studies. Furthermore, none of the studies mentioned above measures sleeping processes physiologically or shows how neuroimaging data and sleep interact in a continuous way across a long period of time. The majority of sleep studies rely only on behavioural and self-reported measurements. Only few of these studies intend to create a consistent measure of EEG data from the same participant and even fewer acquire data during more than a couple of nights per person. In addition, the use of wearables is increasing rapidly but it is still not clear how the information acquired by this technology can be used for bigger projects including neuroimaging techniques.

Our study aims to tackle these problems and investigate how sleep changes over the course of one year. Furthermore, we intend to explore if those sleep measures are associated with any other biological assessments like nutrition, hydration or subjective wellbeing by using wearable devices. To this end, this project was designed to measure a large set of variables from a single participant to allow the assessment of the relationship between psychological, neural and behavioural data. Due to the exploratory approach of the study, no predictions were established in advance. We followed one participant across a whole year and acquired sleep PSG recordings every week along with thirty daily behavioural measurements (including mood, productivity, and stress measurements). This allowed

us to create a rich sleep "phenotype" of one person beyond self-report questionnaires.

Methods

Despite the fact that a tremendous amount of data is available with the current set-up, this thesis mainly focuses on a fraction of the data and possible analyses to comply with text length constraints.

Participant

The experiment assessed a single case: a healthy, right-handed male, aged 29 years at study onset. He did not present any neuropsychiatric disorder prior to or during the study. The participant joined the study after providing written informed consent for EEG usage according to the guidelines of the local ethics committee. The participant offered himself voluntarily and no money rewards was given in exchange for his participation

Design and procedures

PSG recordings were performed on a weekly basis on Monday nights while the behavioural measurements such as mood questionnaires were performed on a daily basis. Collection of EEG started March 20th of 2018 and ended on March 25th of 2019. The aim was to collect weekly EEG sleep data for one calendar year (52 weeks). Due to travel and technical issues, the final amount of 50 sessions was achieved. All statistical analyses were postponed to the end of this period. In order to avoid a bias provoked by self-experimenting, the participant did not have insight into the data throughout the entire collection period (Poldrack et al., 2015). The only information related to sleep known to the participant was sleep self-reports, obtained by phone apps, which indicate sleep quality and how many hours were slept during the previous night. Because the participant was used to observe these phone-acquired reports before the beginning of the study, it was decided that he could continue this routine to increase the ecological validity. When measuring neuroimaging data from participants repeatedly across time, it is useful to keep consistency in data acquisition time of the day to standardize time-of-day effects (Gordon et al., 2017). In our study, data were acquired at the same time and day (Monday night) to keep this consistency. As there were no a priori hypotheses, the analyses approach was exploratory. All the statistical analyses were performed using the

commercially available software "R" (R Core Team, 2013).

PSG procedure

The PSG setup was based on a previous master's thesis (Pottkämper, 2019) and was recorded using a wearable device (SOMNOscreen plus, SOMNOmedics GmbH, Randersacker, Germany). The recording included EEG, EMG, EOG and electrocardiogram (ECG). We acquired EEG data from 15 electrodes, in accordance with the American Academy of Sleep Medicine (AASM) rules (Berry et al., 2015) and the 10-20 system (Cz and the two mastoid electrodes were used as a reference, while Fpz was used as a ground electrode). For further details regarding the positioning of the electrodes and the full protocol, see Appendix B-E from Barry et al. (2015).

During recording, the sleep signal was pre-filtered by the recording system. EEG and EOG were high-pass filtered at 0.3 Hz and low-pass filtered at 30 Hz. EMG was high-pass filtered at 10 Hz and low-pass filtered at 100 Hz. Additionally a notch filter around 50 Hz was applied. ECG was high-pass and low-pass filtered at 0.3 Hz and 70 Hz respectively, and additionally with notch filter around 50 Hz. Sleep was scored offline using the SpiSOP toolbox (Weber, 2013, RRID:SCR_015673), in accordance with the AASM standards. Movement artefacts were manually removed, and the analysis was performed on 30-second epochs. Several variables were obtained with the PSG data, from which one of the most relevant for our analysis is the sleep efficiency, which estimates how much time was slept during the night regarding the total time spent in bed.

Mood measurements

A set of questionnaires were administered during the day by using phone apps. The questionnaires contained information about exercise, activities and mood between others. As a mood assessment, the PANAS Scale was used (Crawford & Henry, 2004). PANAS is a self-report 10 items measuring both positive and negative affect. All data were exported by the participant, without him looking at their contents and then sent to the experimenter for analysis. The data were included as a regressor in the statistical models next to the EEG data in order to assess the relationship between other variables. For the mood measurements, a correlation was assessed with sleep data variables (i.e., sleep efficiency) in

order to assess the relation between mood variables and sleep variables. Of all the subjective measures of well-being measured for the participant's own interest, we focused only on sickness, productivity, happiness, sadness, and tiredness.

PVT (psychomotor vigilance task)

Reaction time was assessed using the PVT (Graeber et al., 1990). The 2-minutes version of the PVT was performed every morning right after waking up to assess objective vigilance by means of the participant's average reaction time.

Correlational analysis

For the correlational analysis, several variables were incorporated to assess their relationship. Two of these variables were obtained during the night recordings (sleep efficiency and total sleep time in minutes) as objective indicators of sleep quality and quantity, while the others were obtained on the following day by means of questionnaires and behavioural tasks (PANAS scores, stress, sickness, tiredness, and PVT reaction time) to indicate subjective and objective measures of well-being. Correlation strength was assessed by using Pearson's r-value. A p-value of 0.05 or lower was considered significant.

Spectral EEG analysis

In this study, we performed a similar analysis to asses "trait-like" characteristics as Ong and colleagues (2019). In brief, this assessed which frequencies vary most across the year while comparing weekly EEG sleep recordings. EEG analyses were performed using the SpiSOP tool (Weber, 2013, RRID:SCR_015673) based on MATLAB 2013b (Mathworks, Natick, USA) and FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011, https://fieldtriptoolbox.org, RRID:SCR_004849). In order to obtain the results, several analysis steps were performed. These include segmentation (epoching data into time segments), calculation (transformation of the signal in each segment using Fast Fourier Transformation), averaging over bands (averaging the data points in order to display them) and normalization (normalize the power estimates to obtain power density estimates; for more information, see SpiSOP web documentation on https://www.spisop.org/documentation/).

Power spectra were calculated on consecutive artefact-free 5-s intervals of Non-REM sleep, which overlapped in time by 4 s. Each interval was tapered by a single Hanning window before applying Fast Fourier Transformation that resulted in interval power spectra with a frequency resolution of 0.2 Hz. Power spectra were then averaged across all

Table 1. Sleep table: Representation of sleep stage time, both in minutes and percentages of the total sample (N = 50). SD refers to standard deviation; SE refers to standard error; Min refers to the minimum value and Max to the maximum value. The difference (Diff) between our participant and the normative group is indicated. Values denoting NA indicate that data were not available. The normative group was taken from Boulos et al. (2019). Asterisks (*) refer to the presence of three atypical data points which could be modifying the outcome.

| Sleep stage parameters	Mean	SD	Min	Max	Range	SE	Normative	Diff
Sleep efficiency (%)	97.31	1.36	93.58	99.29	5.71	0.19	89	8.31
Total sleep time (min)	376.45	36.19	289.50	437.00	147.50	5.12	410.6	-34.15
Sleep onset time (min)	6.06	2.25	2.50	11.00	8.50	0.32	14.3	-8.24
Wake after sleep onset (min)	4.46	4.94	0.00	23.00	23.00	0.70	32.1	-27.64
Stage 1 (min)	8.68	5.79	2.50	22.00	19.50	0.82	NA	NA
Stage 2 [min]	236.01	36.69	168.00	312.00	144.00	5.19	NA	NA
Stage 3 [min]	51.48	16.31	16.00	92.00	76.00	2.31	NA	NA
REM (min)	74.49	24.46	17.00	116.00	99.00	3.46	NA	NA
REM without MA (min)	67.77	22.08	14.50	108.50	94.00	3.12	NA	NA
Stage 1 (%)	2.13	1.34	0.68	5.47	4.79	0.19	6	-3.87
Stage 2 (%)	55.16	9.42	23.69	73.06	49.37	1.33	51.3	3.86
Stage 3 (%)	13.46	4.46	3.45	24.02	20.57	0.63	21.4	-7.94
REM (%)	19.64	5.92	5.21	29.87	24.65	0.84	19.8	-0.16
REM without MA (%)	17.89	5.41	4.45	27.93	23.48	0.76	NA	NA
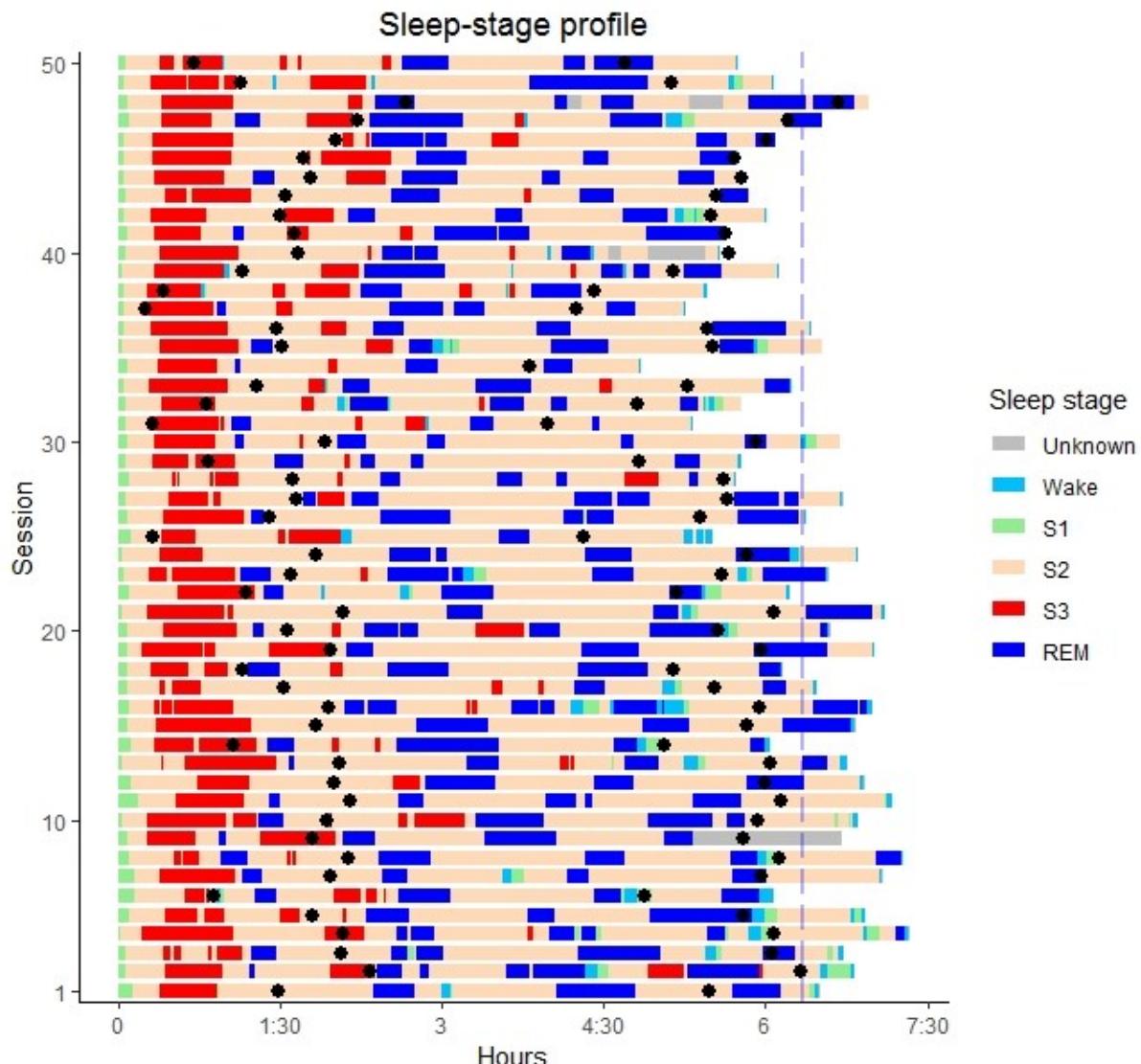
Figure 1. Sleep-stage profile. Each line represents a different session with each stage is coloured differently. The black dots indicate the time of the clock (2 and 6 AM respectively). Vertical dashed line represents the session's average.

blocks (Welch's method) and normalized by the effective noise bandwidth to obtain power spectral density estimates for the whole data (Wang, Weber, Zinke, Inostroza, & Born, 2018). This study follows an exploratory approach; therefore, all frequencies between the ranges of 0.3-30 were included in the analysis. The analysis was performed for all the EEG night recordings in order to compare them. Variation in each frequency band was assessed by calculating the standard error of the mean. For power spectrum analysis the sessions 1, 4, 5, 19, 21, 28, 29, 30, 32, 47, 48, 49 and 50 (final $N = 37$) were excluded because the absolute power values indicated technical problems (e.g., erroneous placement of electrodes).

Results

Sleep composition

In order to summarize the data obtained by the sleep EEG recordings, a descriptive analysis was performed including information about sleep stages, both in minutes and percentages. The results were compared to a normative group, matching the participant's age (healthy adults between 18 and 34 years old) that was taken from a meta-analysis using normal polysomnography parameters (Boulos et al., 2019). Through the comparison of the normative group against the case data, it is visible that our participant fits the expected range of values from the normative group regarding sleep stages length and proportion. Nevertheless, our participant had a high sleep efficiency, together with a short sleep onset duration and small amount of wake after sleep

onset. These results remain even when the lower total sleep time is considered (Table 1).

Sleep-stage profile

Hypnograms were obtained and listed in chronological order for each night of the 50 sessions with colours representing the different sleep stages (Figure 1). Appendix Figure A1 displays the same information ordered based on sleep onset time in order to achieve a different visualization of the data. The start of the sessions was aligned to the sleep onset. This descriptive summary denotes the participant's typical sleep-stage profile across the year. Also, information about when the participant fell asleep, woke up, his sleep length, the number of sleep cycles during the night and the distribution of the sleep stages are marked.

Correlational analysis of sleep parameters and behavioural measures

Association between selected variables of interest was tested with a correlation analysis including sleep efficiency, total sleep time, positive affect score, stress, tiredness, sickness, and PVT reaction time average (Figure 2). The negative affect score did not reveal any strong association with any other variable and therefore was not included in the figure. The detailed correlations are displayed in Figure 3.
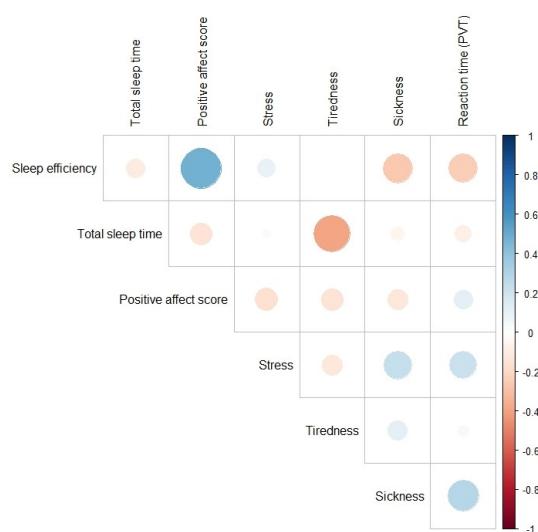
Figure. 2. Sleep quality and quantity association with behavioural measures. Pearson's correlations between the selected variables of interest. Blue dots indicate a positive correlation while the red dots indicate a negative correlation. Higher colour saturation and greater dot size indicate stronger correlation coefficients.

Higher reaction time in the morning from the PVT correlated with higher self-reported sickness of that day ($r(49) = .28$, $p = .047$; Figure 3A), but surprisingly not significantly related with sleep efficiency ($r(49) = -.24$, $p = .095$; Figure 3B). Showing that sleep quality and mood are related, the nights with higher sleep efficiency predicted higher positive affect scores in the self-reported (PANAS) questionnaire the next day ($r(49) = .49$, $p = .021$; Figure 3C). However, sleep was at most weakly associated with less sickness the following day. ($r(49) = -.26$, $p = .065$; Figure 3D).

As expected, mere higher sleep quantity, as measured by the total time of sleep, predicted subjective tiredness the next day ($r(49) = -.4$, $p = .0042$; Figure 3E). Other sleep parameters and behavioural measures were uncorrelated.

Spectral composition of non-REM and REM sleep

Power spectral density was calculated for all non-REM and REM stages separately for every session. Then variability between sessions (standard error of the mean, SEM) for each frequency bin was calculated. In non-REM sleep EEG activity displayed marked higher variability between 2–4 Hz and 12–14 Hz which are typically associated with sleep slow waves/K-complexes and sleep spindles, respectively. This variability was stable after exclusion of outliers due to erroneous recording sessions (Figure 4A). Surprisingly, power-density variability in REM sleep was also higher in the 2–4 Hz range (Figure 4B). By assessing a correlation between power obtained in both frequency ranges with the aforementioned behavioural variables, no significant correlation was found.

Discussion

We measured a single person across one year in order to see how sleep features change across time and how they interact with a diverse set of behavioural and psychological variables. The results show that our participant falls within the normal ranges of his normative group concerning the length of sleep stages. This, however, is not the case with sleep efficiency and duration of sleep. Here, the participant deviates from his normative group; interestingly, we found several relations between the variables that could be assessed in a causal way in future studies: Higher reaction times in the PVT task during the morning seems to predict feeling more sick. In addition, higher sleep efficiency seems
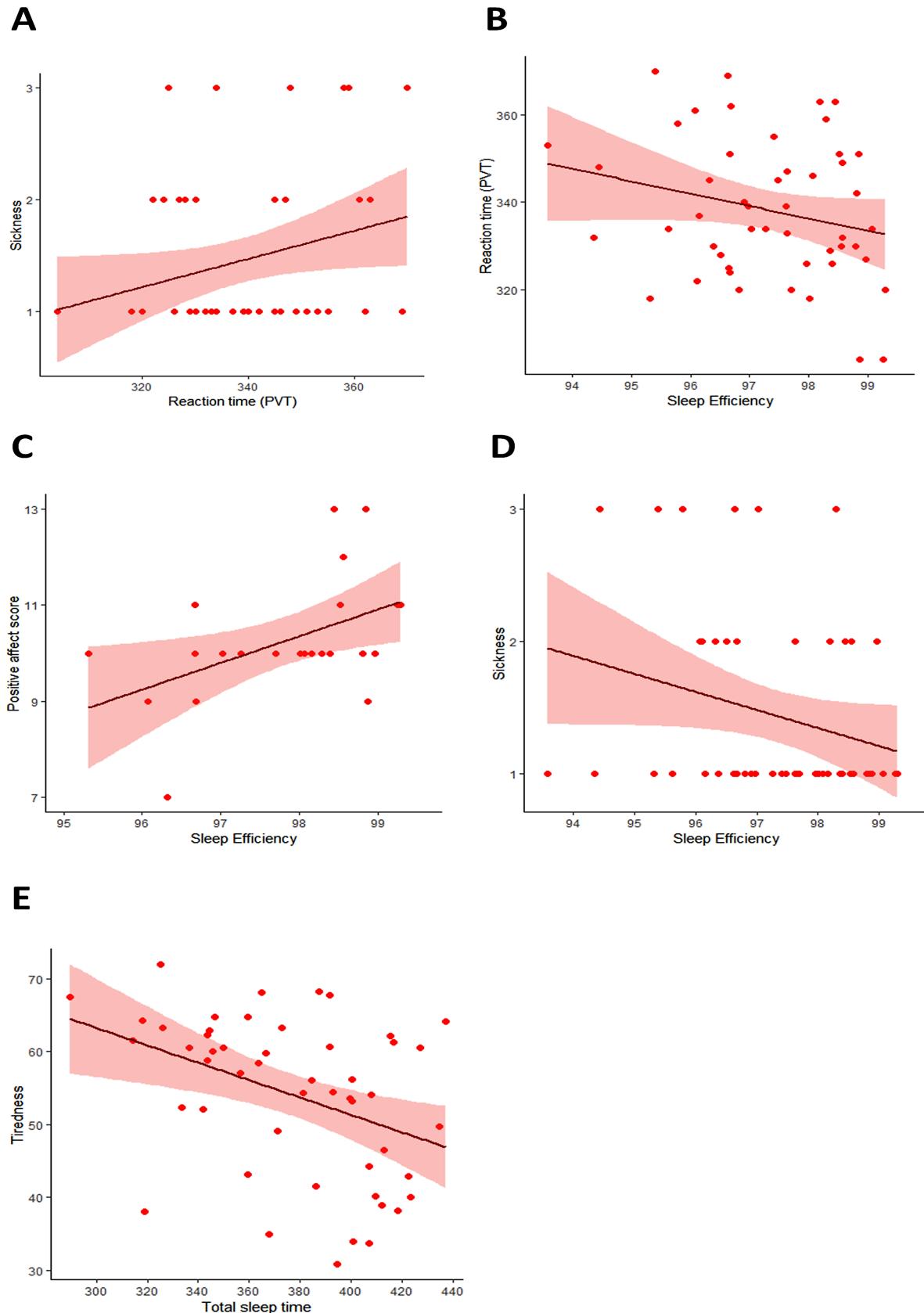
Figure 3. (A) Linear regression between the reaction time (RT) average provided by the PVT and the self-reports of sickness at the day (r(49) = .28, p = .047). (B) Linear regression between sleep efficiency and reaction time average provided by the PVT (r(49) = - .24, p = .095); (C) the Positive Affect score of the PANAS questionnaire (r(49) = .49, p = .021) (D) self-reported sickness on the day after (r(49) = -.26, p = .065), (E) the total sleep time and self-report of tiredness on the day after (r(49) = -.4, p = .004)

to predict reduced feeling of sickness and a more positive mood. Moreover, a longer duration of sleep in minutes during the night seems to predict a lower level of tiredness.

Lastly, a frequency analysis revealed that certain ranges of frequencies (2-4 Hz and 12-14 Hz during Non-REM and only the first range during REM) fluctuated more across time compared to the other frequencies. Nevertheless, these fluctuations were not associated with behavioural changes.

Sleep features

The sleep features found in the current participant are similar to those found in studies with multiple subjects (Mitterling et al, 2015; Boulos et al, 2019). Although the obtained data support most of the results, some findings contradict our expected outcomes based on previous literature. These contradictions involve high sleep efficiency, short sleep onset, low amounts of wake after sleep onset and total sleep time. Our participant displayed a significantly higher sleep efficiency compared to the normative group. He furthermore, did not use an electrode cap system, previously reported to be less comfortable as it restricts participants' movements during the night (Šušmáková, 2004).

Another possible explanation is the fact that our participant slept the nights at his house and not in the lab. Sleeping in a lab is likely more stressful and therefore compromises sleep quality (Johns & Doré, 1978).

Moreover, our participant slept with the EEG device on repeated occasions. He therefore likely habituated to perturbances by wearing the device. Lastly, the sleeping duration during the nights was lower than in the normative group, while the proportion of sleep stages was similar. This could indicate sleep deprivation (Philip et al., 2004). Our participant may have increased his sleep efficiency through higher sleep pressure at sleep onset which also can explain the short sleep onset times.

Associations between sleep features and other variables

As mentioned the association between PVT reaction times and sleep quality is supported by an extensive literature (Graeber et al., 1990; Loh, Lamond, Dorrian, Roach, & Dawson, 2004; Belenky et al, 2003). It is therefore surprising that we found no strong association between PVT and sleep efficiency. This is possibly due to unusually low variability in the measures of sleep quality, which
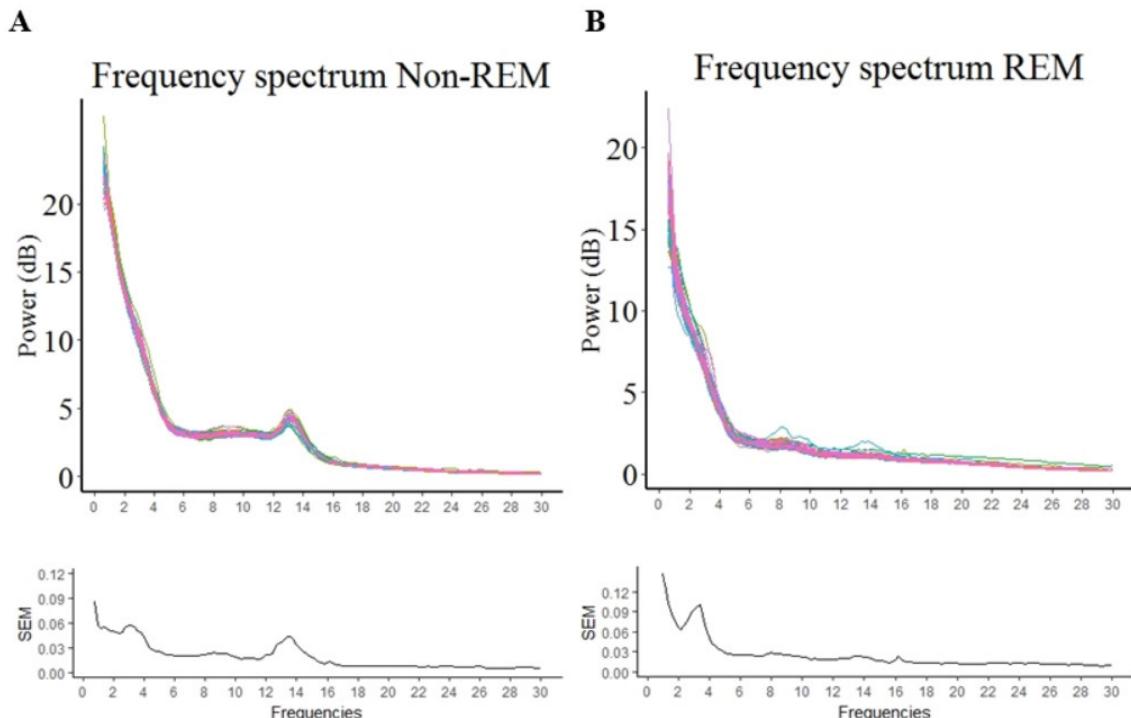

Figure 4. Power analysis of Non-REM stages (A) and REM stages (B). Power spectral density (above) indicating the presence of the different frequency ranges with the corresponding power. Each line corresponds to a different night of recording. The standard error of the mean (SEM; below) was calculated to assess which frequency bins were more variable across time. Average power was calculated from C3 and C4 electrodes as deviations to the linked, averaged mastoids.

was also relatively high. Additionally, we used a 2-minute version of the PVT task when typically, 5- to 10-minute versions are more sensitive to sleepiness during the day. Our subject did clearly not show any improvement or detriment at the PVT task throughout the first year (Figure A2). Thus, the contradiction with the literature gives rise to the question of the sensitivity of the validity of our tasks and assessment of sleep quality and whether they are suitable as valid measures in all populations. Further studies should analyse similar cases to see if the PVT task is useful for this kind of participants, who are possibly confounded by sleep deprivation states. Nevertheless, in the current study, PVT reaction times were at least sensitive to reported sickness as in previous articles (Dinges et al, 1997).

Interestingly, several studies have tried to look for the relation between sleep and mood, indicating a weak correlation between sleep quality and emotional features such as anxiety, anger, etc. (Thomsen, Mehlsen, Christensen, & Zachariae, 2003; Akashiba et al., 2002). The findings reported in the aforementioned studies align with the ones presented in the current study with the difference of the time-frame of data acquisition. Nevertheless, the current study found no correlations while assessing negative affect scores assessed by the PANAS questionnaire.

Most of the studies focused on short timescales such as days/weeks while our study acquired data throughout a whole year. To our knowledge, only one study looked into the sleep-mood relation on a large time scale. In this study, patients with bipolar disorder were measured across 18 months, assessing how sleep duration affected mood changes (Leibenluft, Albert, Rosenthal, & Wehr, 1996). The authors found a relationship between sleep and mood by claiming that decreased sleep duration was the best predictor of mood fluctuations present in mania or hypomania the next day. Nevertheless, the aforementioned study only relies on self-report questionnaires in order to measure sleep features ignoring sleep quality. In the current study, we use the same questionnaires but by adding PSG measurements, we improved sleep assessment while adding an objective and not-biased quantification.

Spectral EEG analysis

Although previous studies included repeated sessions of sleep EEG, they only assessed a few nights per participant (de Genaro et al., 2015; Ong et al., 2019). Based on previous work by Ong et al. (2019), we aimed to assess a trait-like pattern of sleep

EEG power spectra. By doing this, further analysis can evaluate which behavioural/environmental variables are associated with frequency changes in order to establish causal relationships. The aforementioned study assessed sleep during three different nights to see coherent and incoherent patterns in three different conditions. In the current study, we approached this topic in an exploratory way in order to see which frequencies have a high variability throughout a year. Previous research has investigated frequency stability. A high temporal stability of gamma-band activity was found (Keil, Stolarova, Heim, Gruber, & Müller 2003).

However, as the aforementioned study focused on mechanisms underlying complex cognitive tasks, the other frequency (lower than 30 Hz) bands were ignored from the analysis. Taking an exploratory approach, we looked into a broader spectrum of frequencies. Our findings indicate higher variability in lower ranges of frequencies (2-4 Hz) during non-REM and REM stages and the sleep spindle band (12-14 Hz). Furthermore, the analysis made by Keil and colleagues (2003) was based on cognitive task by looking at specific time-windows (moments before and after behavioural response). The same happened with the work published by Zheng, Zhu and Lu (2017) where they find a relatively stable within and between sessions EEG pattern. In both mentioned studies, the analysis was task-related and therefore not comparable to the ones obtained in the current study. Here, we take a new approach by analysing the frequencies during sleep without cognitive tasks.

By assessing the correlation between the overrepresentation of these frequencies in the sleep EEG variability and behavioural measures, no strong association was found. This indicates that these variations may not transfer to subjective and objective measures of well-being (positive affect score, stress, tiredness, sickness and reaction time) during the day. It remains open if they relate to another behaviour at all. Therefore, further studies should look for the variables affected by these variabilities across time.

Additionally, as the current study is a single-case study, it is plausible that our results are highly biased, specifically considering that our participant fell outside of his normative group on several sleep-related variables. Moreover, the effects of gender and age cannot be assessed here. Further studies are needed to determine whether the same results (including the factor of time and variability of the features) are obtained with different populations and in different parts of the world. In order to do that, it would be advisable to consider this study as a pilot

and replicate the methodology proposed here but optimizing certain aspects. For instance, behavioural measurements, which were recorded on a daily basis, could instead be measured on a weekly basis.

A meta-analysis by Gradisar, Gardner and Dohnt (2011) summarized the average sleep patterns for each age range but also classified it by region, country and city. Future studies could compare EEG during sleep and during wakefulness (repeated times) in order to look for consistency within participants. By taking all of these variables into consideration, the development of personalized analysis, diagnostics and treatments are possible. Through their high specificity, individualized interventions would benefit the effectiveness of treatments and could thereby lead to an improvement in life quality.

Future directions

Having sleep features extracted objectively from the sleep EEG recordings (such as sleep efficiency), instead of solely relying on self-report measures, allowed us to successfully look into the relationship between these concepts and other variables such as self-reported tiredness and sickness. Moreover, further analysis could establish causal relations between these variables in order to predict sleep quality based on behavioural variables and vice versa. This could also be achieved by applying specific interventions (such as changing the sleep schedule or length) in order to see the consequences in the behavioural measurements. By doing this, directional relations could be assessed and used for diagnosis and treatments. This approach could allow using sleep to improve behaviour and vice versa.

Additionally, the frequency analysis proposed in the current study could lead to the development of new analyses. With the data of the current study, the relationship between specific frequency bins and behavioural data can be addressed by using the time as a variable factor. To establish causal relationships, future studies could assess whether behavioural interventions can modify these changes in the frequency power to improve sleep quality and therefore, quality of life. Although several software tools are already available for the pre-processing and statistical analysis of neural data, new tools are also being developed. The data provided from the current study can also be used to guide the validity and reliability tests of measures on brain structure and function obtained through newly developed tools. Ultimately, by analysing the relation between sleep quality and mood reports with sufficient data opens the potential for the development of algorithms to predict mood states based on sleep recordings, and vice versa. Finally, the data and results collected and analysed in the current project will be part of an online open-source data. This will allow researchers to validate our analyses as well as perform additional ones.

Conclusions

This study investigated vital subjective and objective parameters of sleep and behaviour from the most extensive longitudinal study of this kind. Sleep quality and quantity were compared to a normative group indicating that our participant revealed similar length of sleep stages but different levels of sleep efficiency and duration of sleep.

The relation between sleep quality and several variables over the course of one year (i.e., physiological and behavioural) showed how sleep efficiency predicted decreased feeling of sickness and a positive mood. Besides, it was revealed that higher reaction times predicted feeling sicker and higher duration of sleep predicted a decreased feeling of tiredness. The current study constitutes the first attempt to create a dense longitudinal sleep phenotype and adds data to the growing open database of the analysed study that can be used for more future interdisciplinary studies.

The current study provides a one-year dataset to enrich this idea by creating a longer electrophysiological "fingerprint". Therefore, ideas related to how consistent are certain electrophysiological features, that were based on theory, can now be assessed empirically. The dependency on behavioural reports can now be modified and complemented with EEG data providing richer and well-founded information. The current study provides new findings about sleep frequency analysis by revealing that certain frequencies vary more than others (2-4 and 12-14 Hz fluctuate more than the others across time) in one individual during the course of a year and in a habituated sleep situation.
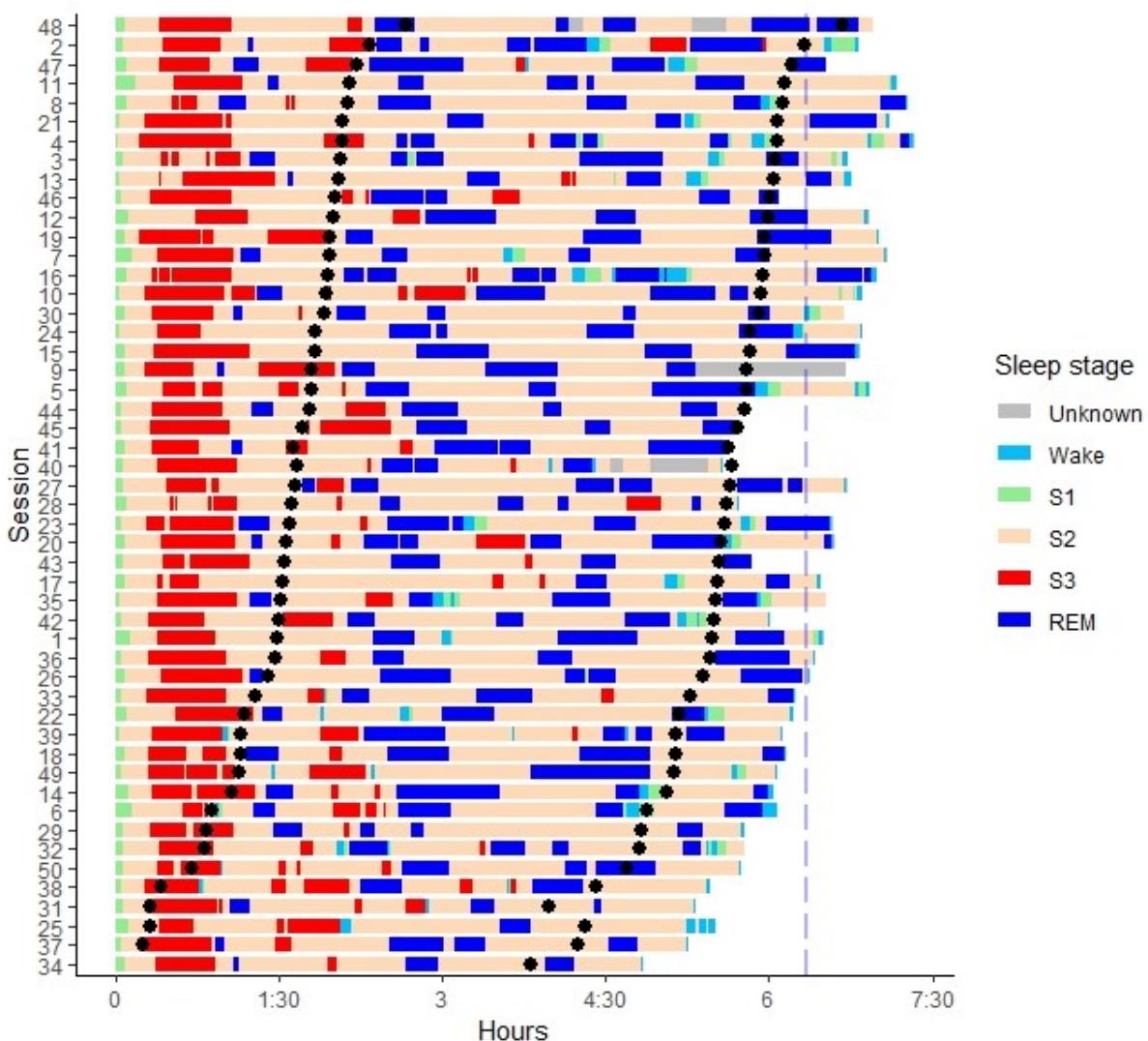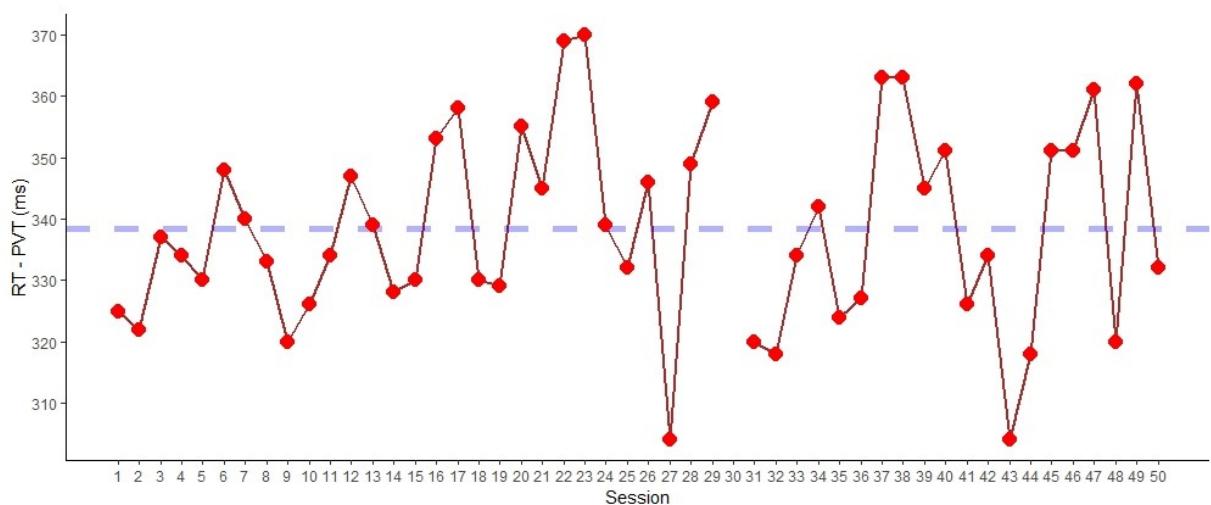
References

Akashiba, T., Kawahara, S., Akahoshi, T., Omori, C., Saito, O., Majima, T., & Horie, T. (2002). Relationship between quality of life and mood or depression in patients with severe obstructive sleep apnea syndrome. *Chest*, *122*(3), 861-865.

Belenky, G., Wesensten, N. J., Thorne, D. R., Thomas, M. L., Sing, H. C., Redmond, D. P., ... & Balkin, T. J. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent

recovery: A sleep dose-response study. *Journal of sleep research*, *12*(1), 1-12.

Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S. M., Lloyd, R. M., Marcus, C. L., & Vaughn, B. V. (2015). The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications, version 2.2. *American Association of Sleep Medicine.*

Boulos, M. I., Jairam, T., Kendzerska, T., Im, J., Mekhael, A., & Murray, B. J. (2019). Normal polysomnography parameters in healthy adults: a systematic review and meta-analysis. *The Lancet Respiratory Medicine*, *7*(6), 533-543.

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non☐clinical sample. *British Journal of Clinical Psychology*, *43*(3), 245-265.

De Gennaro, L., Ferrara, M., Vecchio, F., Curcio, G., & Bertini, M. (2005). An electroencephalographic fingerprint of human sleep. *NeuroImage*, *26*(1), 114–122.

De Zambotti, M., Goldstone, A., Claudatos, S., Colrain, I. M., & Baker, F. C. (2018). A validation study of Fitbit Charge 2TM compared with polysomnography in adults. *Chronobiology International*, *35*(4), 465–476.

Dinges, D. F., Pack, F., Williams, K., Gillen, K. A., Powell, J. W., Ott, G. E., ... & Pack, A. I. (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4–5 hours per night. *Sleep*, *20*(4), 267-277.

Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers*, *17*(6), 652–655.

Filevich, E., Lisofsky, N., Becker, M., Butler, O., Lochstet, M., Martensson, J., … & Kühn, S. (2017). Day2day: investigating daily variability of magnetic resonance imaging measures over half a year. *BMC Neuroscience*, *18*(1), 65.

Finelli, L., Achermann, P., & Borbély, A. A. (2001). Individual "Fingerprints" in Human Sleep EEG Topography. *Neuropsychopharmacology*, *25*(5), S57–S62.

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, *115*(27), E6106–E6115.

Fitbit Official Site for Activity Trackers and More. (2019). Retrieved 18 July 2019, from https://www.fitbit.com/nl/home

Gordon, E. M., Laumann, T. O., Gilmore, A. W., Newbold, D. J., Greene, D. J., Berg, J. J., … & Dosenbach, N. U. F. (2017). Precision Functional Mapping of Individual Human Brains. *Neuron*, *95*(4), 791–807.e7.

Graeber, R. C., Rosekind, M. R., Connell, L. J., & Dinges, D. F. (1990). Cockpit napping. Retrieved from https://ntrs.nasa.gov/search.jsp?R=19910043203

Gradisar, M., Gardner, G., & Dohnt, H. (2011). Recent worldwide sleep patterns and problems during adolescence: a review and meta-analysis of age, region, and sleep. *Sleep medicine*, *12*(2), 110-118.

Johns, M. W. & Doré, C. (1978). Sleep at home and in the sleep laboratory: Disturbance by recording procedures. *Ergonomics*, *21*(5), 325-330.

Keil, A., Stolarova, M., Heim, S., Gruber, T., & Müller, M. M. (2003). Temporal stability of high-frequency brain oscillations in the human EEG. *Brain Topography*, *16*(2), 101-110.

Laumann, T. O., Gordon, E. M., Adeyemo, B., Snyder, A. Z., Joo, S. J., Chen, M.-Y., … & Petersen, S. E. (2015). Functional system and areal organization of a highly sampled individual human brain. *Neuron*, *87*(3), 657–670.

Leibenluft, E., Albert, P. S., Rosenthal, N. E., & Wehr, T. A. (1996). Relationship between sleep and mood in patients with rapid-cycling bipolar disorder. *Psychiatry research*, *63*(2-3), 161-168.

Loh, S., Lamond, N., Dorrian, J., Roach, G., & Dawson, D. (2004). The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 339–346.

Mirchi, N., Betzel, R. F., Bernhardt, B. C., Dagher, A., & Mišic, B. (2019). Tracking mood fluctuations with functional network patterns. *Social Cognitive and Affective Neuroscience*, *14*(1), 47–57.

Mischel, W., Ayduk, O., Berman, M. G., Casey, B. J., Gotlib, I. H., Jonides, J., … & Shoda, Y. (2011). 'Willpower' over the life span: decomposing self-regulation. *Social Cognitive and Affective Neuroscience*, *6*(2), 252–256.

Mitterling, T., Högl, B., Schönwald, S. V., Hackner, H., Gabelia, D., Biermayr, M., & Frauscher, B. (2015). Sleep and Respiration in 100 Healthy Caucasian Sleepers—A Polysomnographic Study According to American Academy of Sleep Medicine Standards. *Sleep*, *38*(6), 867–875.

Ong, J. L., Lo, J. C., Patanaik, A., & Chee, M. W. L. (2019). Trait-like characteristics of sleep EEG

power spectra in adolescents across sleep opportunity manipulations. *Journal of Sleep Research*, e12824.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience*, *2011*, 1.

Philip, P., Taillard, J., Sagaspe, P., Valtat, C., Sanchez-Ortuno, M., Moore, N., ... & Bioulac, B. (2004). Age, performance and sleep deprivation. *Journal of sleep research*, *13*(2), 105-110.

Poldrack, R. A., Laumann, T. O., Koyejo, O., Gregory, B., Hover, A., Chen, M.-Y., …, & Mumford, J. A. (2015). Long-term neural and physiological phenotyping of a single human. *Nature Communications*, *6*(1), 8885.

Pottkämper, J. (2019) *The role of rapid eye movement sleep in emotional schema processing* (Master's thesis).

Rasch, B. & Born, J. (2013). About Sleep's Role in Memory. *Physiological Reviews*, *93*(2), 681–766.

Šušmáková, K. (2004). Human sleep and sleep

EEG. *Measurement science review*, *4*(2), 59-74.

Thomsen, D. K., Mehlsen, M. Y., Christensen, S., & Zachariae, R. (2003). Rumination—relationship with negative mood and sleep quality. *Personality and Individual Differences*, *34*(7), 1293-1301.

Wang, J. Y., Weber, F. D., Zinke, K., Inostroza, M., & Born, J. (2018). More effective consolidation of episodic long-term memory in children than adults—Unrelated to sleep. *Child development*, *89*(5), 1720-1734.

Weber, F. D. (2013). SpiSOP tool(box). https://www.sipsop.org

Zheng, W. L., Zhu, J. Y., & Lu, B. L. (2017). Identifying stable patterns over time for emotion recognition from EEG. *IEEE Transactions on Affective Computing*.

Appendix

Figure A1 above, A2 below.

Task-Potency Investigations in Autism Spectrum Disorder

Tristan Looden[1], Roselyne Chauvin[1], Maarten Mennes[1], Jan Buitelaar[1,3], and Christian Beckmann[1,2]

[1]Donders Institute, Nijmegen
[2]FMRIB, Oxford
[3]Karakter, Nijmegen

Autism Spectrum Disorder (ASD) is a developmental disorder characterised by social and communicative deficits and repetitive behaviours. Research into ASD is challenged by heterogeneity among patients and a high prevalence of co-morbid disorders, such as Attention Deficit Hyperactivity Disorder (ADHD), a developmental disorder characterised by impulsivity, hyperactivity and/or inattention. Both ASD and ADHD are associated with functional brain alterations, observed in resting-state- and task-fMRI (functional magnetic resonance imaging) studies. To better characterise functional connectivity alterations related to performance on a specific task, we propose to use a measure called task-potency – which effectively enables comparing different tasks relative to a common baseline architecture, modelled from the resting state. In this study, we investigate whether modulation of connections during task performance is altered in ASD compared to control participants and whether different effects can be observed within the subpopulation of ASD meeting clinical criteria for ADHD. We observed that the ASD group employed greater modulation amplitude for each task. We further observed that the ASD_{ADHD+} (ASD subjects co-morbid with ADHD) group modulated fewer connections than TD (typically developing) when performing the Hariri and both reward tasks. Conversely, their mean modulation amplitude was greater than TD for every task. Moreover, we observed significant effects for ADHD symptom severity scores in the set of connections shared by the ASD_{TOTAL} (all ASD subjects) and ASD_{ADHD+} groups for Flanker, non-social reward and theory of mind tasks validating a co-morbidity effect. We found that task modulation amplitude in these five tasks is altered in ASD, but the participants without ADHD co-morbidity did not show this effect. Secondly, we found strong differences between ASD_{ADHD+} and TD in both the percentage of connections modulated as well as in the mean amplitude of modulation, and furthermore found effects of ADHD symptom severity on modulation amplitude for a set of connections in multiple tasks. We conclude that task-potency metrics display sensitivity to subgroup effects and therefore show potential for the development of stratification biomarkers.

Keywords: functional connectivity, fMRI, resting state, ASD, ADHD

Corresponding author: Tristan Looden; E-mail: t.looden@donders.ru.nl

Autism spectrum disorder (ASD) is a complex neurodevelopmental disorder marked by social and communicative deficits, as well as repetitive and restricted behaviors and interests. There appears to exist significant behavioral, cognitive and neural heterogeneity within ASD populations (Brunsdon & Happé, 2014; Nunes, Peatfield, Vakorin, & Doesburg, 2018). This has made investigations into any hypothetical underlying commonalities challenging. A further difficulty in ASD research lies where individuals diagnosed with ASD can also meet criteria for co-morbid disorders, most notably Attention Deficit Hyperactivity Disorder (ADHD) (Gargaro, Rinehart, Bradshaw, Tonge, & Sheppard, 2011). ADHD has itself been associated with functional brain alterations (Chauvin, Oldehinkel, Buitelaar, Beckmann, & Mennes, 2018) and it can be important to dissociate these alterations from those that might be related strictly to ASD. Attempts at describing ASD at a cognitive level have led to the emergence of several different theories (Hull et al., 2017). Relevant work has been published on how these various theories about ASD might interrelate at the cognitive level (Brunsdon & Happé, 2014). However, thus far our knowledge about how the neural processes that underlie tasks which tap into these cognitive domains interrelate is limited. The current research aims to advance our understanding of these cognitive interrelations in terms of their associated neural patterns in individuals with ASD and typically developing individuals through a method called 'task-potency'. Because this method incorporates both resting state and task-fMRI (functional magnetic resonance imaging) based measures of functional connectivity, we will start by briefly discussing resting state and task-fMRI functional connectivity findings in ASD before introducing the new angle of attack that task-potency allows us to take on.

Resting state functional connectivity in ASD

Analysis of resting-state fMRI data has proven to be a convenient means to ascertain functional connectivity patterns across brain areas. It provides insight into basal connectivity patterns, while at the same time being relatively easy to conduct independent of participant age and/or cognitive ability. The ASD resting state functional connectivity literature can roughly be divided into two classes of descriptions: broadly speaking that ASD brains exhibit either over-connectivity or under-

connectivity relative to typically developing controls (Hull et al., 2017). The exact interpretation of these descriptions is, however, not trivial. For instance, what counts as over- or under-connectivity may be defined in several different ways. As an example, it can be defined globally (e.g., as an average value across all edges) or locally (e.g., as a value expressing the connectivity between one or several brain regions). Furthermore, the specific pattern of connectivity may be of relevance. For instance, a previous ASD study found diffuse over-connectivity with an aberrant pattern in ASD resting state fMRI data (Di Martino et al., 2011), whereas another resting state study found evidence for globally weaker connectivity, but also with an abnormal organisation (Yerys et al., 2017). Incongruent findings may further be driven by differences in factors such as age, sex and co-morbidity in the populations used, as well as the inherent heterogeneity present in ASD (Hull et al., 2017). These findings already underline the challenge of integrating results across studies and coming to a scientific consensus as to what might be going awry in the brains of individuals diagnosed with ASD. There is another layer of complexity to add in unravelling the neural underpinnings of cognition in ASD. Because even though resting state analyses provide a solid foundation of the functional architecture of the brain (Smith et al., 2009), it cannot tap into the processing underlying specific cognitive work. This means that task-based fMRI measures are crucial to include if the aim is to understand why deficits might be present in some cognitive domains but not so much in others, and where these dysfunctions present themselves.

Task-based functional connectivity in ASD

Task-based fMRI studies may be particularly useful in understanding abnormalities in ASD as they might more directly probe the neural correlates of the cognitive domains affected by the disorder. Studies using pattern recognition algorithms to predict ASD diagnoses from functional connectivity matrices derived from social reward tasks (Just, Cherkassky, Buchweitz, Keller, & Mitchell, 2014) and theory of mind tasks (Deshpande, Libero, Sreenivasan, Deshpande, & Kana, 2013) have reported substantial accuracy. The association of ASD with deficits in different cognitive domains has led to the inclusion of a varied battery of five tasks in the longitudinal European autism project in an attempt to probe a wide range of behavioural

and neural responses in participants (Charman et al., 2017a). However, as we know from the previous paragraph that complex findings have been reported at the level of the resting state, and because functional connectivity during tasks builds upon the ongoing activity architecture that resting state fMRI aims to estimate (Mennes et al., 2010), interpreting task-based functional connectivity findings is not straightforward. This study makes use of the novel task-potency method that provides a tool for disentangling the relative contribution of task-induced functional connectivity from that of the baseline architecture, potentially allowing a more precise interpretation of task-based functional connectivity findings and how these may vary across individuals and groups.

Task-potency

At its core, task-potency can be viewed as an analogue of the well-known fMRI activation designs in the realm of functional connectivity. Similar to where brain activation localization for a task is modelled with respect to a baseline condition, task-potency models the functional connectivity state that exists during a task with respect to the resting-state functional connectivity state. We consider the resting state a suitable baseline for the brain's functional architecture (Mennes et al., 2010). Through this modelling, we attain a new connectivity matrix which can be interpreted as the modulation away from the baseline - or 'potentiation', - that is instantiated in each edge due to engaging in the task, for any particular participant. The task potency framework then allows to isolate the task-based functional connectivity profile at a subject-specific level. In other words, for each subject and each task we attain a matrix of functional connectivity values across brain regions that represents the connectivity modulation away from the subject's resting state baseline that the cognitive demands in a particular task induce. This procedure can be repeated for a number of different tasks for each participant, where each task might have a particular way in which edges in the participant's brain are potentiated. Some tasks might potentiate a greater number of edges than others or potentiate a similar number of edges but with a greater amplitude. We might further investigate if there are edges that are potentiated above a certain threshold in many of the tasks under consideration and discover patterns therein. The tasks included in the current study intend to probe cognitive domains that are affected

in ASD. We will apply the novel method of task-potency to analyse them and provide information about functional connectivity modulation in the ASD brain as compared to controls.

Materials and methods

Participants

The dataset from the EU-AIMS LEAP project was used for the current analyses (Charman et al., 2017). Data obtained from participants with intellectual impairments (IQ < 75) was excluded. Participants performed a resting state fMRI and one or more of the following task fMRI scans: Hariri emotion processing (Emotion) (Hariri, Tessitore, Mattay, Fera, & Weinberger, 2002), Flanker (inhibition) (Meyer-Lindenberg & Weinberger, 2006), non-social reward anticipation (Reward_ns), social reward anticipation (Reward_s) (Delmonte et al., 2012), animated shapes theory of mind (ToM) (Castelli, Frith, Happé, & Frith, 2002; White, Coniston, Rogers, & Frith, 2011). Additionally, each participant completed a T1-weighted anatomical scan for the purpose of registration. fMRI parameters are shown in the next paragraph. We further removed participants from the analysis for data quality using the following criteria. All participants had acceptable overlap (> 94%) with the MNI152 standard brain after image registration. We excluded 57 participants due to poor overlap (< 50%) with one or more regions from the brain parcellation atlas that had been chosen for the analysis (van Oort et al., 2017). Participants were further excluded on the basis of incidental findings, incomplete scans and those in the top 5% in terms of head motion quantified through RMS-FD (Jenkinson, Bannister, Brady, & Smith, 2002). The above criteria resulted in the inclusion of data for analyses from the following participants: 282 participants with autism spectrum disorder (age range: 7.5 - 30.3 years; M = 17.1; SD = 5.4; 72.3% male) and 221 typically developing controls (age range: 6.9 - 29.8 years; M = 17.0; SD = 5.5; 63.8% male) (see Fig. 1).

fMRI scanning parameters

MRI data were acquired on 3T scanners at multiple sites in Europe. Structural images were obtained using a 5.5 min MPRAGE sequence (TR = 2300 ms, TE = 2.93 ms, T1 = 900 ms, voxel size = 1.1 x 1.1 x 1.1 mm, flip angle = 9°, matrix size = 256 x 256, FOV = 270 mm, 176 slices). An 8 - 10 min

resting state fMRI (R-fMRI) scan was acquired using a multi-echo planar imaging (ME-EPI) sequence: TR = 2300 ms, TE = 12 ms, 31 ms, and 48 ms (slight variations are present across centers), flip angle = 80°, matrix size = 64 x 64, in-plane resolution = 3.8 mm, FOV = 240 mm, 33 axial slices, slice thickness/gap = 3.8 mm/0.4 mm, volumes = 200 (UMCU), 215 (KCL, CIMH) or 265 (RUNMC, UCAM). Participants were instructed to relax and fixate on a cross presented on the screen for the duration of the R-fMRI scan. The information was retrieved from Oldehinkel et al. (2019).

fMRI preprocessing

Preprocessing of the fMRI data was performed with tools from FSL 5.0.6. (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012). The first five volumes from each acquisition were removed to allow for equilibration of the magnetization. To correct for head movement, we performed volume realignment to the middle volume using MCFLIRT. This was followed by global 4D mean intensity normalization and smoothing with a 6 mm FWHM (full-width half maximum) kernel. ICA (independent component analysis)-aroma was used to identify and remove secondary motion-related artefacts (Pruim, Mennes, Buitelaar, & Beckmann, 2015; Pruim, Mennes, van Rooij, et al., 2015). Next, signal from white matter and CSF (cerebrospinal fluid) was regressed out and we applied a 0.01 Hz high-pass filter. For each participant, we registered acquisitions to their respective high-resolution T1 anatomical images by means of the Boundary-Based Registration (BBR) tool from FSL-FLIRT (Jenkinson et al., 2002). The high-resolution T1 image belonging to each participant was registered to MNI152 space with FLIRT 12 degrees-of-freedom linear registration and further refined using FNIRT non-linear registration (Andersson, Jenkinson, & Smith, 2007). We used the inverse of these transformations to take a brain atlas to the native space of each participant, where all further analyses were performed.

Behavioural data

We used clinical variables related to ASD and – in order to analyse possible co-morbidity effects – ADHD symptomatology in our analysis. For ASD, these were the total scores for the Social Responsiveness Scale Second Edition (SRS-2), the Repetitive Behaviours Scale - Revised (RBS-R) and the Short Sensory Profile (SSP). The SRS-

2 contains questions about broad, characteristic ASD behaviours. The RBS-R contains questions specifically about repetitive and restrictive behaviours associated with ASD. The SSP scores abnormalities in sensory processing. For our analyses, we summed the SRS-2, RBS-R and SSP scores for each participant in order to produce an ASD symptom sum-score. For ADHD symptomatology, we used parent report scores for the clinical domains of inattentiveness and hyperactivity. We summed the ADHD hyperactivity and inattentiveness score for each participant in order to produce an ADHD symptom sum-score.

Task-potency calculation

We used a hierarchical brain atlas with 168 brain regions (distributed across 11 larger-scale networks) (van Oort et al., 2017) to superimpose on the native space of each subject and for each fMRI acquisition to define the regions. For each participant, we calculated the covariance between the average BOLD time series extracted from each brain region pair. Using Ledoit-Wolf regularization, we estimated partial correlations from the covariance matrix (Ledoit & Wolf, 2012) and consecutively applied the Fisher-Z transformation. This provided us with 168 x 168 connectivity matrices, one resting state matrix and one or more task matrices for each participant. The main Gaussian from a mixture Gaussian-gamma model, that was applied on each individual matrix, supplied us with parameterised information about the distribution of values in each matrix. We used these parameters to normalize the elements in each matrix. In order to produce individual matrices of connectivity modulations induced by the task (i.e., task-potency), we standardized each participant's task matrices by subtracting that participant's resting state matrix. The resulting matrices are interpreted as
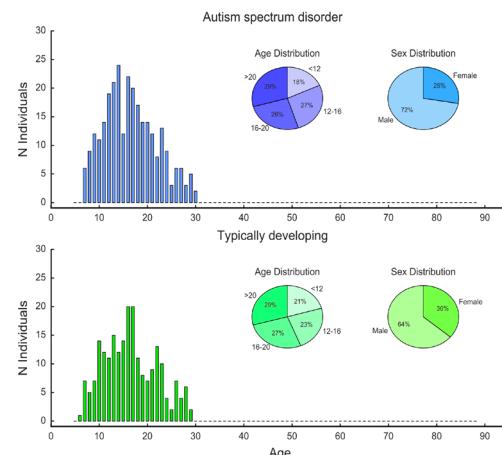

Figure 1. Demographical characteristics of the sample used in this study.
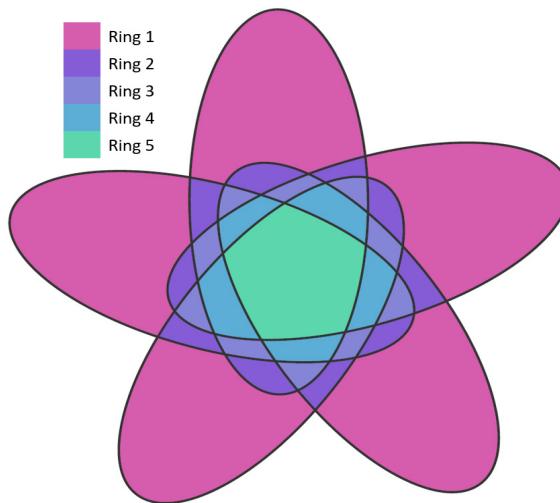
Figure 2. Venn diagram rings representation. Surface areas are not to scale. Each of the five overlapping ovals corresponds to the set of edges selected in one of the five tasks. The rings refer to those sets that are increasingly shared across multiple tasks with 'ring 5' at the centre, being the set that is shared across all of the tasks in the current paradigm

containing the connectivity modulations away from the resting state baseline that the respective task induces in the brain – we refer to these modulations as task-potency (Fig. 3) (Chauvin, Mennes, Llera, Buitelaar, & Beckmann, 2019).

Task-fingerprints

We obtained group-level task-potency matrices by averaging across participant potency matrices for each task within each participant group (ASD and TD). Within these group-level matrices, we identified the most strongly and consistently modulated edges by applying a threshold informed by a mixture modelling procedure (Bielczyk et al., 2018). The resulting binary matrices for each task and diagnostic group are referred to as task-fingerprints and consist of those edges sensitive to their respective task. We refer to the edges passing the threshold as 'selected edges' for that task. Within each diagnostic group, we now have five task-fingerprints, one per task. From these five fingerprint matrices, we can now identify those edges that are modulated by one particular task, as well as those edges that are modulated by two, three, four or all of the five tasks. Any edge can therefore be modulated by a combination of the five tasks, resulting in 32 (including the set of edges modulated by none of the tasks – the null set) possibilities across the tasks. We can report the proportion of edges modulated across tasks as a Venn diagram containing each of the 31 non-null

sets of edges. To simplify our interpretation of these sets, we decided to define five categories as 'levels of modulation', which we can interpret as the level of generalization in the use of an edge. Essentially, we define five concentric 'rings' within the Venn-diagram representation, where each ring consists of all the sets that have the same sensitivity 'level'. For example, ring 1 consists of those five sets that are modulated in only one of the respective five tasks, and ring 5 consists of the one set that is modulated in each of the five tasks (Fig. 2). This representation allows us to make inferences about whether any connectivity differences, which we will test between ASD and TD, might be situated in for example a task-general or a task-specific domain.

Statistical testing

For ASD and TD, we investigated and compared edges that were selected in a particular task or in a particular ring. Four types of behaviour were investigated in the above conditions: 1) the percentage of selected edges, 2) the mean amplitude within the selected edges, 3) the normalized distribution of selected edges across the five sensitivity-levels, 4) the ratio of edges connecting areas within or between networks as defined in our atlas – across the five sensitivity-levels. For point 4, we average information across networks as we do not aim at investigating network-specific differences. Bootstrapping was used to provide a robust estimate of the variance within these measures. Here, 10000 samples were taken from the original data, each including a random 80% of participants of the groups involved, selected with replacement. This allowed us to produce an estimate of the uncertainty of the particular metric within the populations, which we need to quantify the uncertainty as to their difference. Any differences between ASD and TD in the obtained metrics above were then assessed using significance testing, where we assessed how surprising our findings were given the distributions that were obtained from the sample bootstrapping.

Results

Data

We had thus far obtained 2 (groups) x 5 (tasks) = 10 separate mean functional connectivity matrices which were thresholded to reveal the connectivity fingerprints, consisting of a selection of modulated edges for each of these ten groups. We compared
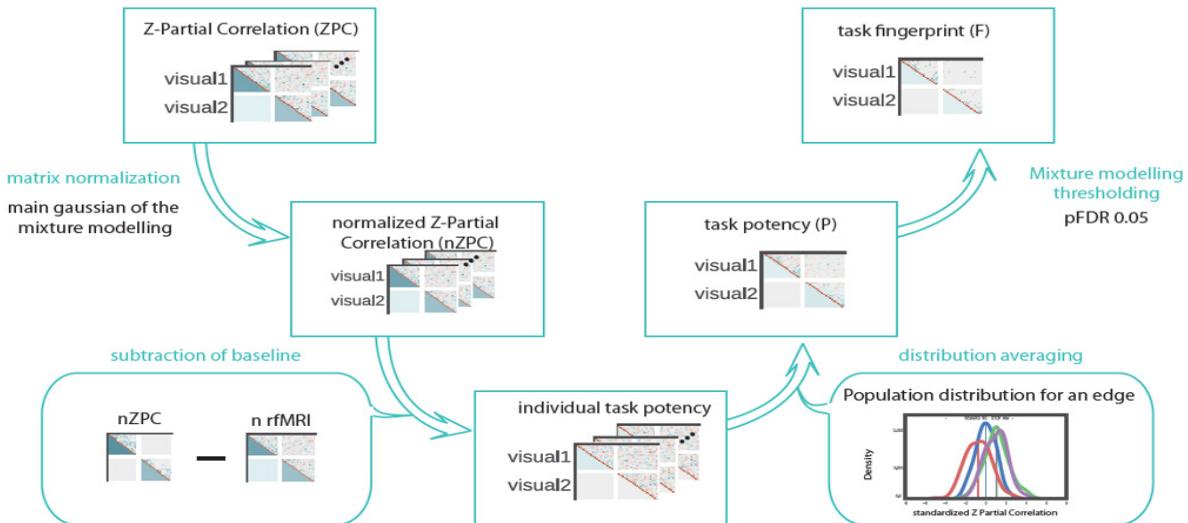
Figure 3. The task-potency pipeline.

the modulation strength, amount and spatial pattern of these selected edges between groups and tasks.

Comparisons between diagnostic groups

As the first step in analysing the data, we took the connectivity fingerprints and compared the amount of selected edges in each of these fingerprints between diagnostic categories for each task. Differences in the amount of selected edges present in the task connectivity fingerprint of diagnostic groups can provide insight into a global connectivity alteration where more or less edges are modulated to achieve the same task-goal. This could indicate more diffuse or more focused network involvement. In this measure, the specific spatial connectivity modulation profile (e.g., which networks connect where) was not taken into consideration. On first inspection, there appeared to be a trend for a greater amount of edges being recruited in performing the tasks in ASD as compared to TD across the tasks (Fig. 4). This could suggest a more diffuse modulation of edges across networks in ASD. However, the overall modulation across selected edges between groups was not significant in any of the five particular tasks (pairwise comparisons with FDR (false discovery rate) correction across tasks). Additionally, we noticed that different tasks modulated a different amount of edges in both diagnostic groups.

As a second step, we analysed the mean amplitude of modulation within the selected edges present in the fingerprints. For each diagnosis x task combination, we identified the mean amplitude of edge modulation across the edges (Fig. 4). After FDR correction across groups, we saw significant differences in the tasks Flanker (p = 0.005), social

reward (p = 0.001), non-social reward (p = 0.002) and theory of mind (p = 0.026). In each of these tasks, the amplitude is lower for the ASD group, suggesting that the brain connectivity profile in the ASD group was not as strongly modulated in the
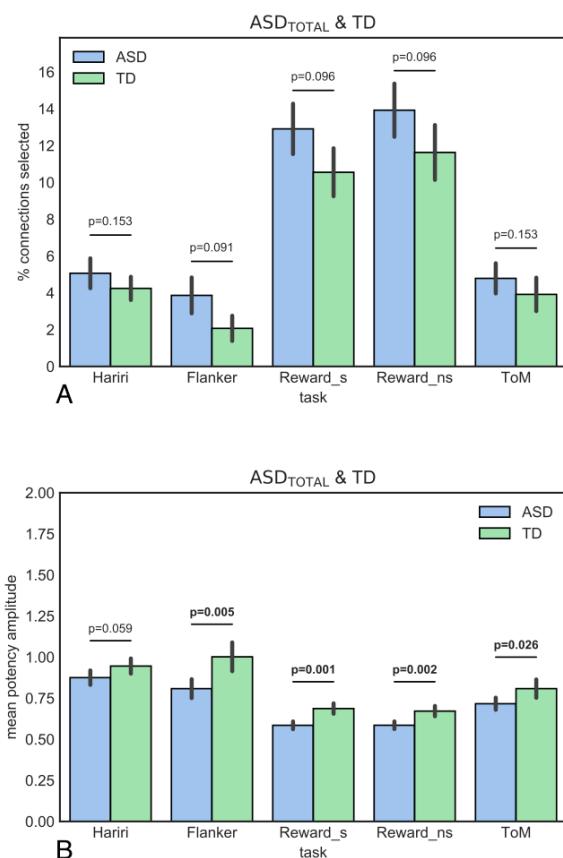

Figure 4. A. The percentage of edges that were consistently and strongly modulated in ASD_{TOTAL}. **B.** The mean potency amplitude of the selected edges in ASD_{TOTAL}.
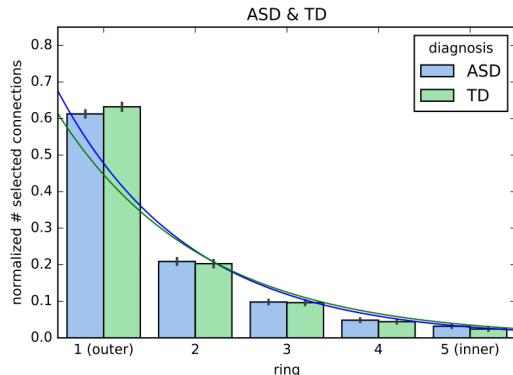
Figure 5. The distribution (normalized) of selected edges across the rings. Ring 1 contains only those edges selected in one task, while ring 5 contains those selected in all tasks.

face of any of these tasks as was the case for the TD group.

Comparisons between tasks

One of the primary goals of this investigation was to attempt to integrate brain connectivity findings across different tasks (i.e., cognitive domains) in the ASD population. As described earlier, for each task x group combination we obtain a task-fingerprint – containing those edges that passed our threshold for being consistently and strongly potentiated in their respective condition. If we consider these 'selected edges' for each of the tasks, we can represent edges that are modulated by one or a combination of tasks in a Venn-diagram. This allows us to describe similarity in terms of connectivity patterns between tasks. We focused our analysis on the level of selectivity across tasks, as in each edge being modulated by one, two, three, four or all five of the tasks (Fig. 2). The five areas defined as 'rings' within this diagram represent advancing levels of generalized modulation across tasks, with ring one being a combination of the five sets of edges that are only modulated in one specific task and ring five consisting of that set of edges which are modulated in each task. We investigated the way in which the amount of selected edges is distributed across these five rings and how this might differ between ASD and TD. For better visualization, the distributions were normalized within diagnosis group – meaning that the sum of the values across the five tasks, within diagnosis group, equals one. Because this makes the data dependent across rings, we modeled the representation of edges across the five rings using a parametric fit. We chose an exponential for this fit, based on visual inspection as well as for model simplicity. Using one fit per bootstrap, we extracted
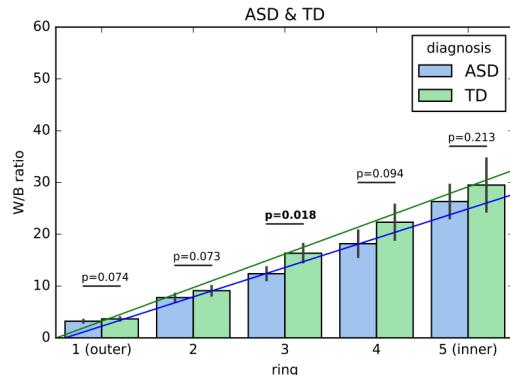

Figure 6. Displays the ratio of the amount of within-network edges to the amount of between-network edges in each ring set. Higher values correspond to greater within network connectivity. The lines indicate the linear fit for the respective diagnosis type of the of ring on the W/B ratio (β = 5.66, SD = 0.60 for ASD; β = 6.49, SD = 0.95 for TD).

the distribution of the exponential coefficients for each group and compared if the representation of shared edges across rings shows a different pattern between the groups (Fig. 5). We find that the first parameter of the exponential fit is estimated as significantly greater in the ASD group (p = 0.04). This parameter indicates the intersect with the y-axis rather than the rate of decay, suggesting that the ASD group has a relatively greater proportion of edges that are specific to one task.

We had furthermore mapped the connectivity profile of each ring of edges, here defined as being edges that connect nodes either within or between networks. We quantified this per ring by taking the fraction of the amount of within-network edges and the amount of between network edges (W/B ratio). Higher values here correspond to greater within-network connectivity. Firstly, we observed that this ratio is always greater than one – replicating the finding that tasks preferentially modulate within network edges (Chauvin, Mennes, Buitelaar, & Beckmann, 2017). We further modeled a mean linear relationship between W/B ratio and the ring level using one fit per bootstrap (β = 5.66, SD = 0.60 for ASD, β = 6.49, SD = 0.95 for TD), where we found that edges shared by more tasks are increasingly more likely to be a within network edge. We did not find evidence of these linear effects differing between ASD and TD (p = 0.14). Visually, the TD group appears to recruit more within-network edges across each of the rings, however, this was only significant in ring 3 (p = 0.018) (FDR corrected across the five rings) (Fig. 6).

ASD and ADHD

In previous work, task-potency measures in an ADHD sample have displayed effects with a different direction than those found in the current study (Chauvin et al., 2018). It is further known that ASD and ADHD are frequently co-morbid. To investigate this possible confound, we separated our sample into 'ASD$_{ADHD+}$', consisting of those ASD subjects that met clinical classification cut-offs (DSM-5) for an unofficial ADHD diagnosis, and 'ASD$_{ADHD-}$', consisting of those subjects that did not meet the ADHD cut-off. In both cases, the cut off is defined as having a score of five or greater on either the hyperactivity or inattentiveness ADHD subscales as measured by parent-report. Out of our total sample of 282 ASD subjects, 201 fell into the ASD$_{ADHD-}$ subgroup and 81 fell into the ASD$_{ADHD+}$ subgroup. We remark that this means that some ADHD co-morbidity may be registered in almost 30% of the participants.

Potency profile of the different subgroups

In order to visualize any posited heterogeneity of the subgroups in relation to the total population with respect to their set of selected edges for each of the five tasks, we computed five Venn-diagrams (Fig. 7A). Inspecting the diagrams, we showed a number of trends occurring across the tasks: 1) Selection sizes appeared to be smaller for the subgroups than they are for the total. 2) The ASD$_{ADHD-}$ and the ASD$_{ADHD+}$ subpopulations appeared to cover large unique areas in the ASD$_{TOTAL}$ (ASD$_{ADHD-}$ and ASD$_{ADHD+}$) selected edges profile (yellow and purple areas in Fig. 7). 3) There was negligible overlap between the ASD$_{ADHD-}$ and the ASD$_{ADHD+}$ subpopulations outside of the ASD$_{TOTAL}$ selected edges profile. In order to check

whether effects might have been driven by co-morbid ADHD, we repeated the above analyses of Figure 4 in both the ASD$_{ADHD-}$ and ASD$_{ADHD+}$ subgroups. As the subgroups do not represent well-posed subpopulations, we only look for similarities or divergences in results compared to our original group.
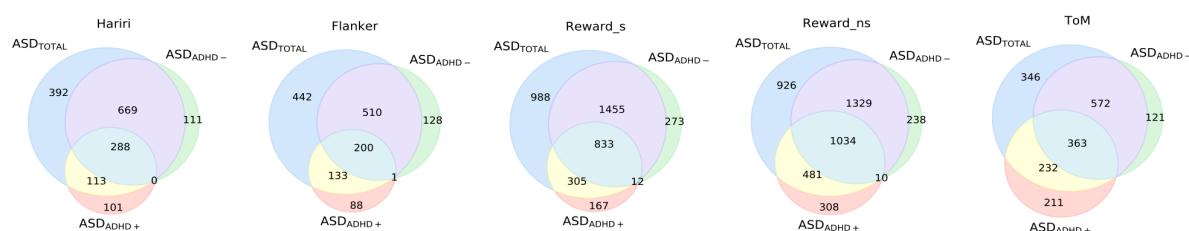
Task-analysis in ASD$_{ADHD-}$

The amount of edges selected by the mixture model as showing strong connectivity modulation in each task appears lower, and hence more similar to TD, when we only use the ASD$_{ADHD-}$ subgroup in the analysis (Fig. 8A). Just as was the case for the comparison for the selection of edges between the ASD$_{TOTAL}$ group and TD, we do not find significant differences for a particular task. Furthermore, when we look at the amplitude present in these edges, we see that the values for the ASD$_{ADHD-}$ subgroup appear closer to normal TD values (Fig. 8B). None of the current comparisons exceed the significance threshold for differences.

Task-analysis in ASD$_{ADHD+}$

We find that the ASD$_{ADHD+}$ group as a whole recruits significantly more edges due to task demands compared to TD in all but the Flanker task (Fig. 9A). Moreover, inspecting the mean potency amplitude of these edges shows us that the potentiation taking place is more powerful in the ASD$_{ADHD+}$ group compared to TD for all the tasks (Fig. 9B).

Modelling potency from symptom scores in ASD and ADHD

	ASD score	ADHD score		ASD score	ADHD score		ASD score	ADHD score		ASD score	ADHD score		ASD score	ADHD score
Potency	0.65	0.79	Potency	0.99	0.34	Potency	0.89	0.99	Potency	0.80	0.34	Potency	0.56	0.98
Potency	0.52	0.34	Potency	0.99	**0.02**	Potency	0.52	0.34	Potency	0.99	0.07	Potency	0.34	0.19
Potency	0.34	0.09	Potency	0.98	**<0.01**	Potency	0.56	0.07	Potency	0.98	**<0.01**	Potency	0.33	**0.02**

Figure 7. Venn diagrams displaying percentage of shared and unique modulated edges for the ASD$_{TOTAL}$ group and its two subgroups, ASD$_{ADHD-}$ and ASD$_{ADHD+}$. The numbers denote the amount of modulated edges in each Venn-set. The table shows significance values for a linear model predicting the mean potency amplitude of edges within the respective Venn-set from ASD and ADHD symptom scores and age.

As a way of investigating whether we can link potency to ASD and ADHD severity, we model the mean potency amplitude of the edges in the respective sets from ASD and ADHD aggregate symptom scores (Fig. 7B). Interestingly, for the Flanker, non-social reward and theory of mind tasks, we find significant positive associations between ADHD symptom scores and the set of edges that was selected in both the ASD_{TOTAL} and ASD_{ADHD+} groups but not the ASD_{ADHD-} group (yellow area in Fig. 7). We furthermore find a significant positive association between ADHD scores and the set of edges selected in all three conditions for the Flanker task. We do not find significant associations between ASD symptom scores and a particular set of edges in any of the current tasks. All significance values are FDR corrected for the 15 possible set x task combinations.

Discussion

Theories on ASD, such as weak central coherence and excitation/inhibition imbalance, suggest pervasive, global alterations in brain connectivity (Hull et al., 2017). For this reason, we have focused our first inquiry on general alterations in the way brain connectivity is modulated in the previously mentioned five tasks. When we compared the percentage of modulated edges away from the baseline architecture of the brain in the overall ASD group (ASD_{TOTAL}) to the TD group, we found no significant differences. However, when we measure the amplitude of modulation, we find that it is lower in the $ASDT_{OTAL}$ group for four of the five tasks under consideration (flanker, social reward, non-social reward, theory of mind). This could mean that even though participants with ASD engage a normal percentage of their connectome in performing
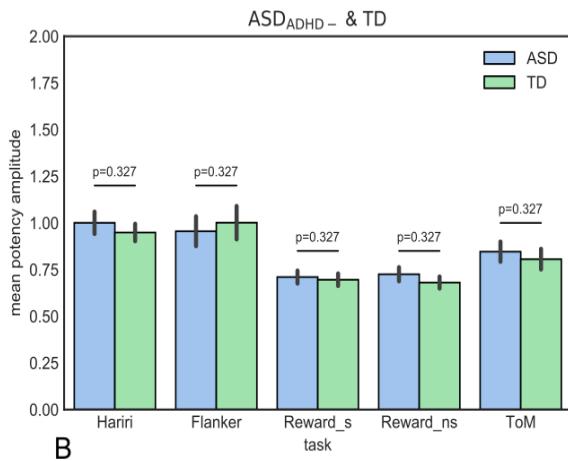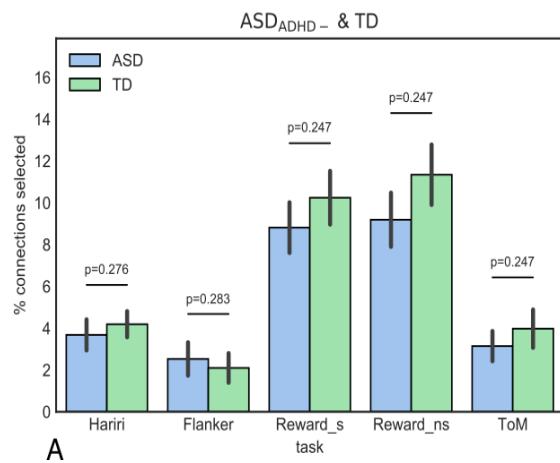

Figure 8A. The percentage of edges that were consistently and strongly modulated in ASD_{ADHD-}. B. The mean potency amplitude of the selected edges in ASD_{ADHD-}.
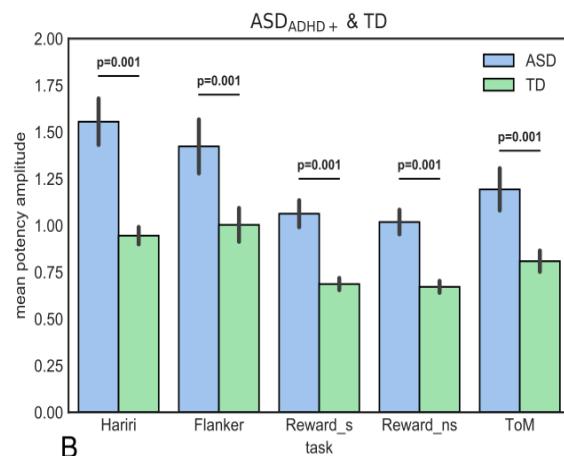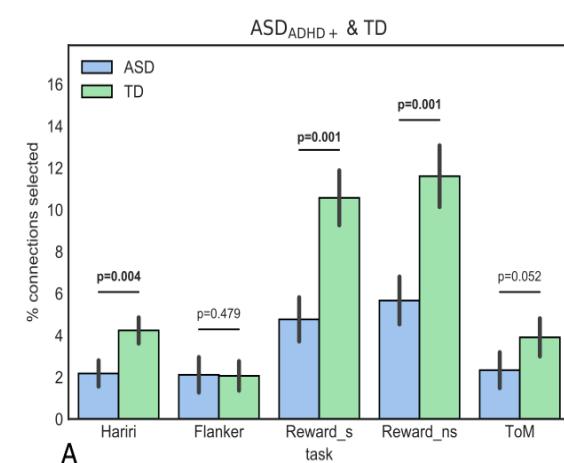
Figure 9A. The percentage of edges that were consistently and strongly modulated in ASD_{ADHD+}. B. The mean potency amplitude of the selected edges in ASD_{ADHD+}.

a task, the altered strength of these modulations could be a reason for their impairments in achieving the respective cognitive goals. We did not find this alteration for the Hariri emotion processing task, suggesting that the lower modulation present in ASD may be specific to certain tasks. Or rather to a cognitive process underlying the above tasks, but not emotion processing. In other words, the task profiles (i.e., which edges are potentiated for a particular task) are similar between ASD and TD, but not the amplitude of their modulation. It is important to reiterate is that the effect of individual differences in base resting state connectivity architecture is explicitly removed from the analyses in this paper. As such, if we do not find differences in task modulation, we might hypothesise that the connectivity alteration previously observed in the resting state (Hull et al., 2017) could be a core alteration influencing the integration of overall information during task performance.

As the tasks involved in this study have been selected for their implication in ASD (Charman et al., 2017), finding altered connectivity in a set of edges common to these tasks could help us in searching for the brain processes at the root of ASD symptoms. This approach was used because it might help us to identify whether any alterations are primarily found in those edges that are specific to certain tasks or rather in edges that are involved in multiple or all the tasks. In order to explore similarities and differences between task modulation profiles, as well as how these might differ between groups, we constructed a Venn-diagram of the sets of edges (task profile) involved in each of the five current tasks. We found for both groups that the majority of edges across the five tasks are task specific, with fewer edges modulated across tasks as we go towards considering sets selected in higher numbers of tasks (i.e., each consecutive ring). Very few edges were common across all five tasks. In order to find out whether this pattern differed between the groups, we modelled this 'decay' from specific to common edges for both groups with an exponential curve. We found that the model showed some difference in the initial state but importantly no difference in the rate of decay, providing no evidence that the task profile in ASD is altered. This contrasts for instance with what has been found for ADHD (Chauvin et al., 2018), where a lack of common processing compared to TD has been identified.

The weak central coherence theory of autism refers to the apparent cognitive deficits individuals with ASD have with fluently integrating/segregating information (e.g., 'not seeing the forest for the trees') (Brunsdon & Happé, 2014). If we make the biologically plausible assumption that adjacent types of information processing occur in adjacent brain regions, this allows us to investigate this cognitive theory aided by fMRI data. Adjacency can be defined, for instance, in terms of graph theoretical distance or Euclidean distance. Indeed, previous evidence in line with the weak central coherence hypothesis of ASD suggests that there may be resting state underconnectivity at short range and overconnectivity at long range in the ASD brain (Happé & Frith, 2006; Hull et al., 2017). Adjacency can, however, also be defined in functional terms (e.g., as resting state networks). The task-potency method allowed us to assess whether such a pattern can also be observed when it comes to the connectivity modulation that occurs during a task. Because we make use of a hierarchical atlas (11 large-scale networks, in which 168 smaller-scale areas are situated), we have the ability to classify edges as being either within or between networks, which served as our measure for adjacency of processing. We found that edges in both groups become more likely to be within-network edges as we look at the sets that are selected in more tasks. In other words, processing that is more general and shared across multiple cognitive domains is much more likely to take place within the same networks when compared to processing that is only implicated in a single task. Contrary to the literature that has associated ASD with short range overconnectivity, we have not found evidence for this in our analyses. We rather find that in ring 3, the TD group has significantly greater within-network oriented edges. We can hypothesise that only observing an effect in ring 3 might be due to a particular interaction between three of the tasks, which fall under this umbrella. More investigation is needed by selecting specific tasks together, as we cannot compare our results to the existing literature on short- and long-range connectivity. Indeed, we analysed modulation of connectivity during tasks rather than resting-state connectivity and we cannot compute the distance in terms of graph theoretical metrics as we focused on a subset of edges modulated during the task and removed the underlying architecture that connects them.

In order to investigate variability and heterogeneity in our ASD sample and because previous findings have shown that the task profile is altered in ADHD (Chauvin et al., 2018), we decided to compare the task profiles of an ASD group, which also scored high on ADHD symptoms, to

the rest of the ASD sample. Because of the high levels of ADHD co-morbidities known to exist in ASD cohorts, we were able to split the ASD group up into two groups, ASD_{ADHD-} and ASD_{ADHD+} with respectable sizes. In the ASD_{ADHD-} group, where participants with ADHD were excluded, we did not find any effect for either percentage of modulated edges or their amplitude. This might suggest that the amplitude effect, that was found in the original ASD_{TOTAL} group, could have been in fact driven by the subjects with ADHD. Though we find a significant amplitude effect in ASD_{ADHD+} for each of the tasks, they have an opposite direction to the effects found for ASD_{TOTAL}, with the ASD_{ADHD+} group rather having greater modulation amplitudes than TD. The fact that the amplitude findings of ASD_{TOTAL} are not a simple weighted average of ASD_{ADHD+} and ASD_{ADHD-} implies that these groups must have had differences in the set of edges from which this amplitude was calculated. This led us to computing the selected edges for each of the five tasks for these new groups and we found that they mostly make use of different sets of edges to perform tasks. What appeared to be a confound due to heterogeneity in the population now suggests that our method is highly sensitive to group differences even if they come from subgroups. Further research is needed to investigate how this generalises to other ways of splitting populations and whether it can be used to correctly classify individual participants. For example, we can try to assess whether observed differences are characteristics of the entire population by using this as input in a normative modelling analysis (Marquand, Wolfers, Mennes, Buitelaar, & Beckmann, 2016). In this approach, participants' deviance from a norm defined in the sample is assessed in order to capture subgroup effects or even individual profiles.

We found a significant positive association of ADHD symptom scores and mean potency amplitude in the Flanker, non-social reward and theory of mind tasks. This was the case when we considered a set of selected edges that was common between ASD_{TOTAL} and ASD_{ADHD+} but not ASD_{ADHD-}. These particular tasks have a strong executive function component, which is known to be impaired in ADHD (Gargaro et al., 2011). This finding may form a steppingstone to a possible clinical/diagnostic application of task-potency in ADHD. Further research can be done in a cohort of ADHD participants to narrow down which edges are responsible for driving this effect and to find out whether we have identified an ADHD-related effect

or rather an ASD-ADHD co-morbidity effect.

In the current research, we take a holistic perspective at brain connectivity modulation. We analyse characteristics of the sets of selected edges independently of their spatial localisation, taking a general view on whole brain connectivity. Even when investigating edges as being within or between networks, where we carry out a slightly finer grained analysis, we still aggregate across tasks and loose some resolution in that dimension. It stands to reason that there may exist many relevant differences between (and within) the groups under consideration with respect to task-potency that we have not detected, which are left to future research. As it is known that ASD has a component of sensory processing abnormalities, it could for instance be interesting to compare and contrast the task profiles of sensory versus higher-order regions.

In summary, we found that participants with ASD less strongly modulate those parts of their functional connectome that are involved in tasks. However, splitting the ASD group with respect to ADHD co-morbidity, we find that for ASD_{ADHD-} these effects disappear whereas for ASD_{ADHD+} we find strong differences in their task profile compared to controls. We furthermore found that as edges are involved in performing multiple tasks, they are proportionally more likely to be within network edges. Lastly, we related ADHD symptom scores to mean potency amplitude in a set of edges but were not able to do the same for ASD symptom scores. We have shown that task-potency is a method of integrating task fMRI data which shows promise in the domains of parsing heterogeneous clinical groups as well as predicting symptom scores from fMRI data.

References

Andersson, J. L. R., Jenkinson, M., & Smith, S. (2007). Non-linear registration, aka spatial normalisation. FMRIB Technial Report TR07JA2. Oxford Centre for Functional Magnetic Resonance Imaging of the Brain, Department of Clinical Neurology, Oxford University, Oxford, UK(June), 22.

Bielczyk, N. Z., Walocha, F., Ebel, P. W., Haak, K. V., Llera, A., Buitelaar, J. K., ... & Beckmann, C. F. (2018). Thresholding functional connectomes by means of mixture modeling. NeuroImage.

Brunsdon, V. E., & Happé, F. (2014). Exploring the 'fractionation' of autism at the cognitive level. Autism, 18(1), 17–30.

Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the

attribution of mental states to animated shapes. Brain: a journal of neurology, 125(8), 1839–1849.

Charman, T., Loth, E., Tillmann, J., Crawley, D., Wooldridge, C., Goyard, D., ... & Buitelaar, J. K. (2017a). The EU-AIMS Longitudinal European Autism Project (LEAP): Clinical characterisation. Molecular Autism, 8(1), 1–21.

Charman, T., Loth, E., Tillmann, J., Crawley, D., Wooldridge, C., Goyard, D., ... & Buitelaar, J. K. (2017b). The EU-AIMS Longitudinal European Autism Project (LEAP): Methods. Molecular Autism, 8(1), 1–19.

Chauvin, R., Oldehinkel, M., Buitelaar, J. K., Beckmann, C. F., & Mennes, M. (2018). Inefficient use of the brain's functional infrastructure in ADHD Abstract. bioRxiv.

Chauvin, R. J., Mennes, M., Buitelaar, J. K., & Beckmann, C. F. (2017). Assessing age-dependent multi-task functional co-activation changes using measures of task-potency. Developmental Cognitive Neuroscience (March), 0–1.

Chauvin, R. J., Mennes, M., Llera, A., Buitelaar, J. K., & Beckmann, C. F. (2019). Disentangling common from specific processing across tasks using task potency. NeuroImage, 184(September 2018), 632–645.

Delmonte, S., Balsters, J. H., McGrath, J., Fitzgerald, J., Brennan, S., Fagan, A. J., & Gallagher, L. (2012). Social and monetary reward processing in autism spectrum disorders. Molecular Autism, 3(1), 1–13.

Deshpande, G., Libero, L. E., Sreenivasan, K. R., Deshpande, H. D., & Kana, R. K. (2013). Identification of neural connectivity signatures of autism using machine learning. Frontiers in Human Neuroscience, 7(October), 1–15.

Di Martino, A., Kelly, C., Grzadzinski, R., Zuo, X. N., Mennes, M., Mairena, M. A., ... & Milham, M. P. (2011). Aberrant striatal functional connectivity in children with autism. Biological Psychiatry, 69(9), 847–856.

Gargaro, B. A., Rinehart, N. J., Bradshaw, J. L., Tonge, B. J., & Sheppard, D. M. (2011). Autism and ADHD: How far have we come in the comorbidity debate? Neuroscience and Biobehavioral Reviews, 35(5), 1081–1088.

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. Journal of Autism and Developmental Disorders, 36(1), 5–25.

Hariri, A. R., Tessitore, A., Mattay, V. S., Fera, F., & Weinberger, D. R. (2002). The amygdala response to emotional stimuli: A comparison of faces and scenes. NeuroImage, 17(1), 317–323.

Hull, J. V., Jacokes, Z. J., Torgerson, C. M., Irimia, A., Van Horn, J. D., Aylward, E., ... & Webb, S. J. (2017). Resting-state functional connectivity in autism spectrum disorders: A review. Frontiers in Psychiatry, 7(JAN).

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage, 17(2), 825–841.

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. NeuroImage, 62(2), 782–790.

Just, M. A., Cherkassky, V. L., Buchweitz, A., Keller, T. A., & Mitchell, T. M. (2014). Identifying autism from neural representations of social interactions: Neurocognitive markers of autism. PLoS ONE, 9(12), 1–22.

Ledoit, O., & Wolf, M. (2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices. Annals of Statistics, 40(2), 1024–1060.

Marquand, A. F., Wolfers, T., Mennes, M., Buitelaar, J., & Beckmann, C. F. (2016). Beyond Lumping and Splitting: A Review of Computational Approaches for Stratifying Psychiatric Disorders. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1(5), 433–447.

Mennes, M., Kelly, C., Zuo, X. N., Di Martino, A., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2010). Inter-individual differences in resting-state functional connectivity predict task-induced BOLD activity. NeuroImage, 50(4), 1690–1701.

Meyer-Lindenberg, A., & Weinberger, D. R. (2006). Intermediate phenotypes and genetic mechanisms of psychiatric disorders. Nature Reviews Neuroscience, 7(10), 818–827.

Nunes, A. S., Peatfield, N., Vakorin, V., & Doesburg, S. M. (2018). Idiosyncratic organization of cortical networks in autism spectrum disorder. NeuroImage.

Oldehinkel, M., Mennes, M., Marquand, A., Charman, T., Tillmann, J., Ecker, C., ... & Buitelaar, J. K. (2019). Archival Report Altered Connectivity Between Cerebellum, Visual, and Sensory-Motor Networks in Autism Spectrum Disorder: Results from the EU-AIMS Longitudinal European Autism Project.

Pruim, R. H., Mennes, M., Buitelaar, J. K., & Beckmann, C. F. (2015). Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. NeuroImage, 112, 278–287.

Pruim, R. H., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J. K., & Beckmann, C. F. (2015). ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. NeuroImage, 112, 267–277.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., ... & Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. Proceedings of the National Academy of Sciences, 106(31), 13040–13045.

van Oort, E. S., Mennes, M., Navarro Schröder, T., Kumar, V. J., Zaragoza Jimenez, N. I., Grodd, W., ... & Beckmann, C. F. (2017). Functional parcellation using time courses of instantaneous connectivity. NeuroImage(November 2016), 1–10.

White, S. J., Coniston, D., Rogers, R., & Frith, U. (2011). Developing the Frith-Happé animations: A quick and objective test of Theory of Mind for adults with

autism. Autism Research, 4(2), 149–154.

Yerys, B. E., Herrington, J. D., Satterthwaite, T. D., Guy, L., Schultz, R. T., & Bassett, D. S. (2017). Globally weaker and topologically different: resting-state connectivity in youth with autism. Molecular Autism, 8(1), 39.

Abstracts

Proceedings of the Master's Programme Cognitive Neuroscience is a platform for CNS students to publish their Master theses. Given the number of submissions, we select the articles that received the best reviews, under recommendation of our editors, for the printed edition of the journal. The abstracts of the other articles are provided below, and for interested readers a full version is available on our website: www.ru.nl/master/cns/journal.

Visual Attention Through Uncertainty Minimization in Recurrent Generative Models

Kai Standvoss

Allocating visual attention through saccadic eye movements is a key ability of intelligent agents. Attention is both influenced through bottom-up stimulus properties as well as topdown task demands. The interaction of these two attention mechanisms is not yet fully understood. A parsimonious reconciliation posits that both processes serve the minimization of predictive uncertainty. We propose a recurrent generative neural network model that predicts a visual scene based on foveated glimpses. The model shifts its attention in order to minimize the uncertainty in its predictions. We show that the proposed model produces naturalistic eye-movements focusing on salient stimulus regions. Introducing the additional task of classifying the stimulus modulates the saccade patterns and enables effective image classification. Given otherwise equal conditions, we show that different task requirements cause the model to focus on distinct, task-relevant regions. The model's saccade statistics correspond well with previous experimental data in humans and provide insights into unsettled controversies in the literature. The results provide evidence that uncertainty minimization could be a fundamental mechanisms for the allocation of visual attention.

Adaptation to Accents by learning from Mistakes - Event Related Potentials during error-based Phonological Learning

Ronny Bujok

If people are exposed to an unfamiliar accent, they can usually learn to understand it and adapt to its specific patterns of pronunciation very quickly. However, the mechanisms involved in this learning process are still not fully understood. In this study we suggest that such phonological learning is based on external feedback processing and internal monitoring. When we make errors in our phonological perception and receive feedback, we process it and feed it into our internal monitoring, which should become better at detecting errors itself over time, and lead to better performance. We tested a group of Dutch native speakers in a novel accent learning task. Participants were presented with accented forms of Dutch words, pronounced according to an artificially created, novel accent. From two visually presented words, which differed from one another and the accented word only in their vowel, participants had to choose the correct word that corresponded to the accented word they had heard. They received corrective feedback on their choice to learn the mapping of the phonemes of the novel accent. We measured the electrophysiological activity (EEG) and analyzed response-locked (ERN & Pe) and feedback-locked (FRN & P300) data to investigate the mechanisms of this learning process. Participants adapted to the accent and they showed internal monitoring and feedback processing throughout the entire task. Internal monitoring was already present in the beginning of the experiment and no change over time could be found, indicating that learning had already taken place in the very beginning of the experiment and thus that phonological learning is very quick. The results suggest that phonological learning can indeed be described in terms of development of specific internal monitoring processes based on external feedback on erroneous responses.

Neural variability quenching during the planning of reaching movements

Debora Nolte

While the reduction (quenching) of trial-by-trial variability of neural activity during repeated visual stimulus presentation determines perceptual performance, it is still unknown whether quenching occurs and has any behavioral relevance in the motor system. We investigated neural variability quenching during delayed reaching movements to targets of two different sizes (large versus small). Previous studies have shown that reaches to large targets induce more task-irrelevant variability than those too small targets, but it is unclear if this can be related to neural activity during movement planning. Using electroencephalography (EEG), the participants' level of neural variability quenching during the planning of reaching movements was assessed at two channels, one covering the occipital cortex, the other the motor cortex. The broadband signal at both of these channels did not show any signs of quenching during the movement planning phase and therefore also no relation of to variability in behavior. However, both the alpha and beta band showed general quenching during the movement preparation at both channels, but again no relation to behavior could be observed. To our knowledge, these findings are the first indication of neural variability quenching in the motor domain, but without a relation to the behavioral performance.

Multimodal imaging of compulsivity across neurodevelopmental disorders: A longitudinal investigation

Viola Hollestein, Jan K. Buitelaar, Daniel Brandeis, Sarah Durston, Steven C. R. Williams & Jilly Naaijen

Background: Autism spectrum disorder (ASD) and obsessive compulsive disorder (OCD) are neurodevelopmental disorders characterized by compulsivity. Deficits in cognitive control have been associated with compulsivity and fronto-striatal alterations, which seem to be regulated by glutamate. In the current study we will link these findings together by investigating fronto-striatal neurochemical and functional alterations during inhibitory control in association with the cross-disorder trait compulsivity.

Methods: Participants were part of the European TACTICS consortium and included adolescents with ASD, OCD and controls aged 8-17. Continuous measures of compulsivity and impulsivity were used to investigate the association with fronto-striatal glutamate (1H-MRS), and functional activity (fMRI) during inhibitory control. These analyses were performed using linear mixed effects models.

Results: We found that ACC glutamate levels decreased significantly more over time in ASD compared to OCD and controls. In the striatum, a decrease in striatal glutamate over time across all groups was found. During successful inhibitory control, increased glutamate in the striatum was associated with decreased ACC BOLD, independent of diagnosis. Additionally, also during successful inhibitory control, increased ACC glutamate was associated with increased striatal BOLD, but in the OCD group only.

Conclusions: Our findings suggest that glutamate in ACC and striatum changes with development and may affect BOLD patterns during inhibitory control differently in ASD and OCD. This suggests that glutamate may play a modulatory role in compulsivity.

A vestibular-derived internal model for task-dependent reaching

Anne Hoffmann

To plan and control arm movements while the body is in motion the brain needs to compensate for inertial forces acting on the arm. Here we investigated, if vestibular information is used to adapt reach-planning by contributing to the formation of an internal model of the arm dynamics. Specifically, we were interested in tracking the development of such an internal model over time. In addition we examined whether vestibular-mediated reach-control is modulated by task-demands. A linear sled was used to passively move participants while they performed reaching movements. Participants learned an association between reaching direction and the direction of a simultaneous sideways body translation. Randomly scattered trials in which the expected body motion was omitted were used to probe learning and induce online corrections. Task-demand was manipulated by using narrow and wide reach-targets. Recorded hand-trajectories were analyzed for adaptation-effects in deviations perpendicular to the reach direction. The analysis of task-modulations on online corrections focused on differences in movement endpoints and durations. We observed an adaptation of hand trajectories during body motion as well as indications for aftereffects when body motion was omitted. Further, both reaches with and without body motion showed increased endpoint variability and shorter durations for wide compared to narrow targets. Our findings support previous research showing that vestibular information is used to adapt reaching movements during whole-body motion. The observed aftereffects suggest that this adaptation relies on the formation of an internal model. Lastly, our results corroborate previous findings showing task-dependent modulations of vestibular feedback gains.

Development of prediction in the ventral visual stream

Ellen Saskia Voorrips

Predicting future states of the environment enables us to act rapidly and efficiently in the world around us. According to predictive coding theories, mismatches between our expectations and incoming sensory inputs are represented by the amount of neural activation. Therefore, predicted events lead to relatively reduced neural activation compared to unpredicted events. Events become more predictable as they often follow a preceding event. It has been established that expected stimuli lead to attenuated neural response throughout the ventral visual stream after statistical learning, although it remains unclear what happens during the acquisition of these associations. By using a two-session MRI experimental setup, the effect of expectation on visual processing is investigated during an initial learning session (session one) and after statistical learning has taken place (session two). Unexpectedly, results indicate no effect of expectation during session two. A trend of expectation enhancement is visible in session one. Interestingly, a general decrease in neural response was found in session two compared to session one. Several explanations are proposed to account for the obtained results, ranging from the multimodal quality of the stimulus used, to the design of the study. Further research is necessary to determine which explanation is driving the absence of an effect of expectation.

An exploratory fMRI study on metonymy and metaphor processing

Sofia Fregni, Karin Heidlmayr, Kirsten Weber, & David Peeters

Metaphor and metonymy are believed to play a key role in human communication, language, and cognition. However, it remains unclear what neural mechanisms allow for their comprehension, and whether the brain regions involved during their processing reflect enhanced linguistic processing or inferential processes. Through the use of fMRI and two localizer tasks, for language and Theory-of-Mind, we investigated the mechanisms underlying the comprehension of metaphoric and metonymic language. Both types of figurative language elicited increased left-lateralized frontotemporal activity, with overlaps across regions belonging to both the Theory-of-Mind network and the language network. Metonymy processing additionally recruited the right inferior frontal gyrus, whereas metaphor comprehension downregulated the frontoparietal control system. Results are discussed in terms of the degree of semantic distance between the metaphoric and metonymic allusions and their referent.

Altered motor imagery-related cerebral activity in patients with Neuralgic Amyotrophy
Central neuroplasticity after peripheral dysfunction

Elze Wolfs

Peripheral nerve disorders often lead to central neuroplasticity and cerebral functional reorganization. We hypothesize that this also occurs in patients with the peripheral neuromuscular disorder Neuralgic Amyotrophy (NA). Patients with NA generally cannot move their arm properly due to damage to brachial plexus nerves. We hypothesize that, when there is damage, motor planning in the central nervous system changes. Fortunately, these peripheral nerves recover after time, but patients can often still not move their affected arm as before. Therefore, there may be maladaptive motor planning. In our study population of 22 patients and 11 age- and gender matched healthy controls, motor planning was assessed by use of Parson's laterality judgement task, an implicit motor imagery task, inside the MRI scanner. All participants were right-handed and all NA patients were affected on the right side. Behaviorally, participants performed in accordance with previous research on this task on reaction time. Participants were slower to identify non-dominant (left) hands compared to dominant (right) hands and they were slower on biomechanically complex movements (indicating motor imagery). NA patients were slower when their posture did not match the stimulus (indicating first-person motor imagery), but healthy controls were equally fast for matching and non-matching stimuli, which could be an indication of third-person motor imagery. All patients were affected unilateral on the right side, but there was no association between errors made for the right or left side. The only behavioral difference found between NA patients and healthy controls was the effect of posture. Cerebrally, all participants on average showed activation in brain areas previously associated with motor imagery, such as the parietal-frontal network, Extrastriate Body Area and several motor areas. In addition, we found that NA patients, in comparison with healthy controls, showed additional activation at the ipsilateral Supramarginal Gyrus/Post-Central Gyrus when performing motor imagery for right hand stimuli compared to left hand stimuli. This could be an indication of a compensatory mechanism that patients use to be able to move their affected hand. However, this activity could also solely be an effect of a noisy baseline. More data is necessary to provide a clearer view on the difference in motor planning between NA patients and healthy controls and to provide insights into future rehabilitation programs.

SpeechTrackinginInfantswith(FamilyHistoryof)Autism
The Relationship of Early Speech Tracking with Language and Autism Development

Katharina Menn

Adults with autism spectrum disorder (ASD) show decreases in speech tracking, the brain's ability to take over the rhythm of speech. On an individual level, the amount of speech tracking is directly related to verbal abilities and autism symptoms. In the current study, we investigate whether differences in speech tracking during infancy can explain the deficit in language acquisition often observed in children developing ASD. We measured speech-brain coherence in a total of 22 infants at risk for developing ASD because of family history at either 10 or 14 months of age. They were compared to a control group of 20 infants without family history of autism. We analyzed the relationship between speech-brain coherence in these infants and their vocabulary knowledge at 24 months (measured by the CDI) as well as autism symptoms at 36 months (measured by the ADOS). Our results showed significant speech-brain coherence in infants. Importantly, speech-brain coherence in two distinct frequencies: The stressed syllable rate (1-3 Hz) and the phonological rate (5-15 Hz) predicted later vocabulary knowledge. This effect was stronger for 10-month-old infants than for 14-month-olds. We found no relationship between speech tracking in infancy and autism symptoms in childhood. Thus, early speech tracking is related to language development in childhood. The mechanism behind this relationship as well as generalizability to all infants are discussed.

Hypomyelination of interneurons in the APO-SUS rat model for schizophrenia

Noor van Vugt

Schizophrenia is a neurodevelopmental disorder that affects about 0,5% of the world population. Symptoms can be categorized into positive, negative, and cognitive clusters. Although most research and treatment currently is focused on the positive symptoms (such as hallucinations), cognitive symptoms represent another core feature of the disorder. The current study investigates the neurobiological mechanism underlying cognitive symptoms of schizophrenia by using the APO-SUS rat model for schizophrenia. Using electron microscopy, we found that the prefrontal cortex (PFC) of APO-SUS rats contains fewer myelinated axons than the PFC of APO-UNSUS rats. This effect is normalized by environmental enrichment of the home cages of the APO-SUS rats. Additionally, using immunofluorescence on 70 nm ultrathin slices of PFC, we were able to visualise individual myelin sheaths and intra-axonal GABA, allowing us to quantify the density of myelinated interneurons. We found that in APO-SUS PFC a lower percentage of myelinated axons was GABAergic than in APO-UNSUS PFC. Together, these data show that in the APO-SUS rat model for schizophrenia interneurons are subject to hypomyelination, and that this defect can be rescued by environmental enrichment.

The effect of an active-controlled 8-week Mindful Eating intervention on food-related attentional control

Demi Ennen

Background: Attentional control is needed to overcome an attentional bias towards food cues. Overcoming this bias is important in avoiding unhealthy eating behaviors, such as overeating. Increasing attentional control will thus help to improve eating behaviors. Attention is an important aspect of Mindfulness Based Interventions (MBIs). While it is shown that these interventions are able to improve eating behaviors, the underlying mechanism is unknown. Improved attentional control towards food cues might play a role.
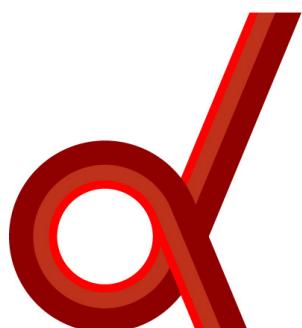
Aim: To investigate whether food-related attentional control is improved by a Mindful Eating intervention and whether activation changes occur in attentional control-related brain areas, thereby exploring whether this might underlie the beneficial effects of MBIs on eating behaviors.

Method: 44 healthy subjects who were motivated to change their eating habits, participated in an 8-week Mindful Eating intervention or Educational Cooking intervention (active control). Attentional control was measured with a food Stroop task in an MRI scanner. For this task, changes in reaction times and BOLD responses from pre- to post-intervention were analyzed.

Results: No differences in reaction times and BOLD responses were found between the intervention groups from pre- to post-intervention, but there were differences when looking over both time points and intervention groups. There was a significant main effect of food interference words (versus positive emotional words) on reaction times. For food interference versus neutral words, the left inferior and superior frontal gyrus and the left middle temporal gyrus showed specific activation. The left inferior frontal gyrus, the left inferior parietal lobule, and the left angular gyrus showed greater activation for food words versus positive emotional words.

Conclusion: There was no difference in changes in attentional control between the Mindful Eating intervention and the active control intervention. This implies that changes in attentional control might not underlie the positive effects of MBIs on eating behavior.

Institutes associated with the
Master's Programme Cognitive Neuroscience

Donders Institute for Brain, Cognition
and Behaviour:
Centre for Cognitive Neuroimaging
Kapittelweg 29
6525 EN Nijmegen

P.O. Box 9101
6500 HB Nijmegen
http://www.ru.nl/donders/

MAX PLANCK INSTITUTE
FOR PSYCHOLINGUISTICS

Max Planck Institute for Psycholinguistics
Wundtlaan 1
6525 XD Nijmegen

P.O. Box 310
6500 AH Nijmegen
http://www.mpi.nl

Radboudumc

Radboudumc
Geert Grooteplein-Zuid 10
6525 GA Nijmegen

P.O. Box 9101
6500 HB Nijmegen
http://www.umcn.nl/

Baby & Child
Research Center

Baby Research Center
Montessorilaan 3
6525 HR Nijmegen

P.O. Box 9101
6500 HB Nijmegen
http://www.babyresearchcenter.nl