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

Contact Information:
Journal CNS
Radboud University
Postbus 9104
6500 HE Nijmegen
The Netherlands
nijmegencns@gmail.com

Table of Contents

From the Editors-in-Chief

Dear Reader,

It is our pleasure to introduce to you the first issue of the 12th volume of the *Proceedings of the Master's Programme Cognitive Neuroscience*. The journal continues in its proud tradition of publishing excellent research articles from graduated master students, providing a perfect opportunity to immerse oneself in the diverse and exciting field that the Cognitive Neuroscience Programme has to offer.

This diversity is a great strength. Not only does our programme unite talented students with multifaceted backgrounds from all across the world, but offers ample opportunity to explore neuroscience to one's heart desires – from Neuroimaging to Neurophilosophy, from Neurodevelopmental Disorders to Brain-Computer Interfacing. The programme also provides state-of-the-art equipment and facilities and allows close contact with experts of the respective fields. The current issue's authors reflect this well: Diverse in their research ideas, techniques and hypotheses, united in their passion for neuroscience. This issue covers aggression, verbal memory development, the effect of the gut microbiome on cognition, post-traumatic stress disorder, familiarity vs. expectation's effect on the brain, and reference processing. Further submissions that could not be included in the printed edition can be found on our website[1] . We congratulate and thank all authors for their high effort achievements.

Next to the authors, we would like to thank our reviewers. Without their insightful, quality-insuring assessment, the journal would surely not hold up to the high standards that it has displayed time and time again. We are especially grateful to the journal team and their continuous dedication to design, review, edit, advertise and do so much more to make the journal you are holding in your hands a pleasant and informative experience. Nothing would get done without their professional acumen and devotion and we could not imagine anyone better to work besides.

Finally, we would like to thank our kind readers for their continued interest. Our readership makes publishing the *Proceedings of the Master's Programme Cognitive Neuroscience* ever so fulfilling.

We hope you enjoy reading this issue as much as we enjoyed working on it!

Nijmegen, February 2017

Kim Fricke & Yvonne Visser

Editors-in-Chief

[1] http://www.ru.nl/master/cns/journal

From the Director of the Max Planck Institute for Psycholinguistics

It is the Facts, Stupid

The major scientific discovery in 2016 was undoubtedly the registration of the gravitational waves. Their existence was predicted exactly 100 years ago by Albert Einstein. Einstein, however, had also predicted that detecting these waves would never be possible. His first prediction was correct, his second was wrong. The joint efforts (team science) of more than 1000 scientists resulted in measuring gravitational waves for the first time on September 14, 2015, that happened as a result of a collapse of two black holes 1.3 billion years ago. As a result of this event for a period of 0.2 seconds the 1000 kilometer distance between Berlin and Paris changed in length by the absolutely minute amount of 10^{-13} cm. That this was measurable is a triumph for the ingenuity of the men and women involved, and of the theoretical precision of Einstein's theory of general relativity.

For me personally this was a decisive moment in 2016, not only because of the triumph for a scientific theory, but also, and maybe even more so, because it illustrates that there are facts out there that can be measured with great precision and reliability.

This is comforting in a period in human history in which postmodernist philosophers have told us that there are no objective facts, but only subjective constructions of the mind. In line with their arguments these days it is a commonly heard statement that "science is also only just an opinion", not better than phantasies or statements without supporting evidence. "Climate change is a Chinese lie, but not a reality"; "HIV does not exist"; "Autism is caused by vaccination". Statements such as these are made without the felt urge to back them up by evidence.

This CNS Master's journal is exemplary for the opposite point of view. Despite differences in nationality, gender, culture, age, etc., we can agree on what the facts are. In science there are universally accepted rules of establishing the evidence for our statements. We know on how to decide about right and wrong. This is a great good for mankind that should be defended at all costs. Of course, the game of science is not perfect, the rules are not always followed with the required adequacy; in some cases the rules need to be improved. But these are minor issues compared to the overall good of science, which is the hunt for truth. Bill Clinton won the 1992 presidential campaign against George H.W. Bush with the slogan, "It is the economy, stupid".

In this time, we should stick to our guns and tell the world "It is the facts, stupid". The discovery of the gravitational waves confirmed that there are facts, that they can be established with great precision, that they are backed up by evidence, and moreover, that they matter. So, whatever the specific qualities are that you acquired during your training, this is the lesson that I hope all of you learned and that will guide you during the rest of your lives, whatever the ideological currency of the day is: "It is the facts, stupid".

Nijmegen, February 2017

Prof. Dr. Peter Hagoort

Director of the Donders Centre for Cognitive Neuroimaging,
Director of the Max Planck Institute for Psycholinguistics

Beyond Aggression: Characterising the Phenotype of the BALB/cJ Mouse

Sabrina van Heukelum[1]
Supervisors: Martha N. Havenith[2], Amanda Jager[2], and Shaghayegh Abghari[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

Children with conduct disorder (CD) show high levels of aggression. Common comorbidities of CD are attention-deficit-hyperactivity disorder (ADHD) and autism spectrum disorder (ASD). Here, we further characterised the behavioural and physiological phenotype of BALB/cJ mice, known for their increased levels of aggression. We investigated if they show symptoms of ASD (social withdrawal) and ADHD (hyperactivity and low temperature). In experiment 1, social withdrawal was investigated with a modified version of the three-chamber social interaction test. In experiment 2, telemetric devices were implanted to measure locomotion and body temperature for a period of 86 hours. In the social interaction test, BALB/cJ mice showed less interest in an unfamiliar mouse compared to BALB/cByJ mice. Experiment 2 showed that BALB/cJ mice had increased locomotor activity during the active period and a lower body temperature in the non-active period compared to control mice. In this study, we have further characterised the behavioural and physiological phenotype of BALB/cJ mice, demonstrating that these mice show symptoms of CD, and its associated comorbidities, ASD and ADHD. The model can be used to study brain structures that might give rise to the linked symptoms of CD, ADHD and ASD.

Keywords: conduct disorder, autism spectrum disorder, attention-deficit hyperactivity disorder, comorbidities, animal model, BALB/cJ mouse

Corresponding author: Sabrina van Heukelum; **E-mail:** Sabrina.vanHeukelum@radboudumc.nl

From an evolutionary perspective, aggression increases an individual's chance of survival. It facilitates the competition for resources and protection of the individual or its offspring (Coppens, De Boer, Buwalda, & Koolhaas, 2014). However, to avoid negative consequences, aggressive behaviour needs to be proportionate to the provocation or context and is subjected to strong inhibitory mechanisms (De Boer, van der Vegt, & Koolhaas, 2003). If these inhibitory mechanisms fail, aggressive behaviour can quickly escalate and lose its adaptive function in social interaction (De Boer, Caramaschi, Natarajan, & Koolhaas, 2009). Aggressive behaviour is among the most common causes of referrals to child and adolescent psychiatrists (Gurnani, Ivanov, & Newcorn, 2016) and often, conduct disorder (CD) is the diagnosis (Blair, 2013; Finger et al., 2012; Serper, Beech, Harvey, & Dill, 2008).

CD is characterised by a persistent pattern of disruptive behaviour, violating the basic rights of others and societal norms (Loeber, Burke, Lahey, Winters, & Zera, 2000). Children with CD show high amounts of aggression, fighting, bullying or being cruel to others and animals (Finger et al., 2012). They are less empathic, force someone into sexual activity or run away from home. Often, they already show these symptoms at a young age (Blair, 2013). Prognosis is poor; children diagnosed with CD show high rates of domestic violence, unemployment and homelessness in adulthood, and about 54% receive a diagnosis of antisocial personality disorder (APD) later in life (Blair, 2013; Loeber et al., 2000; Noordermeer, Luman, & Oosterlaan, 2016). Multiple brain circuits have been implicated in the aetiology and maintenance of CD, for instance, the basic threat circuitry (amygdala–hypothalamus–periaqueductal grey [PAG]), the hypothalamus–pituitary–adrenal system (HPA axis), as well as frontal circuits (anterior cingulated cortex [ACC] and ventromedial prefrontal cortex [vmPFC]). However, not all patients with CD demonstrate the same pathophysiology as different symptom clusters and different comorbidities can be present.

Common comorbidities of CD are autism spectrum disorder (ASD) and attention-deficit hyperactivity disorder (ADHD). About 30% of children with CD receive a diagnosis of ASD and even more than 50% receive a diagnosis of ADHD. ASD is characterised by severe and pervasive deficits in social interaction and repetitive or stereotyped behaviours (Glickman, 2010; McDougle, Stigler, & Posey, 2003). ADHD is characterised by a deficit in attention, increased impulsivity and hyperactivity, and in children with ADHD that develop CD, the age of onset of CD is considerably lower than in children with CD only (Loeber et al., 2000). Furthermore, children with comorbid ADHD demonstrate higher levels of physical aggression, delinquency and more severe symptoms of both CD and ADHD than children with CD or ADHD only (Noordermeer et al., 2016).

In the last two decades, increasingly more knowledge on structural and functional abnormalities in CD, ADHD and ASD has been generated but effective treatments are sparse and long-term prognosis is poor (Casanova, 2007; Esbensen, Greenberg, Seltzer, & Aman, 2009; Matthys, Vanderschuren, & Schutter, 2012; Nestler & Hyman, 2010). It is hypothesised that CD, ADHD and ASD share a common underlying aetiology and this is supported by genetic studies demonstrating that about 50–72% of the contributing genetic factors overlap (Leitner, 2014; Thapar, Harrington, & McGuffin, 2001). Furthermore, several brain structures, such as amygdala, ACC and prefrontal circuits, are affected in all three disorders (Blair, 2013; Brieber et al., 2007). However, it is unknown if these brain structures underlie the co-occurrence of CD, ADHD and ASD. Human neuroimaging studies can guide in identifying neural structures and neuro-circuitry that co-occur in all three disorders but causal relationships between neural structures and behaviour cannot be examined in such studies. Animal models enable us to experimentally manipulate neural structures, observe the effects on behaviour and gain detailed insights into pathophysiological mechanisms (Markou, Chiamulera, Geyer, Tricklebank, & Steckler, 2009; Nestler & Hyman, 2010). An animal model that shows symptoms of CD, ADHD and ASD would enable us to investigate which neural structure(s) might underlie the comorbidity of these three disorders.

The BALB/cJ mouse model of CD

BALB/cJ mice have been repeatedly used as a model for aggressive behaviour (Dow et al., 2011; Velez, Sokoloff, Miczek, Palmer, & Dulawa, 2010). These mice were derived from an initial BALB/c stock, which was established in 1935. Several other laboratories acquired mice of the initial BALB/c stock, maintained them and bred them as independent stocks including BALB/cJ, BALB/cN, and BALB/cByJ. Due to breeding errors introducing new alleles and/or spontaneous mutations, the substrains started to exhibit genetic and phenotypic differences (Velez et al., 2010). For example,

BALB/cJ and BALB/cByJ show differences in eleven copy number variants (CNVs) and 38 mRNAs (Jager et al., unpublished data; Velez et al., 2010). BALB/cJ and BALB/cByJ mice also demonstrate differences in their behavioural phenotype; BALB/cJ mice are more aggressive than BALB/cByJ mice. They show high levels of intermale aggression, a shorter latency to attack and a higher incidence of attack in comparison to BALB/cByJ mice and other mouse strains in the resident intruder paradigm (Dow et al., 2011; Velez et al., 2010). In comparison to BALB/cByJ mice, BALB/cJ mice demonstrate decreased structural connectivity and decreased gamma-aminobutyric acid (GABA) inhibition in the ACC, changes that have been linked to aggressive behaviour (Jager et al., 2015). The BALB/cJ mouse model thus reproduces the core symptom of CD – increased aggression – and it also shows brain pathology comparable to the human situation (Teng et al., 2016). However, symptoms of common comorbidities of CD, such as ASD and ADHD, have been insufficiently studied in BALB/cJ mice. Here, we further characterised the behavioural and physiological phenotype of BALB/cJ mice by investigating if these mice show social withdrawal and hyperactivity, major symptoms of ASD and ADHD.

Experiment 1. In order to verify social withdrawal behaviour as a symptom of ASD in the BALB/cJ mouse model, we tested them in a social interaction test. We used a modified version of the three-chamber social preference test. The original test consists of a rectangular arena with three chambers and in one of the end chambers a stimulus mouse is restrained within a clear Plexiglas cylinder. For a period of five minutes a second mouse, the test mouse, can explore the whole arena. In the past, BALB/cJ mice have been reported to spend less time staying close to the cylinder with the stimulus mouse compared to other mouse strains; this behaviour has been interpreted as social withdrawal (Brodkin, Hagemann, Nemetski, & Silver, 2004; Fairless et al., 2008). A reduced size of the corpus callosum and increased brain size in these mice has been linked to social withdrawal (Brodkin, 2007; Fairless et al., 2008). However, there are two specific problems associated with previous studies. First, BALB/cJ mice have never been compared to other BALB/c substrains with a similar genetic background. It is known that mice of the BALB/c strain are less social than C57BL/6 mice; therefore, comparing BALB/cJ mice to C57BL/6 mice might lead to an overestimation of effects. Second, the use of a three-chamber apparatus might have a large influence on the behaviour of BALB/cJ mice. The mice are placed into a novel environment and are exposed to an unfamiliar stimulus mouse. Both the stress and novelty may alter the social behaviour of the test mouse, and this may be particularly of importance in BALB/cJ mice, as these mice display high levels of anxiety and are more sensitive to stress than other mice (Crawley et al., 1997; Fairless et al., 2013). Therefore, we created a modified version of the three-chamber test that consists of an arena with a single chamber and cylinders to the right and left side. This arena is similar to the homecage of mice and allows testing of social behaviour in a more familiar environment. We tackled the following hypothesis: BALB/cJ mice demonstrate less social interest than BALB/cByJ mice, a behaviour that can be interpreted as social withdrawal.

Experiment 2. In this experiment, we investigated if BALB/cJ mice show signs of hyperactivity commonly associated with ADHD, more specifically, increased locomotor activity. We implanted radio telemetry transmitters that could measure locomotor activity, temperature and heart rate in freely moving animals (Butz & Davisson, 2001). It has been observed that children with ADHD have a decreased core body temperature compared to healthy control children (Bijlenga et al., 2013; Dahl & Lewin, 2002). Therefore, we also investigated if a decreased core body temperature is seen in BALB/cJ mice. We aimed to tackle the following two hypotheses: (1) BALB/cJ mice demonstrate increased locomotor activity in comparison to BALB/cByJ mice, and (2) BALB/cJ mice show a decreased core body temperature in comparison to BALB/cByJ mice.

Methods

General methods

Housing conditions. All mice were housed individually in an enriched environment (High Makrolon® cages with Enviro Dri® bedding material and Mouse Igloo®) and had free access to dry food and water. They were kept at a reversed 12/12 day/night cycle with sunrise at 7.00 pm. Efforts were taken to restrict the number of mice and to keep the discomfort as minimal as possible. All animal procedures, including behavioural tests and surgical procedures were conducted in strict compliance with the European regulations for animal experimentation. The study was approved by the Ethics Committee on Animal Experimentation of Radboud University (RU- DEC).

Experiment 1: Social withdrawal

Animals. Fifteen-week-old male BALB/cJ (n = 5) and BALB/cByJ (n = 4) mice were obtained from the Jackson Laboratory (Bar Harbor, ME, USA) and used as test mice. Male C57BL/6 (n = 2, Charles River Laboratories, Erkrath, Germany) were used as stimulus mice.

Experimental procedure. Social behaviour was assessed with a modified version of the three-chamber test. The arena (50 cm × 43 cm) consisted of a single chamber with cylinders at the centre of the right and left side (see Fig. 1). Testing was done in the dark and behaviour was video-recorded. At the start of the test a stimulus mouse was randomly placed in one of the cylinders ("social cylinder"). We randomly assigned a stimulus mouse to each test mouse. The test mouse was then placed in the middle arena and was allowed free exploration of the arena and cylinders for a period of five minutes. Both cylinders had many holes, so that the test mouse could sniff the stimulus mouse. The side of the arena with the stimulus mouse was labelled as "social side" and the side with an empty cylinder as "non-social side". There was no barrier or line between the sides. Mice were sacrificed two weeks after the social interaction test.

Data Analysis. We manually scored "social cylinder investigation", defined as the amount of time (in seconds) that the test mouse sniffed, reared against, and climbed on the walls of the cylinder with the stimulus mouse inside. Climbing on the walls occurred very rarely in both groups and sniffing of the social cylinder was the predominant behaviour. For manual scoring The Observer XT software (Noldus Information Technology BV, Wageningen, The Netherlands) was used. Social cylinder investigation was analysed with a one-way ANOVA. All statistical analyses were performed using SPSS21- software (SPSS inc., Chicago, USA).

Experiment 2: Hyperactivity and temperature

Animals. Six-week-old BALB/cJ (n = 5) and BALB/cByJ (n = 6) mice were obtained from the Jackson Laboratory (Bar Harbor, ME, USA). Upon arrival, mice were housed individually. Surgery was performed at seven weeks of age and mice were allowed two weeks of recovery. Experimental recordings were performed at nine weeks of age.

Surgical Procedure. Transmitters were purchased from Data Sciences International (DSI™, St. Paul, MN, USA). The transmitter model PhysioTel ETA-F10 was used, allowing simultaneous recordings of heart rate, temperature and locomotor activity. These transmitters operate on radio frequency, have a battery life of about 2 months and are placed intraperitoneally. The transmitters consist of an insulated red (positive) and an insulated white (negative) electrode wire. Before surgery mice were weighed and transferred to a clean homecage. Surgery was performed under sterile conditions. Anaesthesia was induced and maintained using isoflurane (3% and 1.5-1.8%, respectively). The mouse was fixated with tape on all four legs ventral side up on a temperature-controlled surgery stage, adjusted to 36.5 °C. The abdomen and chest of the mice were shaved and the skin was disinfected

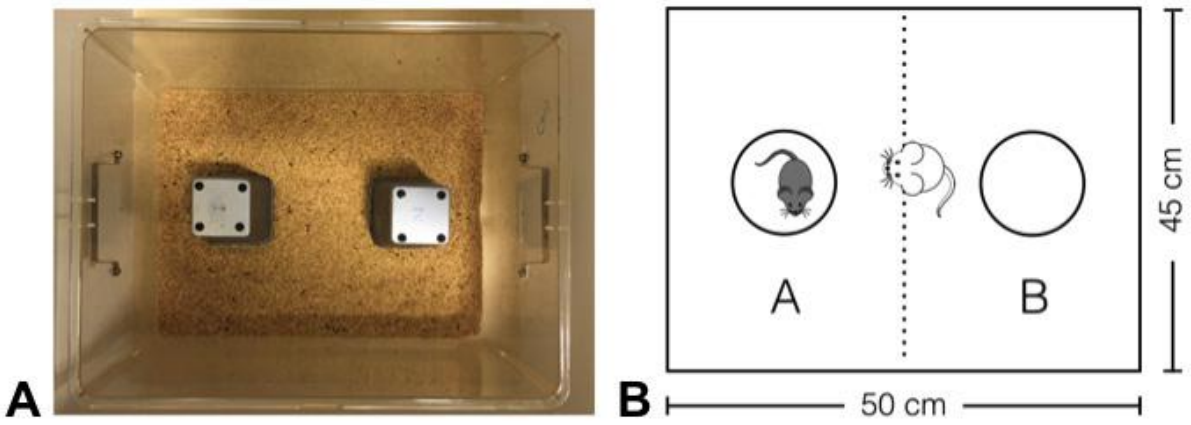


Fig. 1. A. The arena consisted of a single chamber and two metallic cylinders. The bottom was covered in corn pops to create a more familiar environment. **B.** Schematic of the arena with the stimulus mouse restrained in the cylinder in the social side (A) and an empty cylinder in the non-social side (B). In this example, the test mouse would be scored as being in the social side.

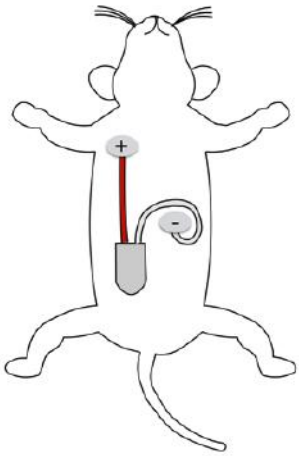


Fig. 2. Schematic of the transmitter placement. The transmitter was placed abdominally. The positive wire (red) was placed at chest-level; the negative wire (light-grey) was placed in a loop just below the costal arch.

with iodine sterilisation solution. Next, a longitudinal and medial abdominal skin incision of 1.8-2 cm was performed. The incision extended from the navel area to about 1 cm below the caudal tip of the sternum. Using a blunt metal probe moistened with physiological solution (0.9% sterile NaCl), skin and underlying muscle tissue were separated. Then, a subcutaneous tract towards the right front leg and the left hind leg was opened for the electrode wires. According to the first skin cut, a cut of muscle tissue was performed (about 1.5 cm long). The transmitter was then placed into the abdominal cavity with the electrode wires positioned anteriorly. Next, using a sharp probe, a little opening left from the abdominal cut, at chest-level, was made (through the muscle tissue). Through this opening the red electrode wire was guided and non-absorbable threads were used to fixate the electrode on the muscle tissue. The muscle tissue was then closed with absorbable threads. The white wire was placed in a loop under the skin and fixated to the muscle tissue below the costal arch (using non-absorbable threads; see Fig. 2 for a schematic). The two plastic fixation straps of the transmitter were connected to the muscle tissue using non-absorbable threads. Then, the muscle tissue was closed with absorbable threads. Before closing the skin, the transmitter was switched on with a magnet to assure proper function and signal. The skin was closed using three to four wound clips. Analgesia was provided by subcutaneously injecting rimadyl (Carprofen 5% with Ethanol 10%; 5-10 mg/kg) directly after surgery and twice a day for two days post-surgery. After surgery, mice were kept overnight in a warming chamber (38.5 °C) for recovery. The weight and the general condition of the mice were checked daily. Wound clips were removed 10 days post-surgery.

Experimental Procedure. Two weeks after surgery, mice were transferred to a separate room (same reversed day/night cycle as in the housing room, with sunrise at 7.00 pm). Their cages were placed on receiver plates (PhysioTel®, DSI™) that collected the signal from the transmitters. To initiate data collection, the transmitter was switched on by touching the mouse with a magnet (possible through the cage). Ponemah Software (DSI™) was used for the detection, collection and initial analysis of signals. This program collects data signals sent to the computer from the receiver plates via a Data Matrix (Matrix 2.0, DSI™). Data were collected at regular intervals (every 5 seconds). Recordings were started at 5.00 pm and lasted 86 hours. During this period no one was allowed entrance to the room, preventing any effect of the experimenter on the mice. Mice were provided with enough food and water for this period. At the end of the experiment, mice were brought back to their original room and were sacrificed two days after the recordings.

Data Analysis. The first two hours of the recordings were not analysed to account for possible stress due to transportation. Data were analysed with start of the non-active phase (7.00 pm, lights on) for a total of 84 hours (3.5 days). Ponemah Software (DSI™) initially pre-processed the data in data intervals of five minutes (the user can manually change these intervals). Temperature was measured in °C per minute and locomotor activity in counts per minute. Initially, for each hour we calculated a mean for temperature and locomotor activity. The data were analysed with repeated-measures ANOVAs (hour as within-subject factor and group as between-subject factor) and post-hoc tests were done when necessary. All statistical analyses were performed using SPSS21- software (SPSS inc., Chicago, USA).

Results

Experiment 1: Social withdrawal

BALB/cJ mice showed decreased (M = 42.47, SD = 10.36) social cylinder investigation compared to BALB/cByJ mice (M = 75.28, SD = 14.12, $F(1, 7) = 16.28$, $\eta^2 = .69$ $p = .004$). The data are presented in Figure 3.

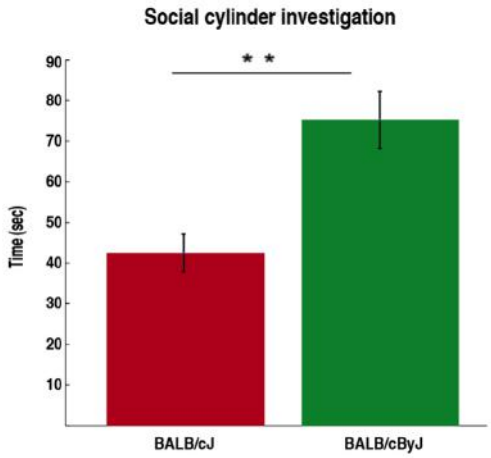


Fig. 3. Bar diagram showing the mean and standard error of the mean (SEM) of social cylinder investigation for BALB/cJ and BALB/cByJ mice. BALB/cJ mice spent significantly less time investigating the social cylinder than BALB/cByJ mice. $^{**}p < .01$

Experiment 2: Hyperactivity and temperature

Locomotor activity. In both groups, locomotor activity was higher during the active than the non-active hours ($F(23, 207) = 7.55$, $\eta^2 = .46$, $p = .000$). There was an interaction effect of hour × group ($F(23, 207) = 2.53$, $\eta^2 = .22$, $p = .036$), demonstrating that BALB/cJ mice were more active than BALB/cByJ mice during the active phase (all $p < .05$). The data are presented in Figure 4.

Temperature. In both groups, temperature was lower during the non-active hours than the active hours ($F(23, 207) = 9.81$, $\eta^2 = .53$, $p = .000$). BALB/cJ mice showed a lower temperature than BALB/cByJ mice (trend, $F(1, 9) = 3.29$, $\eta^2 = .26$, $p = .07$). There was an interaction effect of hour × group ($F(23, 207) = 2.99$, $\eta^2 = .25$, $p = .025$), demonstrating a lower temperature of BALB/cJ mice compared to BALB/cByJ mice in the non-active phase (all $p < .05$). The data are presented in Figure 5.

Discussion

Experiment 1: Social withdrawal

The results of this experiment confirm our hypothesis that BALB/cJ show social withdrawal. We demonstrated that BALB/cJ mice are less interested in a stimulus mouse than BALB/cByJ mice. By comparing BALB/cJ mice to a proper control group, the BALB/cByJ strain, we avoided an overestimation of effects as frequently observed in literature. BALB/cJ mice are usually compared to C57BL/6 mice, although it is known that C57BL/6 mice are more social than most other strains and that mice of the BALB/c strain in general show less social interest than other mice (Brodkin, 2007; Fairless et al., 2013). The arena we used closely resembled the homecage environment and did not require the mice to explore three different chambers. It has been repeatedly demonstrated that BALB/cJ mice

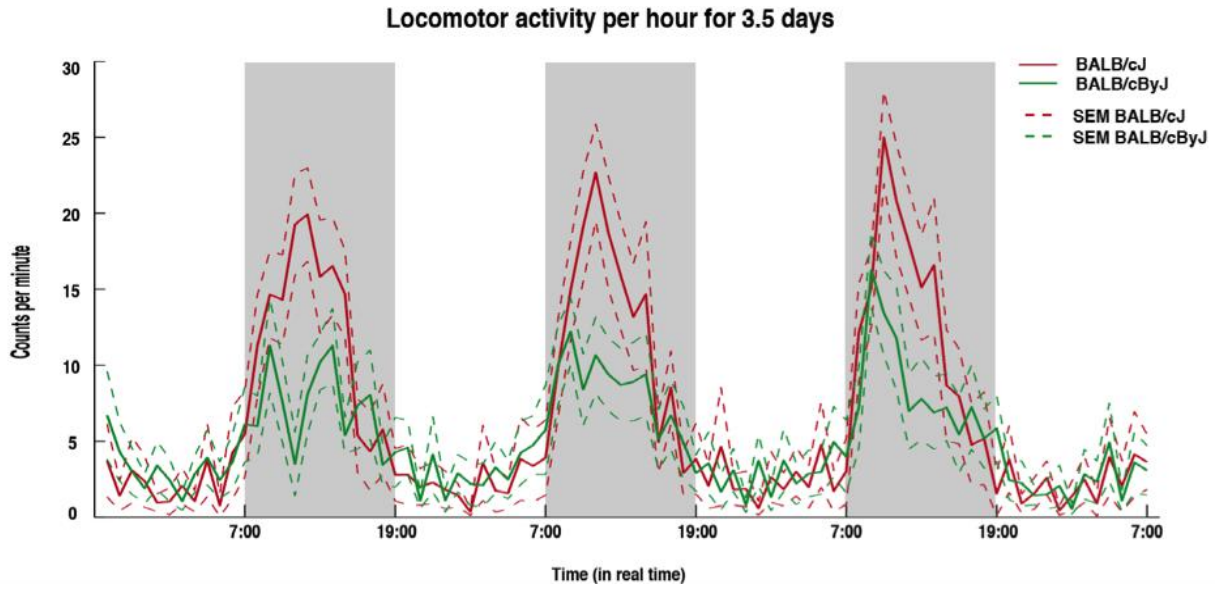


Fig. 4. Graph showing the mean and SEM of locomotor activity per hour for 3.5 days. Grey shaded areas indicate the active hours (lights off). Both groups show a higher locomotor activity in the active hours than the non-active hours. BALB/cJ mice are more active in the active hours than BALB/cByJ mice.

show high levels of anxiety (e.g., Brodkin, 2007). Therefore, we aimed to reduce anxiety induced by a new environment. The fact that BALB/cJ mice did not differ from BALB/cByJ mice in exploration behaviour (both groups investigated both sides of the arena), illustrates that BALB/cJ mice were not too anxious to investigate the arena. In this way, we were able to demonstrate that reduced social interest of BALB/cJ mice, interpreted as social withdrawal, is not (only) due to being too anxious to investigate the arena.

We need to note that apart from the utilised arena, there are a few other methodological differences between our study and previous reports. We housed both BALB/cJ and BALB/cByJ mice individually while previous studies housed mice in groups. By housing them individually we prevented high levels of intermale aggressive behaviour in BALB/cJ mice and to account for individual housing effects we also housed our control group individually. It is known that individual housing can increase aggressive behaviour but it is unknown how individual housing affects behaviour in tests where test and stimulus mouse cannot interact freely, as the three-chamber social preference test and the arena used in this study (Beery & Kaufer, 2015). Follow-up studies should investigate if individual housing has any effects on social behaviour in tests as the three-chamber social preference test. Previous studies usually tested BALB/cJ mice at young age (around 4 weeks, corresponding to pre-pubescence) whereas we have chosen to test our mice at a later stage (15 weeks), corresponding to adulthood (Kumar et al., 2012). In patients with ASD, deficits in social interaction persist into adulthood (Bejerot, Eriksson, & Mörtberg, 2014). Here, we have not only demonstrated that BALB/cJ mice show social withdrawal but also that they demonstrate decreased social interest in adulthood, recapitulating the human situation.

Experiment 2: Hyperactivity and temperature

In experiment 2 we investigated if BALB/cJ mice show hyperactivity (increased locomotor activity) and a lower body temperature, symptoms observed in patients with ADHD. We implanted transmitters and took measurements for a period of 86 hours. We found that BALB/cJ mice showed increased locomotor activity in comparison to BALB/cByJ mice and a lower (nocturnal) body temperature than BALB/cByJ. Multiple lines of research point toward a role of dopamine in ADHD, and more concretely, a dopamine dysfunction in the mesocortical, mesolimbic, and nigrostriatal pathways (Sonuga-Barke, 2005). In patients with ADHD, there seems to be increased dopamine reuptake by dopamine transporters (DAT), resulting in decreased extracellular dopamine (Gold, Blum, Oscar-Berman, & Braverman, 2014; Volkow et al., 2001). Symptoms of hyperactivity have been related

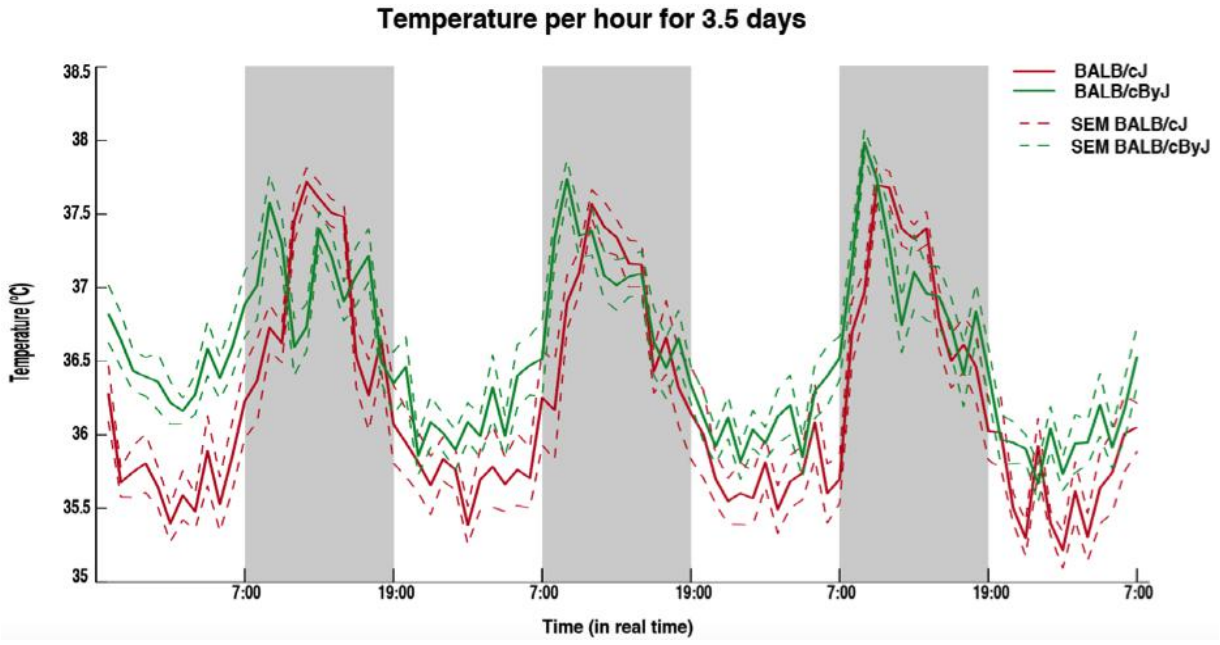


Fig. 5. Graph showing the mean and SEM of temperature per hour for 3.5 days. Grey shaded areas indicate the active hours (lights off). Both groups show a lower temperature in the non-active hours than the active hours. BALB/cJ mice have a lower temperature in the non-active hours than BALB/cByJ mice.

to the nigrostriatal pathway, which is involved in cognitive and voluntary movement control (Aguiar, Eubig, & Schantz, 2015) and projects from the substantia nigra to the striatum (Cho, Baek, & Baek, 2014).

Dopamine levels can also be regulated by serotonin, and serotonin dysfunction has been related to ADHD and hyperactivity (Quist et al., 2003). For example, studies in rodents and humans demonstrated that polymorphisms in the 5-HT1B receptor are associated with hyperactivity (Guimarães et al., 2009; Quist et al., 2003). Deletion of the 5-HT1B receptor in mice resulted in hyperactivity, increased exploratory activity and increased aggressive behaviour. Patients with ADHD frequently have the G861C polymorphism of the 5-HT1B receptor, which is associated with reduced 5HT1B receptors (Huang, Grailhe, Arango, Hen, & Mann, 1999; Quist et al., 2003). Interestingly, polymorphisms in the TPH2 gene, coding for the rate-limiting enzyme in the synthesis of 5-HT, have been associated with ADHD and BALB/cJ mice have a polymorphism in the TPH2 gene as well (Biskup et al., 2012). Therefore, it might be that a serotonin dysfunction in BALB/cJ mice influences the dopamine system, resulting in symptoms of ADHD, such as increased locomotor activity. Furthermore, the lower body temperature we observed in BALB/cJ mice is also observed in children with ADHD and has been linked to low levels of serotonin and dopamine (Catalina, Milewich, Frawley, Kumar, & Bennett, 2002). Children with ADHD often have sleep problems (Stein, 1999) and waking up during a period of low body temperature has been related to sleep deprivation and attention deficits (Dahl & Lewin, 2002).

Increased locomotor activity in BALB/cJ mice might be a method to augment their low body temperature, instead of being a sign of hyperactivity. However, one would then expect to see increased levels of locomotor activity especially during the non-active phase, as the temperature of BALB/cJ mice is significantly lower during the non-active phase. In the non-active phase there is no difference in locomotor activity between BALB/cJ mice and BALB/cByJ mice. It is known that stimulant medication used for the treatment of ADHD decreases symptoms of hyperactivity and there are indications for increases in body temperature in patients taking stimulant medication (Lakhan & Kirchgessner, 2012; Schacher, Tannock, Cunningham, & Corkum, 1997). However, patients that misuse their prescribed stimulants (e.g., take increased doses) show increased hyperactivity and increased body temperature at the same time (Lakhan & Kirchgessner, 2012), and healthy individuals that make use of stimulant medication also show increased activity and increased body temperature (Pigeau et al., 1995). Therefore, hyperactivity in ADHD does not seem to be a method to augment body temperature. To definitely test whether a low body temperature causes increased locomotor activity in BALB/cJ mice, one could house BALB/cJ mice in a heat chamber and observe if they still show signs of hyperactivity.

General discussion

The current study served to characterise the behavioural and physiological phenotype of BALB/cJ mice. These mice are highly aggressive, and aggression is a cardinal symptom of CD (Blair, 2013; Velez et al., 2010). However, it was unknown whether BALB/cJ mice also demonstrate symptoms of ADHD and ASD, common comorbidities of CD. Therefore, we further characterised the behavioural and physiological phenotype of BALB/cJ mice and designed two experiments that enabled us to explore if BALB/cJ mice show symptoms of ASD and ADHD. Our results show that BALB/cJ mice not only demonstrate increased aggression, a core symptom of CD, but also symptoms of common comorbidities such as ASD and ADHD in form of social withdrawal, hyperactivity and a lower body temperature. Having validated the behavioural and physiological phenotype of BALB/cJ mice, we can use this model to study brain structures that might give rise to the linked symptoms of CD, ADHD and ASD. Ultimately, this could aid in the discovery of new treatments for children and adolescents that suffer from CD with comorbid ADHD and/or ASD.

Children with ADHD that develop CD, show a considerably earlier onset of CD symptoms and more severe symptoms than children with CD only (Loeber et al., 2000). Furthermore, children with ASD that also receive a diagnosis of ADHD, have a higher chance to develop CD (Montes & Halterman, 2007). This suggests that symptoms of ADHD, such as impulsivity, hyperactivity or inattention, might play a role in the link between CD, ADHD and ASD. The focus has long been on impulsivity, as high levels of impulsivity in children with ADHD or CD contribute strongly to the risk of criminal involvement, even more than early symptoms of CD alone (Babinski, Hartsough, & Lambert, 1999). However, to date it is unknown if high levels of impulsivity are a common cause of ADHD, ASD and CD, explaining the high

comorbidity of these three disorders. More recently, the focus has been shifted toward inattention as a possible mediator of the relation between CD, ADHD and ASD. Attention enables us to selectively concentrate on certain aspects of information, suppressing distracting or irrelevant information (Kim, Ährlund-Richter, Wang, Deisseroth, & Carlén, 2016). Social situations in general, and even more so ambiguous social situations, require high levels of attention. If a person is non-attentive to subtle cues during interactions (e.g., tone and facial expression), situations can be interpreted as hostile leading to an (unprovoked) outburst of aggression (Evans, Fite, Hendrickson, Rubens, & Mages, 2015). It is known that patients with ASD, ADHD as well as CD have difficulty understanding social cues and that they tend to interpret ambiguous social situations as hostile (Evans et al., 2015). It might be that being inattentive to subtle cues leads to a misinterpretation of a situation, which results in aggressive behaviour. Indeed, causal modelling in a population of ADHD patients suggests that inattention is causal to aggression (Heskes, unpublished data). This also implies that inattention should precede aggression, being in line with the fact that children with ADHD, who develop CD, have an earlier onset of CD symptoms than children with CD only. Preliminary results in our group indicate that the BALB/cJ mouse model shows signs of inattention, enabling us to utilise this model to further study the role of inattention in aggression.

Future Directions: The BALB/cJ mouse and attention

Recently, Kim et al. (2016) demonstrated that gamma-aminobutyric acid (GABA-ergic) interneurons expressing parvalbumin (PV) in prefrontal cortex are involved in sustaining and directing attention. Mice performed a 3-choice-serial reaction-time task (modification of the 5-choice-serial reaction time task) and neuronal responses in prefrontal cortex were recorded. In the 3- (or 5-) choice task, the animal needs to sustain and divide its attention across a row of three (or five) screen locations to detect and respond to a brief visual stimulus in order to receive a reward. At the start of each trial the activity of the PV neurons increased and this heightened activity was sustained during the whole delay period (i.e., until presentation of the stimulus). The activity increased even more when the animal was about to perform correctly. This means that a high and sustained activity of PV neurons at the start of a trial predicts whether the animal will perform correctly, more than 2.5 seconds before presentation of the stimulus. The activity of the PV neurons neither correlated to the motivational state of the animal or motor behaviour and it can be concluded that prefrontal PV neurons are involved in attentional control. Decreased activity during tasks involving attentional processing (e.g., measures of sustained attention comparable to the 5-choice-serial reaction time task as used in rodents) in prefrontal cortex has also been observed in patients with ADHD and comorbid CD and/or ASD (Dickstein, Bannon, Castellanos, & Milham, 2006). Furthermore, prefrontal cortex is heavily interconnected and regions such as vmPFC and ACC have also been implicated in the control of aggressive behaviour, and both ACC and vmPFC have been found to be hypofunctioning in ADHD, ASD and CD (Blair, 2013; Hare, Rakimi, & Rangel, 2014). Possibly, malfunctioning prefrontal circuits cannot communicate efficiently with each other and cannot exert control over the basic aggression circuitry running from medial amygdala to medial hypothalamus and to the dorsal half of the PAG (Blair, 2013), which in turn could lead to aggressive behaviour. Future studies can utilise the BALB/cJ mouse model to investigate the role of prefrontal circuits in attention and aggression.

Acknowledgements

First of all, I would like to express my gratitude to my supervisors Amanda Jager, Shaha Abghari and Dr. Martha N. Havenith. Amanda, Shaha and Martha, thank you for your continuous support and for giving me the opportunity to work on a variety of different projects. Amanda, thank you for encouraging me to present my research (and your research!) on several congresses, I have learned a lot from that. Shaha and Martha, I would particularly like to thank you, as you supervised me throughout all the stages of my internship and thesis writing. Both of you have always been a source of calmness and wisdom when I was worried. Without your feedback and great ideas this thesis would be of a much lesser quality. I also would like to thank Dr. Jeffrey C. Glennon – Jeffrey, thank you for giving me the opportunity to do my internship in your group.

References

Aguiar, A., Eubig, P., & Schantz, S. L. (2015). Attention deficit/hyperactivity disorder: a focused overview for children's environmental health researchers. In C. Croft (Ed.), *Environmental Hazards and Neurodevelopment:*

Where Ecology and Well-Being Connect (pp. 223). Oakville, Canada: Apple Academic Press, Inc.

Babinski, L. M., Hartsough, C. S., & Lambert, N. M. (1999). Childhood Conduct Problems, Hyperactivity-Impulsivity, and Inattention as Predictors of Adult Criminal Activity. *Journal of Child Psychology and Psychiatry, 40,* 347-355.

Beery, A. K., & Kaufer, D. (2015). Stress, social behavior, and resilience: insights from rodents. *Neurobiology of stress, 1,* 116-127.

Bejerot, S., Eriksson, J. M., & Mörtberg, E. (2014). Social anxiety in adult autism spectrum disorder. *Psychiatry research, 220,* 705-707.

Bijlenga, D., Van Someren, E. J., Gruber, R., Bron, T. I., Kruithof, I. F., Spanbroek, E. C., & Kooij, J. J. (2013). Body temperature, activity and melatonin profiles in adults with attention-deficit/hyperactivity disorder and delayed sleep: a case–control study. *Journal of sleep research, 22,* 607-616.

Biskup, C. S., Sánchez, C. L., Arrant, A., Van Swearingen, A. E., Kuhn, C., & Zepf, F. D. (2012). Effects of acute tryptophan depletion on brain serotonin function and concentrations of dopamine and norepinephrine in C57BL/6J and BALB/cJ mice. *PLoS One, 7,* e35916.

Blair, R. J. R. (2013). The neurobiology of psychopathic traits in youths. *Nature Reviews Neuroscience, 14,* 786-799.

Brieber, S., Neufang, S., Bruning, N., Kamp-Becker, I., Remschmidt, H., Herpertz-Dahlmann, B., ... & Konrad, K. (2007). Structural brain abnormalities in adolescents with autism spectrum disorder and patients with attention deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry, 48,* 1251-1258.

Brodkin, E. S. (2007). BALB/c mice: low sociability and other phenotypes that may be relevant to autism. *Behavioural Brain Research, 176*(1), 53-65.

Brodkin, E. S., Hagemann, A., Nemetski, S. M., & Silver, L. M. (2004). Social approach–avoidance behavior of inbred mouse strains towards DBA/2 mice. *Brain Research, 1002*(1), 151-157.

Butz, G. M., & Davisson, R. L. (2001). Long-term telemetric measurement of cardiovascular parameters in awake mice: a physiological genomics tool. *Physiological genomics, 5*(2), 89-97.

Casanova, M. F. (2007). The neuropathology of autism. *Brain Pathology, 17,* 422-433.

Catalina, F., Milewich, L., Frawley, W., Kumar, V., & Bennett, M. (2002). Decrease of core body temperature in mice by dehydroepiandrosterone. *Experimental Biology and Medicine, 227,* 382-388.

Cho, H. S., Baek, D. J., & Baek, S. S. (2014). Effect of exercise on hyperactivity, impulsivity and dopamine D2 receptor expression in the substantia nigra and striatum of spontaneous hypertensive rats. *Journal of Exercise nutrition & Biochemistry, 18,* 379-384.

Coppens, C. M., De Boer, S. F., Buwalda, B., & Koolhaas, J. M. (2014). Aggression and aspects of impulsivity in wild-type rats. *Aggressive Behavior, 40,* 300-308.

Crawley, J. N., Belknap, J. K., Collins, A., Crabbe, J. C., Frankel, W., Henderson, N., ... & Wehner, J. M. (1997). Behavioral phenotypes of inbred mouse strains: implications and recommendations for molecular studies. *Psychopharmacology, 132*(2), 107-124.

Dahl, R. E., & Lewin, D. S. (2002). Pathways to adolescent health sleep regulation and behavior. *Journal of Adolescent Health, 31,* 175-184.

De Boer, S., Caramaschi, D., Natarajan, D., & Koolhaas, J. (2009). The vicious cycle towards violence: focus on the negative feedback mechanisms of brain serotonin neurotransmission. *Frontiers in Behavioral Neuroscience, 3,* 52.

De Boer, S. F., Van der Vegt, B. J., & Koolhaas, J. M. (2003). Individual variation in aggression of feral rodent strains: a standard for the genetics of aggression and violence? *Behavior Genetics, 33,* 485-501.

Dickstein, S. G., Bannon, K., Xavier Castellanos, F., & Milham, M. P. (2006). The neural correlates of attention deficit hyperactivity disorder: An ALE meta-analysis. *Journal of Child Psychology and Psychiatry, 47,* 1051-1062.

Dow, H. C., Kreibich, A. S., Kaercher, K. A., Sankoorikal, G. M. V., Pauley, E. D., Lohoff, F. W., ... & Brodkin, E. S. (2011). Genetic dissection of intermale aggressive behavior in BALB/cJ and A/J mice. *Genes, Brain and Behavior, 10*(1), 57-68.

Esbensen, A. J., Greenberg, J. S., Seltzer, M. M., & Aman, M. G. (2009). A longitudinal investigation of psychotropic and non-psychotropic medication use among adolescents and adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39,* 1339-1349.

Evans, S. C., Fite, P. J., Hendrickson, M. L., Rubens, S. L., & Mages, A. K. (2015). The role of reactive aggression in the link between hyperactive–impulsive behaviors and peer rejection in adolescents. *Child Psychiatry & Human Development, 46,* 903-912.

Fairless, A. H., Dow, H. C., Toledo, M. M., Malkus, K. A., Edelmann, M., Li, H., ... & Brodkin, E. S. (2008). Low sociability is associated with reduced size of the corpus callosum in the BALB/cJ inbred mouse strain. *Brain Research, 1230,* 211-217.

Fairless, A. H., Katz, J. M., Vijayvargiya, N., Dow, H. C., Kreibich, A. S., Berrettini, W. H., ... & Brodkin, E. S. (2013). Development of home cage social behaviors in BALB/cJ vs. C57BL/6J mice. *Behavioural Brain Research, 237,* 338-347.

Finger, E. C., Marsh, A., Blair, K. S., Majestic, C., Evangelou, I., Gupta, K., ... & Sinclair, S. (2012). Impaired functional but preserved structural connectivity in limbic white matter tracts in youth with conduct disorder or oppositional defiant disorder plus psychopathic traits. *Psychiatry Research: Neuroimaging, 202,* 239-244.

Glickman, G. (2010). Circadian rhythms and sleep in children with autism. *Neuroscience & Biobehavioral Reviews, 34,* 755-768.

Gold, M. S., Blum, K., Oscar–Berman, M., & Braverman, E. R. (2014). Low dopamine function in attention

deficit/hyperactivity disorder: should genotyping signify early diagnosis in children? *Postgraduate Medicine, 126,* 153-177.

Guimarães, A. P., Schmitz, M., Polanczyk, G. V., Zeni, C., Genro, J., Roman, T., ... & Hutz, M. H. (2009). Further evidence for the association between attention deficit/hyperactivity disorder and the serotonin receptor 1B gene. *Journal of Neural Transmission, 116,* 1675-1680.

Gurnani, T., Ivanov, I., & Newcorn, J. H. (2016). Pharmacotherapy of Aggression in Child and Adolescent Psychiatric Disorders. *Journal of Child and Adolescent Psychopharmacology, 26*(1), 65-73.

Hare, T. A., Hakimi, S., & Rangel, A. (2014). Activity in dlPFC and its effective connectivity to vmPFC are associated with temporal discounting. *Frontiers in Neuroscience, 8,* 50.

Huang, Y. Y., Grailhe, R., Arango, V., Hen, R., & Mann, J. J. (1999). Relationship of psychopathology to the human serotonin1B genotype and receptor binding kinetics in postmortem brain tissue. *Neuropsychopharmacology, 21,* 238-246.

Jager, A., Amiri, H., Oomen, C. A., Buitelaar, J. K., Kozicz, L. T., Aschrafi, S. A., & Glennon, J. C. (2015). P. 1. h. 028 Structural and neurochemical changes in the anterior cingulate cortex (ACC) are associated with aggression and global attention deficits: A study in the BALB/cJ mouse. *European Neuropsychopharmacology, 25,* S296-S297.

Kim, H., Ährlund-Richter, S., Wang, X., Deisseroth, K., & Carlén, M. (2016). Prefrontal Parvalbumin Neurons in Control of Attention. *Cell, 164,* 208-218.

Kumar, M., Kim, S., Pickup, S., Chen, R., Fairless, A. H., Ittyerah, R., ... & Poptani, H. (2012). Longitudinal in-vivo diffusion tensor imaging for assessing brain developmental changes in BALB/cJ mice, a model of reduced sociability relevant to autism. *Brain Research, 1455,* 56-67.

Lakhan, S. E., & Kirchgessner, A. (2012). Prescription stimulants in individuals with and without attention deficit hyperactivity disorder: misuse, cognitive impact, and adverse effects. *Brain and Behavior, 2,* 661-677.

Leitner, Y. (2014). The co-occurrence of autism and attention deficit hyperactivity disorder in children–what do we know? *Brain Development and the Attention Spectrum, 80,* 268.

Loeber, R., Burke, J. D., Lahey, B. B., Winters, A., & Zera, M. (2000). Oppositional defiant and conduct disorder: a review of the past 10 years, part I. *Journal of the American Academy of Child & Adolescent Psychiatry, 39,* 1468-1484.

Markou, A., Chiamulera, C., Geyer, M. A., Tricklebank, M., & Steckler, T. (2009). Removing obstacles in neuroscience drug discovery: the future path for animal models. *Neuropsychopharmacology, 34*(1), 74-89.

Matthys, W., Vanderschuren, L. J., Schutter, D. J., & Lochman, J. E. (2012). Impaired neurocognitive functions affect social learning processes in oppositional defiant disorder and conduct disorder: Implications for interventions. *Clinical Child and Family Psychology Review, 15,* 234-246.

McDougle, C. J., Stigler, K. A., & Posey, D. J. (2003). Treatment of aggression in children and adolescents with autism and conduct disorder. *The Journal of Clinical Psychiatry, 64,* 16-25.

Montes, G., & Halterman, J. S. (2007). Bullying among children with autism and the influence of comorbidity with ADHD: A population-based study. *Ambulatory Pediatrics, 7,* 253-257.

Nestler, E. J., & Hyman, S. E. (2010). Animal models of neuropsychiatric disorders. *Nature Neuroscience, 13,* 1161-1169.

Noordermeer, S. D., Luman, M., & Oosterlaan, J. (2016). A systematic review and meta-analysis of neuroimaging in oppositional defiant disorder (ODD) and conduct disorder (CD) taking attention-deficit hyperactivity disorder (ADHD) into account. *Neuropsychology Review, 26,* 44-72.

Pigeau, R., Naitoh, P., Buguet, A., McCann, C., Baranski, J., Taylor, M., ... & Mack, I. (1995). Modafinil, d-amphetamine and placebo during 64 hours of sustained mental work. I. Effects on mood, fatigue, cognitive performance and body temperature. *Journal of Sleep Research, 4,* 212-228.

Quist, J. F., Barr, C. L., Schachar, R., Roberts, W., Malone, M., Tannock, R., ... & Kennedy, J. L. (2003). The serotonin 5-HT1B receptor gene and attention deficit hyperactivity disorder. *Molecular Psychiatry, 8,* 98-102.

Serper, M., Beech, D. R., Harvey, P. D., & Dill, C. (2008). Neuropsychological and symptom predictors of aggression on the psychiatric inpatient service. *Journal of Clinical and Experimental Neuropsychology, 30,* 700-709.

Sonuga-Barke, E. J. (2005). Causal models of attention-deficit/hyperactivity disorder: from common simple deficits to multiple developmental pathways. *Biological Psychiatry, 57,* 1231-1238.

Stein, M. A. (1999). Unravelling sleep problems in treated and untreated children with ADHD. *Journal of Child and Adolescent Psychopharmacology, 9,* 157-168.

Thapar, A., Harrington, R., & McGuffin, P. (2001). Examining the comorbidity of ADHD-related behaviours and conduct problems using a twin study design. *The British Journal of Psychiatry, 179,* 224-229.

Teng, B. L., Nikolova, V. D., Riddick, N. V., Agster, K. L., Crowley, J. J., Baker, L. K., ... & Moy, S. S. (2016). Reversal of social deficits by subchronic oxytocin in two autism mouse models. *Neuropharmacology, 105,* 61-71.

Velez, L., Sokoloff, G., Miczek, K. A., Palmer, A. A., & Dulawa, S. C. (2010). Differences in aggressive behavior and DNA copy number variants between BALB/cJ and BALB/cByJ substrains. *Behavior Genetics, 40,* 201-210.

Volkow, N. D., Wang, G., Fowler, J. S., Logan, J., Gerasimov, M., Maynard, L., ... & Franceschi, D. (2001). Therapeutic doses of oral methylphenidate significantly increase extracellular dopamine in the human brain. *The Journal of Neuroscience, 21*(2), RC121.

Effects of Structural and Functional Prefrontal Cortex Maturation on Verbal Memory Development

Nadia Klijn[1]
Supervisors: Nils Müller[1,2], James McQueen[1,3,4], Guillén Fernández[1,2]

[1] Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands
[2] Radboud University Medical Centre Nijmegen, Donders Institute of Brain, Cognition and Behaviour, The Netherlands
[3]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[4] Behavioural Science Institute, Radboud University Nijmegen, The Netherlands

The prefrontal cortex, including core regions of our mnemonic system, is characterised by protracted structural and functional maturation from late childhood into early adulthood. We hypothesize that this relatively late maturation contributes to the steep increases in verbal memory performance observed during adolescence. To test this, we compared verbal memory scores of children (10-12 years old), adolescents (18 years old) and adults (25-32 years old). As a measure of structural maturation, we investigated whether cortical thickness of the left inferior frontal gyrus (LIFG, important for verbal memory encoding) mediates the effect of age on verbal memory performance. Furthermore, we investigated whether a developmental increase in specialisation (i.e., lateralisation) of functional connectivity between the inferior frontal gyrus (IFG) and the left medial temporal lobe mediates the effect of age on verbal memory performance. Firstly, our results show increased verbal memory performance and leftward lateralisation of functional connectivity, and decreased cortical thickness with age. More importantly, our results show that LIFG cortical thickness does, but LIFG functional specialisation does not significantly mediate the increase in verbal memory performance with age. This indicates that structural maturation is indeed an important contributor to verbal memory development during adolescence, whereas functional LIFG specialisation seems to play a less significant role.

Keywords: prefrontal cortex, maturation, adolescence, verbal memory, cortical thickness, functional connectivity

Corresponding author: Nadia Klijn; **E-mail:** nadiaklijn@hotmail.com

An intriguing question in the cognitive neuroscience of memory is how the maturation of neural structure and function contributes to the development of an adult-like declarative memory system. One central characteristic of declarative memory is our capacity to form new semantic representations, for instance by acquiring new words. But why exactly does our ability to remember words improve so drastically when we grow up, while our brain becomes less plastic at the same time? Brain regions supporting memory encoding undergo substantial functional and anatomical changes throughout the lifespan (e.g., Johnson, 2001; Simons & Spiers, 2003). Therefore, we expect that structural and functional differences between the developing and the adult brain underlie the differences in memory performance between these two. We investigated whether this is the case by comparing verbal memory scores of children (10-12 years old), adolescents (18 years old) and adults (25-32 years old). We examined whether these scores can be predicted by structural thinning and functional specialisation of brain areas that are related to verbal memory performance in adults.

Verbal memory: brain and behaviour

Our capacity to successfully remember information improves drastically across development with declarative memory only reaching full functionality in early adulthood (Baltes, Lindenberger & Staudinger, 2006; Finn et al., 2016). Within the declarative memory system, there is a transition from favouring perceptual, episodic-like representations towards favouring more abstract, semantic representations (Ofen & Shing, 2013). In line with this, verbal learning studies have revealed an increase in verbal memory performance across development (e.g., Vakil & Blachstein, 1997; Vakil, Blachstein & Sheinman, 1998; Blachstein & Vakil, 2016; Davis et al., 2013). Using the Rey Auditory Verbal Learning Test, these studies measured immediate and delayed recall after intentional encoding trials. Results of these studies indicate an increase in verbal memory performance during childhood and adolescence that peaks in early adulthood, after which performance levels decrease again with age. However, in these studies participants were aware that their memory for the items would be tested later on in the experiment. Therefore, differences in memory performance across age groups are not only driven by differences in mnemonic abilities, but are also influenced by differences in strategic abilities between children and adults (Shing, Werkle-Bergner, Li & Lindenberger, 2008). To truly assess the pure mnemonic advances in verbal memory development, it is important to disentangle the (neural) maturational changes underlying true mnemonic abilities from these strategic abilities. To avoid the confound of unequal strategy use in the current study, we investigated the developmental differences in performance on a verbal memory task in which encoding is incidental and thus strategy-free.

The maturational trajectories of neural regions involved in verbal memory performance in adults can provide us with insights on changes in structural and functional properties that potentially underlie differences in memory performance across development. Studies investigating verbal memory in the adult brain found that subsequent memory for verbal material is primarily associated with left lateralized activation in the medial temporal lobe (MTL) and the prefrontal cortex (PFC), more specifically the left inferior frontal gyrus (LIFG) (e.g., Wagner et al., 1998; Golby et al., 2001; Opitz, Mecklinger & Friederici., 2000; Duverne, Motamedinia & Rugg, 2009; see for a meta-analysis Kim, 2011). Activation in the anterior extent of the LIFG (Brodmann's area [BA] 45/47) is assumed to reflect semantic control processes mediating internal representations. These processes then optimize memory encoding when interacting with core storage regions in the MTL (Thompson-Schill, D'Esposito, Aguirre & Farah, 1997; Bokde, Tagaments, Friedman & Horwitz, 2001; Simons & Spiers, 2003; Kim, 2011). Furthermore, activation of MTL regions is assumed to be involved in binding together the features of an event into a compound representation and in consolidating this representation for later use (Werkle-Bergner, Müller, Li & Lindenberger, 2006). Structural maturation of these regions and the maturation of their interaction with each other, could therefore potentially explain differences in verbal memory performance across development. Nevertheless, it is good to keep in mind here that the correlational evidence mentioned above only provides us with limited information on the cause of the reported effects.

Structural maturation – Cortical thickness

Trajectories of structural PFC maturation differ substantially from that of other cortical regions, such as the MTL (Sowell et al., 2003; Lenroot & Giedd, 2006; Raz & Rodrigue, 2006). Anatomical evidence suggests that most MTL regions mature at a relatively fast rate before late childhood, whereas

the PFC and associated neural networks undergo serious age-related changes well into adolescence and early adulthood (e.g., Sowell et al., 2003; Sowell et al., 2004; Gogtay et al., 2004; Gogtay et al., 2006). Because of this differential pattern of maturation between the MTL and the PFC, the increase in verbal memory performance from late childhood into early adulthood as suggested by the behavioural literature is likely driven by ongoing maturational processes in the PFC. One protracted structural change in the PFC that could be of relevance in explaining developmental differences in verbal memory performance, is the gradual decline in cortical thickness in frontal association cortices until early adulthood (Sowell et al., 2003;Sowell et al., 2004; Gogtay et al., 2004; Koolschijn & Crone, 2013; Wierenga, Langen, Oranje & Durston, 2014; Amlien et al., 2014; Ducharme et al., 2016). Decreased cortical thickness over time is usually associated with increased synaptic pruning from childhood into adulthood (Petanjek et al., 2011). Because of this, cortical thinning can be indirectly associated with more efficient processing, which makes it a relevant measure in relation to gains in memory performance during development (Østby, Tamnes, Fjell & Walhovd, 2011 ; Sowell, Delis, Stiles & Jernigan, 2001). Sowell et al. (2001) for instance showed that an increase in verbal memory performance across development is associated with decreasing cortical thickness of the PFC, but less so of the MTL. However, Sowell et al. (2001) used an intentional verbal learning task to test their hypotheses. Therefore, based on this study we cannot conclude whether PFC structural maturation is specifically beneficial for verbal memory development itself, or whether the effect is driven only by improved strategic abilities that also rely on the PFC. Since our task is strategy-free, we would be able to draw such conclusions when testing the hypothesis that protracted cortical thinning of the PFC allows for increasing verbal memory performance.

Functional maturation – Specialisation of connectivity

On a functional level, protracted synaptic pruning of the PFC from childhood into early adulthood (Petanjek et al., 2011) potentially causes a decrease in short-range, inter-regional, functional connectivity (segregation) while simultaneously causing long-range connectivity to keep increasing due to task-related co-activation, known as integration. (Jolles, Van Buchem, Crone & Rombouts, 2011). This is in line with the finding that children show more diffuse functional connectivity patterns and increased short-range functional connectivity compared to adults, whereas adults show increased long-range functional connectivity patterns (Kelly et al., 2009). Therefore, protracted synaptic pruning of the PFC, as structurally reflected in protracted cortical thinning across development, is potentially also associated with specialisation of brain regions within larger functional networks (Durston & Casey, 2006; Durston et al., 2006; Fair et al., 2007;Fair et al., 2009; Supekar, Musen & Menon, 2009; reviewed in Johnson, 2011). In studies using language-related tasks, this functional specialisation of brain regions is often observed as decreases in bilaterality of task-related fMRI activity towards a more left lateralized activation pattern from childhood to early adulthood. This lateralisation is then associated with an increase in performance on verbal tasks (Holland et al., 2001; Szaflarski, Holland, Schithorst & Byars, 2006a; Szwaflarski et al., 2006b: Ressel et al., 2008; Everts et al., 2009; reviewed in Holland et al., 2007). Interestingly, Szaflarski et al. (2006a) show an inverted u-shape for language lateralisation similar as to what has been shown in behavioural studies investigating verbal memory across the lifespan (Blachstein & Vakil, 2016; Davis et al., 2013), where the learning curves increase until age 20, plateaus around age 25 and slowly decreases afterwards. However, most of these studies used a verb generation task which does not allow us to draw conclusions on whether maturing lateralisation of language functions predicts verbal memory development. Contrarily to previous language-related studies exploring developmental lateralisation of task-related brain activity, we examined lateralisation of resting-state functional connectivity between IFG – left MTL regions with age. This enabled us to investigate whether increasing functional specialisation (i.e., lateralisation) of the PFC between childhood and early adulthood is related to increasing verbal memory capacities without the fMRI measurements being confounded by differences in task performance between groups. Clearly, these measures do not allow us to make direct inferences about specialisation in response to the task. Resting-state functional connectivity is however believed to reflect recent experience as well as neuroanatomy and thus provides insight on both levels (Uddin, Supekar & Menon, 2010; Tavor et al., 2016). This makes it an appropriate measure to detect whether maturing functional specialisation of core mnemonic regions contributes to the development of an adult-like memory system.

Current study

In this study we are taking the behavioural, structural and functional maturational patterns as described above together. To do so, we carried out an implicit verbal memory task, and both structural and functional (resting-state) MRI scans in children (10-12 years old), adolescents (18 years old) and adults (25-32 years old). This selection of groups is most optimal since it provides us with participants in the beginning, in the middle and at the end of the protracted cortical thinning process of the PFC (Sowell et al., 2003; Sowell et al., 2004; Gogtay et al., 2004). With this data, our main goal was to investigate whether protracted structural and functional maturation of the PFC from late childhood into early adulthood can (at least partially) explain the steep increase in verbal memory performance across these developmental stages.

Our first research question involves the behavioural level: can we replicate the effect of age group on verbal memory scores?

The second question we aim to answer is whether cortical thickness of the left anterior IFG (BA 45/47, further referred to as LIFG) indeed decreases decreases as a function of age between the groups, as an indirect measure of the amount of synaptic pruning that has taken place allowing for greater processing efficiency. We also investigated whether this decrease in cortical thickness mediates the hypothesised increase in verbal memory performance with age.

Our third and last question is whether the specialisation (i.e., lateralisation) of associated LIFG functional connectivity indeed increases with age group, and whether this increase mediates the hypothesised increase in verbal memory performance with age. Lateralisation of LIFG involvement was measured as its resting-state functional connectivity with the already matured left MTL relative to the resting-state functional connectivity of the RIFG with the left MTL.

Materials and Methods

Participants

Ninety right-handed native Dutch-speaking volunteers participated in this study. Thirty of them were adults aged between 25-32 years old (M_{age} = 26.9 years, *SD* = 21.9 months, 12 male), 29 were adolescents aged 18 (M_{age} = 18.5 years, *SD* = 3.1 months, 10 male) and 31 were children aged between 10-12 years old (M_{age} = 11.0 years *SD* = 8.8 months, 8 male). All subjects had normal hearing and normal or corrected-to-normal vision. Furthermore, participants were required to have no history of injury or disease known to affect the central nervous system function (including neuropsychological disorders such as dyslexia, autism and ADHD) and to not have MRI contraindications. Adults and adolescents were recruited from the student population of Radboud University, Nijmegen, and from the surrounding community. Children were recruited through presentations and flyers at local schools. The study was approved by the local ethics committee, the CMO Arnhem – Nijmegen. Written informed consent was obtained prior to participation from all participants who were at least age 18 and from both parents of participants under age 18.

Of all 90 participants, 89 participants ultimately satisfied the inclusion criteria for the behavioural analysis (one child had to be excluded due to dyslexia). Of these, 84 had adequate processed and quality-checked MRI data for the cortical thickness analyses (30 adults, 26 adolescents and 28 children) and 84 had adequate data for the functional connectivity analyses (28 adults, 28 adolescents and 28 children).

Experimental stimuli

The experimental stimuli comprised recordings of 72 bi- and trisyllabic Dutch verbs geared to the children's level of interaction with the world. These words were selected based on a short pilot in which seven children aged between 10-12 years old got a list of 78 hand and arm action verbs. These words were selected from a list of Dutch verbs based on their meaning (a relatively basic action carried out by our hands or arms, like 'throwing') and on the amount of syllables (two or three). On this list of 78 verbs, the children had to circle the words for which they did not know the meaning. If an action verb was encircled twice, we did not include it in our stimulus list. All verbs included in our study were therefore assumed to be well known for both our adult participants as well as for the children.

These stimuli were recorded by a male and a female speaker to introduce variation and create a more natural linguistic encoding situation during the experiment. We chose an auditory stimulus presentation to prevent an encoding advantage towards the older age groups due to better reading abilities. For each stimulus, three different versions with different intensities were created to enable a loudness rating task in the encoding phase. This resulted in one version of 65 decibel (dB), one

version of 67.5 dB and one version of 70 dB for each word. We counterbalanced the presentation of one of these versions for a specific word across subjects. Furthermore, word frequency, voice and the amount of syllables were balanced across target ('old', to be remembered words) and filler ('new', words added in the recognition phase) conditions. The distribution of the individual words over these two conditions was randomized across subjects. In the encoding phase 36 verbs were presented. In the recognition phase the same 36 verbs as well as 36 new, unstudied, verbs were presented to the participants. For the lists of the recognition phase, words were roughly equally spaced relative to their first occurrence in the encoding phase, and the sequences of words were pseudorandomly ordered such that no more than three items belonging to the same condition (old versus new) occurred sequentially.

Verbal memory task

The verbal memory task consisted of two separate blocks run in Presentation (Version 16, www.neurobs.com). The first block consisted of an incidental, shallow encoding task. We chose a shallow encoding phase (loudness rating task) to ensure that children did not have a disadvantage at the level of task performance during the study block that might be present in a more difficult task, such as a deep encoding semantic decision task. More difficulty with the encoding task could namely confound children's mnemonic abilities. Furthermore, we chose incidental encoding to rule out the possibility that differences in memory performance across groups were mainly attributable to differences in strategic encoding abilities (Shing et al., 2010). The encoding phase was immediately followed by a second block in which recognition memory of the verbs was tested.

Before the start of the first block, participants were instructed that words would be presented to them via headphones and that their task was to rate the loudness of the presented stimulus in comparison to the loudness of the previous stimulus (the system's sound volumes were kept constant across participants). To rate the loudness of the verbs, participants could use six different responses presented on a computer screen also reflecting their degree of confidence, namely: 'surely louder', 'louder', 'maybe louder', 'maybe softer', 'softer', 'surely softer'. Whether the options for 'louder' and 'softer' were presented on the left or on the right was counterbalanced across subjects. The participants were instructed to use their left and right index, middle and ring fingers and let them rest on the response buttons on the keyboard (buttons 1-6 in the upper-left corner). Instructions emphasized the need for both speed and accuracy. They were not informed that their memory for the verbs would be tested afterwards. The incidental encoding phase consisted of 36 unique trials in total, such that each target verb was encountered only once before test.

Immediately after completing this encoding phase, the participants got the instructions for the recognition task. They were told that they would hear the words again together with words they had not heard in the first block of the experiment. All words in the recognition phase were presented at the 67.5 dB level. We explained to the participants that their task was to judge which words they already encountered ('old') and which words they did not ('new'), and to indicate their degree of confidence on this decision in the same way as during the encoding phase. The assignment of old and new responses to the left or right hand was counterbalanced across participants. Furthermore, the participants were instructed to respond as in the first block, and again as fast and as accurate as possible. The recognition phase consisted of 72 trials in total, of which 36 were old verbs and 36 were new (counterbalanced across subjects).

Both in the encoding and the recognition phase, each trial comprised of a white fixation cross in the center of the screen presented for 200 ms, followed by a visual presentation of the response continua for 300 ms before presentation of the stimulus. After the onset of the word, participants could press a response button. If this response was given within 1500 ms after stimulus onset, the participant had to wait until 1500 ms after stimulus onset had passed before the fixation cross was presented again to make sure participants could not go through the whole experiment without listening to the words. The inter-trial interval during which the fixation cross was presented was 500 ms.

Procedure

The verbal memory task reported here was implemented within a bigger study investigating the role of prior knowledge in schema memory development (Müller et al., in preparation). For this study, a battery of tests was performed. On Day 1, participants came to the Donders Centre for Cognitive Neuroimaging and started in the behavioural lab with a practice fractal N-back task (Ragland et al., 2002), after which the described verbal memory task was performed. Immediately

after completing the verbal memory task, participants were taken to the MRI lab where we first made their resting-state fMRI scan. Then the fractal N-back task that was practiced before was performed in the scanner and a structural scan was conducted. Afterwards, participants performed five sessions of a memory game in the lab and at home, after which they came back to the Donders Centre for Cognitive Neuroimaging on Day 8. Then they performed the recall session of the memory game in the MRI scanner after which another structural scan was made. For our analyses, we used either the structural scan from Day 1, the structural scan from Day 8, or an average of the two, based on which of these three options allowed for the best quality (i.e., least motion artifacts). After the second MRI session, the procedure continued with a Wisconsin Card Sorting Task (Heaton, Chelune, Talley, Kay & Curtiss, 1993), a Digit Span task (Alloway, 2007) and a long-delay follow-up on the memory game for adolescents and adults.

fMRI data acquisition

Scanning was performed with a Siemens Magnetom Skyra 3 tesla MR scanner equipped with a 32-channel phased array head coil. First, the resting-state scans were acquired during which the participants were instructed to keep their eyes fixated on a black fixation cross in the center of a white screen. Furthermore, we told them to try to think of nothing in particular. A total of 900 blood oxygen level-dependent scans were acquired using a T_2*-weighted gradient-echo, whole-brain echo planar imaging (EPIs) sequence with the following parameters: time repetition (TR) = 657 ms; time echo (TE) = 36.80 ms; flip angle = 50°; multiband factor = 8; matrix size = 88 x 88; field-of-view (FOV) = 210 mm; slice thickness = 2.4 mm; no slice gap; 64 slices acquired interleaved. The first ten scans were discarded to allow for equilibration of T_1 saturation effects. Furthermore, to control for the effect of field inhomogeneities, we acquired a fieldmap with the same parameters as the resting-state scans (TE1 = 4.54 ms, TE2 = 7.00 ms).

After acquiring the functional data, a high resolution T1-weighted anatomical scan was made to enable registration and the analysis of cortical thickness. We used a magnetization prepared, rapid-acquisition gradient echo sequence (parameters: TR = 2300 ms; TE = 3.03 ms; flip angle = 8°; matrix size = 256 x 256; FOV = 256 mm; slice thickness = 1 mm; 192 sagittal slices).

Data analysis

Behavioural data analysis – Verbal memory. On a behavioural level, we wanted to investigate whether we could replicate the effect of age group on verbal memory scores as described in the literature.

Statistical tests were performed using IBM SPSS Statistics (23.0, SPSS Inc., Chicago, USA). As a first control analysis, we tested whether the participants could generally discriminate successfully between old and new verbs. For this, the recognition memory data was first analysed using a 2 x 6 ANOVA with the 'Condition' of the items as the first factor (old versus new), the confidence rating as the other six-level factor ('Response') and the proportion of responses given for this combination of factors as the dependent variable (Liu, Qin, Rijpkema, Luo & Fernandez, 2010). We then applied post-hoc paired samples t-tests to verify for each response bin whether participants could discriminate old from new items. If the difference between the proportions of old and new responses for a certain confidence bin was proven significant, we concluded that most of the responses in this bin reflected true memory and decided to include responses from that bin for further analyses.

To achieve a representative score for memory performance per participant, we calculated the z-scores of both hit (old items recognized as old) and false alarm rates (new items recognized as old). Then, the z-scores for false alarms were subtracted from the z-scores for the hits (MacMillan & Creelman, 2005), leading to a d', representing the memory performance for each subject while controlling for response biases. This d' served as the dependent variable in a one-way ANOVA and pairwise comparison with age group as an independent variable, to test whether there is a significant difference between age groups on performance in a verbal memory task. To add meaning to the development of d' memory scores, we also performed one-way ANOVA's and pairwise comparison to investigate whether and how hit rates and false alarm rates differ across age groups.

Structural MRI data analysis – Cortical thickness.

Analysis of the structural MRI data served two purposes. Firstly, we wanted to determine whether cortical thickness of the LIFG indeed decreases with age group (as an indirect measure of the amount of synaptic pruning that has taken place, allowing for greater processing efficiency). Furthermore, we hypothesized that this decrease in cortical thickness mediates the increase in verbal memory performance with age group.

Preprocessing. High resolution T_1-weighted anatomical scans were analysed to estimate cortical thickness using FreeSurfer 5.1 software (http://surfer.nmr.mgh.harvard.edu/). This procedure provided us with a measure of cortical thickness for each participant at each point of the reconstructed surface allowing for the detection of submillimeter differences between age groups (e.g., Fjell et al., 2010; Fischl & Dale, 2000). The automated reconstruction involved motion correction, removal of non-brain tissue (Clarkson et al., 2011), intensity normalization (Sled, Zijdenbos & Evans, 1998), tessellation of gray/white matter boundary automated topology correction (Segonne, Pacheco & Fischl, 2007), surface deformation, and registration, segmentation of subcortical white matter and deep brain structures (Hutton, Draganski, Ashburner & Weiskopf, 2009). Estimates of cortical thickness on the tessellated surface were obtained by calculating the closest distance between representations of the cortical surface and the gray/white matter border (Fischl & Dale, 2000). After construction of surface-based maps for each participant's anatomical scan, cortical structures of each individual scan were registered to a spherical atlas and labeled using the Desikan-Killiany atlas (Desikan et al., 2006). This resulted in estimates of average cortical thickness in millimeters for every labeled cortical structure per participant.

Analyses. Cortical thickness analyses were performed, extracting each participant's cortical thickness of the anterior extent of the LIFG (BA45/47). This anterior extent including both the pars triangularis and the pars orbitalis which have been shown to contribute to verbal memory performance and to play a key role in semantic control (Bokde et al., 2001). Therefore, we extracted both the pars triangularis and the pars orbitalis of the LIFG (BA45/47). These regions of the anterior extent had to be extracted separately due to the arrangement of the Desikan-Killiany atlas used by FreeSurfer 5.1. Based on the described literature we were interested in the anterior extent of the LIFG as a whole, we collapsed these Regions of interest (ROI) resulting in a mean of the average cortical thickness estimates of both regions. The result of this is our measure of average anterior LIFG cortical thickness for each participant.

An ANOVA and pairwise comparison were performed to test the effects of age group on cortical thickness. We also performed a simple linear regression to test whether there was an overall effect of cortical thickness on d' memory scores (uncorrected for age group). Finally, a mediation analysis with the cortical thickness of the LIFG and age group as regressors and verbal memory score (d') as a dependent variable was performed using PROCESS for SPSS (Hayes, 2013). This was to investigate whether the effect of age group on d' memory scores would be mediated by hypothesized decreasing cortical thickness of the LIFG with age.

Resting-state fMRI data analysis – Functional specialisation. The resting-state fMRI data were analysed to determine whether the specialisation of LIFG functional connectivity indeed increases with age group, and whether this increase also mediates the hypothesised increase in verbal memory performance with age.

Preprocessing. Preprocessing steps and analyses were performed in FSL (Jenkinson, Beckmann, Behrens, Woolrich & Smith, 2012) and consisted of several steps using a combination of fMRI Expert Analysis Tool Version 6.00 and custom MATLAB (8.6, The MathWorks Inc., Massachusetts, USA) scripts. Registration to high resolution structural space images was carried out using FLIRT (Jenkinson, Bannister, Brady & Smith, 2002). Registration from high resolution structural to standard space was then further refined using FNIRT nonlinear registration (Andersson, Jenkinson & Smith, 2007a; Andersson, Jenkinson & Smith, 2007b). We applied motion correction using MCFLIRT (Jenkinson et al., 2002), nonbrain removal using BET (Smith, 2002), spatial smoothing using a Gaussian kernel of 5 mm and grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor. To correct for motion, we then used an Independent Component Analysis-based Automatic Removal of Motion Artifacts (ICA-AROMA) allowing us to preserve the autocorrelation structure of fMRI time-series and to avoid heteroscedasticity in group-level statistics (Pruim et al., 2015). ICA-AROMA uses an ICA decomposition on the partly preprocessed fMRI data. In this decomposition, a set of spatial and temporal features and a classification procedure were used to identify independent components representing motion artifacts. Finally, the selected components were removed from the fMRI time-series through linear regression. Residual noise was then removed by the regression of white matter (WM) and cerebrospinal fluid (CSF). The mean WM and CSF signals were extracted using masks obtained by multiplying a participant-specific tissue prior with an MNI152-derived tissue prior, both thresholded at 95% tissue

probability (Pruim et al., 2015). Hereafter, we used high-pass temporal filtering to remove slow drifts (sigma = 100 s). Resting-state scans were co-registered to the participant's structural T_1-weighted image.

Analyses. Functional connectivity analyses were performed measuring the correlations between the following ROIs as explained earlier: 1) between the LIFG (-50,25,12 in MNI space, BA 45/47, see Wagner et al., 1998) and the left MTL (737 voxels, center of gravity = (74,46,35)) and 2) between the right anterior IFG (from now on RIFG, 50,25,12) and the left MTL. For the left and right anterior IFG, we created an 8 mm sphere around the MNI coordinates from Wagner et al. (1998) and their right counterparts that served as a mask to warp onto the functional resting-state data of each subject. As a target in the left MTL, we used the center of an MTL cluster which exhibited a strong resting-state connectivity to the LIFG ($r \geq 0.2$) in an analysis of 1000 subjects (Yeo et al., 2011; via www.neurosynth.org).

We extracted the first eigenvariates for each ROI in native space by warping the three masks from MNI space into functional space. As a measure of functional connectivity, we calculated the Pearson correlations between the resting-state time series of the LIFG and the left MTL and between the RIFG and the left MTL for each participant. This resulted in two connectivity values per participant that we used to compute a laterality index (LI) for each participant. This LI was defined as the difference between the correlational value of the LIFG with the left MTL and the correlational value of the RIFG with the left MTL (r left – r right) (Nielsen, Zielinkski, Ferguson, Lainhart & Anderson, 2013). An ANOVA and pairwise comparison were performed to test the effects of age group on LI. We also performed a simple linear regression to test whether there was an overall effect of LI on d' memory scores (uncorrected for age group). If such an effect would exist, we would perform a mediation analysis with LI and age group as regressors and verbal memory score (d') as a dependent variable using PROCESS for SPSS (Hayes, 2013). In this way, we wanted to investigate whether the effect of age group on d' memory scores is mediated by possibly increasing lateralisation of functional connectivity with age. Additionally, a complementary simple linear regression was performed to investigate whether the assumed relationship between cortical thinning and functional specialisation as assumed in the literature indeed exists (Johnson, 2011).

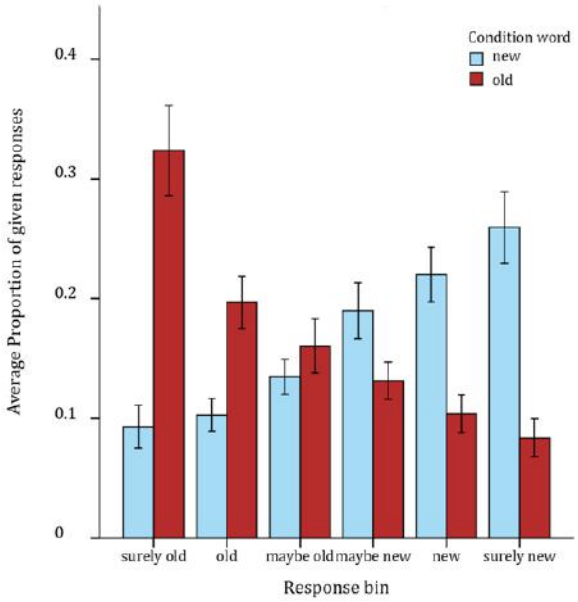


Fig. 1. Mean proportion of responses given for 'old' and 'new' conditions in each response bin.

Results

Behavioural results – Verbal memory

An interaction effect between the factors 'Condition' (old or new) and 'Response' on the proportion of responses given (2 x 6 ANOVA) showed that participants were overall well able to distinguish between old and new verbs in the recognition phase of the verbal memory test ($F(5, 1068) = 89.41$, $p < .001$). See Table 1 and Figure 1 for the mean proportions of responses given in 'old' and 'new' conditions for each response bin.

Post-hoc pairwise comparisons corrected for multiple comparisons (Bonferroni) revealed that the proportion of trials with responses 'surely old'

and 'old' was higher for old verbs than for new verbs (all $p < .001$). On the contrary, responding with 'maybe new', 'new' and 'surely new' occurred more often during trials with new verbs than with old verbs (all $p < .001$). A non-significant difference between proportions of old and new trials receiving the response 'maybe old' ($p > .05$) indicates that participants could not discriminate significantly between old and new verbs when they used this response button. Because of this, verbs that were responded to with 'maybe old' were categorized as trials of no interest and were excluded from further analyses consistent with previous studies (Liu et al., 2010; Qin et al., 2009).

Mean false alarm rates, hit rates and d' memory

Table 1.

Mean proportion of responses given for 'old' and 'new' conditions in each response bin, with standard deviations.

Response button	Condition 'new'	Condition 'old'
Sure old	0.093(0,085)	0.324(0,180)
Old	0.103(0,065)	0.197(0,104)
Maybe old	0.135(0,070)	0.160(0,108)
Maybe new	0.190(0,112)	0.131(0,075)
New	0.220(0,109)	0.104(0,075)
Sure new	0.259(0,143)	0.084(0,076)

scores for the different age groups are shown in Table 2. Firstly, we wanted to investigate whether there is a main effect of age group on mean d' scores using a one-way ANOVA. We found that age group indeed affects mean d' scores ($F(2, 86) = 5.15$, $p < .01$, see Table 2). Pairwise comparisons showed that mean d' scores for children were significantly lower than those for adolescents and adults ($p = .003$ and $p = .017$, respectively), whereas adolescents' mean scores were not significantly higher than those for adults ($p > .05$, see also Fig. 2A). These effects seem to be driven mostly by differences in mean false alarm rates between age groups ($F(2, 86) = 8.20$, $p = .001$), whereas mean hit rates did not differ significantly between children, adolescents and adults ($F(2, 86) = 0.10$, $p = 091$). Pairwise comparisons confirmed that mean false alarm rates follow an inverted pattern of the d' scores where children's false alarm rates were significantly higher than those for adolescents and adults ($p \leq .01$), whereas adolescents' mean scores were lower than those for adults ($p > .05$), but again this difference was not significant.

Structural MRI results – Cortical thickness

In a mediation analysis we wanted to investigate the research question whether cortical thickness of the LIFG predicts memory scores across age groups, as an indirect measure of the amount of synaptic pruning that has taken place allowing for greater processing efficiency. We already showed that age group predicts memory scores. Then, to check whether the decrease in cortical thickness in the LIFG from childhood into adulthood as reported in previous studies (e.g., Amlien et al., 2014) was also applicable to our sample, we performed a one-way ANOVA revealing an effect of age group on the absolute cortical thickness of the LIFG ($F(2, 81) = 18.67$, $p < .001$, see Table 2). This effect reflected a decrease in cortical thickness with increasing age group (see Fig. 2B). Pairwise comparisons showed that cortical thickness of the LIFG was significantly higher for children than for adolescents and adults ($p < .005$ and $p < .001$, respectively). There was also a significant decrease from adolescence into adulthood ($p < .01$).

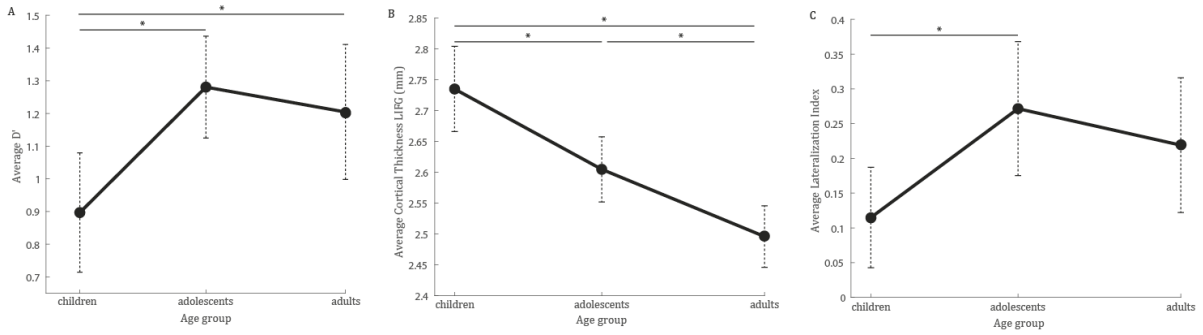


Fig. 2A. Mean d' per age group. Children's scores are significantly lower than that of adolescents and adults. **B.** Mean cortical thickness of the LIFG per age group. Cortical thickness significantly decreases with age across all age groups. **C.** Mean laterality indices of IFG-left MTL connectivity per age group. Children's IFG – left MTL connectivity is significantly less left-lateralized than adolescents.

Table 2.

Means and standard deviations for each variable per age group.

	Children	Adolescents	Adults
Hit rate	0.608(0.123)	0.617(0.145)	0.624(0.139)
False alarm rate	0.288(0.123)	0.180(0.082)	0.213(0.108)
d' memory scores	0.897(0.488)	1.281(0.409)	1.204(0.554)
Cortical thickness LIFG (mm)	2.735(0.178)	2.605(0.131)	2.496(0.133)
Laterality index	0.115(0.186)	0.272(0.249)	0.219(0.250)
LIFG – left MTL connectivity	0.178(0.179)	0.377(0.203)	0.282(0.235)

A simple linear regression showed that cortical thickness of the LIFG predicts d' memory scores irrespective of age group ($F(1, 82) = 9.25$, $p < .005$, $R^2 = .10$, $\beta = -.93$).

Because we found main effects of age group on d' and cortical thickness and of cortical thickness on d' (uncorrected for group), we performed a mediation analysis to investigate whether increasing age group leads to a better verbal memory performance because of the mediating effect of maturing LIFG cortical thickness. This showed that age group was no longer a significant predictor of d' scores after controlling for the mediating factor cortical thickness of the LIFG (age group: $\beta = .07$, $SE = 0.08$, $p > .05$; Cortical thickness: $\beta = -.74$, $SE = 0.37$, $p < .05$). The indirect effect of age group on d' scores through LIFG cortical thickness was tested using a bootstrap estimation approach with 1000 samples using PROCESS (Shrout & Bolger, 2002; Hayes, 2013). The results indicated the indirect coefficient was significant ($\beta = .09$, $SE = 0.05$, 95% confidence interval (CI) = 0.003, 0.21). This is an indication that cortical thickness of the LIFG mediates the effect of age group on d' scores. Approximately 11% of the variation was accounted for by age group and cortical thickness of the LIFG ($R^2 = 0.11$).

Resting-state fMRI results – Functional specialisation

In another mediation analysis we wanted to investigate the research question whether the degree of specialisation (i.e., lateralisation) of functional LIFG involvement predicts memory scores across age groups. To check whether the lateralisation of IFG functional connectivity actually increases from childhood into adulthood, we performed a one-way ANOVA revealing an significant effect of age group on the lateralisation index of functional connectivity of the IFG ($F(2, 81) = 3.36$, $p < .05$, see Table 2). Pairwise comparison showed that there was a significant increase in leftward lateralisation of IFG functional connectivity from childhood into adolescence ($p < .05$, see Fig. 2C). Adults also showed higher lateralisation indices than children, but this difference was not significant ($p > .05$), as was the decrease in lateralisation indices between adolescents and adults ($p > .05$).

Furthermore, a mediation analysis testing our hypothesis that maturation of functional specialisation (as reflected in an increasing lateralisation of IFG-left MTL connectivity) underlies the increase in verbal memory performance with age also requires an effect of laterality index on d'. However, a simple linear regression that was calculated to predict d' memory scores based on lateralisation of IFG-left MTL connectivity did not reach significance ($F(1, 82) = 0.41$, $p > .05$, $R^2 = 0.01$, $\beta = .16$). This means that we did not find evidence that lateralisation of IFG functional connectivity mediates the effect of age group on memory scores. Since cortical thickness did mediate verbal memory development, we performed a complementary simple linear regression to check whether the hypothesized relationship between decreasing cortical thickness of the LIFG and increasing specialisation (i.e., lateralisation) of IFG functional connectivity actually exists. There was a trend towards a significant effect from LIFG cortical thickness on laterality indices of IFG functional connectivity ($F(1, 79) = 3.82$, $p = .054$, $R^2 = 0.05$, $\beta = -.28$).

As an exploratory analysis, we further investigated whether LIFG connectivity with the left MTL in general could possibly be a predictor for developmental differences in verbal memory scores. There was an effect of age group on LIFG – left MTL connectivity ($F(2, 81) = 6.48$, $p < .01$, see Table 2), showing the same pattern as the lateralisation indices. This means there was an increase in LIFG

– left MTL functional connectivity from childhood into adolescence ($p < .05$). Adults also showed higher connectivity values than children, but this difference was not significant ($p > .05$), as was the decrease in connectivity between adolescents and adults ($p > .05$). Again, a simple linear regression showed no significant effect of LIFG – left MTL connectivity on d' memory scores ($F(1, 82) = 0.74$, $p > .05$, $R^2 = 0.10$, $\beta = .22$).

Discussion

Verbal memory

In this study, we aimed to investigate whether the maturation of neural structure and function is associated with behaviourally enhanced verbal memory performance from childhood into adulthood. Previous studies in the domain of verbal memory development that used intentional encoding tasks and a free recall test phase already reported that verbal memory performance increases during childhood and adolescence and flattens off in early adulthood, after which it decreases in older ages (e.g., Vakil & Blachstein, 1997; Vakil, Blachstein & Sheinman, 1998; Blachstein & Vakil, 2016; Davis et al., 2013). The results of our incidental verbal recognition memory task, reflecting strategy-free mnemonic abilities, are in line with these findings. Children between 10-12 years of age performed worse than 18 year old adolescents and 25-32 year old adults. Adults however did not differ significantly from adolescents, which is consistent with the flattening of the curve around these ages as shown in previous studies. Additionally, we showed that the increase in memory performance from childhood into adulthood is most probably driven by a decrease in false alarm rates between these age groups. This is consistent with early developmental word recognition studies showing a decrease in false memories throughout childhood and adolescence (e.g., Brainerd & Reyna, 1996; Brainerd, Reyna & Kneer, 1995; but see Brainerd, Reyna & Ceci, 2008 for an extensive discussion on this topic).

Cortical thickness

In examining the neural changes underlying memory development, we first investigated whether a developmental decrease in cortical thickness of the left inferior frontal gyrus (LIFG) mediates the effect of age group on verbal memory performance. We hypothesised that the protracted thinning of the PFC into early adulthood (e.g., Amlien et al., 2016), as an indirect measure of the amount of synaptic pruning that has taken place (Petanjek et al., 2011), would allow for greater processing efficiency in the LIFG over development. Verbal memory performance would benefit from this increased efficiency because the LIFG is associated with semantic control and has shown to be of particular importance in verbal memory tasks (Bokde et al., 2001; Kim, 2011). Our findings confirm this hypothesis.

Firstly, there was an effect of age group on average cortical thickness of the LIFG (BA 45/57). Children's LIFG were thicker than that of adolescents and adults, and LIFG cortical thickness kept decreasing from adolescence into adulthood. This linear pattern is consistent with several recent studies examining the developmental trajectories of cortical thickness in different brain areas in longitudinal and continuous cross-sectional samples (e.g., Koolschijn & Crone, 2013; Wierenga et al., 2014; Amlien et al., 2014; Ducharme et al., 2016). Although some studies found cubic or quadratic developmental trajectories of cortical thickness instead of a linear decline (e.g., Shaw et al., 2008; Raznahan et al., 2011), peaks always occurred before late childhood (8-10 years old) while after this period thickness only decreases, similar to our observation.

Secondly, cortical thickness of the LIFG predicted verbal memory scores in our sample. This is consistent with previous findings showing that there is a relationship between cortical thickness and verbal memory performance in adults (Walhovd et al., 2006; Dickerson et al., 2008), although the directionality of this relationship is different in our developmental study. In our sample, a general decrease in cortical thickness was associated with a general increase in verbal memory performance. In a sample with adults, individual differences in verbal memory abilities are positively correlated to measures of cortical thickness because such a sample is not characterised by maturational differences in cortical thinning. In another developmental study by Østby et al. (2011), focusing on visual memory instead of verbal memory (which has different developmental implications) found results similar to ours in which an increase in performance correlates with a decrease in cortical thickness. A developmental study most similar to our design (Sowell et al., 2001) already showed that an increase in verbal memory performance across development is associated with decreasing cortical thickness of the PFC. Since Sowell et al. (2001) used an intentional verbal learning task to test their hypotheses we could not

dissociate whether PFC structural maturation truly affected mnemonic components of verbal memory development, or whether the effect was confounded by improved strategic abilities relying on the PFC as well. Based on our incidental, strategy-free encoding task, we can discard the latter possibility and conclude that PFC thinning most probably aids the mnemonic processes themselves.

More importantly, besides the replication of the effect of age on cortical thickness and the general effect of cortical thickness on verbal memory scores, we also confirmed our main hypothesis that cortical thickness mediates the developmental effect of age group on verbal memory scores. This mediation has to our knowledge not explicitly been tested in the literature before, and indicates that the development of verbal memory abilities with age is at least partially driven by maturational differences in cortical thickness of the LIFG. Our interpretation of this is that the cortical thinning, which we assume to indirectly reflect synaptic pruning (Petanjek et al., 2011), indeed allows for more efficient processing in the LIFG. In our sample, this efficiency then appears to be beneficial for a verbal memory task requiring semantic control processes that rely on the LIFG (Bokde et al., 2001; Badre & Wagner, 2007).

Functional specialisation

Another level of neural maturation that potentially underlies changes in mnemonic abilities over the course of development, is that of functional connectivity between relevant brain regions. We expected that due to protracted cortical thinning (i.e., assumed synaptic pruning) in the PFC, relevant changes in short- and long-range functional connectivity would occur within PFC regions during late childhood and adolescence (Jolles et al. 2011). This could for instance allow for specialisation of the IFG within larger functional networks. We hypothesised that in relation to a linguistic task as our verbal memory task, associated functional specialisation could be reflected in increased leftward lateralisation of relevant IFG – left MTL resting-state connectivity from childhood into early adulthood (Szaflarski et al., 2006 a; Holland et al., 2007). Therefore, we investigated whether increasingly left lateralized IFG – left MTL functional connectivity with increasing age group also mediates the effect of age group on verbal memory performance. This hypothesis could not be confirmed by our data. We did find an increase in leftward lateralisation of IFG – left MTL connectivity from childhood to adolescence, which is in line with previous studies, showing increasingly left lateralized task-related fMRI activity during development (reviewed in Holland et al., 2007). However, although mean laterality indices and mean memory scores develop along a similar trajectory, increased leftward lateralisation of IFG – left MTL connectivity with age did not significantly predict the increase in memory performance across development in our sample. This means that specialisation (i.e., lateralisation) of LIFG functional connectivity indeed increases with age group, but most likely does not mediate the increase in verbal memory performance with age.

Taken together, structural maturation of the LIFG appears to be important for verbal memory development, while there is no evidence that the development of left lateralized functional connectivity between the IFG and the left MTL influences the development of mnemonic abilities. This could mean that the maturation of the LIFG itself as a hub for semantic control specifically contributes to gains in performance with age, while the maturation of its interaction with the left MTL might not be of particular relevance. In accordance with the latter, we neither found an effect of mere LIFG – left MTL functional connectivity on verbal memory performance, although this maturation also showed a developmental trajectory similar to that of our memory scores. One argument to justify why only the specific maturation of the LIFG itself would matter for verbal memory development and not its interaction with the left MTL, would be that the MTL is assumed to play an important role in binding processes (Werkle-Bergner et al., 2006). Although most verbal memory studies report task-related MTL activity (see for a meta-analysis Kim, 2011), this region is more involved if the study material consists of verbal associations instead of single items. It might be the case that the relation between the design of our item-recognition task and the reliance on the left MTL is not strong enough to reach a significant association of LIFG – left MTL resting-state functional connectivity on task performance. Because this association is less direct than the task-related fMRI activity our hypotheses were based upon and would therefore profit from a more sensitive design (see 'Limitations'). This also means that specialisation or laterality of IFG – left MTL connectivity in particular might not have been the most optimal measure for our laterality index after all. Therefore, the current results on the laterality index do not rule out the possibility that specialisation of LIFG functional connectivity plays a role in increasing verbal memory performance with age group.

Another possibility to explain why we did find a mediation of LIFG cortical thickness maturation on the effect of age group on verbal memory scores, but not of lateralisation of IFG – left MTL functional connectivity, would be that the relationship between these two measures is not as strong as we assumed. Like previous studies (reviewed in Johnson, 2011), we assumed that cortical thinning and functional specialisation were driven by the same maturational mechanism, namely synaptic pruning leading to increased efficiency, and that both measures would therefore be highly negatively correlated. However, previous studies did not test this relationship directly. Interestingly, in our sample there was indeed a trend towards a negative correlation between cortical thickness of the LIFG and laterality indices of IFG – left MTL connectivity. Nonetheless, this relationship did not reach significance and thus does not provide us with conclusive evidence that it indeed exists. This means there is also a likelihood that cortical thinning and functional specialisation (i.e., lateralisation) during development might be driven by different underlying processes. The finding that cortical thinning mediates the effect of age group on memory score can also be due to other underlying processes than synaptic pruning, such as white matter myelination (Sowell et al., 2003). Intracortical myelination could impact measures of cortical thickness by changing the signal contrasts such that the boundary between white and gray matter, influencing the measure of cortical thickness, moves outward with increasing age (Mills & Tamnes, 2014). If myelination would indeed underlie the age and mediation effects we found for cortical thickness of the LIFG, this would also allow for more efficient processing in the LIFG, but does not necessarily affect its (lateralisation of) functional connectivity with the left MTL as was assumed for synaptic pruning (although it might).

Limitations

Laterality indices and memory scores seem to develop along a similar trajectory on a group level. However, we were unable to detect a significant effect of lateralisation of IFG – left MTL connectivity on memory performance across development in our sample. This could of course be due to a lack of such a relationship as we described above, but it is also possible that this relationship does exist. In the latter case, the design of our study was not sensitive enough to detect it. This could for instance be due to the distribution of our cross-sectional sample, which consists of three different age groups with large gaps in between. Hereby, information is lost because of the use of averages to create and compare group means, which is not optimal when investigating mediations. A continuous distribution however, in which all ages relevant for PFC maturation (i.e., between 10 and 32 years old) were represented, would have allowed us to truly model and compare neural and behavioural developmental trajectories. Another limitation of the use of a cross-sectional design is the amount of participants that is necessary to detect an effect of brain structure or function on behaviour (Steen, Hamer & Lieberman, 2007). In our cross-sectional study, inter-individual differences other than age contribute to the variation in our sample, which makes it more difficult to detect subtle maturational differences and effects than in longitudinal studies where there is no such variation. A future study would thus benefit from either a more continuous distribution of different age samples, or a longitudinal design. Thus, the most optimal design to investigate whether the maturation of neural structure and function mediates the development of verbal memory would be a longitudinal one, or a cross-sectional design with continuous sampling of different age groups and a substantial amount of subjects to represent each age bin. One valuable addition could also be a within-subjects contrast, for instance by examining both item recognition as well as an associative memory component. A primary reason for this would be that the assumed reliance on MTL structures is bigger for associative memory tasks, which would increase the chance of finding a relation between task performance and resting-state LIFG – left MTL connectivity. Besides this, such a within-subjects contrast would allow us to not only draw inferences on quantitative aspects of development, but also on a qualitative level: do different kinds of mnemonic representations develop differently from childhood into adulthood?

Conclusions

We aimed to investigate whether the maturation of neural structure and function is associated with behaviourally enhanced verbal memory performance from childhood into adulthood. In summary, we replicated previous studies showing an increase in verbal memory performance from childhood into adolescence and adulthood. Furthermore, we have shown that cortical thinning of the LIFG across the three age groups mediates the effect of age group on mnemonic abilities. This we interpreted as a reflection of the gains

in semantic control efficiency of the LIFG after synaptic pruning has taken place during adolescence. There was no significant effect of lateralisation of IFG – left MTL functional connectivity on verbal memory performance. This indicates that structural LIFG maturation is indeed an important contributor to verbal memory development during adolescence, whereas functional LIFG specialisation at least seems to play a less significant role.

References

Alloway, T. P. (2007). *Automated Working: Memory Assessment: Manual*. Pearson.

Amlien, I.K., Fjell, A.M., Tamnes, C.K., Grydeland, H., Krogsrud, S.K., Chaplin, T.A., Rosa, M.G.P. & Walhovd, K.B. (2014). Organizing principles of human cortical development—thickness and area from 4 to 30 years: insights from comparative primate neuroanatomy. *Cerebral Cortex, 26,* 257-267.

Andersson, J.L.R., Jenkinson, M. & Smith, S. (2007a). Non-linear optimisation. *FMRIB technical report TR07JA1*.

Andersson, J.L.R., Jenkinson, M. & Smith, S. (2007b). Non-linear registration, aka Spatial normalisation. *FMRIB technical report TR07JA2*

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

Baltes, P.B., Lindenberger, U. & Staudinger, U.M. (2006). Lifespan theory in developmental psychology. In: Damon, W. (Series Ed.), & Lerner, R.M. (Vol. Ed.), *Handbook of Child Psychology: vol. 1. Theoretical Models of Human Development,* 6. Wiley, New York.

Blachstein, H. & Vakil, E. (2016). Verbal learning across the lifespan: an analysis of the components of the learning curve. *Aging, Neuropsychology and Cognition, 23*(2), 133-153.

Brainerd, C.J. & Reyna, V.F. (1996). Mere memory testing creates false memories in children. *Developmental Psychology, 32,* 467–476.

Brainerd, C.J., Reyna, V.F., & Ceci, S.J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin, 134*(3), 343-382.

Brainerd, C.J., Reyna, V.F. & Kneer, R. (1995). False-recognition reversal: When similarity is distinctive. *Journal of Memory & Language, 34,* 157–185.

Bokde, A.L.W., Tagaments, M.A., Friedman, R.B. & Horwitz, B. (2001). Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. *Neuron, 30,* 609-617.

Clarkson, M.J., Cardoso, M.J., Ridgway, G.R., Modat, M., Leung, K.K., Rohrer, J.D., Fox, N.C. & Ourselin, S. (2011). A comparison of voxel and surface based cortical thickness estimation methods. *NeuroImage, 57*(3), 856-865.

Davis, H.P., Klebe, K.J., Guinther, P.M., Schroder, K.B., Cornwell, R.E. & James, L.E. (2013). Subjective Organization, Verbal Learning and Forgetting Across the Life Span: From 5 to 89. *Experimental Aging Research, 39,* 1-26.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, M.A., Maguire, R.P., Hyman, B.T., Albert, M.S. & Killiany, R.J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage, 31*(3), 968-980.

Dickerson, B.C., Fenstermacher, E., Salat, D.H., Wolk, D.A., Maguire, R.P., Desikan, R., Pacheco, J., Quinn, B.T., Van der Kouwe, A., Greve, D.N., Blacker, D., Albert, M.S., Killiany, R.J. & Fischl, B. (2008). Detection of cortical thickness correlates of cognitive performance: reliability across MRI scan sessions, scanners, and field strengths. *NeuroImage, 39*(1), 10-18.

Ducharme, S., Albaugh, M.D., Nguyen, T.V., Hudziak, J.J., Mateos-Pérez, J.M., Labbe, A., Evans, A.C., Karama, S. & Brain Development Cooperative Group. (2016). Trajectories of cortical thickness maturation in normal brain development—The importance of quality control procedures. *NeuroImage, 125,* 267-279.

Durston, S. & Casey, B.J. (2006). What have we learned about cognitive development from neuroimaging? *Neuropsychologia, 44*(11), 2149-2157.

Durston, S., Davidson, M.C., Tottenham, N., Galvan, A., Spicer, J., Fossella, J.A. & Casey, B.J. (2006). A shift from diffuse to focal cortical activity with development. *Developmental Science, 9*(1), 1-8.

Duverne, S., Motamedinia, S. & Rugg, M.D. (2009). The Relationship between Aging, Performance and the Neural Correlates of Successful Memory Encoding. *Cerebral Cortex, 19,* 733-744.

Everts, R., Lidzba, K., Wilke, M., Kiefer, C., Mordasini, M., Schroth, G., Perrig, W. & Steinlin, M. (2009). Strengthening of laterality of verbal and visuospatial functions during childhood and adolescence. *Human Brain Mapping, 30*(2), 473-483.

Fair, D.A., Dosenbach, N.U., Church, J.A., Cohen, A.L., Brahmbhatt, S., Miezin, F.M., Barch, D.M., Raichle, M.E., Petersenm S.E. & Schlaggar, B. L. (2007). Development of distinct control networks through segregation and integration. *Proceedings of the National Academy of Sciences, 104*(33), 13507-13512.

Fair, D.A., Cohen, A.L., Power, J.D., Dosenbach, N.U., Church, J.A., Miezin, F.M., Schlaggar, B.L. & Petersen, S.E. (2009). Functional brain networks develop from a "local to distributed" organization. *PLoS Computational Biology, 5*(5), e1000381.

Finn, A.S., Kalra, P.B., Goetz, C., Leonard, J.A., Sheridan, M.A. & Gabrieli, J.D.E. (2016). Developmental dissociation between the maturation of procedural memory and declarative memory. *Journal of Experimental Child Psychology, 142,* 212-220.

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of*

Sciences, 97(20), 11050-11055.

Fjell, A.M., Walhovd, K.B., Westlye, L.T., Østby, Y., Tamnes, C.K., Jernigan, T.L., Gamst, A. & Dale, A.M. (2010). When does brain aging accelerate? Dangers of quadratic fits in cross-sectional studies. *NeuroImage, 50*(4), 1376-1383.

Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences USA, 101*(21), 8174–8179.

Gogtay, N., Nugent, T.F., Herman, D.H., Ordonez, A., Greenstein, D., Hayashi, K.M. (2006). Dynamic mapping of normal human hippocampal development. *Hippocampus, 16*(8), 664–672.

Golby, A.J., Poldrack, R.A., Brewer, J.B., Spencer, D., Desmond, J.E., Aron, A.P. & Gabrieli, J.D.E. (2001). Material-specific lateralization in the medial temporal lobe and prefrontal cortex during memory encoding. *Brain, 124,* 1841-1854.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Heaton, R.K., Chelune, G.J., Talley, J.L., Kay, G.G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual: Revised and expanded*. Odessa, FL: Psychological Assessment Resources.

Holland, S.K., Plante, E., Byars, A.W., Strawsburg, R.H., Schmithorst, V.J. & Ball, W.S. (2001). Normal fMRI brain activation patterns in children performing a verb generation task. *NeuroImage, 14*(4), 837-843.

Holland, S.K., Vannest, J., Mecoli, M., Jacola, L.M., Tillema, J.M., Karunanayaka, P.R., Schmithorst, V.J., Yuan, W., Plante, E. & Byars, A. W. (2007). Functional MRI of language lateralization during development in children. *International Journal of Audiology, 46*(9), 533-551.

Hutton, C., Draganski, B., Ashburner, J. & Weiskopf, N. (2009). A comparison between voxel-based cortical thickness and voxel-based morphometry in normal aging. *NeuroImage, 48*(2), 371-380.

Jenkinson, M., Bannister, P., Brady, M. & Smith, S. (2002). Improved Optimisation 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., Woolrich, M.W. & Smith, S.M. (2012). FSL. *NeuroImage, 62,* 782-790.

Johnson, M.H. (2001). Functional brain development in humans. *Nature Reviews Neuroscience, 2,* 475-483.

Johnson, M.H. (2011). Interactive specialization: a domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience, 1*(1), 7-21.

Jolles, D.D., Van Buchem, M.A., Crone, E.A. & Rombouts, S.A. (2011). A comprehensive study of whole-brain functional connectivity in children and young adults. *Cerebral Cortex, 21*(2), 385-391.

Kelly, A.C., Di Martino, A., Uddin, L.Q., Shehzad, Z., Gee, D.G., Reiss, P.T., Margulies, D.F., Castellanos, F.X. & Milham, M.P. (2009). Development of anterior cingulate functional connectivity from late childhood to early adulthood. *Cerebral Cortex, 19*(3), 640-657.

Kim, H. (2011). Neural activity that predicts subsequent memory and forgetting: A meta-analysis of 74 fMRI studies. *NeuroImage, 54,* 2446-2461.

Koolschijn, P.C.M. & Crone, E.A. (2013). Sex differences and structural brain maturation from childhood to early adulthood. *Developmental Cognitive Neuroscience, 5,* 106-118.

Lenroot, R.K. & Giedd, J.N. (2006). Brain development and adolescents: insights from anatomical magnetic resonance imaging. *Neuroscience and Biobehavioral Reviews, 30,* 718-729.

Liu, X., Qin, S., Rijpkema, M., Luo, J. & Fernández, G. (2010). Intermediate Levels of Hippocampal Activity Appear Optimal for Associative Memory Formation. *PLoS ONE, 5*(10), e13147.

MacMillan, N.A. & Creelman, C.D. (2005). *Detection Theory: A User's Guide*. Psychology Press, 7.

Maril, A., Avital, R., Reggev, N., Zuckerman, M., Sadeh, T., Ben Sira, L., et al. (2011). Event congruency and episodic encoding: a developmental fMRI study. *Neuropsychologia, 49*(11), 3036–3045.

Mills, K.L., & Tamnes, C.K. (2014). Methods and considerations for longitudinal structural brain imaging analysis across development. *Developmental Cognitive Neuroscience, 9,* 172-190.

Müller et al. (in preparation).

Nielsen, J.A., Zielinski, B.A., Ferguson, M.A., Lainhart, J.E. & Anderson, J.S. (2013). An Evaluation of the Left-Brain vs. Right-Brain Hypothesis with Resting State Functional Connectivity Magnetic Resonance Imaging. *PLoS ONE, 8*(8), e71275.

Ofen, N. & Shing, Y.L. (2013). From perception to memory: Changes in memory systems across the lifespan. *Neuroscience and Biobehavioral Reviews, 37*(9B), 2258–2267.

Opitz, B., Mecklinger, A. & Friederici, A.D. (2000). Functional asymmetry of human prefrontal cortex: Encoding and retrieval of verbally and nonverbally coded information. *Learning and Memory, 7,* 85-96.

Østby, Y., Tamnes, C.K., Fjell, A.M., & Walhovd, K.B. (2011). Dissociating memory processes in the developing brain: the role of hippocampal volume and cortical thickness in recall after minutes versus days. *Cerebral Cortex, 22,* 381-390.

Petanjek, Z., Judaš, M., Šimić, G., Rašin, M.R., Uylings, H.B., Rakic, P. & Kostović, I. (2011). Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proceedings of the National Academy of Sciences, 108*(32), 13281-13286.

Pruim, R.H.R., 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.

Qin, S., Rijpkema, M., Tendolkar, I., Piekema, C.,

Hermans, E. J., Binder, M., Petersson, K.M., Luo, J. & Fernández, G. (2009). Dissecting medial temporal lobe contributions to item and associative memory formation. *NeuroImage, 46*(3), 874-881.

Ragland, J.D., Turetsky, B.I., Gur, R.C., Gunning-Dixon, F., Turner, T., Schroeder, L., Chan, R. & Gur, R. E. (2002). Working memory for complex figures: an fMRI comparison of letter and fractal n-back tasks. *Neuropsychology, 16*(3), 370-379.

Raz, N. & Rodrigue, K.M. (2006). Differential aging of the brain: patterns, cognitive correlates and modifiers. *Neuroscience and Biobehavioral Reviews, 30,* 730-748.

Raznahan, A., Shaw, P., Lalonde, F., Stockman, M., Wallace, G.L., Greenstein, D., Clasen, L., Gogtay, N. & Giedd, J.N. (2011). How does your cortex grow? *The Journal of Neuroscience, 31*(19), 7174-7177.

Ressel, V., Wilke, M., Lidzba, K., Lutzenberger, W. & Krägeloh-Mann, I. (2008). Increases in language lateralization in normal children as observed using magnetoencephalography. *Brain and Language, 106*(3), 167-176.

Ségonne, F., Pacheco, J. & Fischl, B. (2007). Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *Medical Imaging, IEEE Transactions on, 26*(4), 518-529.

Shaw, P., Kabani, N.J., Lerch, J.P., Eckstrand, K., Lenroot, R., Gogtay, N., Greenstein, D., Clasen, L., Evans, A., Rapoport, J.L., Giedd, J.N. & Wise, S.P. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *The Journal of Neuroscience, 28*(14), 3586-3594.

Shing, Y.L., Werkle-Bergner, M., Brehmer, Y., Müller, V., Li, S.C. & Lindenberger, U. (2010). Episodic memory across the lifespan: The contributions of associative and strategic components. *Neuroscience & Biobehavioral Reviews, 34*(7), 1080-1091.

Shing, Y.L., Werkle-Bergner, M., Li, S.C. & Lindenberger, U. (2008). Associative and strategic components of episodic memory: A life-span dissociation. *Journal of Experimental Psychology General, 137,* 495-513.

Shrout, P.E. & Bolger, N. (2002). Mediation in Experimental and Nonexperimental Studies: New Procedures and Recommendations. *Psychological Methods, 7*(4), 422-445.

Simons, J.S. & Spiers, H.J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience, 4,* 637-648.

Sled, J.G., Zijdenbos, A.P. & Evans, A.C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *Medical Imaging, IEEE Transactions on, 17*(1), 87-97.

Smith, S. (2002). Fast Robust Automated Brain Extraction. *Human Brain Mapping, 17*(3), 143-155.

Sowell, E.R., Delis, D., Stiles, J., & Jernigan, T.L. (2001). Improved memory functioning and frontal lobe maturation between childhood and adolescence: a structural MRI study. *Journal of the International Neuropsychological Society, 7*(3), 312-322.

Sowell, E.R., Peterson, B.S., Thompson, P.M., Welcome, S.E., Henkenius, A.L. & Toga, A.W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience, 6,* 309-315.

Sowell, E.R., Thompson, P.M., Leonard, C.M., Welcome, S.E., Kan, E., Toga, A.W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *Journal for Neuroscience, 24*(38), 8223–8231.

Staresina, B.P., Duncan, K.D., Davachi, L. (2011). Perirhinal and parahippocampal cortices differentially contribute to later recollection of object- and scene-relatedevent details. *Journal for Neuroscience, 31*(24), 8739–8747.

Steen, R.G., Hamer, R.M., & Lieberman, J.A. (2007). Measuring brain volume by MR imaging: impact of measurement precision and natural variation on sample size requirements. *American Journal of Neuroradiology, 28*(6), 1119-1125.

Supekar, K., Musen, M. & Menon, V. (2009). Development of large-scale functional brain networks in children. *PLoS Biology, 7*(7), e1000157.

Szaflarski, J.P., Holland, S.K., Schmithorst, V.J., & Byars, A.W. (2006a). fMRI study of language lateralization in children and adults. *Human Brain Mapping, 27*(3), 202-212.

Szaflarski, J.P., Schmithorst, V.J., Altaye, M., Byars, A.W., Ret, J., Plante, E. & Holland, S. K. (2006b). A longitudinal functional magnetic resonance imaging study of language development in children 5 to 11 years old. *Annals of Neurology, 59*(5), 796-807.

Tavor, I., Jones, O.P., Mars, R.B., Smith, S.M., Behrens, T.E. & Jbabdi, S. (2016). Task-free MRI predicts individual differences in brain activity during task performance. *Science, 352*(6282), 216-220.

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 USA, 94,* 14792-14797.

Uddin, L.Q., Supekar, K. & Menon, V. (2010). Typical and atypical development of functional human brain networks: insights from resting-state FMRI. *Frontiers in Systems Neuroscience, 4,* 21.

Vakil, E. & Blachstein, H. (1997). Rey ALVT: Developmental norms for adults and the sensitivity of different memory measures to age. *The Clinical Neuropsychologist, 11,* 356-369.

Vakil, E., Blachstein, H. & Sheinman, M. (1998). Rey ALVT: Developmental norms for children and the sensitivity of different memory measures to age. *Child Neuropsychology, 15,* 21-39.

Wagner, A.D., Schacter, D.L., Rotte, M., Koutstaal, W., Maril, A., Dale, A.M., Rosen, B.R. & Buckner, R.L. (1998). Building Memories: Remembering and Forgetting of Verbal Experiences as Predicted by Brain Activity. *Science, 281,* 1188-1191.

Walhovd, K.B., Fjell, A.M., Dale, A.M., Fischl, B., Quinn, B.T., Makris, N., Salat, D. & Reinvang, I. (2006). Regional cortical thickness matters in recall after

months more than minutes. *NeuroImage, 31*(3), 1343-1351.
Werkle-Bergner, M., Müller, V., Li, S.C. & Lindenberger, U. (2006). Cortical EEG correlates of succesful memory encoding: Implications for lifespan comparisons. *Neuroscience and Biobehavioral Reviews, 30,* 839-854.
Wierenga, L.M., Langen, M., Oranje, B. & Durston, S. (2014). Unique developmental trajectories of cortical thickness and surface area. *NeuroImage, 87,* 120-126.
Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., Fischl, B., Liu, H. & Buckner, R.L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology, 106*(3), 1125-1165.

Aberrant Hippocampal Morphology and Function in a Mouse Model for Post-Traumatic Stress Disorder

Leonie J. C. Madder[1]
Supervisors: Marloes Henckens[1], Judith Homberg[1], Tamas Kozicz[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

Post-Traumatic Stress Disorder (PTSD) is a debilitating disorder that affects approximately 8% of the population. It can be caused by experiencing a traumatic event and is characterised by symptoms like insomnia, hypervigilance and hyperarousal, and intrusive memories (i.e., flashbacks and nightmares) of the event which severely deteriorate quality of life. Overgeneralisation of the trauma memory has been suggested to underlie these intrusive memories triggered by indiscriminate environmental factors, and has been related to deficits in hippocampal function. Particularly pattern separation, the process by which memories are stored as unique representations resistant to confusion, mediated by the dentate gyrus (DG), may be compromised. However, since most evidence originates from patient studies that are done retrospectively, it is unknown whether this abnormal hippocampal functioning is in fact part of pathology or constitutes a predisposition to PTSD. Here, we used a validated mouse model for PTSD induction to induce a PTSD-like phenotype in part of the mice, whereas others are resilient and do not display any PTSD-related symptoms. This model perfectly mimics the human situation, in which only 20 - 25% of the individuals experiencing a trauma will ultimately develop the disorder, whereas the majority stays healthy. Using various methods, we assessed fear generalisation, and monitored hippocampal activity using arterial spin labelling, and the PTSD-associated neuroendocrine changes (i.e., corticosterone levels) over the course of PTSD development in PTSD-like compared to resilient mice. Moreover, Golgi staining enabled us to assess spine density in the ventral DG, to assess the potential for synaptic connectivity at the brain's site for pattern separation. In line with literature, PTSD-like animals displayed a suppressed corticosterone response to stress but no differential fear generalisation compared to resilient animals. Moreover, we observed reduced hippocampal activity in PTSD-like mice, but only after trauma induction, indicating the reflection of pathology rather than a predisposition. Lastly, spine density in the ventral DG was significantly reduced in PTSD-like mice. These results indicate that the hippocampal dysfunction associated with PTSD-like symptomatology in mice is a consequence of PTSD development, and that high stress-induced corticosterone levels might be protective for the development of PTSD.

Keywords: PTSD, fear generalization and pattern separation, ventral dentate gyrus, hippocampus, spine density, cerebral blood flow, corticosterone

Corresponding author: Leonie Madder; **E-mail:** leonie.madder@live.nl

Post-Traumatic Stress Disorder (PTSD) is a psychiatric disorder that affects approximately 8% of the population (Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995). It can develop following exposure to a trauma that threatens the physical or psychological integrity of oneself or others (American Psychiatric Association, 2000). Symptoms that characterise the disorder are insomnia, hypervigilance and hyperarousal, flashbacks of the events and recurrence of memories of the trauma (Pitman et al., 2012), which can severely deteriorate one's quality of life. Its pathophysiological basis is poorly understood, making current treatments symptomatic and only effective for fewer than half of the patients. Hence, there is an urgent need to advance the understanding of its pathophysiology. Symptoms of flashbacks and recurrences of memories experienced by PTSD patients are thought to be caused by the process of fear generalisation (Greenberg, Carlson, Cha, Hajcak, & Mujica-Parodi, 2013), the process of generalising context-dependent fear to other (differing) contexts. Overgeneralisation of the trauma memory makes cues that are only remotely similar to those experienced during trauma capable of inducing a full-blown fear response (Golub, Mauch, Dahlhoff, & Wotjak, 2005; Greenberg et al., 2013), which could in turn relate to the typical avoidance symptoms observed in PTSD. Heightened fear generalisation has been linked to impairments in the hippocampus during the recollection of contextual fear memories in PTSD (Golub et al., 2005). In line with this, abnormal hippocampal functioning has been associated with PTSD (American Psychiatric Association, 2000; Bonne et al., 2001; Bonne et al., 2008; Shin, Rauch, & Pitman, 2006; Shin et al., 2004), with human magnetic resonance imaging (MRI) studies reporting on reduced hippocampal volumes and abnormal functioning such as memory deficits (Shin et al., 2004) in PTSD patients (Bonne et al., 2001; Shin et al., 2006). Animal models for PTSD have also reported abnormal hippocampal neuronal morphology after PTSD development. However, since these studies are usually done retrospectively (i.e., after PTSD diagnosis), it is unknown whether these abnormalities are the respective cause or consequence of the disease (American Psychiatric Association, 2000).

Fear generalisation is thought to originate in a particular subregion of the hippocampus, namely the dentate gyrus (DG; Ji & Maren, 2008). The DG is implicated in pattern separation, the process of extracting unique information from events that have overlapping or similar representations in order to make distinctions between internal representations of these similar memories (Aimone, Deng, & Gage, 2011; Sahay, Wilson, & Hen, 2011). Impairments in pattern separation have been related to an excessive generalisation of fear (Sahay et al., 2011), and it is thought to be an endophenotype for anxiety disorders like PTSD (Kheirbek, Klemenhagen, Sahay, & Hen, 2012). Granule cells are the primary cell type of the DG (Amaral, Scharfman, & Lavenex, 2007).These cells play a critical role in the hippocampal circuitry of health and disease, because these cells regulate the information flow into the hippocampus (Zohar et al., 2011) and thereby control contextual learning and features of anxiety (Kheirbek et al., 2013). The DG can be functionally segmented into the dorsal and ventral DG. The dorsal DG is primarily involved in information processing and cognitive functions needed for exploration, navigation, and locomotion, whereas the ventral DG is involved in regulating affect, motivational, and emotional behaviour and the hormonal stress response (Fanselow & Dong, 2010). In addition, the dorsal DG is implicated in encoding but not retrieval of contextual fear memories while the ventral DG is powerful in suppressing innate anxiety (Kheirbek et al., 2013). Also, abnormal function of the dentate gyrus in particular seems to be associated with PTSD (American Psychiatric Association, 2000; Kheirbek et al., 2012; Kheirbek et al., 2013; McEwan, 1999; Sahay et al., 2011; Zohar et al., 2011).

However, the DG is highly sensitive to stress (Kavushansky, Vouimba, Cohen, & Richter-Levin, 2006) and elevated corticosterone levels are especially harmful to the DG (McEwan, 1999). Also, a relationship between the density of spine in the dentate gyrus and stress is found in the literature and density changes are implicated in post-traumatic stress disorder (Adamec, Hebert, Blundell, & Mervis, 2012; Cohen, Kozlovsky, Matar, Zohar, & Kaplan, 2014; Diamond et al., 2006; Dias et al., 2014; Shors & Leuner, 2004; Zohar et al., 2011). Spines play a role in synaptic plasticity (Nimchinsky, Sabatini, & Svoboda, 2002), and are involved in the establishment and maintenance of connections with axons from other neurons (Bohlen & Halbach, 2009). Spines are important for neural processes, and changes in density might reflect adjustments in synaptic transmission (Bohlen & Halbach, 2009). These changes in spine density, and subsequently the changes in synaptic transmission are thought to underlie psychopathological and pathophysiological alterations (Bohlen & Halbach, 2009). Many studies have reported the existence of a relationship between stress and spine density in the dentate gyrus. However, this literature is often conflicting.

Dias et al. (2014) showed that spine density in the rat DG is increased in a model for generalised anxiety disorder (Dias et al., 2014), while others have shown a reduction in spine density in the rat DG in response to acute stress (Adamec et al., 2012; Cohen et al., 2014; Zohar et al., 2011). Additionally, PTSD is associated with alterations in the hypothalamic-pituitary-adrenal (HPA) axis, resulting in hypocortisolism by increased negative feedback (Van Zuiden et al., 2012). Although generally reduced cortisol secretion has been observed in PTSD patients (Yehuda, Teicher, Trestman, Levengood, & Siever, 1996), animal studies trying to find a causal link between suppressed corticosterone signalling and PTSD susceptibility are conflicting. In these studies, the effects of corticosterone administration following trauma (to increase the potentially suppressed corticosterone response contributing to PTSD development) are investigated. On the one hand, Zohar et al. (2011) found that the administration of a high dose corticosterone immediately after trauma exposure reduced the risk of PTSD development, and prevented the trauma-induced reduction in dorsal DG spine density as observed in the extreme responding animals. However, on the other hand, Kaouane et al. (2012) found that corticosterone, when administered following severe stress, induced PTSD-like fear memory generalisation, and increased fear responses when administered after relatively mild stress. Because of these contractor findings, the exact contribution of corticosterone to PTSD development and DG neuronal morphology remain elusive and further research is needed to elucidate its exact role in PTSD. Here, in a prospective study, we wanted to elucidate the causal relationship between neuroendocrine parameters (i.e., corticosterone levels), fear generalisation, and hippocampal function and the susceptibility to PTSD. Therefore, mice are exposed to a validated mouse model for the induction of PTSD (Lebow et al., 2012), in which mice are first exposed to a severe stressor (i.e., intense electric foot shock) in a certain context (context A), followed by a mild stressor (i.e., weak electric foot shock) the next day in a distinct context (context B) which is different in terms of spatial, auditory, and olfactory cues. By stressing the animals in two distinct contexts, this protocol actively stimulates the process of fear generalisation (of context A to B), and has been shown to reliably induce PTSD-like symptomatology (i.e., hypervigilance, compulsivity, impaired attention, compromised risk assessment, and insomnia) in a subset of mice, whereas others do not show any of these behaviours and are resilient. This observation perfectly mimics the human situation in which only a relatively small fraction (15 - 20%) of the individuals exposed to a traumatic event develops PTSD (Breslau, 2001), but is typically ignored in animal studies implementing very severe traumas to (artificially) warrant PTSD development (Servatius, Ottenweller, & Natelson, 1995). It is, however, unknown what causes this differential vulnerability, and an increased understanding would open up new avenues for treatment. Fear generalisation was measured by the amount of freezing during re-exposure of a traumatic experience (i.e., during trigger exposure). Basal corticosterone levels were measured both at the start and at the end of the protocol, as well as trauma- and trigger-induced corticosterone stress response and corticosterone responses to restraint stress after the development of pathology. To monitor hippocampal function, we obtained two MRI scans (i.e., arterial spin labelling [ASL]; Zerbi et al., 2014), before and after PTSD induction, to identify whether changes in hippocampal activity might relate to PTSD development. Following PTSD induction, we analysed spine density in the ventral DG of PTSD-like and resilient mice to investigate potential differences in the DG morphology following PTSD development. We hypothesised that PTSD-like animals would display reduced hippocampal activity, suppressed corticosterone stress responses and a reduced ventral DG dendritic spine density following PTSD induction compared to their resilient counterparts. The aim of the study was to elucidate whether these factors indicate possible vulnerability to PTSD or rather pathology as a consequence of PTSD.

Methods

Animal background and maintenance

92 C57BL/6J mice were housed in groups of four animals per cage with ad libitum access to food and water in a pathogen-free, temperature-controlled (22 °C ± 1) room with a humidity of 50 - 52%, and on a reversed 12/12 h light/dark cycle (lights on at 20:00 h). The animals were housed in the Central Animal Laboratory animal facility (Centraal Dierenlaboratorium [CDL], Nijmegen, the Netherlands), where the experiments took place. The study consisted of two separate experiments, experiment 1 (n = 32) to assess ventral DG spine density, and experiment 2 (n = 60) for the monitoring of hippocampal activity using

MRI scanning. Different group sizes were chosen based on power analyses on the distinct read-out measures of the experiments (morphology vs. blood flow), characterised by differential sensitivity and variability. As the proposed PTSD model was only validated in male mice, and stress responses in females differ from those in males and are related to the phase of the oestrous cycle (Ter Horst, De Kloet, Schachinger, & Oitzl, 2012), we only include male mice. At the start of the experiment, the animals were 10 weeks old. The experimental protocol was approved by the Animal Experiment Committee Radboud University (Dierexperimenten commissie Radboud Universiteit Nijmegen; DEC) in Nijmegen, and it was in accordance with the guidelines for the Care and Use of Mammals in neuroscience and Behavioural Research (National Research Council 2003), the principles of laboratory animal care, as well as the Dutch law concerning animal welfare. All experiments were carried out in accordance with the guidelines of the European Communities Council Directive.

Experimental design

See Figure 1 below.

PTSD induction protocol

PTSD induction. The PTSD induction protocol used in the experiment is based on the previously validated mouse model for PTSD, established by Lebow et al. (2012), to induce a PTSD-like phenotype in a subset of susceptible mice. In this model, on the first day of PTSD induction (day 15/16 in the experiment), mice receive 14 shocks of 1 mA, 1 s in duration with a continuous pulse over 85 minutes in variable intervals, representing the "trauma". For trauma induction, mice were placed in a fear conditioning apparatus (TSE Systems, Bad Homburg, Germany), using a certain context A. This context consisted of a transparent, Plexiglas cage (21 × 20 × 36 cm) with a metal grid floor and illumination of 10 lux with 70 dB background noise generated by the ventilation system. Inside the cage, a white paper was placed under the metal grid and a semi-circular white opaque screen was placed in the back of the cage to create a white, rounded environment. The mice were individually transferred to the experimental room in fresh cages in the light and the lights in the experimental room were kept on. Between experiments, the cage and grid were cleaned with 70% ethanol solution and sprayed with 1% acetic acid solution to create a scent. On the second day (day 16/17 in the experiment), the same mice received another 5 shocks of 0.7 mA, 1 s in duration over 5 minutes in fixed intervals, representing the "trigger", in a different context B. Context B existed of a black, solid cage (21 × 20 × 36 cm) with a triangular black screen placed over the cage to create a dark environment. There was no illumination inside the box and the experimental room except for infrared light. The mice were transferred to the experimental room in small dark carton cages in pairs. Between sessions, the cage and the metal grid were cleaned with 70% ethanol solution and additionally sprayed with 70% ethanol solution to create a scent.

Freezing behaviour. Video recordings made during trauma and trigger exposure were analysed for freezing behaviour of the animals by The Observer 5.0 (Noldus, Wageningen, The Netherlands). Freezing behaviour was scored by an observer, blinded to the experimental groups. Freezing was scored as the complete immobility of the animal for more than 1 second. To screen for potential

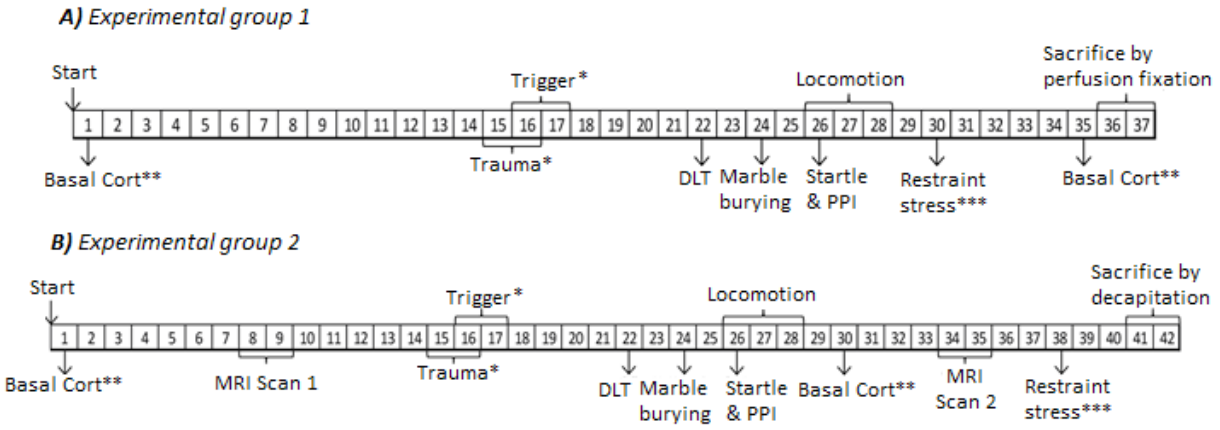


Fig. 1. Experimental design experimental group 1 and 2. DLT = dark/light transfer test, PPI = pre-pulse inhibition, Startle = startle response, CORT = corticosterone. *10 µl blood samples taken of the animals.

differences in both the initial response to the trauma, as well as its ultimate behavioural effect, the duration of freezing during trauma exposure was measured during the first 5 minutes (at t = 3 min, the first shock occurred and the second shock at t = 5 min) and the last 5 minutes of trauma exposure (at t = 82 min, the fourteenth [last] shock occurred, after which a 4 minute interval without shocks was presented until the end of the trauma session).

Behavioural assessment. All behavioural testing was performed during the dark phase, which is the active phase of the animals. The testing was performed at least four hours after the lights were turned off (from 12:00 h onwards) to ensure stable and low corticosterone plasma levels during testing. Following the PTSD induction protocol, the mice were subjected to five behavioural tests: the dark/light transfer test to assess risk taking behaviour (on day 22), the marble burying test to assess hypervigilance (day 24), acoustic startle response and pre-pulse inhibition to assess hyperarousal (day 26), and total light activity as an index for insomnia (days 26 - 28). The experimental setups were cleaned with 70% ethanol between every testing session. All animals from the same cage were tested consecutively and only after all animals from the same cage were tested, the mice were reunited in their home cage and returned to their housing room.

Risk assessment. In order to measure risk assessment, the dark/light transfer test was conducted. The test took place in a box with two compartments, with a partition in between. The dark compartment (15 × 25 × 25 cm) was covered with a dark plate and the light compartment (30 × 25 × 25 cm) was brightly illuminated by 1000 – 1100 lux. A small passage (3 × 6 cm) at the bottom of the partition connected the two compartments with a retractable door that opened the passage to initiate the test. At the start of the 5 minute test session, the mouse was placed in the dark compartment, and recordings of the test were made by a small camera installed above the apparatus. Data were analysed by Ethovision 10.1 xt (Noldus, Wageningen, The Netherlands). A risk assessment zone surrounding the opening of the passage in the light zone (3 × 6 cm) was programmed into the software tracking measurements.

The latency to enter the light compartment, the number of visits and the time spent in the risk assessment area were analysed. Percentage risk assessment time was calculated as the amount of time spent in the risk assessment arena as a percentage of total time spent in the light area outside of the risk assessment zone.

Marble burying. The mice were placed in a black box (30 × 27 × 26 cm), without a roof, with an illumination of 10 lux. The box contained 5 cm autoclaved bedding, and on top of the bedding 20 marbles were centrally arranged 4 by 5. The mice were filmed for 25 minutes by a camera mounted above the box. The videos were manually scored by counting the number of buried marbles at 25 minutes (Lebow et al., 2012).

Acoustic startle and pre-pulse inhibition. The acoustic startle response (ASR) (StartleResponse, TSE-systems, Bad Homburg, Germany) protocol was adapted from Lebow et al. (2012). A mouse was placed in a small Plexiglas and wire mesh cage in a sound-attenuated, ventilated chamber.

The cage was placed on top of a vibration-sensitive platform with a high-precision sensor detecting movements. Audio stimuli inside the chamber were produced by two high-frequency loudspeakers. 70 dB white background noise was maintained throughout the whole test. The ASR session started with a 5 minute acclimatisation period. After this acclimatisation session, randomly selected twenty-four startle stimuli of 120 dB, 40 ms in duration with a variable ITI of 12 – 30 ms were executed. Alternated with these startle stimuli, thirty-six additional startle stimuli with randomly preceding pre-pulses (20 ms) of either 75 dB, 80 dB, or 85 dB were applied. Maximal ASR, latency to peak startle amplitude, and the amount of pre-pulse inhibition (PPI) were individually analysed following the ASR protocol. The latency to peak startle was defined as the time from onset of the startle sound until the peak amplitude of the first startle. The time of latency to peak startle was averaged over all startle response trials. The percentage PPI was defined as the percent difference of the maximal ASR to the startle stimuli preceded by pre-pulses compared to the startle stimuli without pre-pulses, which was then averaged over all pre-pulse intensities.

Day-night locomotion. Mice were individually housed for 72 hours in Noldus Phenotyper 4500 (Noldus, Wageningen, The Netherlands) consisting of Plexiglas cages (45 × 45 × 60 cm) with a camera mounted on top. The first 24 hours was considered as a habituation period to the individual housing condition, after which the animals were tracked for 48 hours, consisting of two light and two dark cycles. The locomotion measurements were collected in 10 minute intervals and

analysed for total light and total dark locomotion by the Noldus software Ethovision 10.1 xt.

PTSD categorisation. The aforementioned behavioural tests were attributed to a particular amount of points according to factor analysis established by Lebow et al. (2012). The tests were divided into three groups: 1) percentage risk assessment and latency to peak startle amplitude, 2) percentage PPI, and 3) total light activity and marble burying. This resulted in the following scoring of extreme behavioural outcomes: Top 20% of animals showing high marble burying and high total light activity were attributed 1 point, bottom 20% of animals showing extremely low risk assessment were attributed 3 points, bottom 20% showing extreme latencies to peak startle were attributed 3 points, bottom 20% of abnormally low PPI were attributed 2 points (see Table 1). The scores per animal were tallied. This resulted in a behavioural phenotype that could be used to classify the animals in resilient or PTSD-like. Animals were considered resilient when they obtained a total score of 0 (and thus displayed no extreme behaviour in any of the tests), and as PTSD-like when the total score was 4 or more (reflecting extreme behaviour on multiple tests).

Table 1.

Inclusion criteria PTSD-like and resilient phenotype.

Behavioural measure		**Score**
Marble buying	Top 20%	1
Total locomotion light phase		1
Latency to peak startle		3
% Prepulse inhibition	Bottom 20%	2
% Risk assessment		3
	PTSD-like	≥ 4
	Resilient	<1

Neuroendocrine assessment

Blood sample collection during the experiment. To assess potential differences in corticosterone levels between PTSD-like and resilient mice, 10 µl blood samples were collected on various moment during the experiment. Blood sampling through tail bleed is in itself only mildly aversive to the animal, as it is done rather quickly (typically < 30 s) and does not entail any further restraint. Therefore, it can be conducted multiple times without negative consequences. Before PTSD induction, baseline peak and trough diurnal levels were assessed by two blood samples taken at 7:30 h (30 minutes before the start of the active phase to capture the circadian peak) and at 20:30 h (30 minutes after the end of the active phase to assess the circadian trough). Blood samples were also collected immediately after trauma and 10 min post trigger exposure. Three more samples were taken to assess the corticosterone stress response to restraint after PTSD induction (day 30 for group 1, day 38 for group 2). During restraint, animals were put in plastic restrainers for 25 minutes in a brightly lit room. Blood samples were collected by tail bleed before (t = 0 min, baseline), immediately after (t = 25 min, stress-induced peak levels) and 90 minutes after stress initiation (following recovery). Finally, towards the end of the experiment, another two basal levels were taken to assess the circadian peak and trough levels.

Corticosterone collection and measurement. 10 µl blood samples of were collected by tail bleed. After collection, blood samples were briefly stored on ice until they were centrifuged (3500 rpm for 20 min at 4 °C), and plasma was extracted. The plasma samples were stored at -20 °C and later assayed for corticosterone using the Corticosterone Double Antibody RIA Kit (MP Biomedicals, Orangeburg, NY, USA).

MRI procedure. All animals from the second experiment group, n = 60, underwent two MRI sessions, before and after PTSD induction protocol. The mice were anaesthetised with 3.5% isoflurane and transferred to the MRI platform, where anaesthesia levels of isoflurane were reduced to 2%, administered by inhalation through a nosetube. The mice subsequently received a bolus of medetomidine (Dexdomitor, Pfizer, 0.05 mg/kg) subcutaneously (Grandjean, Schroeter, Batata, & Rudin, 2014). After five minutes, the isoflurane was further reduced to 1%, and another five minutes later it was lowered to 0.5%, and infusion of medetomidine (0.1 mg/kg/h) started (Grandjean et al., 2014), which was maintained throughout the scanning session to maintain the superficial sedation level.

The mice were placed in a MR-compatible stereotactic device and immobilised with earplugs and a tooth holder. Body temperature was measured using a rectal thermometer and maintained at 37 °C using a heated air flow device. The respiration rate

was registered using a breathing pad. To protect the eyes from dehydration, eye ointment was used.

The animals underwent approximately 2 hours of MRI scanning. Two ASL scans to assess the cerebral blood flow (CBF) in the hippocampus were acquired at 45 to 60 minutes after the medetomidine bolus. After completion of scanning, the animals were removed from the apparatus, halting the administration of isoflurane and the medetomidine infusion, and a bolus (0.25 mg/kg) of antisedan (Atipamezole, Pfizer) was administered subcutaneously to antagonise the medetomidine and ensure quick recovery (Adamczak, Farr, Seehafer, Kalthoff, & Hoehn, 2010).

MRI acquisition. MRI measurements were performed on an 11.7 T BioSpec Avance III small animal MR system (Bruker BioSpin) equipped with an actively shielded gradient set of 600 mT/m and operated by Paravision 6.0 software. A circular polarised volume resonator was used for signal transmission and an actively decoupled mouse brain quadrature surface coil was used for signal reception (Bruker BioSpin).

To study brain perfusion under resting conditions, a flow-sensitive alternating inversion recovery arterial spin labelling (FAIR ASL) technique was used (Zerbi et al., 2014; Field of View = 25 × 25 mm; matrix size = 512 × 512; slice thickness = 1 mm; repetition time [TR]/echo time [TE] = 12 s/10.077 ms; spatial resolution = 0.195 × 0.260 × 1 mm/pixel). Sixteen images with increasing inversion times (TIs; 100 – 1600 ms) were obtained for the T1 calculations, adding up to a total scan time of 12 minutes 48 seconds. Inversion recovery data from the imaging slice were acquired after sequential selective inversion and nonselective inversion. Slices were positioned based on Bregma coordinates; centered around Bregma = -1.94 mm for the hippocampus (Allen Mouse Brain Atlas).

MRI analyses. For each mouse, the FAIR images with different TIs were realigned over the first TI. For this, a rigid-body model was used, implemented in Statistical Parametric Mapping (SPM). Determination of $T1_{selective}$ and $T1_{nonselective}$ was performed by fitting the averaged signal intensities in each region of interest (ROI) with a three-parameter monoexponential T1 relaxation curve. Using the following equation, CBF was determined in hippocampus:

$$\frac{CBF}{\lambda} = \frac{T1_{nonselective}}{T1_{blood}}\left(\frac{1}{T1_{selective}} - \frac{1}{T1_{nonselective}}\right)$$

In this equation, λ is the blood/tissue partition coefficient for water (assumed to be 0.9 ml/g) and T1blood was assumed to be 2.75 s at 11.7 T (Zerbi et al., 2014).

The regional cerebral blood flow was determined by drawing regions of interest in the retrieved ASL MRI scans in MatLab. The MRI procedure was performed by two independent researchers, who were blinded to the experimental conditions.

Brain tissue collection

Mice from experimental group 1 (n = 32) were sacrificed 20 days after trauma exposure by perfusion with 4% paraformaldehyde (PFA) in 0.1 M PBS. The brains were removed immediately after death and post-fixed for 24 hours by PFA immersion. After post-fixation, the brains were stored at 4°C. At a later stage, Golgi staining was performed on the brains.

Mice from experimental group 2 (n = 60) were sacrificed 25 days after trauma exposure by decapitation. Immediately after death, brains were removed, quick-frozen on dry ice and stored at -80 °C.

Golgi staining

Tissue preparation. The brains of animals from experimental group 1 that had undergone the PTSD induction protocol and were scored with either PTSD-like (n = 5) or resilient (n = 6) phenotype, were processed for rapid Golgi-Cox staining (the brains of mice with an intermediate phenotype were not included). The Golgi-Cox-based kit (FD Rapid GolgiStain™ FDNeurotechnologies, Inc. Ellicott City, MD, USA) was used for Golgi-Cox tissue impregnation. Brain tissues were immersed in a solution consisting of a mixture of mercury chloride, potassium dichromate and potassium chromate, and kept in the dark for 14 days. After impregnation, the tissues were transferred into a cryoprotectant solution for 72 hours to equilibrate and sink. After this, the tissues were cut into 140 µm coronal sections using a freezing-stage sledge microtome (Microtom HM440E, GMI Inc., Ramsey, MN, USA), and mounted on gelatine-coated slides. The sections were left to air-dry for 3 days before the sections were washed twice for 4 minutes in dH_2O, and incubated in an ammonium hydroxide solution for ten minutes. Then, the sections were washed twice in dH_2O for 4 minutes, placed for 4 minutes in ascending grades of alcohol (70%, 95% and four times 100%) and placed three times for 4 minutes each in

xylene for clearance. As a last step, the sections were coverslipped with Permount™ mounting medium (Fisher Chemicals, Leicestershire, UK) and dried for 1 week until the sections were processed for analysis.

Spine density. Prior to the quantitative analysis of neuronal spine density, a set of criteria was determined in order to obtain accurate measurement of neuronal morphology. Neurons were included into analyses if they fulfilled the following selection criteria: neurons should be 1) granule cells obtained from the granule cell layer of the ventral dentate gyrus (-3.28 to -3.80 mm from Bregma) in both the right and left hemisphere 2) completely impregnated, and 3) granule cells containing a clearly stained cell body and primary dendritic branch that was distinguishable from neighbouring neurons. Spine counting was performed on tertiary (3), quaternary (4), quinary (5), senary (6) and septenary (7) order dendrites.

Spine density measures were obtained using a Zeiss Axioskop FS microscope (Oberkochen, Germany). Dendritic spines were captured using a 100x oil immersion objective lens (plan-NEOFLUAR, Zeiss). The dendritic spines were traced using Neurolucida 11.09 (MBF Bioscience, Microbrightfield Inc. Williston, VT, USA) by an investigator that was blind to the treatment conditions. Firstly, the whole apical dendritic tree was traced. Secondly, the spines were counted per selected segment of at least 50 µm in length. Spine density was calculated as the number of spines per 10 µm of dendrite. On average, six segments per cell and five cells per animal were used for analysis. It should be noted that the spine density values obtained are likely to be an underestimation of the actual density of the dendritic spines because spines protruding beneath or above the dendritic segment were not visible and thus not accounted for.

Statistical analysis

Statistical analyses were computed with IBM SPSS statistics 23 (IBM corporation, Armonk, NY, USA). Group differences between PTSD-like and resilient animals in terms of behavioural measures, corticosterone levels and spine density were tested for using independent-samples t-tests. Hippocampal CBF was analysed using a repeated measures ANOVA, followed up by paired and independent t-tests. Spearman's rho correlations were performed for correlational analyses. Results were considered significant if $p < .05$.

Results

Behavioural assessment of PTSD-like phenotype in mice

Experimental group 1. As expected, PTSD-like animals showed impaired risk assessment during the dark/light transfer test ($t(9) = -2.78$, $p < .05$; Fig. 2A), and attenuated pre-pulse inhibition ($t(9) = -2.23$, $p = .05$; Fig. 2C). However, differences in marble burying behaviour ($t(9) = 0.92$, $p = .38$; Fig. 2E), latency to peak startle amplitude ($t(9) = -0.27$, $p = .79$; Fig. 2B), and activity during the light phase, ($t(9) = 1.47$, $p = .17$; Fig. 2D) failed to reach significance.

Experimental group 2. As expected, PTSD-like animals showed impaired risk assessment during the dark/light transfer test ($t(2) = -2.17$, $p < .05$; Fig. 3A), had a reduced latency to peak startle response ($t(22) = -3.45$, $p < .01$; Fig. 3B), and buried significantly more marbles ($t(22) = -2.83$, $p < .01$; Fig. 3E). However, PTSD-like mice did not significantly differ for the resilient ones in terms of pre pulse inhibition ($t(22) = -0.95$, $p = .34$; Fig. 3C), and their increased activity during the light phase just failed to reach significance ($t(22) = 1.78$, $p = .09$; Fig. 3D).

Behavioural response during the PTSD induction protocol: Trigger exposure

Experimental group 1. Freezing behaviour during trigger exposure was measured to assess fear generalisation. There was no significant difference in latency to freeze ($t(5) = 1.69$, $p = .15$; Fig. 4A) between PTSD-like and resilient mice. Also the duration of freezing was not significantly different between the two groups ($F(1,6) = 0.52$, $p = .50$; Fig. 4B). Both innate fear generalisation, prior to shock exposure ($t(9) = -0.82$, $p = .44$) and triggered fear generalisation, after first shock ($F(1,6) = 0.52$, $p = .45$) were not significantly different between the two groups, although the freezing duration over time was significantly increased ($F(1,6) = 173.58$, $p < .001$; Fig. 4B).

Experimental group 2. PTSD-like animals had a significantly shorter latency to freeze than resilient animals ($t(20) = -3.31$, $p < .005$; Fig. 5A). The duration of freezing did not significantly differ between the groups ($F(1,21) = 0.2$, $p = .90$; Fig. 5B). Both innate fear generalisation, prior to shock exposure ($t(22) = 1.64$, $p = .12$) and triggered fear generalisation, after first shock ($F(1,22) = 0.21$, $p = .65$) were not

significantly different between the two groups. However, there was a significant increased freezing effect over time ($F(1,21) = 350.03$, $p < .001$; Fig. 5B).

HPA-alterations in PTSD-like versus resilient mice

Experimental group 1. PTSD-like mice displayed a significantly increased pre-exposure circadian peak corticosterone level ($t(6) = 5.08$, $p < .005$; Fig. 6A), although circadian trough level was not significantly different between the groups ($t(9) = -1.49$, $p = .17$; Fig. 6A). The post-exposure circadian peak ($t(9) = -0.50$, $p = .63$; Fig. 6B) and trough ($t(9) = -0.53$, $p = .61$; Fig. 6B) levels were not significantly different between the groups. As expected, the PTSD-like mice had a significantly suppressed corticosterone level during restraint stress at $t = 25$ min ($t(8) = -3.14$, $p < .05$; Fig. 6D) and trend-level significant increase during restraint stress baseline measures ($t(8) = -2.13$, $p = .06$; Fig. 6B), which replicates the results of Lebow et al. (2012). Interestingly, corticosterone levels after trigger exposure ($t(9) = -2.91$, $p < .05$; Fig. 6C) are also suppressed in PTSD-like mice.

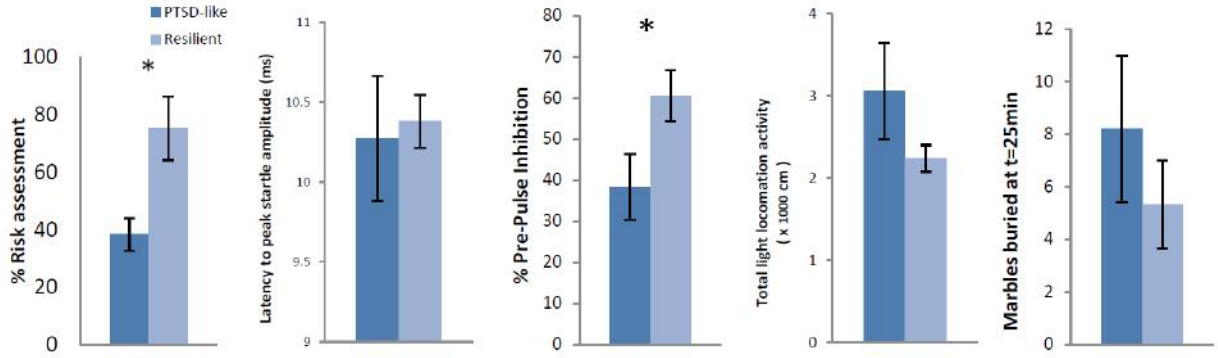


Fig. 2. Behavioural assessment. PTSD-like mice spent significantly less time engaged in risk assessment behaviour ($p = .02$) compared to resilient mice. No significant differences were detected ($p = .79$) for the latency to peak startle between the groups. Pre-pulse inhibition was significantly impaired ($p = .05$) in PTSD-like mice compared to resilient. PTSD-like mice tended to be more active in the light phase compared to resilient mice, but this difference failed to reach significance ($p = .18$). PTSD-like mice did not bury significantly more marbles ($p = .38$) then resilient mice. *$p < 0.05$. Error bars indicate standard errors of the mean (*SEM*).

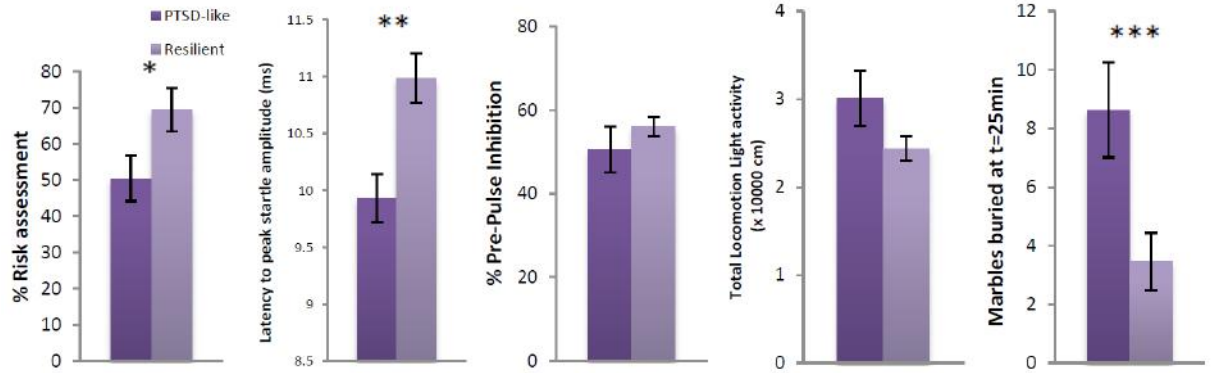


Fig. 3. Application of inclusion criteria for group 2. PTSD-like mice spent significantly less time engaged in risk assessment behavior ($p = .04$) compared to resilient mice. Significantly reduced latency to peak startle for PTSD-like animals ($p = .002$) was found. Pre-pulse inhibition was not significantly impaired ($p = .35$) in PTSD-like mice compared to resilient. PTSD-like mice were significantly not more active in the light phase ($p = .09$) compared to resilient mice. PTSD-like mice buried significantly more marbles ($p = .01$) compared to resilient mice. *$p < .05$, **$p < .01$. Error bars indicate *SEM*.

Experimental group 2. No significant differences between the groups were found during pre-exposure circadian peak ($t(21) = 1.42$, $p = .17$; Fig. 7A) and trough ($t(20) = 0.33$, $p = .75$; Fig. 7A) levels, and post-exposure circadian ($t(22) = -0.64$, $p = .53$; Fig. 7B) peak and trough ($t(21) = 0.26$, $p = .79$; Fig. 7B) levels. PTSD-like mice had a significantly increased baseline corticosterone level during restraint stress ($t(20) = 3.02$, $p < .01$; Fig. 7D). However, resilient mice had a significantly increased corticosterone level during recovery of restraint stress ($t(21) = -1.93$, $p = .07$; Fig. 7D).

Ventral DG spine density in PTSD-like versus resilient mice

Spine density analysis on the ventral DG revealed a significant reduction in spine density in PTSD-like animals ($t(9) = -2.22$, $p < .05$; Fig. 8C) compared to resilient animals (Fig. 8).

Hippocampal blood flow before and after PTSD induction protocol in PTSD-like versus resilient mice

Prior to PTSD induction, no significant differences ($t(20) = 0.96$, $p = .89$; Fig. 9C) were detected between PTSD-like and resilient mice. However, after PTSD induction, the hippocampal blood flow was significantly reduced ($t(18) = -2.09$, $p < .05$; Fig. 9C) in PTSD-like mice compared to resilient mice. A trend-level significant time x group interaction in hippocampal CBF levels ($F(1,17) = 3.30$, $p = .09$) was observed, without an overall effect of time ($F(1,17) = 2.29$, $p = .15$). Post-hoc analysis revealed that this group x time interaction was caused by a relative decrease in hippocampal blood flow in the PTSD-like group only ($t(6) = 2.28$, $p = .06$; Fig. 9C).

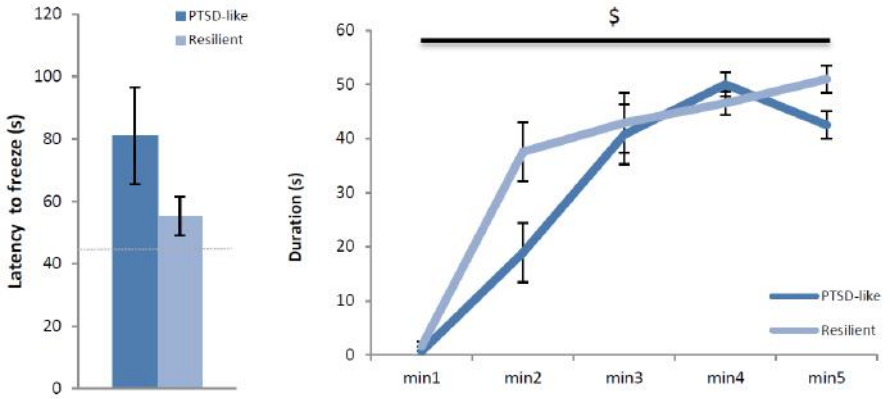


Fig. 4. Freezing behaviour during trigger exposure in mice from experiment 1. No significant difference was found in latency to freeze. Dashed line indicates the timing of the first shock (after 60 seconds). No significant difference was detected in the freezing duration between PTSD-like and resilient mice. However, there was a significant freezing effect over time ($p < .001$). Error bars indicate *SEM*.

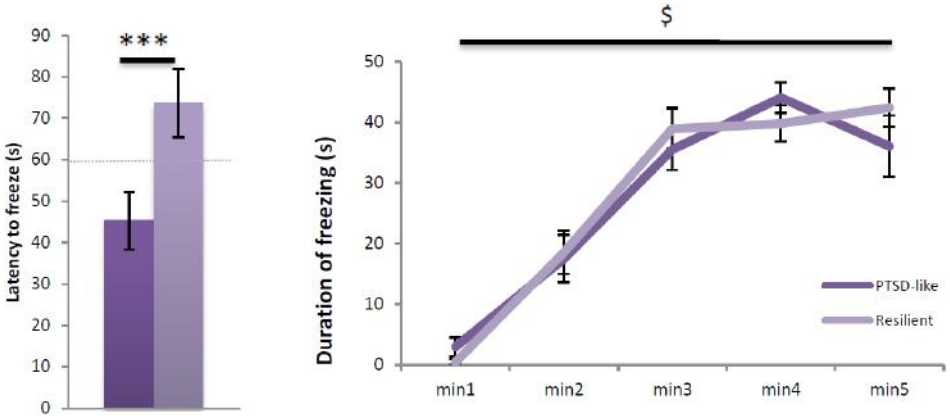


Fig. 5. Freezing behaviour during trigger exposure in mice from experiment 2. Resilient mice had an increased latency to freeze ($p = .003$) compared to PTSD-like mice. There are no differences in freezing duration ($p = .90$), although there is a significant increase in freezing over time ($p < .001$). ***$p < .005$. Error bars indicate *SEM*.

Correlational analyses

To gain additional understanding about the relationship between corticosterone levels, freezing behaviour, spine density and hippocampal blood flow, correlations were computed between measures.

A positive correlation was found between PTSD-like mice from group 1 for the spine density and post-PTSD development basal morning corticosterone levels ($r(5) = 0.90$, $p < .05$), and between trigger freezing latency and post-PTSD development basal morning corticosterone levels ($r(5) = 0.90$, $p < .05$).

Discussion

Symptoms of PTSD, like flashbacks and recurrence of memories, are thought to be caused by abnormal hippocampal functioning, and more specifically by exaggerated fear generalisation. Here, we investigated the contribution of hippocampal function and fear generalisation to PTSD. We particularly aimed to address the existence of potential differences in hippocampal function reflecting vulnerability for PTSD predicting its development. We used a validated mouse model for PTSD induction by a trauma exposure on day 1 and a trigger exposure on day 2, to induce a PTSD-like phenotype in part of the mice, whereas others are resilient and do not display any PTSD-related symptoms. No difference in fear generalisation (assessed by freezing behaviour to a novel context) between PTSD-like and resilient mice was found prior to PTSD development. However, PTSD-like mice did display suppressed corticosterone responses to stress prior to PTSD development. In line with previous research reporting on abnormal hippocampal dysfunction in PTSD patients, we found that hippocampal blood flow, representing hippocampal activity, was significantly reduced in mice who had acquired a PTSD-like phenotype, which was not observed in resilient animals. Moreover, spine density in the ventral DG – critical for the process of pattern separation, preventing generalisation of memories – was significantly decreased in mice with a PTSD-like phenotype compared to resilient mice. Though we observed several indications for hippocampal dysfunction as pathology once a PTSD phenotype had been developed, none of the findings indicate hippocampal dysfunction as vulnerability factor before PTSD development (i.e., it is not a predisposition).

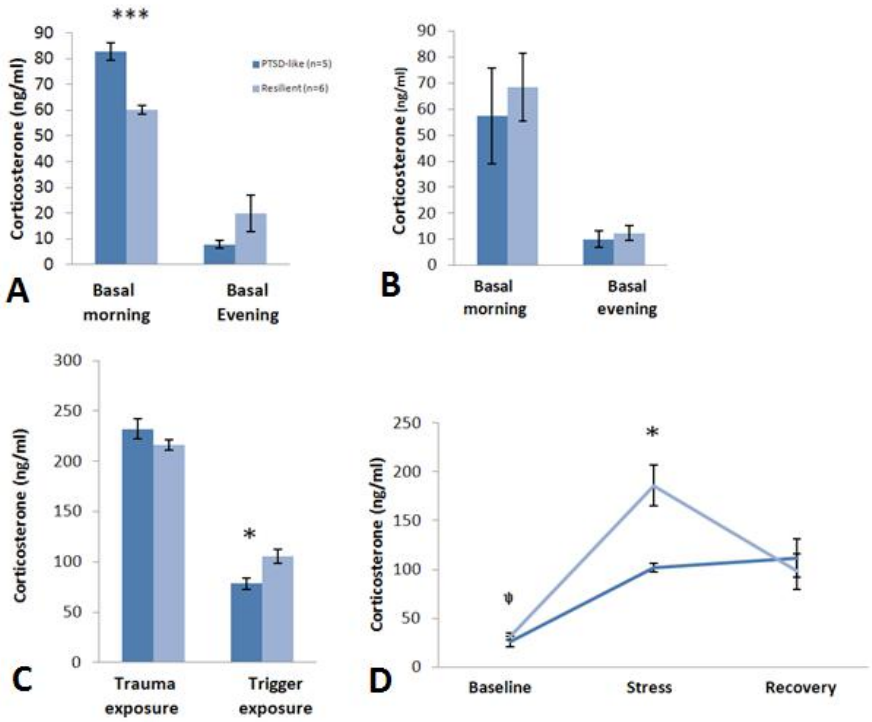


Fig. 6. Corticosterone levels in mice from experiment 1. Top graphs: Basal measurements before PTSD induction (**A**), and after PTSD induction (**B**). Bottom graphs: Stress-induced CORT levels (**C** and **D**). PTSD-like mice had an increased corticosterone level during pre-exposure morning ($p = .002$) measurement (D). Resilient mice had a significantly increased corticosterone level during trigger exposure ($p = .02$; C), during restraint stress ($p = .01$; D) and a trend-level significant increase during baseline measure of restraint stress ($p = .06$; D). *$p < .05$, ψ$p = .06$. Error bars indicate standard errors of the mean (*SEM*).

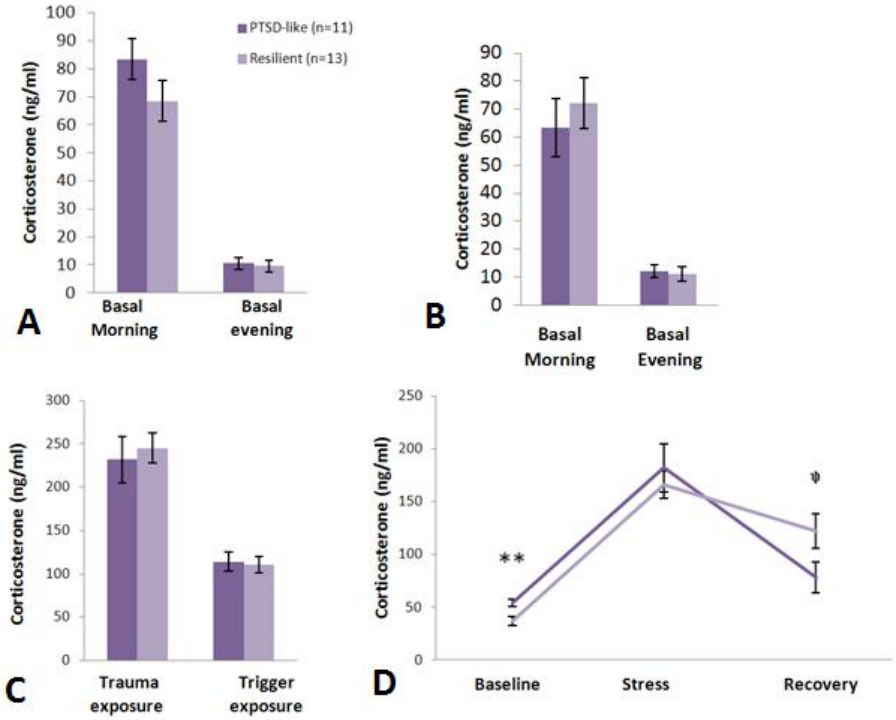


Fig. 7. Corticosterone levels from experimental group 2. Top graphs: Basal measurements before PTSD induction (**A**), and after PTSD induction (**B**). Bottom graphs: Stress-induced CORT levels (**C** and **D**). PTSD-like mice had a significantly increased baseline restraint stress corticosterone level (p = .006; D). However, resilient mice had a trend-level significant increase in corticosterone during restraint stress recovery (p = .06; D). **p < .01, ψp = .06. Error bars indicate *SEM*.

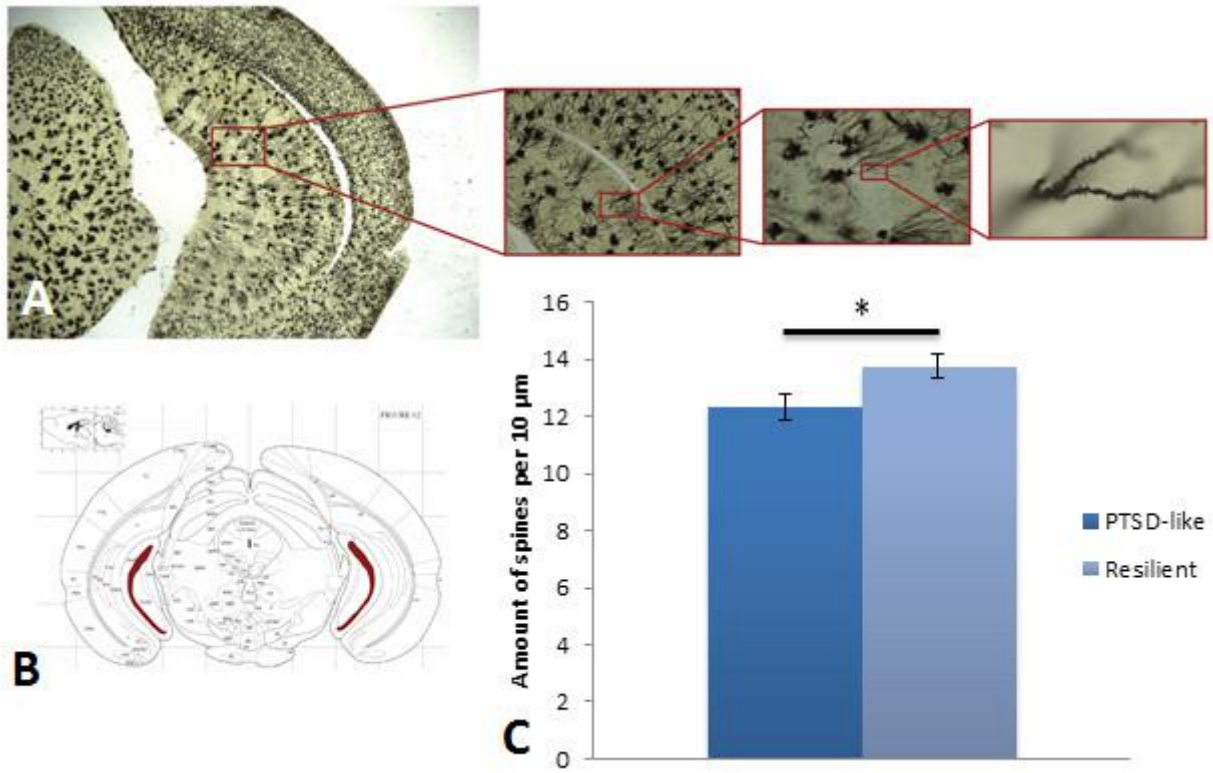


Fig. 8A. Golgi-Cox staining in brain tissue (from left to right: 5×, 20×, 40×, and 100× magnification). **B.** Mouse Brain Atlas drawing of ventral DG. Bregma coordinates = -3.28:-3.80 mm. **C.** Spine density per 10 µm in the ventral DG. Spine density was significantly reduced (p = .05) in PTSD-like mice compared to resilient mice. *p = .05. Error bars indicate *SEM*.

During the PTSD induction protocol, we exposed the animals to a trauma on the first day and a trigger on the next day. To measure fear generalisation, we measured the freezing response to the trigger context, which was an unfamiliar context to the animals, but contained a feature from the trauma context (i.e., the grid). We found that all mice of the experimental groups froze faster (during the first minute of trigger exposure, before the first shock) compared to a control group who underwent the same protocol but did not receive shocks during trauma exposure (M = 114.15, SD = 100.44; Henckens et al., unpublished data). This might indicate that the mice are in general more fearful in new environments (novelty-induced anxiety), or that the mice associate the trigger context with the trauma exposure context of the previous day. However, in this study, no reliable differences could be found between the PTSD-like and resilient animals. Therefore, we conclude that there is no indication of fear generalisation during PTSD induction in either PTSD-like or resilient animals. This is in line with research from Golub et al. (2005), in which they did not find an increase in freezing in a novel context containing a feature from the conditioning context (grid). They explain that cues most proximal to the trauma are often not directly associated with it (Ehler, Hackmann, & Michael, 2004; Golub et al., 2005).

Furthermore, in the first experimental group, we found a suppressed stress induced corticosterone level in PTSD-like animals, to restraint stress once the pathology had established. The findings of suppressed stress-induced corticosterone levels are in line with findings from previous animal work (Lebow et al., 2012), and patient studies (Yehuda, Giller, Southwick, Lowy, & Mason, 1991; Yehuda et al., 1996) on PTSD, where it is suggested that an enhanced negative feedback of the HPA-axis is causing these decreased cortisol (corticosterone) levels (Yehuda et al., 1991). Such enhanced negative feedback is also in line with the findings of lower corticosterone levels in PTSD-like animals during the recovery phase of restraint stress in group 2. Interestingly, we also found initial evidence for a suppressed corticosterone response to trigger exposure in PTSD-like mice compared to their resilient counterparts prior to PTSD development. Findings that PTSD-like mice display suppressed corticosterone responses already during PTSD induction seem to fit the findings of Zohar et al.

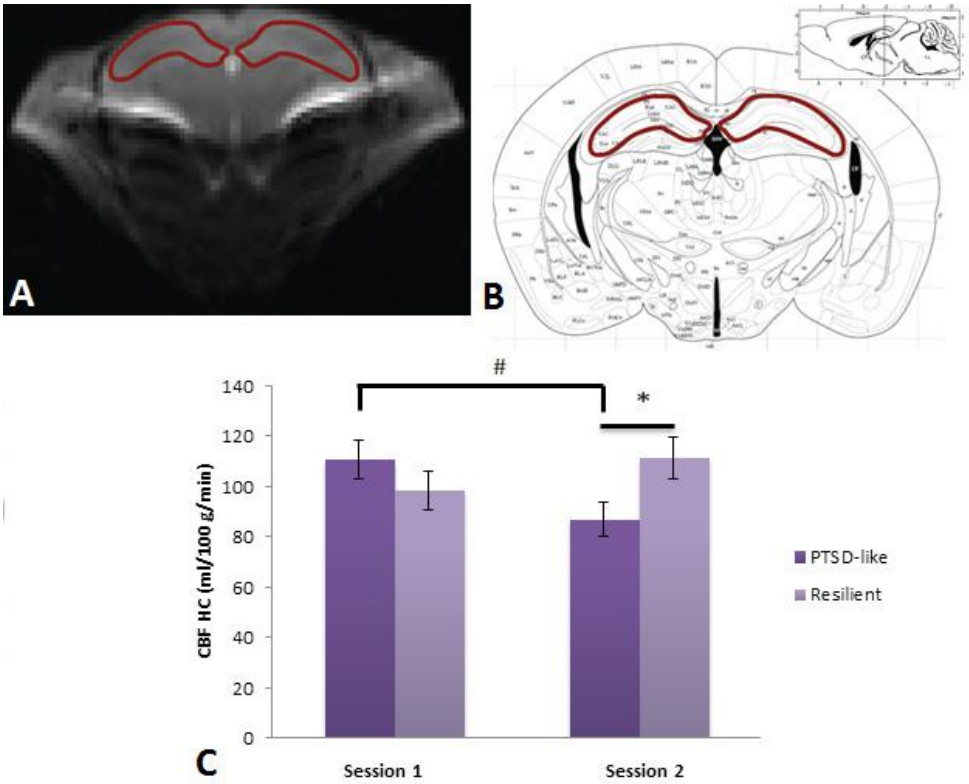


Fig. 9A. ASL MRI scan of the hippocampus, Bregma coordinate = -1.94 mm. **B.** Mouse Brain Atlas showing drawn region of interest (hippocampus), Bregma coordinate = -1.94 mm. **C.** Cerebral blood flow in the hippocampus in PTSD-like and resilient mice prior to and following PTSD induction. Blood flow was significantly reduced in PTSD-like mice after PTSD induction compared to resilient mice (p = .05). Furthermore, there was a trend-level significant time × group interaction in hippocampal CBF levels due to the decrease in CBF in PTSD-like only (p = .06). *p = .05 (t-test), #p = .06 (repeated-measures ANOVA). Error bars indicate *SEM*.

(2011), which suggested that the administration of a single high dose of corticosteroids immediately after trauma exposure can actually reduce symptoms of acute stress and post-traumatic stress. It was found that high-dose corticosterone in polytrauma patients administered immediately after trauma, reduced symptoms of PTSD, and in the animal study, high-dose corticosterone stimulated dendritic outgrowth and spine density in the DG (Zohar et al., 2011). Thus, immediately after trauma, corticosterone can counteract the symptomatological and morphological effects of PTSD. These findings were explained by a proposed role for high corticosterone levels immediately after trauma in recalibrating the HPA-axis in order to ensure appropriate regulation of the HPA-axis following trauma (Zohar et al., 2011). Also, corticosterone has been reported to facilitate the acquisition of contextual memory (Cordero, Venero, Kruyt, & Sandi, 2003). So, this indicates that corticosterone improves context-dependent memory and counteracts fear generalisation (Cordero et al., 2003; Pugh, Tremblay, Fleshner, & Rudy, 1997). This is also shown by Albrecht, Caliskan, Oitzl, Keinemann, & Stork (2013), who found that high levels of corticosterone infused in the ventral hippocampus following fear exposure have an anxiolytic effect following fear exposure (Albrecht et al., 2013). This might be in line with our result of increased corticosterone levels in resilient mice during trigger exposure. However, another seminal paper by Kaouane et al. (2012) reported that infusion of corticosterone in the hippocampus after fear conditioning actually induces PTSD-like memory impairments, which was reflected in a flawed memory for the correct predictor of the threat in mice (which thus generalised their fear).

Unfortunately, findings on suppressed stress-induced corticosterone levels were not replicated in the second experimental group, which could be the consequence of the differential experimental time line and repeated stress exposure in this group. Although maximal effort was put into reducing the stress induced by the anaesthesia required for MRI scanning, animals most likely still experienced moderate stress. Moreover, isoflurane, the anaesthesia required for MRI scanning, has been shown to induce a significant corticosterone response (Altholtz, Fowler, Badura, & Kovacs, 2006). Also, the restraint and basal corticosterone measures in this group were obtained at slightly different time points during the protocol. In sum, we observed a suppressed stress-induced corticosterone level in PTSD-like mice which could be the consequence and a predictor of an enhanced negative feedback of the HPA-axis in PTSD.

In addition to fear generalisation and neuroendocrine changes associated with PTSD development, we also assessed ventral DG neuronal morphology in PTSD-like versus resilient mice. The spine density analysis focused on the ventral DG, because it is associated with pattern separation and fear generalisation, and because it is involved in regulating affect, motivational, and emotional behaviour and the hormonal stress response (Fanselow & Dong, 2010; Kheirbek et al., 2013). We observed a significant decrease in ventral DG spine density in PTSD-like mice, suggesting a decreased excitatory input (Amaral et al., 2007) on the dendrites of the ventral DG granule cells in these animals compared to their resilient counterparts. These alterations in morphology might be a maladaptive response to trauma and can underlie the physiological and behavioural changes observed due to decreased excitatory input into the ventral DG. Previous research has produced some conflicting data on the effects of stress and trauma on spine density in the dorsal DG. On one hand, a rather acute increase in spines density in the dorsal DG has been found in response to acute stress (1 hour platform stress) after which the rats were immediately sacrificed (Sebastian, Estil, Chen, Schrott, & Serrano, 2013), while other studies have reported decreased dorsal DG spine density in response to stress on the longer time scale (8 days following single acute predator scent stress) in the DG (Cohen et al., 2014; Zohar et al., 2011). The current study is the first to describe the long-term effects of PTSD-like symptoms in the ventral DG.

The observed reduction in spine density could have several causes. Firstly, spine loss can be the result of a widespread loss of neurons in the amygdala and hippocampus (Fanselow & Dong, 2010; Kheirbek et al., 2012) and their axonal inputs on spines. However, compensatory increases in spine density in response to the loss of synaptic input have also been found (Fiala, Spacek, & Harris, 2006). Secondly, it might be caused by a local loss of neurogenesis or axonal sprouting (Fiala et al., 2006), known to be affected by stress (Altman & Das, 1967; McEwan, 1999; Redila & Christie, 2006). The exact reasons for spine density reduction in this study thus remain unclear. Further research must be conducted to investigate the origin of this loss of spines in the ventral DG in animals with a PTSD-like phenotype. Also, as these invasive measures of neuronal morphology could only be obtained after sacrificing the animals, it remains unknown whether the examined reductions in spine density in PTSD-

like mice are a cause or a consequence of PTSD development.

In PTSD patients, a relationship has been reported between reduced hippocampal activity and severity of PTSD symptoms (Astur et al., 2006). However, literature is conflicting on whether the reduced hippocampal activity is a predisposition or pathology (Hayes et al., 2011). Prospective studies where large groups of subjects (for example combat soldiers) are followed over time (Admon, Milad, & Hendler, 2013) provided initial evidence that reduced hippocampal activity in PTSD is a pathology rather than predisposition (Hayes et al., 2011). Here, using a mouse-model for PTSD, we show that differences in hippocampal blood flow as a measure of hippocampal activity (Zerbi et al., 2014), are not apparent before trauma induction, but become significant only after the PTSD induction protocol. This might indicate that neuronal alterations taking place following trauma and trigger exposure are responsible for this endophenotype. Potentially, this reduction of hippocampal blood flow following PTSD development might be caused by the physiological effects of the chronic stress experienced due to PTSD symptomatology, since chronic stress has been related to increased susceptibility to excitatory amino acids such as glutamate and the inhibition of growth factors like brain-derived neurotrophic factor (BDNF; Carrion, Haas, Garrett, Song, & Reiss, 2010).

Some limitations to this study should be mentioned as well. Findings on corticosterone levels in experiment 1 and experiment 2 were not replicated between the groups. As mentioned before, this might be the result of differences in the experimental design of the two groups. The restraint corticosterone and post-exposure basal corticosterone blood sample collections were obtained on different time points in both groups (they were reversed). For experiment 1, we strictly adhered to the previously validated timeline of PTSD induction, testing, and time of sacrifice (Lebow et al., 2012; Henckens et al., unpublished data). For experiment 2 we decided to deviate from this protocol, as we considered the second MRI session the main measurement of interest, which is why it should be performed immediately after the completion of the assessment of the behavioural phenotype (i.e., around the time of sacrifice in experiment 1). We did not want to delay this measurement, nor confound it by the additional exposure to restraint stress. The corticosterone stress response to restraint was considered of lesser importance to this experiment, which is why it was delayed. Also, additional stress caused by the anaesthesia used during the MRI scanning of group 2 may be responsible for the apparent inconsistencies. Further research is needed to replicate the association between stress-induced corticosterone levels and PTSD symptomatology, and to examine whether the differences found in the corticosterone results are in fact time-dependent (i.e., dependent on the specific moment of sampling). Another limitation is that we only investigated the density of spines (i.e., total amount of spines), but we did not investigate differences in the spine structure itself.

Changes in spine structure and morphology, for example spine size or shape alterations, might indicate reductions in functional integrity and synaptic activity (Fiala et al., 2006). Future research should examine whether there are apparent differences in spine structure and morphology after PTSD development. Furthermore, in this study, we only focused on spine density in the ventral DG, because the ventral DG is associated in pattern separation and regulation of affect, motivation and emotional behaviour. However, others also found changes in spine density in other areas of the hippocampus in the rodent brain in response to stress. An increase in spine density has been found in the hippocampal CA1 (Adamec et al., 2012; Diamond et al., 2006) and CA3 (Dias et al., 2014; Redila & Christie, 2006) after acute stress and in the amygdala in response to chronic stress (Astur et al., 2006), also a reduction in spine density was found in the medial prefrontal cortex in response to chronic stress (Hayes et al., 2011). Therefore, future research should assess whether the reduction in spine density found in the ventral DG in PTSD-like animals are also apparent in other hippocampal subregions (dorsal DG, CA3, etc.) or other PTSD-related brain regions. Lastly, we only assessed blood flow through the complete hippocampus, without separating the several subnuclei. However, the hippocampus consists of several subnuclei with different functions (Amaral & Lavenex, 2006), and others found reduced hippocampal activity in the posterior and, to a lesser extend in the anterior, hippocampus in response to traumatic pictures in PTSD patients (Hayes et al., 2011). Therefore, the specificity of the findings remains to be elucidated. Also, blood flow in subregions of the hippocampus, and especially the ventral and dorsal DG should be examined in future studies. As it has been found that in both humans and mice, sex differences exist in stress susceptibility and anxiety-like behaviour after trauma exposure, where females seem to be more susceptible to develop stress-related disorders

(Adamec, Head, Blundell, Burton, & Berton, 2006; Shansky, 2015).

Here, we chose to study the population (i.e., males) displaying the most robust and stable response to physical stressors as a proof of principle (Ter Horst et al., 2012). Future studies should assess whether similar effects in terms of functional and morphological changes in the hippocampus as a consequence of PTSD are present in the female brain as well. Lastly, we studied inter-individual differences in PTSD vulnerability in animals with a nearly identical genetic background. Although clear inter-individual differences were observed, previous reports have indicated that gene-environment interactions render certain individuals more vulnerable to the development of PTSD after trauma exposure than others (Mehta & Binder, 2012), a factor currently neglected in this study. Future assessment of such gene-trauma interactions is of critical importance to enhance our insight in the underlying mechanisms of aberrant hippocampal function.

Conclusion

Here, we investigated the association between hippocampal function and associated fear generalisation in PTSD, to find out when exactly in disease development aberrant hippocampal functioning emerges. We found no indication of differential fear generalisation during the PTSD induction protocol predicting the subsequent development of PTSD-like symptomatology. The only index predicting PTSD development was the reduced corticosterone response to trigger exposure. These data suggest that corticosterone secretion during trauma exposure might be a compensatory mechanism (which is operational in resilient animals) to protect against the harmful effects of trauma and trigger exposure. This protective effect of increased corticosterone after trauma might be by recalibration of the HPA-axis in order to ensure appropriate regulation of the HPA-axis and proper memory for the trauma context, and thus, counteracting fear generalisation. In line with literature, we observed reduced hippocampal blood flow in PTSD-like compared to resilient animals once pathology had been established. However, this difference only became apparent after PTSD development, indicating the reflection of pathology rather than a predisposition. Moreover, spine density in the ventral DG was reduced in PTSD-like mice. This might indicate reduced excitatory input due to PTSD development, which might cause impairments in pattern separation and ultimately cause the phenotype of fear generalisation in PTSD. Although further research is needed to obtain more insight in the exact mechanisms underlying the observed effects, these findings indicate that the hippocampus plays a significant role in the development of PTSD and aberrant fear generalisation, and therefore significantly contribute to the symptomatology of PTSD.

Acknowledgments

The completing of this research paper could not have been made possible without the contribution of several people to whom I would like to express my gratitude. Firstly, I would like to thank Dr. Marloes Henckens for her supervision and support during this research project. Furthermore, I would like to thank Gil Schulte and Miranda van Bodegom for their experimental work during MRI acquisition, and Prof. Dr. Tamas Kozicz as the second reader of the thesis. Lastly, I would like to thank Dr. Judith Homberg for the opportunity to do my MSc research internship at her department.

References

Adamczak, J., Farr, T., Seehafer, J., Kalthoff, D., & Hoehn, M. (2010). High field BOLD response to forepaw stiumulation in the mouse. *NeuroImage*, *51*, 704-712.

Adamec, R., Head, D., Blundell, J., Burton, P., & Berton, O. (2006). Lasting anxiogenic effects of feline predator stress in mice: Sex differences in vulnerability to stress and predicting severity of anxiogenic response from the stress experience. *Physiology and Behaviour, 88*(1-2), 12-29.

Adamec, R., Hebert, M., Blundell, J., & Mervis, R. (2012). Dendritic morphology of amygdala and hippocampal neurons in more and less predator stress responsive rats and more and less spontaneously anxious handled controls. *Behavioural Brain Research, 226*(1), 133-146.

Admon, R., Milad, M., & Hendler, T. (2013). A causal model of post-traumatic stress disorder: disentangling predisposition from acquired neural abnormalities. *Trends in Cognitive Sciences, 17*(7), 337-347.

Aimone, J., Deng, W., & Gage, F. (2011). Resolving new memories: a critical look at the dentate gyrus, adult neurogenesis and pattern separation. *Neuron, 70*(4), 589-596.

Albrecht, A., Caliskan, G., Oitzl, M., Keinemann, U., & Stork, O. (2013). Long-lasting increase of corticoterone after fear memory reactivation: anxiolytic effects and network activity modulation in the ventral hippocampus. *Neuropsychopharmocology, 38*, 386-394.

Altholtz, L., Fowler, K., Badura, L., & Kovacs, M. (2006). Comparison of the stress response in rats to repeated isoflurane or CO2:O2 anesthesia used for restraint during serial blood collection via jugular vein. *Journal of the American Association for Laboratory Animal Science, 45*(3), 17-22.

Altman, J., & Das, G. (1967). Postnata neurogenesis in the guinea-pig. *Nature, 214*, 1098-1101.

Amaral, D., & Lavenex, P. (2006). Hippocampal neuroanatomy. In D. Amaral, & P. Lavenex (Eds.), *The Hippocampus Book*. Oxford, England: Oxford University Press.

Amaral, D., Scharfman, H., & Lavenex, P. (2007). The dentate gyrus: Fundamental neuroanatomical organization (dentate gyrus for dummies). *Progress in Brain Research, 163,* 3-22.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.). Washington, DC: American Psychiatric Press.

Astur, R., St.Germanin, A., Tolin, D., Ford, J., Russell, D., & Stevens, M. (2006). Hippocampus function predicts severity of post traumatic stress disorder. *CyberPsychology and Behaviour, 9*(2), 234-240.

Bohlen, V., & Halbach, O. (2009). Structure and function of dendritic spines within the hippocampus. *Annals of Anatomy, 191*, 518-531.

Bonne, O., Brandes, D., Gilboa, A., Mosche Gomori, J., Shenton, M., Pitman, R., & Shalev, A. (2001). Longitudinal MRI study of hippocampal volume in trauma survivors with PTSD. *American Journal of Psychiatry, 158*, 1248-1251.

Bonne, O., Vythilingam, M., Ingaki, M., Wood, S., Neumeister, A., Nugent, A., ... Chamey, D. (2008). Reduced posterior hippocampal volume in posttraumatic stress disorder. *Journal of Clinical Psychiatry, 69*(7), 1087-1091.

Breslau, N. (2001). The epidemiology of posttraumatic stress disorder: What is the extent of the problem? *Journal of Clinical Psychiatry, 62*, 16-22.

Carrion, V., Haas, B., Garrett, A., Song, S., & Reiss, A. (2010). Reduced hippocampal activity in youth with posttraumatic stress symptoms: an fMRI study. *Journal of Pediatric Psychology, 35*(5), 559-569.

Cohen, H., Kozlovsky, N., Matar, M., Zohar, & Kaplan, Z. (2014). Distinctive hippocampal and amygdalar cytoarchitectural changes underlie specific patterns of behavioural disruption following stress exposure in an animal model of PTSD. *European Neuropsychopharmacology, 24*, 1925-1944.

Cordero, M., Venero, C., Kruyt, N., & Sandi, C. (2003). Prior exposure to a single stress session facilitates subsequent contextual fear conditioning in rats, evidence for a role of corticosterone. *Hormones and Behaviour, 44*, 338-345.

Diamond, D., Campbell, A., Park, C., Woodson, J., Conrad, C., Bachstetter, A., & Mervis, R. (2006). Influence of predator stress on the consolidation versus retrieval of long term spatial memory and hippocampal spinogenesis. *Hippocampus, 16*, 571-576.

Dias, G., Bevilaqua, M., Da Luz, A., Fleming, R., Carvalho, L., Cocks, G., ... Gardina, P. (2014). Hippocampal biomarkers of fear memory in an animal modal of generalized anxiety disorder. *Behavioural Brain Research, 263,* 34-45.

Ehler, A., Hackmann, A., & Michael, T. (2004). Intrusive re-experiencing in post traumatic stress disorder: Phenomenology, theory and therapy. *Memory, 12*(4), 403-415.

Fanselow, M., & Dong, H. (2010). Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron, 65*, 7-19.

Fiala, J., Spacek, J., & Harris, K. (2006). Dendritic structure and complexity in adult hippocampal dentate gyrus. *Neuroscience, 137*, 1299-1307.

Golub, Y., Mauch, C., Dahlhoff, M., & Wotjak, C. (2005). Consequences of extinction training on associative and non-associative fear in a mouse model of posttraumatic stress disorder (PTSD). B*ehavioural Brain Research, 205*, 544-549.

Grandjean, J., Schroeter, A., Batata, I., & Rudin, M. (2014). Optimization of antesthesia protocol for resting-state fMRI in mice based on differential effects of anesthetics on functional connectivity patterns. *NeuroImage, 102*, 838-847.

Greenberg, T., Carlson, J., Cha, J., Hajcak, G., & Mujica-Parodi, L. (2013). Ventromedial prefrontal cortex reactivity is altered in generalized anxiety disorder during fear generalization. *Depression and Anxiety, 30*, 242-250.

Hayes, J., Labar, K., McCarthy, G., Selgrade, E., Nasser, J., Dolcos, F., & Morey, R. (2011). Reduced hippocampal and amygdalar activity predicts memory distortions for trauma reminders in combat-related PTSD. *Journal of Psychiatric Research, 45*, 660-669.

Ji, J., & Maren, S. (2008). Differential roles for hippocampal areas CA1 and CA3 in the contextual encoding and retrieval of extinguished fear. *Learning & Memory, 15*, 244-251.

Kaouane, N., Porte, Y., Valle, M., Brayda-Bruno, L., Mons, N., Calandreau, L., ... Desmedt, A. (2012). Glucocorticoids can induce PTSD-like memory impairments in mice. *Science, 335*, 1510-1513.

Kavushansky, A., Vouimba, R., Cohen, H., & Richter-Levin, G. (2006). Activity and plasticity in CA1, the dentate gyrus and the amygdala following controllable vs. uncontrollable water stress. *Hippocampus, 16*, 35-42.

Kessler, R., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. (1995). Posttraumatic Stress Disorder in the National Comorbidity Survey. *Archives of General Psychiatry Journal, 52*, 1048-1060.

Kheirbek, M., Drew, L., Burghardt, N., Costantini, D., Tannenholz, L., Ahmari, S., ... Hen, R. (2013). Differential control of learning and anxiety along the dorsoventral axis of the dentate gyrus. *Neuron, 77,* 955-968.

Kheirbek, M., Klemenhagen, K., Sahay, A., & Hen, R. (2012). Neurogenesis and generalization: a new approach to stratify and treat anxiety disorders. *Nature*

Neuroscience, 15, 1613-1620.

Lebow, M., Neufeld-Cohen, A., Kuperman, Y., Tsoory, M., Gil, S., & Chen, A. (2012). Susceptibility to PTSD-like behaviour is mediated by corticotropin-releasing factor receptor type 2 levels in the bed nucleus of stria terminalis. *The Journal of Neuroscience, 32*(20), 6906-6916.

McEwan, B. (1999). Stress and hippocampal plasticity. *Annual Review of Neuroscience, 22*, 105-122.

Mehta, D., & Binder, E. (2012). Gene x environment vulnerability factors for PTSD: The HPA-axis. *Neuropharmocology, 62,* 654-662.

Nimchinsky, E., Sabatini, B., & Svoboda, K. (2002). Structure and function of dendritic spines. *Annual Review of Physiology, 64*, 313-353.

Pitman, R., Rasmusson, A., Koenen, K., Shin, L., Orr, S., Gilbertson, M., ... Liberzon, I. (2012). Biological Studies of Post-traumatic Stress Disorder. *Nature Reviews Neuroscience, 13,* 769-787.

Pugh, R., Tremblay, D., Fleshner, M., & Rudy, J. (1997). A selective role for corticosterone in contextual-fear conditioning. *Behavioural Neuroscience, 111*(3), 503-511.

Redila, V., & Christie, B. (2006). Exercise-induced changes in dendritic structure and complexity in adult hippocampal dentate gyrus. *Neuroscience, 137*, 1299-1307.

Sahay, A., Wilson, D., & Hen, R. (2011). Pattern separation: A common function for new neurons in the hippocampus and olfactory bulb. *Neuron, 70*, 582-588.

Sebastian, V., Estil, J., Chen, D., Schrott, L., & Serrano, P. (2013). Acute physiological stress promotes clustering of synaptic markers and alters spine morphology in the hippocampus. *Plos One, 8*(10), e79077.

Servatius, R., Ottenweller, J., & Natelson, B. (1995). Delayed startle sensitization distinguishes rats exposed to one or three stress sessions: Further evidence towards an animal model of PTSD. *Biological Psychiatry, 38*, 589-546.

Shansky, R. (2015). Sex differences in PTSD resilience and susceptibility: Challenges for animals models of fear learning. *Neurobiology of Stress, 1*, 60-65.

Shin, L., Rauch, S., & Pitman, R. (2006). Amygdala, medial prefrontal cortex and hippocampul function in PTSD. *Annals of New York Academy of Sciences, 1071*, 67-79.

Shin, L., Shin, P., Heckers, S., Krangel, Y., Macklin, M., Orr, S., ... Rauch, S. (2004). Hippocampal function in posttraumatic stress disorder. *Hippocampus, 14*, 292-300.

Shors, T. J., & Leuner, B. (2004). The opposite effects of stress on dendritic spines in male vs. female rats are NMDA receptor-dependant. *European Journal of Neuroscience, 19*(1), 145-150.

Ter Horst, J., De Kloet, E., Schachinger, H., & Oitzl, M. (2012). Relevance of stress and female sex hormones for emotion and cognition. *Cellular and Molecular Neurobiology, 32*, 725-735.

Van Zuiden, M., Geuze, E., Willemen, H., Vermetten, E., Maas, M., Amarouchi, K., ... Heijnen, C. (2012). Glucocorticoid receptor pathway components predict posttraumatic stress disorder symptom development: a prospective study. *Biological Psychiatry, 71*, 309-316.

Yehuda, R., Giller, E., Southwick, S., Lowy, M., & Mason, J. (1991). Hypothalamic-pituitary-adrenal dysfunction in posttraumatic stress disorder. *Biological Psychiatry, 30*, 1031-1048.

Yehuda, R., Teicher, M., Trestman, R., Levengood, R., & Siever, L. (1996). Cortisol regulation in posttraumatic stress disorder and major depression: a chronobiological analysis. *Biological Psychiatry, 40,* 79-88.

Zerbi, V., Wiesman, M., Emmerzaal, T., Jansen, D., Van Beek, M., Mutsaers, M., ... Kiliaan, A. (2014). Resting-state functional connectivity changes in aging apoE4 and apoE-KO mice. *Neurobiology of Disease, 34*(42), 13963-13975.

Zohar, J., Yahalom, H., Kozlovsky, N., Cwikel-Hamzany, S., Matar, M., Kaplan, Z., ... Cohen, H. (2011). High dose hydrocortisone immediately after trauma may alter the trajectory of PTSD: Interplay between clinical and animal studies. *European Neuropsychopharmacology, 21*, 796-809.

How Do Familiarity and Expectation Influence the Human Brain Signal?

Mariya E. Manahova[1]
Supervisors: Floris de Lange[1], Pim Mostert[1]

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

Familiarity (i.e., whether an observer has seen an image before) and expectation (i.e., whether an observer can predict which image will follow based on the current image) can influence the processing of visual information in the brain. The electrophysiological signal from the brain can indicate how visual processing changes between familiar and novel as well as between expected and unexpected images. An important question arising from previous research is whether familiarity has an effect on neural activity that is separable from the effect of expectation. In order to address this issue, we adopted an experimental design which manipulated familiarity separately from expectation, and we used magnetoencephalography (MEG) to record brain activity in humans. We found that familiarity, unaffected by expectation, has a genuine influence on the brain signal, such that novel images induced significantly higher amplitude than familiar images. In addition, expectation also had an effect on the neural response: unexpected images were accompanied by significantly higher amplitude than expected images. These outcomes demonstrate that both visual familiarity and expectation influence the human electrophysiological signal, and they do so in similar ways. These findings improve our understanding of how visual processing changes with the amount and type of experience the brain has had with a visual stimulus.

Keywords: familiarity, expectation, vision, perception, MEG

Corresponding author: Mariya E. Manahova; **E-mail:** m.manahova@donders.ru.nl

Our environment is characterised by complex visual features, making visual perception an arduous task. The human visual system has evolved to parse this complicated conglomeration of colours, orientations, and textures and deliver a stable percept of the world. In order to achieve this computationally heavy process, the visual system allots more resources to the processing of some types of information than others. For instance, input can be represented in more detail in visual cortex if it comes from the fovea compared to the periphery of the eye (Engel, Glover, & Wandell, 1997). Attention can also affect how information is prioritised in the brain: input can be weighted more strongly if it is attended compared to unattended (Lu & Dosher, 1998; Yeshurun & Carrasco, 1998; Baldassi & Burr, 2000; Carrasco, Penpeci-Talga, & Eckstein, 2000; Cameron, Tai, & Carrasco, 2002).

Another factor affecting perception is how much experience the visual system has had with a particular object or feature. When an individual observes a specific scene multiple times, the brain learns about this environment. The next time the person encounters this scene, the visual system uses its knowledge about the features characteristic of this scene and does not process it in the same way it would a novel scene (Li, Miller, & Desimone, 1993). Specifically, when an image is presented repeatedly, the brain activity in response to it diminishes (Miller, Li, & Desimone, 1991; Grill-Spector Henson, & Martin, 2006). This shows that the visual system readily adapts to repeated input.

When an observer views an image repeatedly, this renders the image familiar. Li et al. (1993) investigated how the electrophysiological signal measured in macaque inferior temporal (IT) cortex changes as a visual stimulus becomes familiar. The researchers presented streams of images; on multiple occasions, they showed a single image which was initially novel but over time became familiar. They found that, for a subset of neurons, the spike rate gradually diminished as the initially novel stimulus became familiar to the monkey. Then, when a novel stimulus was presented, those same neurons reached a higher spiking rate, signifying that the reduction in activity for familiar stimuli was not simply due to neural fatigue.

Studies in humans also show that familiar images are associated with lower brain activity than novel images. In a functional magnetic resonance imaging (fMRI) study (Rossion, Schiltz, & Crommelinck, 2003), the BOLD response was measured while participants saw familiar and novel faces. In the fusiform face area (FFA) and in the occipital face area (OFA) of the right hemisphere, which are especially sensitive to facial features, novel stimuli elicited higher activation than familiar stimuli. Even though the BOLD signal in human sensory cortex and the spike rate of monkey IT neurons are two very different measures, in this case they are in agreement and demonstrate that familiar visual stimuli induce lower brain activity than novel images.

The presence of a sequence in visual information can also affect visual processing. If an individual observes images in a specific temporal sequence, then it is possible to predict which image will be presented next. In order to investigate this type of visual expectation, Meyer and Olson (2011) recorded the firing rate of neurons in monkey IT. They presented pairs of images (A and B), one image following the other, so the monkeys learned to expect the second image (B) once they saw the first (A). Sometimes the first image (A) was followed by the trained second one (B), rendering the second image expected. At other times, an image from another pair (C) was shown after A, and in this case the second image (C) was unexpected based on the first one (A). Then, the researchers recorded the neural response to the two images. The firing rate of IT neurons was lower for an expected image than for an unexpected one. Moreover, the response to an expected image was reduced, or truncated, more rapidly than the response to an unexpected one.

Similar results have been acquired in human studies. In an experiment examining expectation with fMRI, participants first heard an auditory cue which predicted the orientation of a visual grating (Kok, Jehee, & de Lange, 2012). Then, a grating was shown, and the BOLD response to that visual stimulus in V1 was analysed. Activity in V1 was lower when the tone correctly predicted the grating orientation, rendering it expected, compared to when the tone induced an incorrect prediction about the grating, making it unexpected. fMRI investigations in humans and electrophysiological experiments in monkeys jointly suggest that expected stimuli are associated with lower activity than unexpected ones, a phenomenon called expectation suppression.

These studies show that both familiarity and expectation can lead to decreased neural activation. This makes sense since familiarity and expectation are related phenomena: conceptually, familiarity is highly relevant for expectation. While observing a visual scene, the brain becomes familiar with the scene's features. This acquired knowledge can be used to predict the upcoming input, which results in expectation (Clark, 2013). Therefore, it seems plausible that people need to

be familiar with something in order to expect it. In terms of brain activity, it is possible that any effect of expectation is actually due to familiarity. Expectation suppression may not be caused by the expectation itself but rather by the familiarity, which is necessary for the expectation to occur. However, the electrophysiological study in monkey IT with expected and unexpected stimuli by Meyer and Olson (2011) suggests that this is not the case. The researchers presented pairs of stimuli, and the leading image could render the trailing one expected or unexpected. Both types of trailing images had been shown during training, so they were familiar to the monkeys. The neural response to expected and unexpected differed even though both types of images were familiar, suggesting that expectation has an effect which cannot be explained by familiarity.

Familiarity has also been investigated with single-cell recordings from IT (Meyer, Walker, Cho, & Olson, 2014). Monkeys first saw single images multiple times and became familiar with them. Later, while neural activity was being recorded, the monkeys saw the images in sequences. Pairs of images were alternated, such that on a single trial the macaques would see image A, then image B, then A, B, A, B. Sometimes these sequences were comprised of familiar images, and other times of novel ones which the monkeys had never seen before. Images were presented very rapidly, a single image lasting for 120 ms and immediately followed by the next image. The recordings showed that the firing rate for familiar images was truncated more quickly than for novel images. In addition, familiar stimuli elicited responses with a larger dynamic range (larger peak-to-trough difference) than novel stimuli.

According to this study (Meyer et al., 2014b), familiarity's influence on the neural response is similar to that of expectation (i.e., truncation of the response) as found by Meyer and Olson (2011). However, based on the experimental design, it is possible that expectation also contributes to this observed effect. In particular, each familiar image was used during 6.5 sessions on average; in contrast, each novel image was only used during one session. Since the monkeys had been exposed to the familiar images extensively but had never seen the novel images before, it is plausible that, when observing a sequence of familiar images, they could predict an upcoming familiar image much more easily than an image in a novel sequence. It is also possible that the same pair of familiar stimuli was observed on multiple occasions, leading to more sequence learning for the familiar images than for the novel ones. In other words, upcoming images were more predictable in the familiar condition than in the novel condition. Therefore, the way in which familiarity was manipulated also influences expectation, making it possible that the observed effect is not solely caused by familiarity but also by expectation.

The current research project

In order to determine whether familiarity has a genuine effect on the neural signal even when stimuli are not predictable, we conducted a study in which we manipulated familiarity separately from expectation. We used magnetoencephalography (MEG) to record brain activity in humans. Similarly to Meyer et al.'s (2014b) design, participants saw rapidly presented six-image sequences, and the images depicted objects. Subjects were familiarised with images by means of a training session. Then, during the MEG recording session, participants saw the familiar images as well as novel images which they had not seen before. In order to induce expectation, we showed subjects two sets of six images during the training session. For one set, they saw the images in a specific sequence, making it possible to come to expect that sequence (familiar sequenced condition). For the other set, they always saw the images in a shuffled order, so they could not learn to expect a specific sequence (familiar unsequenced condition). During the MEG session, the familiar sequenced images were sometimes presented in the learned order (expected condition) and sometimes in a shuffled order (unexpected condition). The unsequenced familiar images were again presented in shuffled orders during the MEG session (for a diagram depicting the full list of conditions, see Fig. 1.). Thus, familiarity referred to whether the subjects had seen an image during the training session, while expectation referred to whether they could predict the sequence of the upcoming images. With this design, we could isolate the effect of familiarity from that of expectation because we manipulated familiarity separately from expectation. Notably, familiarity still had a significant influence on the neural response even without the influence of expectation.

Method

Participants

Twenty-nine healthy human volunteers (15 female, 14 male, M_{age} = 24.17 years, SD = 3.80 years) with normal or corrected-to-normal vision, recruited from the university's

participant pool, completed the experiment and received either monetary compensation or study credits. The study was approved by the local ethics committee (CMO Arnhem-Nijmegen, Radboud University Medical Center) under the general ethics approval ("Imaging Human Cognition", CMO 2014/288), and the experiment was conducted in compliance with these guidelines. Written informed consent was obtained from each individual.

Procedure

First, participants completed a behavioural training session in which they only saw the familiar images (sequenced and unsequenced). They saw the familiar sequenced images always in the same order, while the familiar unsequenced images were shown in shuffled orders. Importantly, the order for the sequenced images was circular: each of the six images could be presented first, and they all had equal predictive values. Familiar sequenced images comprised 50% of trials, and familiar unsequenced images comprised the other 50%. Participants performed an oddball detection task by pressing the spacebar when they saw an image of a rubber duck. Images of duckies were presented on 10% of trials as one of the six images in the sequence. The duckies were of eight different colours and there were two viewpoints per colour for a total of 16 images of rubber ducks. Multiple images of ducks were used to avoid confounding the oddball stimulus with a particular colour (e.g., yellow) or a specific viewpoint. We chose to include a visual oddball task in order to keep participants' attention on the visual stimuli even though subjects did not actually have to perform a task on those stimuli.

During the behavioural training session, participants completed 10 blocks of 80 trials each for a total of 800 trials. Each block lasted for 4.9 minutes, so the whole experiment lasted for about one hour. At the end of the behavioural training session, participants' knowledge of the order of familiar sequenced images was assessed. Participants were shown one of the six sequenced images, and they had to indicate which of the five remaining images was most likely to follow it. This was done for each of the six images in the familiar sequenced set. The assessment took about three minutes.

One or two days later, participants completed the MEG testing session in which they saw familiar and novel images. In contrast with the training session, the familiar sequenced images were sometimes presented in the learned order (expected) and other times in a shuffled order (unexpected). The shuffled sequences for unexpected trials were chosen in such a way that each image in the sequence was followed by an unpredicted image; in other words, none of the images were followed by the image they predicted. Similarly to the training session, the familiar unsequenced images were shown in shuffled orders. Familiar unsequenced images comprised one third of trials, and familiar sequenced images also comprised one third of trials, half of those being expected and half unexpected (see Fig. 1). Participants also saw novel images which they had not seen before, and these comprised the remaining one third of trials. Different novel images were used for every trial, so each novel image was only shown once during the whole experiment. Similarly to the training session, participants had to perform an oddball task: they had to respond when they saw a rubber duck, and these were presented on 10% of trials (Fig. 1 shows a diagram of the experimental conditions). During the MEG testing session, participants completed 8 blocks of 120 trials each for a total of 960 trials. Each block lasted for 7.4 minutes, so the whole experiment lasted about one hour. At the end of the MEG testing session, participants' knowledge of the familiar images was assessed. Participants saw 60 images, the twelve familiar ones and 48 randomly selected from the novel images participants had been shown, and subjects had to indicate whether the image was familiar or novel. 'Familiar' referred

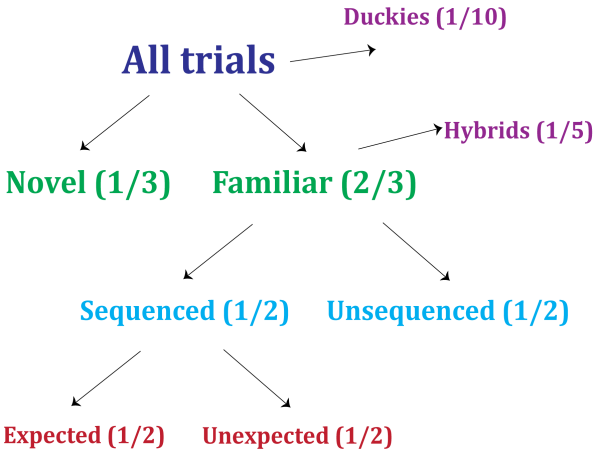


Fig. 1. A diagram of the different conditions included in the experiment. The trials were divided into familiar (2/3 of all trials) and novel (1/3 of all trials). One half of familiar trials were sequenced (1/3 of all trials) and the other half were unsequenced (1/3 of all trials). One half of sequenced trials were expected (1/6 of all trials) and the other half were unexpected (1/6 of all trials). Out of all familiar trials, 20% were hybrids. Out of all trials, ducks were presented in 10% of the cases.

to images seen repeatedly during the behavioural training session as well as during the MEG testing session, while 'novel' referred to images seen only once during the MEG testing session.

Trial structure

For the behavioural training session as well as for the MEG testing session, each trial began with a fixation dot. Fixation dots were of the type "bull's eye": a small black dot in the middle surrounded by a larger, thin white circle, which was in turn surrounded by a larger, thin black circle. This type of fixation dot has been shown to improve participants' ability to fixate (Jehee, Brady, & Tong, 2011). For a diagram of the trial structure, see Fig. 2. The fixation dot was presented for a period between 500 and 750 ms; the exact duration was determined randomly per trial. Then an image was shown for 180 ms, immediately followed by another image, also lasting for 180 ms. This was repeated until all six images in the trial were presented. Afterwards, if an oddball was presented during the trial and a response was given, the fixation dot turned green for 500 ms. If the response was incorrect, the fixation dot turned red for 500 ms. A response was considered incorrect on three occasions: 1) if the subject pressed the button during a trial with an oddball stimulus but before the oddball was presented, 2) if the participant pressed the button on a trial where no oddball was presented, or 3) if the subject did not press the button on a trial where an oddball was presented. If no oddball was presented and no response was given, the change in colour of the fixation dot was omitted, and the white-and-black fixation dot remained on the screen for 750 ms. Then, a blank screen was presented for 1250 ms, and participants were encouraged to blink during this period. After this, the next trial began with a fixation dot.

Materials

The stimuli were shown on monitors with a resolution of 1920x1080 pixels using MATLAB (The Mathworks, Inc., Natick, Massachusetts, United States) and the Psychophysics Toolbox extensions (Brainard, 1997). A refresh rate of 100 Hz was used in order to ensure that the presentation of the stimuli lasted for exactly 180 ms. We chose the duration of 180 ms for image presentation based on data from Meyer et al. (2014b) suggesting that this induces a larger response truncation effect in humans than a duration of 120 ms. In both the behavioural and MEG sessions, the images subtended four degrees of visual angle, again following Meyer et al.'s (2014b) paradigm. For the behavioural training session, 24-inch monitors were used. For the MEG session, the screen was 53 cm in width and 41 cm in height, and a PROpixx projector was used to project the images on the screen; the projector had a resolution of 1920x1080 and an aspect ratio of 16:9.

Stimuli

Participants viewed images from the set provided at http://cvcl.mit.edu/MM/uniqueObjects.html. A different object was represented in each image, and all objects were shown against a white background. A total of 2377 images were available to be presented, and from those 2377, 2054 were presented for each participant. Familiar images were randomly selected for each pair of participants and manually inspected before presentation in order to avoid any striking pictures. Each pair of participants saw different familiar images, and the familiar sequenced and familiar unsequenced images were counterbalanced within a pair of participants. Specifically, if for participant 1, set 1 comprised the familiar sequenced images and set 2 comprised the familiar unsequenced images, the opposite was true for participant 2: set 2 comprised the familiar sequenced images, and set 1 comprised the familiar unsequenced images.

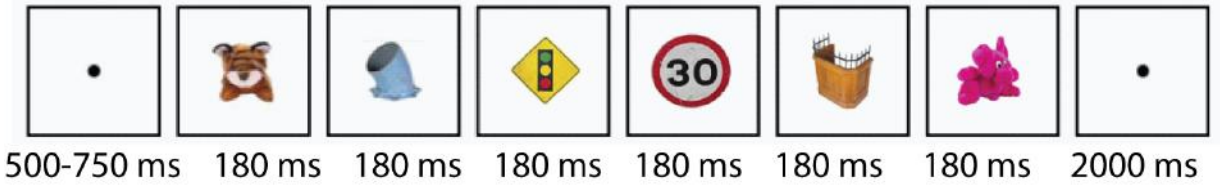


Fig. 2. A diagram of the trial structure. At the beginning of each trial, a fixation dot appeared for a randomly selected time period between 500 and 750 ms. The first image was presented for 180 ms. Immediately after that, the second image was shown for 180 ms. This continued until all six images had been displayed. The last image was followed by a fixation dot present for 2000 ms.

MEG recordings

Brain activity was recorded using a 275-channel MEG system with axial gradiometers (VSM/CTF Systems, Coquitlam, BC, Canada) in a magnetically shielded room. During the experiment, head position was monitored online and corrected if necessary (Stolk, Todorovic, Schoffelen, & Oostenveld, 2013). This method uses three coils: one placed on the nasion, one in an earplug in the left ear, and one in an earplug in the right ear. To aid in the removal of eye- and heart-related artifacts, horizontal and vertical electrooculograms (EOG) as well as an electrocardiogram (ECG) were recorded. A reference electrode was placed on the left mastoid. The sampling rate for all signals was 1200 Hz. A projector outside the magnetically shielded room projected the visual stimuli onto a screen in front of the participant via mirrors. Participants gave their behavioural responses via an MEG-compatible button box. Participants' eye movements and blinks were also monitored by an eye-tracker system (EyeLink, SR Research Ltd., Mississauga, Ontario, Canada).

MRI Recordings

Anatomical magnetic resonance imaging (MRI) scans were acquired for the purpose of source localization analysis. Anatomical images were collected for each participant or retrieved from a database when available. The images were acquired using a 3T MRI system (Siemens, Erlangen, Germany). These recordings were not used for the current work.

MEG data analysis

The MEG data were preprocessed offline using the FieldTrip software (Oostenveld, Fries, Maris, & Schoffelen, 2010) (www.fieldtriptoolbox.org). Trials with high variance were manually inspected and removed if they contained excessive and irregular artifacts. Independent component analysis (ICA) was applied to identify regular artifacts such as heartbeat and eye blinks. The independent components for each participant were then correlated to the horizontal and vertical EOG signals and to the ECG signal. In this way, it was possible to identify which components most likely corresponded to the heartbeat and eye blinks. Furthermore, the data were baseline-corrected on the interval starting at 200 ms before stimulus onset until stimulus onset (0 ms). A low-pass filter at 30 Hz was applied to the data. Trials where oddball stimuli were presented and/or a response was given were removed from analysis. This was done because oddballs and responses elicited neural activity unrelated to the research question. A planar transformation and event-related field (ERF) analysis were applied to the MEG data.

Statistical analysis

For the behavioural results, mean reaction time and accuracy were first calculated within participant per condition. Paired-samples t-tests (p-value = .05, two-tailed) were applied to the data of all participants within the two relevant conditions for a comparison.

In order to statistically test the MEG results and control for multiple comparisons, we applied cluster-based permutation tests (Maris & Oostenveld, 2007), as implemented by FieldTrip (Oostenveld et al., 2010). The tests were carried out on the time period between 0 and 1200 ms, 0 ms being the onset of the first stimulus, over all sensors, and 1000 permutations were used per contrast. For each sensor over multiple time points, the MEG signal was compared between two conditions, yielding a t-value. A sensor could potentially contribute to a cluster at a certain time point if its corresponding p-value was lower than .05 (two-tailed). Temporally adjacent time points with such p-values were grouped into positive and negative clusters. Cluster-level statistics were calculated by summing the t-values within a cluster, and a cluster was considered significant if its p-value was smaller than .05. The standard error of the mean was computed within participants, as described by Cousineau (2005) and with the correction suggested by Morey (2008).

Results

Behavioural results

The participants' task was to press a button whenever they saw an oddball stimulus, in this case an image of a rubber duck. Participants correctly identified almost all oddballs and refrained from responding when no oddball was presented (M = 99.35%, SD = 0.35%). Participants' accuracy was not significantly influenced by familiarity ($t(28) = -1.20$, $p = .24$), sequence ($t(28) = 1.17$, $p = .25$), or expectation ($t(28) = 1.04$, $p = .31$). This high accuracy was not driven by the fact that participants did not respond regardless of trial type. On oddball trials when they had to press a

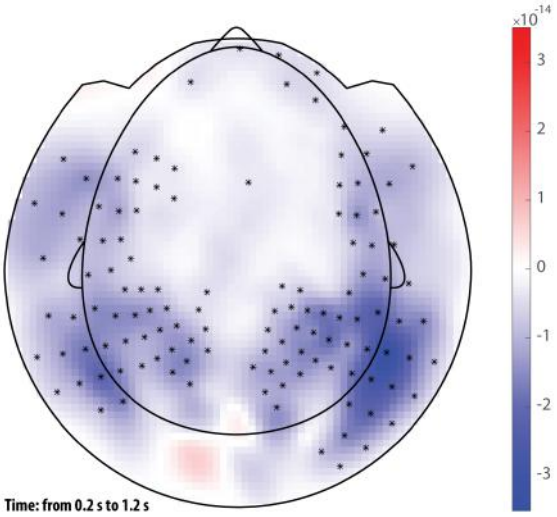


Fig. 3. Topography of the difference between the familiar unsequenced and novel conditions. Black asterisks mark sensors that contribute to the significant cluster for at least half of the time period from 200 ms to 1200 ms.

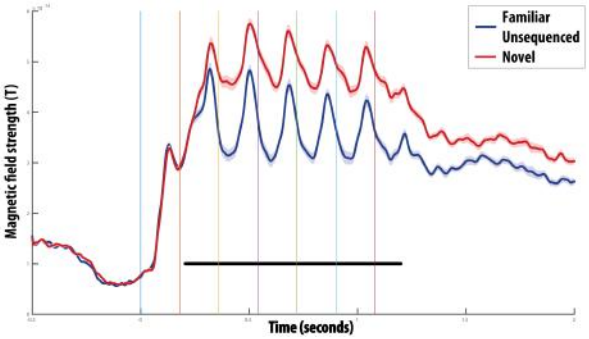


Fig. 4. Activity over time for the familiar unsequenced condition (blue) and the novel condition (red). Activity is averaged over the sensors highlighted in Figure 3. The shaded areas are error bars illustrating the within-subject SEM for the unsequenced familiar (light blue) and the novel (light red) conditions. The horizontal black bar at the bottom shows that at least one of the selected sensors contributes to the significant cluster at this time point. The coloured vertical lines denote the onset of each image (1-6), and the last one denotes the offset of the last image.

button, their accuracy remained high (M = 94.90%, SD = 2.88%). Participants' accuracy on oddball trials was not significantly affected by familiarity ($t(28)$ = -0.81, p = .42), sequence ($t(28)$ = 0.07, p = .95), or expectation ($t(28)$ = 0.57, p = .57). Furthermore, participants' reaction times were not significantly affected by condition either. Familiarity did not affect reaction times ($t(27)$ = -1.74, p = .09) and neither did sequence ($t(27)$ = -0.30, p = .77) nor expectation ($t(27)$ = -0.04, p = .97). Reaction time data were analysed for 28 out of the 29 participants because an error in data acquisition rendered accuracy and reaction times from the first participant unusable.

At the end of the behavioural training session, participants' knowledge of the order of the familiar sequenced images was assessed. On average, when participants were shown an image and had to report which image followed it, they selected the correct image 25% of the time (SD = 19.7%), which was not significantly different from chance level, i.e., 20% ($t(27)$ = 1.29, p = .21). This suggests that subjects were mostly unaware of the sequence, which is in agreement with their informal verbal reports.

At the end of the MEG session, participants' knowledge of the familiar images was assessed. On average, when participants had to report whether an image was familiar or novel, they did so correctly with a mean accuracy of 91.9% (SD = 5.8%). Apparently, subjects were aware which images were familiar and which were novel.

MEG results

To look at the difference between familiar and novel items without any influence of the expectation manipulation, we tested the difference between the familiar unsequenced vs. novel conditions since participants did not learn a sequence for the images in the unsequenced condition. A significant difference emerged for the cluster shown in Figure 3 from approximately 200 ms until 1200 ms (p = .001). The black asterisks in the figure denote sensors that contribute to the significant cluster for at least half of the time period from 200 ms to 1200 ms (Fig. 4). Clearly, there is a significant difference in the amplitude of brain activity between familiar and novel items. A visual inspection of the signal demonstrates that the dynamic range (peak-to-trough difference) for familiar images is larger than for novel ones. Also, there appears to be a sharper reduction, i.e., truncation, of the signal for familiar than for novel images.

To look at the difference between expected and unexpected items when familiarity was held constant, we tested the difference between the expected vs. unexpected conditions. A significant difference emerged for the cluster shown in Figure 5 from approximately 500 ms until 900 ms (p = .005). The black asterisks in the figure denote sensors that

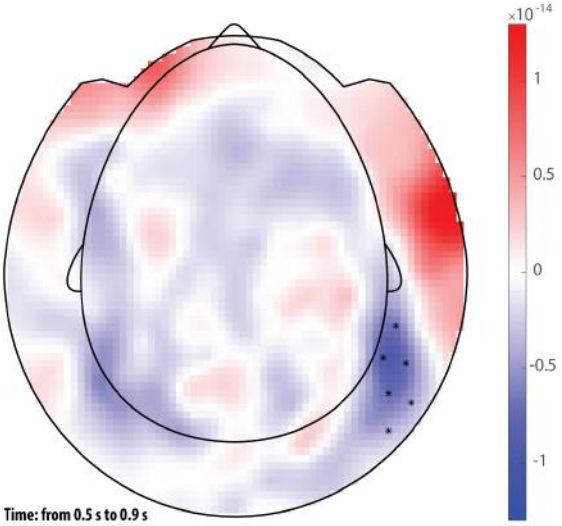


Fig. 5. Topography of the difference between the expected and unexpected conditions. Black asterisks mark sensors that contribute to the significant cluster for at least half of the time period from 500 ms to 900 ms.

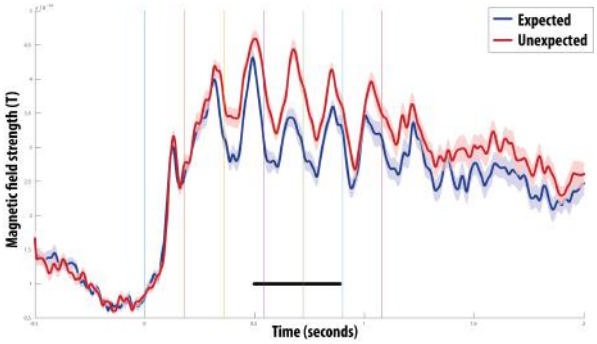


Fig. 6. Activity over time for the expected condition (blue) and the unexpected condition (red). Activity is averaged over the sensors highlighted in Figure 5. The shaded areas are error bars illustrating the within-subject SEM for the expected (light blue) and the unexpected (light red) conditions. The horizontal black bar at the bottom shows that at least one of the selected sensors contributes to the significant cluster at this time point. The coloured vertical lines denote the onset of each image (1-6), and the last one denotes the offset of the last image.

contribute to the significant cluster for at least half of the time period from 500 ms to 900 ms (Fig. 6). Evidently, there is a significant difference in the amplitude of neural activity between expected and unexpected items. A visual inspection of the signal does not reveal a clear difference in dynamic range or truncation between the two types of responses.

Moreover, we were interested in the expected vs. unsequenced comparison as well as the unexpected vs. unsequenced comparison, since the former could illustrate the effect of a confirmed expectation and the latter could demonstrate the effect of a violated expectation. However, neither contrast yielded a significant difference.

Discussion

Our findings demonstrate that familiarity and expectation both affect brain activity. Notably, the effects of familiarity and expectation on the electrophysiological signal are similar. Both lead to a significant reduction in the amplitude of the signal, and in both cases this decrease in activity is strongest in posterior, right-lateralised areas. A visual inspection of the timecourse for the familiar-novel comparison reveals a truncated response with a higher dynamic range for familiar stimuli compared to novel ones. This effect is not so clearly present in the expected-unexpected comparison.

We qualitatively observe a larger dynamic range and more truncation of the signal for familiar than novel stimuli, which is in agreement with Meyer et al.'s (2014b) findings. The topography maps for the differences in neural activity between familiar and novel and between expected and unexpected point to a posterior, right-lateralised brain area. It is possible that this area reflects the ventral visual stream and specifically inferior temporal (IT) cortex, which would be in accordance with the fact that Meyer et al. (2014b) recorded neural firing from monkey IT. The amplitude effect we found for familiar compared to novel items has not been reported by Meyer et al. (2014b). Perhaps the measures used by Meyer et al. (2014b), spike rate and local field potentials, did not reveal this difference in amplitude, while the MEG signal did. Our finding fits with other studies showing that familiarity reduces neural activity (Li et al., 1993; Rossion et al., 2003).

In our experimental design, familiarity was manipulated separately from expectation, allowing us to distinguish between the effects of these two factors. The results suggest that familiarity, defined as having seen an image in the past, has a genuine effect on brain activity even when the stimuli occur in an unpredictable sequence. Moreover, expectation, defined as being able to predict the upcoming image, influences the neural response even when familiarity is held constant. This is noteworthy because Meyer et al. (2014b) show how the electrophysiological signal differs between familiar and novel images

presented with a rapid presentation design, but in their design, the way familiarity was manipulated also influenced expectation. Therefore, the neural effect they observed may have partially occurred because monkeys were able to predict the upcoming stimulus better when the images were familiar than when they were novel. In our design, which also featured rapidly presented image sequences of six objects, the familiarity manipulation was unaffected by expectation. We isolated the effect of familiarity by comparing the brain activity for familiar unsequenced images and for novel images. This comparison did not include any influence from expectation because familiar unsequenced images always appeared in unpredictable orders and because each novel image was only shown once, so it was not possible to form expectations about novel items. The comparison between familiar unsequenced and novel images still showed a significant difference, indicating that familiarity has a genuine effect on the amplitude of the signal and on the qualitatively observed truncation and dynamic range of the response. Furthermore, the expectation manipulation was in addition to the familiarity one: participants saw familiar sequenced images, which could be expected (shown in the learned sequence) or unexpected (shown in a shuffled order). The comparison between expected and unexpected images also showed a significant difference in terms of amplitude. These results demonstrate the separable effects of familiarity and expectation.

Interestingly, the difference between familiar and novel items becomes significant around 200 ms, while the difference between expected and unexpected images attains significance later, around 500 ms. This may be caused by a distinction between the two types of experimental manipulations. With respect to the familiar vs. novel comparison, participants know whether the trial will be comprised of familiar or novel images once they see the first image. Regarding the expected vs. unexpected comparison, however, the first image is uninformative; based on the second image, the visual system can tell whether the current trial is expected or unexpected. Since the presentation of one image lasts for 180 ms, this difference between the manipulations could explain why the expectation effect becomes significant 180 ms later than the familiarity effect. In our data, however, the expectation effect attains significance approximately 300 ms later than the familiarity effect. This further delay could be due to noise, or it may occur because the visual system needs to accumulate more information before detecting a violation of expectations than before detecting novel input, resulting in a delayed latency for the expectation effect compared to the familiarity effect.

It is intriguing that neither the expected-unsequenced comparison nor the unexpected-unsequenced comparison yielded a significant difference. Perhaps this is the case because the difference in neural processing between a confirmed expectation (expected) and a lack of expectation (unsequenced) was not substantial enough in our dataset to produce a significant result. Likewise, the difference between a violated expectation (unexpected) and a lack of expectation (unsequenced) was not prominent enough to bring about a significant outcome. Apparently, in our dataset, the difference in neural processing between a confirmed expectation (expected) and a violated one (unexpected) was larger than in the previous two cases and was substantial enough to induce a significant result.

Importantly, in our experiment we did not have a complete orthogonal manipulation of familiarity and expectation. Such a design would have required a novel, expected condition which we did not have. This is because it is difficult to build an expectation for which image is going to come next without being familiar with the images. Usually, familiarity is necessary for expectation, and this is also the case in our operationalization of these two concepts. Future studies could examine how to manipulate expectation without familiarity. Perhaps this can be done if expectation is based on an abstract rule: an image of class A (e.g., animal) is followed by an image of class B (e.g., fruit). Participants would then expect the type of image coming next even if they are not familiar with the exact image, i.e., even if they have not seen this specific exemplar from the general category that is expected. This would manipulate expectation on a more conceptual level than the low-level sequence-based expectation we implemented in our experiment because participants' expectations would refer to categories of objects instead of specific items.

It is of interest that we observed an effect of expectation although participants were not consciously aware of having learned a sequence for the familiar sequenced images. When asked, participants said that they did not notice any specific order for the images. The behavioural assessment of sequence knowledge also showed that participants' performance was very low when they were shown an image and had to report which image should follow. Nevertheless, the neural response distinguished between expected and unexpected conditions. It appears that the observed difference represents

low-level sequence learning which may occur only in sensory cortical areas and may not reach widespread recurrent processing, thus not entering conscious awareness (Lamme, 2006). Moreover, this effect of expectation was induced by only a one-hour training session one or two days before the MEG recording session, which is very little time compared to the extensive training monkeys underwent (Meyer et al., 2014b). The fact that participants were unaware of the image sequence and the short training period make it even more remarkable that we found a significant difference in neural activity between expected and unexpected trials.

In our study, expectation operated on a stimulus-to-stimulus basis. In the learned sequence, image A predicted image B, image B predicted image C, image C predicted image D, etc. Importantly, the sixth image, F, predicted the first image, A, meaning that the sequence was circular. This ensured that image A did not carry the predictive value for the whole sequence; rather, each image predicted the next one. When the image order was shuffled, each image was followed by an unexpected image; therefore, the stimulus-to-stimulus expectation was violated by every image in the sequence. In accordance with this, the significant difference in the electrophysiological signal between expected and unexpected trials was sustained over the presentation of multiple images, suggesting that each of those images was unexpected based on the previous one. If the prediction value had been carried by a single image only, when the expectation was violated, we would have expected to see a significant difference only for the image after that one. In this case, expectation would have operated on a trial-to-trial basis: as soon as the brain discovers that the images in this trial will appear in a shuffled order, expectations about following images are not employed. However, the fact that we found a sustained difference over multiple stimuli suggests that expectation operated on a stimulus-to-stimulus basis in our experiment. Meyer, Ramachandran, and Olson (2014) revealed a similar outcome when they presented monkeys with sequences of three images. The authors found that the expectation for the current image was based on the immediately preceding image and not on the first image in the sequence. Perhaps rapid presentation designs such as theirs and ours preclude the categorization of a trial as expected or unexpected, thus preventing trial-to-trial expectation, but rather permit low-level expectation effects on a stimulus-to-stimulus basis.

We aimed to determine the neural effects of familiarity and expectation by strictly defining these concepts in terms of experimental manipulations. Obviously, being familiar with visual input can be much more nuanced than simply having seen an image during a training session, and expecting a visual feature can take many other forms besides predicting which image will be presented next. In reality, these concepts are much richer than how they are defined by our experimental manipulations, so the division between familiarity and expectation may not be so clear-cut. In our everyday environments, we usually become familiar with images because they appear more often, which means that we also expect to see them more often. In this sense, familiarity and expectation are necessarily intertwined; perhaps familiarity can even be construed as a type of expectation. A noteworthy difference remains between the two, however: familiarity refers to the fact that the system has knowledge of certain past visual input, while expectation implies that the system is making predictions about upcoming visual information. This focus on past or future input may indeed be a true distinction between familiarity and expectation, or it may only be a semantic difference. It remains unclear whether this distinction is actually implemented in neural processes or whether a single neural mechanism underlies both familiarity and expectation.

Conclusion

This study aimed to determine whether the effects of familiarity and expectation can be dissociated and, specifically, whether familiarity can influence brain activity even when stimuli cannot be predicted. We found that familiarity has a genuine effect on the amplitude of the electrophysiological signal in the human brain. Moreover, expectation also influences the amplitude of the neural response when familiarity is held constant. These findings give rise to numerous possibilities for future investigations. An important direction is to explore the corresponding brain activity when participants have expectations about upcoming images without being familiar with the stimuli. Another intriguing option is to investigate how the neural signal changes as the contingencies between stimuli change, so that images are not simply expected or unexpected but rather they can be expected to different degrees. Alternatively, the familiarity with the visual input can be graded, so stimuli can be familiar to different extents. These research trajectories can enhance our understanding of how familiarity and expectation influence the brain signal, so that ultimately we can discover whether these two concepts refer to the same or to distinct neural phenomena.

References

Baldassi, S. & Burr, D.C. (2000). Feature-based integration of orientation signals in visual search. *Vision Research, 40,* 1293–1300.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433-436.

Cameron, E. L., Tai, J., & Carrasco, M. (2002). Covert attention affects the psychometric function of contrast sensitivity. *Vision Research, 42,* 949–967.

Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. (2000). Spatial covert attention increases contrast sensitivity along the CSF: Support for signal enhancement. *Vision Research 40,* 1203–1215.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181-204.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology, 1*(1), 42-45.

Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex, 7*(2), 181-192.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Science, 10,* 14-23.

Jehee, J. F. M., Brady, D. K., & Tong, F. (2011). Attention improves encoding of task-relevant features in the human visual cortex. *The Journal of Neuroscience, 31*(22), 8210-8219.

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron, 75*(2), 265-270.

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences, 10* (11), 494-501.

Li, L., Miller, E. K., & Desimone, R. (1993). The representation of stimulus familiarity in anterior inferior temporal cortex. *Journal of Neurophysiology, 69*(6), 1918-1929.

Lu, Z.-L. & Dosher, B.A. (1998). External noise distinguishes attention mechanisms. *Vision Research, 38,* 1183–1198.

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

Meyer, T. & Olson, C. R. (2011). Statistical learning of visual transitions in monkey inferotemporal cortex. *Proceedings of the National Academy of Sciences of the USA, 108*(48), 19401-19406.

Meyer, T., Ramachandran, S., & Olson, C. R. (2014). Statistical learning of serial visual transitions by neurons in monkey inferotemporal cortex. *The Journal of Neuroscience, 34*(28), 9332-9337.

Meyer, T., Walker, C., Cho, R., & Olson, C. R. (2014). Image familiarization sharpens response dynamics of neurons in inferotemporal cortex. *Nature Neuroscience, 17*(10), 1388-1394.

Miller, E. K., Li, L., & Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science, 29,* 1377-1379.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology, 4*(2), 61-64.

Oostenveld R., Fries P., Maris E., & Schoffelen J.-M. (2010). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*, e156869.

Roisson, B., Schiltz, C., & Crommelinck, M. (2003). The functionally defined right occipital and fusiform "face areas" discriminate novel from visually familiar faces. *NeuroImage, 19*(3), 877-883.

Stolk, A., Todorovic, A., Schoffelen, J.-M., & Oostenveld, R. (2013). Online and offline tools for head movement compensation in MEG. *Neuroimage, 68*, 39-48.

Yeshurun, Y. & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature, 396,* 72–75.

The Gut-Brain Axis: Impact of a Probiotic Intervention on Neurocognitive Measures of Emotion and Cognitive Control

Franziska Michels[1]
Supervisor: Esther Aarts[1]

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

The human gut microbiome plays a key role in human well-being. Specifically, differences have been found between the gut microbiome compositions of healthy individuals compared with patient groups suffering from affective psychiatric disorders. Research has indicated that probiotics can have beneficial effects on the gut microbiome and its functioning as well as on the brain. In the present study we investigated the effects of a multispecies probiotic on neurocognitive measures of emotion and cognitive control. We were interested in whether the effects of a probiotic intervention can extend beyond effects on affective measures and associated brain regions, also including cognitive control (of emotion). Analyses were carried out on a first, small subset of participants (n = 11) who were divided into a placebo (n = 5) and a probiotic group (n = 6). In this study all participants were healthy women (mean age 23.3 years) with a body mass index (BMI) ranging from 19.2 – 24.8. The study consisted of two measurement points separated by a four-week probiotic vs. placebo intervention. In this preliminary study we did not find significant effects of the probiotic intervention on the brain. However, on exploratory thresholds, findings indicated an effect of the probiotic intervention on the amygdala during emotion regulation as well as on lateral frontal regions during general cognitive control processes, in line with our hypotheses. From these preliminary results we can conclude that probiotic effects tend to extend beyond modulating affective processes, also tending to affect prefrontal cortex and associated cognitive control processes.

Keywords: gut microbiome, affective psychiatric disorders, probiotic, amygdala, prefrontal cortex, hypothalamic-pituitary-adrenal (HPA) axis, inflammation

Corresponding author: Franziska Michels; **E-mail:** f.michels@donders.ru.nl

The human gut microbiome plays a key role in human well-being (Dinan, Stanton, & Cryan, 2013; Rook, Lowry, & Raison, 2013). The intestinal microbiota, collectively referred to as gut-microbiome, is a metabolic ecosystem consisting of microorganisms that outnumber the total amount of human cells in the body by far (Dinan et al., 2013). The gut microbiome is crucial for digestion and is involved in the development of the immune system, helping protect the body against pathogens (e.g., Shreiner, Kao, & Young, 2015; Smith, McCoy, & Macpherson, 2007; Wu & Wu, 2012). Additionally, it has a role in regulating endocrinological processes as well as gastrointestinal nerve activity (Allin, Nielsen, & Pedersen, 2015) and has been shown to influence several aspects of behaviour, stress, mood and cognition (Cryan & Dinan, 2012; Desbonnet, Clarke, Shanahan, Dinan, & Cryan, 2014; Mayer, 2011; Neufeld, Kang, Bienenstock, & Foster, 2011).

Bercik et al. (2011) provided evidence for the involvement of gut microbiota in mice's behaviour by showing that a short-term perturbation of the microbiota with an administration of antimicrobials[1] (ATM) increased exploratory behaviour in mice. Additionally, germ-free mice, ones without a gut microbiome or only a small amount of it therefore characterised by an undeveloped immune system, did not show altered behaviour in response to ATM. These findings indicate that the microbiome is necessary for these behavioural effects to take place rather than ATM working on the brain directly (Bercik et al., 2011). Furthermore, when mice showing timid behaviour received a gut microbiome transplant of mice that were relatively outgoing, a personality shift towards more outgoing behaviour could be observed (Bercik et al., 2011).

The gut microbiome has also been indicated to be crucial for stress regulation, since it is involved in the development of the hypothalamic-pituitary-adrenal (HPA) axis (i.e., an important stress regulation mechanism). Germ-free mice, for instance, displayed an increased HPA reaction to stress indicating a malfunction of the HPA axis and associated hormone secretion (Sudo et al., 2004). Additionally, early life stress has been linked to alterations in gut microbiome composition and was found to be a risk factor for the development of depression later in life (Juruena, 2014; O'Mahony et al., 2009). Excessive amounts of stress signals have generally been linked to various changes in the body, for instance altered hormone levels (Hargreaves, 1990) that can in turn lead to changes in neuroendocrine processes such as the HPA response to stress (e.g., Sudo et al., 2004). Additionally, inflammations can follow from stress signals that can, for example, take place in the gut. These inflammations are able to disrupt the functioning of the intestinal epithelial barrier. This barrier consists of a single layer of mucosa and restricts access to the gut to water, nutrients and electrolytes, preventing toxins and bacteria from entering. Its dysfunction, not only caused by stress but multiple factors including heredity, diet, exercise and drugs, can lead to increased amounts of lipopolysaccharides (LPS), parts of the cell membrane of gram-negative bacteria, entering the blood (Söderholm & Perdue, 2001; Santos et al. 2001; Van Hemert & Ormel, 2014). Big quantities of LPS can in turn provoke increased immune signalling causing new inflammatory reactions (Van Hemert & Ormel, 2014). Food allergies, diabetes, chronic fatigue and chronic intestinal disorders such as irritable bowel syndrome (IBS), associated with impaired epithelial barrier function, can follow from these inflammatory reactions (Dinan et al., 2013; Messaoudi et al., 2011).

The pro-inflammatory cytokines which are released by the immune system as a reaction to increased levels of LPS are also able to enter the central nervous system with the potential of influencing various processes in the brain by interacting with its cytokine network (Capuron & Miller, 2011). Increased levels of pro-inflammatory cytokines may, for instance, have effects on the neuroendocrine system in the context of depression; alterations such as stimulation of the corticotrophin-releasing hormone and thereby HPA activity, which is important for stress regulation (Miller, 1998), can occur. Therefore, a link between the gut microbiome and affective psychiatric disorders has gained recent interest. Naseribafrouei and colleagues (2014) have shown that healthy individuals exhibited a different gut microbiome composition relative to individuals suffering from depression. Additionally, differences in immune functioning as well as related increased levels of inflammation markers in the blood and brain have been linked to different psychiatric disorders (mainly depression) as well as to declines in cognitive functions such as working memory and learning performance (Capuron & Miller, 2011; Gimeno et al., 2009; Rook et al., 2013; Sparkman et al., 2006). The high co-morbidity between affective psychiatric symptoms (e.g., anxiety or depression) and various chronic intestinal disorders such as IBS provides additional evidence for a role of the intestinal microbiota in gut-brain communication (Bercik et al., 2011).

1 Agent that can kill microbiota.

Probiotics[2] are one kind of treatment indicated to have the potential of beneficially influencing gut-functioning and gut microbiome composition in animals and humans and has been found to reduce symptom severity in patients with (chronic) intestinal disorders (e.g., Kajander, Hatakka, Poussa, Färkkilä, & Korpela, 2005; Moayyedi et al., 2010). By means of probiotics, it was possible to investigate the role of the intestinal microbiota in the regulation of anxiety, mood, cognition, pain and behaviour in rodents (Chen, D'Souza, & Hong, 2013; Cryan & Dinan, 2012; Foster & McVey Neufeld, 2013). Probiotics have been shown to lower levels of systemic inflammatory cytokines and up-regulate plasma IL-10 levels, a cytokine suggested to have anti-inflammatory properties, in vivo (in mice) and in vitro. These two findings are of importance considering the link between elevated pro-inflammatory cytokine levels and depression (Ghosh, Van Heel, & Playford, 2004; Kopp et al., 2008; Logan & Katzman, 2005). In a study with human participants a reduction in sad mood by means of a self-reported questionnaire has been shown after a probiotic intervention (Steenbergen, Sellaro, Van Hemert, Bosch, & Colzato, 2015). Additionally, decreased activity in different affective (including amygdala and insula), viscerosensory and somatosensory brain regions in response to an emotional face matching paradigm after intake of a fermented milk product versus no intervention has been indicated in the first neuroimaging (i.e., functional Magnetic Resonance Imaging [fMRI]) study in this field (Tillisch et al., 2013). Overall, existing evidence suggests various positive effects of probiotics on the gut and brain, yet without uncovering the neurocognitive mechanisms underlying these effects. The study by Tillisch and colleagues was a first attempt to investigate probiotic effects on the brain in humans. Nonetheless, due to a number of limitations — very small group sizes (ranging from 10 to 12 subjects) and specific findings showing effects only in comparison with a no-intervention group instead of placebo, results should be interpreted with caution.

To investigate the effects of a multispecies (i.e., consisting of multiple strains of bacteria) probiotic and to unravel its underlying neurocognitive mechanisms (i.e., the gut-brain mechanism) with respect to emotion processing, emotion regulation and cognitive control processes, an extended design was used in the present study. Specifically, we added two extra paradigms in addition to the one used by Tillisch et al. (2013). The design was extended since affective psychiatric disorders are not solely characterised by differences in emotion processing but also by emotion regulation and associated cognitive control processes in which the prefrontal cortex (PFC) plays an important role (Joormann & Gotlib, 2010). In a study by Bishop et al. (2004) for instance, decreased activity in PFC regions has been found in patients suffering from anxiety in a task used to study processing of threat-related distractors. These findings indicate a dysfunction of cognitive control in the context of threat in this patient population. In line with this finding, functional connectivity between PFC and amygdala has been found to increase when healthy participants were presented with unexpected threat-stimuli, which suggests an important role of the PFC to control amygdala in order to maintain goal-directed behaviour in the context of threat (Gold, Morey, & McCarthy, 2015). Additionally, since the effects of a probiotic specifically designed to improve epithelial barrier function were investigated in the current study, it is unlikely that effects are being specific to emotion processing alone since immunological mechanisms (here one of the mechanisms assumed to underlie the effects) can affect the whole system including various brain regions (Van Hemert & Ormel, 2014). For instance, cytokines, which are important in immune defense, have been shown to be able to act on the central nervous system with diverse consequences such as changes in gastric function, induction of fever, increased metabolism as well as changes in behaviour (Rothwell & Hopkins, 1995). Furthermore, as inflammations can lead to a decline in cognitive control functions such as working memory (see Sparkman et al., 2006), probiotics that improve the intestinal barrier and thereby presumably decrease inflammation, might also increase such cognitive control functions by acting on brain regions other than those involved in emotion.

In our design we included the emotional face matching paradigm (see methods for detailed task descriptions) as used by Tillisch et al. (2013) in order to study effects of the probiotic on affective brain regions (e.g., amygdala) involved in emotion processing. Furthermore, an emotional face Stroop paradigm as used by Etkin, Egner, Peraza, Kandel, & Hirsch (2006) has been chosen to investigate intervention-induced effects on emotion regulation. This paradigm has been shown to capture activation of PFC (medial and lateral PFC [mPFC/lPFC], including supplementary motor areas) and affective

2 Probiotics are defined by the World Health Organisation as "live micro-organisms that, when administered in adequate amounts, confer a health benefit on the host.

regions (e.g., amygdala) in particular (Etkin, Egner, & Kalisch, 2011). In addition, investigating intervention-induced effects on cognitive control processes in the absence of emotion, widely associated with dorso-medial (dm) frontal regions (e.g., anterior cingulate cortex [ACC] and [pre-] supplementary motor area [SMA]) and lateral PFC (e.g., inferior frontal gyrus [IFG] and middle frontal gyrus [MFG]) was implemented by means of a classic colour-word Stroop task (Courtney, Petit, Haxby, & Ungerleider, 1998; Stroop, 1935).

In the present study we put forth several predictions: based on the findings by Tillisch et al. (2013) we expected to find (1) intervention-induced *decreases* in the activity of affective brain regions (e.g., amygdala) in response to the emotional face matching paradigm. In addition, we predicted to find (2) intervention-induced increases in PFC activity, especially ventro-medial (vm) and dorso-medial (dm) PFC (including [pre-]SMA) in response to the emotional face Stroop paradigm along with (3) *decreased* activity in affective brain regions (e.g., amygdala). Beyond that we predicted (4) an increase in the connectivity between prefrontal and affective cortices in this task, suggesting intervention-induced enhances in cognitive control over affective cortices (not further described here). In the classic colour-word Stroop task we also expected to find (5) intervention-induced increases in PFC activity, especially dmPFC and lPFC, suggesting enhancement of cognitive control processes in absence of emotion.

This part of the study includes only a small, first subsample (n = 11) of the data (the final sample will include sixty participants) in order to test the validity of the paradigms used in the study and to explore some of the hypotheses stated above.

Methods

Participants

In total, 11 of the planned 60 participants, randomly (double-blind) divided into the probiotic intervention group (n = 6, age M = 21.3 years, SD = 2.3) and the placebo group (n = 5, age M = 25.6 years, SD = 3.8) were analysed in this study. All participants were healthy women (age M = 23.3 years, SD = 3.6), with a body mass index (BMI) ranging from 19.2 – 24.8 (M = 22.0, SD = 2.0), a range considered to be healthy taking age and gender into account, both in the probiotic (BMI M = 21.1, SD = 1.4) and the placebo group (BMI M = 23.1, SD = 2.2). Except for two participants who graduated from university already, all participants were university students. All participants took hormonal contraceptives and were not in the stop week during test sessions to ensure similar hormone levels between both sessions across participants. They were screened for medical conditions (including neurological, psychiatric, gastrointestinal or endocrine disorders) and relevant medical history. Furthermore, participants were screened for MRI compatibility, probiotics and prebiotic use, diet, alcohol and smoking behaviour. In order to ensure good task comprehension and clear understanding of the neuropsychological questionnaires (not described here), all participants exhibited sufficient knowledge of Dutch. The study was conducted following the Declaration of Helsinki with human subjects and the complete procedure was approved by the local Ethics Committee CMO Arnhem-Nijmegen. Written informed consent was obtained from each participant.

Procedure

We employed a double-blind, randomised, placebo-controlled, between-subject design. The study consisted of two sessions separated approximately five weeks in time. During four (28 days) of the five weeks an intervention consisting of daily probiotic or placebo intake was implemented. Both sessions of the experiment were conducted at the Donders Centre for Cognitive Neuroimaging in Nijmegen, the Netherlands. Feces samples were taken, one before the start of the intervention and one after taking the last probiotic/placebo (within a time window of approximately 29 – 40 days). At the beginning of the first test session, the experimental procedure was explained and informed consent was obtained from the subjects. Physical measurements including height, weight, waist circumference and blood pressure were taken. Participants practised the tasks that were performed in the scanner at a later stage and were asked to fill out different questionnaires (not described here). An MRI part of 75 minutes including acquisition of anatomical, functional and resting state images followed during which participants had to perform the tasks they had practised earlier (starting with an emotional face matching paradigm, followed by an emotional Stroop task and ending with a classical colour-word Stroop task). The MRI part was followed by another session outside of the scanner consisting of neuropsychological and dietary questionnaires participants had to fill in as well as a stress test (not reported here). At the end of the first test session,

subjects were provided with the probiotics/placebo and instructions on how to take it as well as with the toolkit for the feces samples. After the intervention, participants came to the centre for the second test session, which resembled the first one.

fMRI tasks

In this study, three tasks were chosen to be performed in the MRI scanner, including the emotional face matching paradigm, an emotional Stroop task and a classic colour-word Stroop paradigm. The experiment was performed using Presentation® software (Version 0.70, www.neurobs.com). Trial sequences of each task were pseudo-randomised in order to guarantee equal numbers of presentations of each stimulus type. A different version of each task was performed in the first and the second test session; the order was counterbalanced.

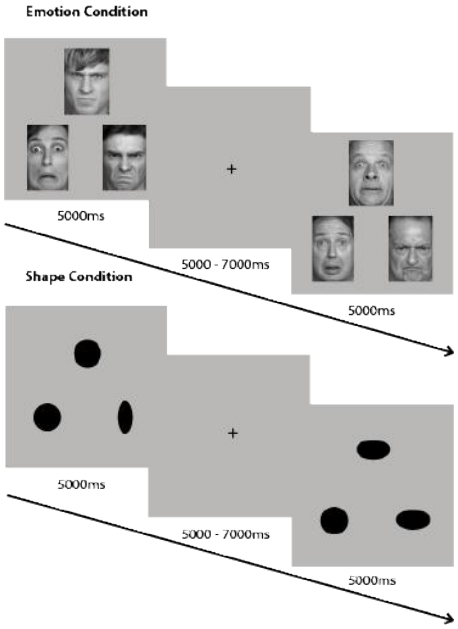


Fig. 1. Emotional face matching paradigm: example stimuli.

Emotional face matching paradigm. This paradigm was chosen to investigate intervention-induced changes in emotional processing or 'reactivity' (Hariri, Bookheimer, & Mazziotta, 2000). A block design was used for this task with a total of 18 blocks consisting of three stimuli each. The task included two different conditions, a control and an emotion condition. In the control condition subjects had to match one of two geometric shapes presented at the bottom to a target shape presented at the top of the screen. The experimental condition involved subjects choosing one of two emotional (angry or fearful) faces presented at the bottom of the screen as best matched the emotional expression of a face seen at the top of the screen (see Fig. 1 for example). The condition was kept constant over a block duration of 17 seconds, but was randomised between blocks. Participants were asked to react as fast and accurately as possible. The total duration of the task amounted to seven minutes.

Emotional face Stroop paradigm. A Dutch version of the emotional face Stroop task (Etkin et al., 2006) was used to assess intervention-induced differences in cognitive control in the face of emotional distractors. During this task, participants were presented with pictures of male faces expressing fear or happiness. On top of the faces, the Dutch words for happy (i.e., *blij*) and fearful (i.e., *bang*) were presented in prominent red letters (see Fig. 2 for example). The emotions described by the words were either congruent with the emotion of the face or incongruent and subjects had to indicate the

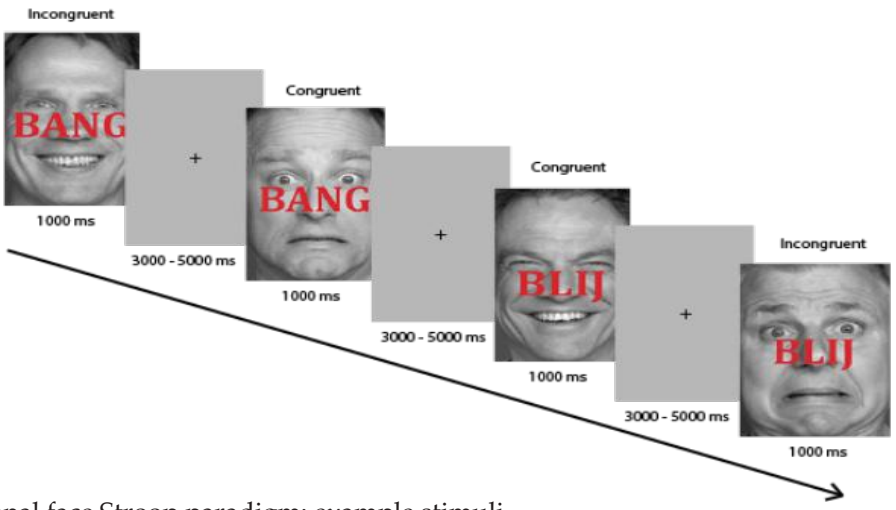


Fig. 2. Emotional face Stroop paradigm: example stimuli.

emotion of the face ignoring the emotion word. In total, stimulus presentations amounted to 148 presentations of happy or fearful male faces. The order of stimulus presentation was pseudo-randomised. The total duration of the task added up to 15 minutes.

Classic colour-word Stroop paradigm. A Dutch version of the classic colour-word Stroop task (Stroop, 1935) was used to assess intervention-induced differences in general cognitive control in absence of emotional stimuli. During this task, participants were presented with four different colour words written either in the same ink colour as the word (e.g., red written in red ink) or in an incongruent colour (e.g., red written in blue ink, see Fig. 3 for example). The task was to indicate the ink colour of the word by pressing a button mapped to that colour, and ignore the word meaning. The task consisted of 80 stimulus presentations in total. As in the other tasks, participants were asked to react as fast and accurately as possible. Colour-button mappings were randomised across subjects, but kept constant between the two sessions of each subject. The total duration of the task amounted to approximately 10 minutes, depending on participants' performances on the practice trials.

MRI data acquisition

MRI data were acquired using a 3T MAGNETOM Prisma system, equipped with a 32-channel head coil. During the three tasks, 3D echo planar imaging (EPI) scans (using a T2*-weighted gradient echo multi-echo Echo Planar Imaging (EPI) sequence [Poser, Versluis, Hoogduin, & Norris, 2006]) were acquired (voxel size = 3.5 × 3.5 × 3 mm isotropic; repetition time [TR] = 2070 ms; echo time [TE] = 9 ms, 19.25 ms, 29.5 ms, 39.75 ms; field of view [FoV] = 224 mm). The slab positioning and rotation (average angle of 14 degrees to AC axis) optimally covered both prefrontal and deep brain regions (i.e., including affective brain regions such as amygdala). During the tasks, thirty dummy volumes were discarded immediately before the main scan to allow for the weight calculations of the four echoes used for image reconstruction. The total scan duration was about 75 minutes. Whole-brain high-resolution T1-weighted anatomical scans were acquired using an MPRAGE sequence (voxel size = 1.0 × 1.0 × 1.0 mm isotropic, TR = 2300 ms, TE = 3.03 ms, 192 slices).

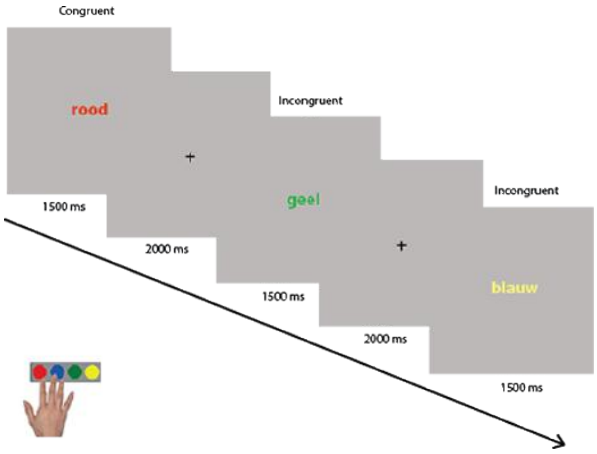


Fig. 3. Classic colour-word Stroop paradigm: example stimuli.

fMRI data preprocessing

Processing of the data was implemented using Statistical Parametric Mapping (SPM8, Wellcome Department of Imaging Neuroscience, London). Volumes for each echo-time were realigned using six rigid body spatial transformations (translations and rotations). Thirty volumes acquired before the tasks were used to combine the four echo images into a single MRI volume using an echo weighting method known as PAID-weighting (Poser et al., 2006). Resulting combined functional (EPI) images were slice-time corrected by realigning the time series for each voxel to the time of acquisition of the reference slice (here slice 17). Subject-specific structural and functional data were subsequently co-registered to a standard structural or functional stereotactic space respectively, to Montreal Neurological Institute (MNI) templates. A unified segmentation approach was then used to segment the structural images, which were subsequently spatially co-registered to the mean of the functional images. The transformation matrix resulting from the segmentation step was used to normalise the structural and functional images to MNI space, resampled at a voxel size of 2 × 2 × 2 mm. In a final step, normalised functional images were spatially smoothed using an 8-mm full-width at half maximum (FWHM) Gaussian kernel.

Statistical analyses

First level fixed effects analyses of fMRI data were performed using an event-related approach for both Stroop paradigms. The statistical model for event-related fixed effects analyses contained two

regressors of interest for the classic colour-word Stroop paradigm representing the different task conditions (correct incongruent and congruent trials) and four regressors of interest for the emotional face Stroop paradigm (incongruent trials followed by congruent trials, congruent trials followed by incongruent trials, congruent trials followed by congruent trials and incongruent trials followed by incongruent trials). Miss and incorrect trials were taken into account in a regressor of no interest for both of these paradigms. First level analyses of the emotional face matching paradigm were carried out using a block-design fMRI approach with block duration of 17 seconds. Onsets of the independent regressors for the event-related Stroop paradigms were modelled as a stick function convolved with the canonical hemodynamic response function (HRF) (Friston et al., 1998). Additionally, twelve regressors of no interest were added in order to account for motion artifacts consisting of twelve rigid-body transformation parameters (i.e., movement regressors consisting of three translations and rotations and their linear derivatives) obtained during realignment. A high-pass filter with a cut-off of 128 seconds was applied to the time-series of the functional images to remove low-frequency drifts. By applying an autoregressive AR(1) model, correction for serial correlations was carried out.

Three GLMs were run as random effect second level analyses based on the different contrast images of the contrasts applied in the first level analyses of each of the three tasks. For the emotional face Stroop, a GLM with the contrast images of adaptation minus non-adaptation ([incongruent – congruent, congruent – incongruent] > [congruent – congruent, incongruent – incongruent]) was run. For the classic colour-word Stroop paradigm, incongruent minus congruent contrast images were used (incongruent > congruent) for the GLM and the GLM for the emotional face matching paradigm was based on the contrast images of emotion minus shape (emotion > shape). Analysis of variance (ANOVA) was performed in a full-factorial design with the above-specified contrast images from first level analyses and two additional factors were added at second level analyses, 'Group' (probiotic, placebo) as a between-subject factor and 'Time' (pre- or post-intervention) as a within-subject factor.

In the presented figures, results are displayed at exploratory thresholds of $p < .001$ (uncorrected) and $p < .005$ (uncorrected). Whole-brain corrected results at p *corr (FWE)* $< .05$ (cluster-level, with intensity threshold $p < .001$) are reported in the tables. MarsBaR was used to extract regionally-averaged beta weights at $p < .001$ (uncorrected) and $p < .005$ (uncorrected) for two brain regions in the Stroop paradigms for illustrative purposes.

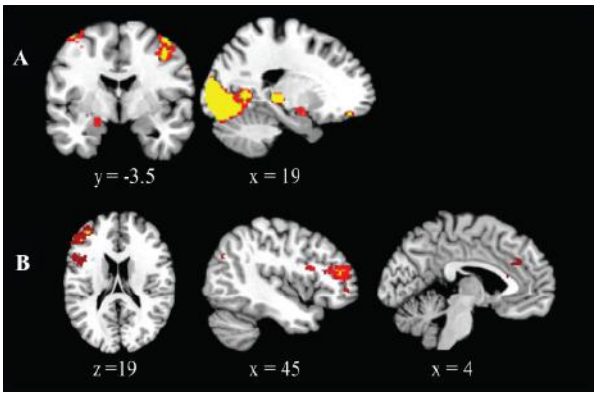


Fig. 4. Emotional face matching: positive task effects and effects of probiotic intervention on brain activity. **A.*** Positive task effects of affective matching (emotion > shape). **B.*** Effects of the probiotic intervention on the brain (Probiotic > Placebo, Pre > Post). Here regions are displayed that showed more activity for the probiotic group relative to the placebo group, in the pre- compared with the post-session of the study.
* Results are displayed at exploratory thresholds in *yellow* at $p < .001$ (uncorrected) and in *red* at $p < .005$ (uncorrected). Images are shown in radiological orientation, left = right (MNI coordinates).

Results

Imaging data

Emotional face matching paradigm. A main task effect of emotion processing, matching affect (emotion > shape) was shown in various brain regions including occipital, temporal and frontal regions as displayed in Table 1. On an exploratory threshold of $p < .005$ (uncorrected) amygdala activation could be observed as well (see Fig. 4A). We did not find a significant effect of the probiotic (Probiotic > Placebo, Pre > Post) on brain regions during emotional face matching at the stringent threshold of p *corr (FEW)* $< .05$ for this paradigm. Even on exploratory thresholds of $p < .001$ (uncorrected) and $p < .005$ (uncorrected) we did not observe significant amygdala deactivation, yet a deactivation of a medial frontal region ($p < .005$, uncorrected) as well as IFG could be observed when the probiotic group was compared with the placebo group (Pre > Post, $p < .001$, uncorrected) (see Fig. 4B).

Table 1.

Main task effect of emotion processing (emotion – shape) and (shape – emotion). Clusters showing greater BOLD* activity for the match emotion condition compared to the match shape condition, (whole brain corrected *pFWE* < .05). MNI stereotactic coordinates of local BOLD maxima.

Region	cluster *p* (FEW-corrected)	cluster equivk	peak T	x, y, z {mm}
Emotion processing				
(emotion > shape)				
Right inferior occipital gyrus	.000	13661	26.00	44, -80, -10
Left inferior occipital gyrus			20.74	-18, -98, -8
Right cuneus			19.42	16, -96, 6
Right hippocampus	.001	489	8.19	22, -30, -2
Left thalamus			7.55	-20, -28, -2
Midbrain			4.41	-4, -32, -4
Right inferior frontal gyrus pars opercularis	.000	746	6.37	38, 14, 26
Right middle frontal gyrus			5.40	44, 32, 18
Right inferior frontal gyrus pars triangularis			5.21	48, 14, 24
Right superior temporal gyrus	.004	363	5.50	50, -40, 14
Right middle temporal gyrus			4.28	50, -48, 6
Left inferior frontal gyrus pars triangularis	.021	248	4.56	-44, 14, 26
Left inferior frontal gyrus pars triangularis			4.01	-44, 20, 18
Right pre-supplementary motor area	.038	212	4.55	4, 16, 52
Left pre -supplementary motor area			3.91	-2, 24, 48
(shape > emotion)				
No significant clusters				

*BOLD = Blood oxygenated level dependent

Emotional face Stroop paradigm. Main task effects of emotional Stroop adaptation revealed significant medial and lateral frontal cortex activation as expected at *p corr (FEW)* < .05 (see Table 2, Fig. 5A). Additionally, on an exploratory threshold of $p < .005$ (uncorrected) amygdala deactivation could be observed. We did not observe any significant effects of the probiotic intervention on brain regions during emotional Stroop adaptation. Yet at an exploratory threshold of $p < .005$ (uncorrected),

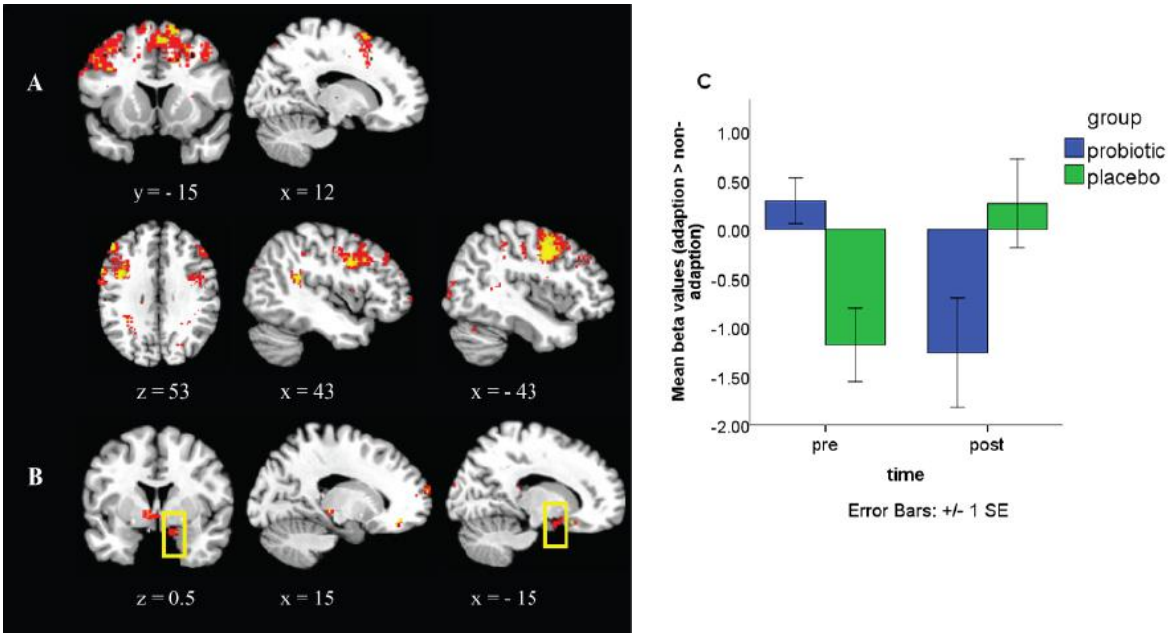


Fig. 5. Emotional face Stroop: positive task effects and effects of probiotic intervention on brain activity. **A.*** Positive task effects of emotional Stroop adaptation (adaptation > non-adaptation). Here, brain regions are displayed showing more activity during adaptation compared with non-adaptation trials. **B.*** Effects of probiotic intervention on the brain during emotional Stroop adaptation (adaptation > non-adaptation) (Probiotic > Placebo, Pre > Post). Here, brain regions are displayed showing more activity for the probiotic relative to the placebo group in the pre- compared with the post-session.
C. Extracted mean beta values (adaptation > non-adaptation) from an ROI created for the amygdala (MNI coordinates: -16 4 -22) based on the functional images received from the Group × Time interaction contrast on (Probiotic > Placebo, Pre > Post) at $p < .001$ (uncorrected).
* Results are displayed at exploratory thresholds in *yellow* at $p < .001$ (uncorrected) and in *red* at $p < .005$ (uncorrected). Images are shown in radiological orientation, left = right (MNI coordinates).

regions in the frontal cortex including vmPFC were shown to be *deactivated* for the probiotic versus placebo group after the intervention, contrary to our hypothesis. Minor amygdala *deactivation* could be observed as well ($p < .001$, uncorrected; more clearly at $p < .005$, uncorrected) when groups were compared after intervention, being in line with our hypothesis (see Fig. 5B). For illustrative purposes we created a region of interest (ROI) for this region of activation (see Fig. 5C) from functional images received from the Group × Time interaction contrast (Probiotics > Placebo, Pre > Post; $p < .001$, uncorrected). Mean beta values (adaptation > non-adaptation) were extracted showing a *decrease* in amygdala activity for the probiotic group (Pre > Post), as expected. The opposite could be observed for the placebo group.

Classic colour-word Stroop paradigm. Main task effects of the Stroop effect revealed significant activations for regions in left lateralised frontal regions (see Table 3 and Fig. 6A). We did not observe significant effects of the probiotic intervention on brain regions during this paradigm, however at an exploratory threshold of $p < .005$ (uncorrected), activation of lPFC (i.e., IFG) for the probiotic group compared with the placebo group (Post > Pre) could be observed in line with our hypothesis. Additionally, we found vmPFC activity for this contrast (see Fig. 6B). For illustrative purposes we created an ROI for the area in the lPFC (see Fig. 6C) based on the functional images received from the Group × Time interaction contrast (Probiotic > Placebo, Post > Pre; $p < .005$, uncorrected). Mean beta values were extracted showing an increase in activity for the probiotic group, as expected (Post > Pre). The placebo group showed effects in the opposite direction.

Discussion

In the present study our aim was to investigate the effects of a multispecies probiotic on brain functioning and its underlying mechanisms. We were particularly interested in whether the probiotic effects would extend beyond emotion processing and associated brain regions, also affecting brain regions involved in emotion regulation and general cognitive control processes.

Table 2.

Main task effects of emotional Stroop adaptation (adaptation – non-adaptation) and (non-adaptation – adaptation). Clusters showing greater BOLD* activity for adaptation trials compared to non-adaptation trials (whole brain corrected $p_{FWE} < .05$). MNI stereotactic coordinates of local BOLD maxima.

Region	cluster *p* (FWE-corrected)	cluster equivk	peak T	x, y, z {mm}
Emotional Stroop adaptation				
(adaption > non-adaptation)				
Left precentral gyrus	.000	441	6.24	-38, -2, 44
Left precentral gyrus			5.96	-46, -2, 38
Left precentral gyrus			5.25	-38, 0, 36
Right precentral gyrus	.001	189	5.30	42, 0, 32
Right inferior frontal gyrus pars opercularis			5.28	38, 16, 32
Right inferior frontal gyrus pars opercularis			4.31	46, 14, 34
Right pre-supplementary motor area	.002	167	4.93	6, 4, 68
Left pre-supplementary motor area			4.47	0, 8, 54
Left pre-supplementary motor area			4.36	-8, 16, 54
(non-adaptation > adaptation)				
No significant clusters				

*BOLD = Blood oxygenated level dependent

Main effects of tasks

Despite the small sample size studied here, we found sufficient task-related brain activations in line with findings of earlier studies, thus we can confidently conclude that the tasks used in our design functioned as expected. This is not very surprising since we used three robust paradigms in order to ensure that we could measure the effects we were interested in. The significant activation of thalamus during the emotional face matching task for instance, a region known to influence amygdala activity, is in line with findings by Hariri et al. (2000). Additionally, activation in bilateral inferior occipital gyri (IOG), a region shown to be involved in face processing, also indicated that this task worked well since it was found when brain activation for matching of emotional faces was compared with matching of shapes (Haxby, Hoffman, & Gobbini, 2000). Significant task-related brain activations in the emotional face Stroop task in medial and later PFC, particularly right IFG and bilateral pre-SMA, as well as *deactivation* of the vmPFC when adaptation trials were compared with non-adaptation trials were also in line with earlier studies using similar paradigms (Roberts & Hall, 2008; Etkin et al., 2006). Finally, the classic colour-word Stroop task activated medial and lateral frontal areas as well, here also including left pre-SMA and left IFG, regions often found to be activated during (Stroop) conflict paradigms (e.g., Roberts & Hall, 2008; Zoccatelli, Beltramello, Alessandrini, Pizzini, & Tassinari, 2010).

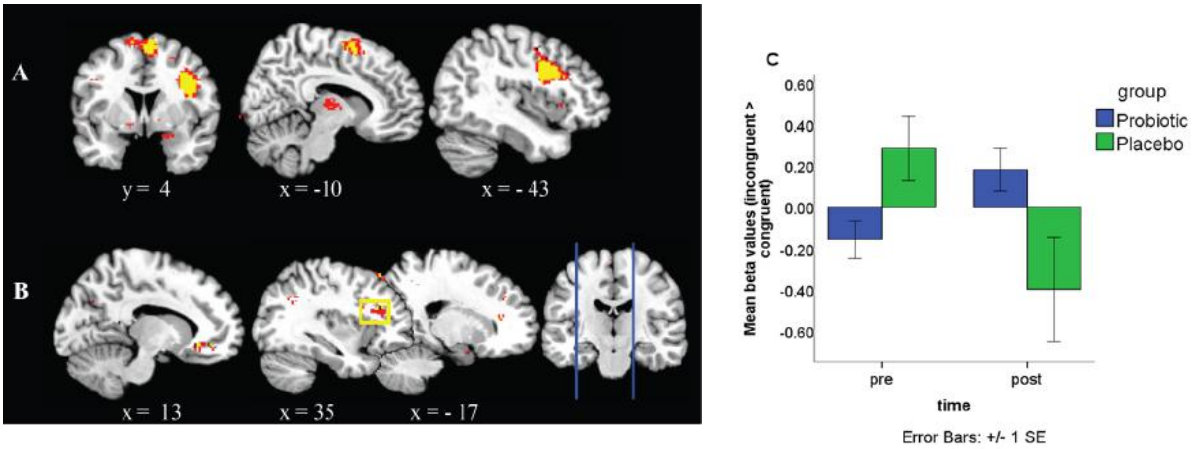


Fig. 6. Classic colour-word Stroop: positive task effects and effects of probiotic intervention on brain activity. **A.*** Positive task effects of Stroop effect (incongruent > congruent). Here, brain regions are displayed showing more activation during incongruent versus congruent trials. **B.*** Effects of probiotic intervention on the brain during Stroop effect (incongruent > congruent) (Probiotic > Placebo, Post > Pre). Here, brain regions are displayed showing more activation for the probiotic relative to the placebo group in the post- compared with the pre-session.
C. Extracted mean beta values (incongruent > congruent) from an ROI created for a region in the LPFC (MNI coordinates: 34 26 20) based on the functional images received from the Group × Time interaction contrast on (Probiotic > Placebo, Post > Pre) at $p < .005$ (uncorrected).
* Results are displayed at exploratory thresholds in *yellow* at $p < .001$ (uncorrected) and in *red* at $p < .005$ (uncorrected). Images are shown in radiological orientation, left = right (MNI coordinates).

Probiotic intervention effects

The probiotic intervention did not show significant effects on brain regions during the utilised paradigms, which is not surprising considering the very small sample size of $n = 6$ in the probiotic group and $n = 5$ in the placebo group. Yet, with exploratory, uncorrected thresholds indications of probiotic effects on brain activation in all three tasks could be seen of which some were in line with our hypotheses. Analyses showed an intervention-induced decrease in amygdala activity for the probiotic group compared with the control group only for the emotional face Stroop paradigm. Earlier studies have found the amygdala to be crucial for experiencing emotional conflict (Etkin et al., 2006). This preliminary finding could thus indicate that participants in the probiotic group experienced less emotional conflict in this task after the intervention compared with the placebo group. Additionally, this indication of *decreased* amygdala activity fits with what Tillisch et al. (2013) described in their results. However, they found this *deactivation* only in comparison with a non-intervention group than in comparison between the experimental group and the control group. This result might be due to the small group sizes investigated in their study or could indicate effects of the fermented milk product (here delivery vehicle) itself. In case our preliminary finding reaches significance with the final sample in our study, we could provide evidence for effects of probiotics on this affective brain region, unlikely induced by the delivery vehicle of the probiotics since here the probiotic and placebo group were compared.

In addition to decreased activity in the amygdala we expected to find enhanced activations in lPFC and mPFC areas during both Stroop paradigms due to the probiotic intervention. However, intervention-induced enhancements in lateral frontal activation, here left IFG, could be observed only for the classic colour-word Stroop paradigm. The IFG has been shown to be involved in cognitive control processes and to be important for inhibitory control (e.g., Swick, Ashley, & Turken, 2008). This preliminary finding could thus indicate an improvement of inhibitory control in the probiotic group by means of the probiotic intervention. Intervention-induced enhancement of vmPFC activation was also found in this paradigm for the probiotic group compared with the placebo group after the intervention. This finding is not in line with our hypotheses since this region has been shown to be more involved in cognitive control processes involving emotion whereas this paradigm does not include emotion stimuli (e.g., Winecoff et al., 2013). Nevertheless, in addition to many studies indicating its importance with respect to decisions involving

emotions or varying degrees of (un)certainty, it has also been indicated that this region is crucial for general decision making (Fellows & Farah, 2007). Therefore, these preliminary findings could hint towards an improvement of general decision making, for instance fewer impulsive decisions due to the probiotic-intervention. Intervention-induced *deactivations*, rather than the expected enhancement of vmPFC, were found in the emotional face Stroop paradigm as well as the emotional face matching paradigm; for the latter we did not predict any probiotic effects on prefrontal regions. Even though these preliminary results are not in line with our hypotheses, all regions found here were close to or overlapping with the regions activated during main task effects. This overlap indicates an intervention-induced *decrease* of activity of task-related regions during emotion processing and regulation.

Nevertheless, these preliminary brain-related findings, whether unexpected or in line with our hypothesis, need to be interpreted with caution due to the small sample size studied here, which is underlined by the findings of ROI beta value extraction. The mean beta values extracted from the ROIs created from the classic colour-word Stroop (lPFC) and the emotional face Stroop paradigm (amygdala) showed effects in line with our hypothesis for the probiotic group for both brain regions. Differences in activation in these regions could be seen when the baseline activation was compared to the activation at the post-session for the two groups, revealing group differences. However, group differences at baseline could also be observed. It is thus difficult to interpret these preliminary findings. With the final sample, unbiased statistical ROI analyses wherein ROIs are created based on anatomical images, will be performed in order to see by which group and time point the effects are driven.

Underlying gut-brain mechanisms

As we had a small sample size of eleven participants and therefore examined at exploratory, uncorrected thresholds, it is difficult to draw significant conclusions about underlying neurocognitive mechanisms in this study. With the complete sample of participants we hope to find effects of the probiotic on brain regions extending beyond emotion regions, which have already been indicated with this small sample size. In case these findings reach significance in the final sample it would be plausible to suggest a general gut-brain mechanism to underlie the effects such as immunological mechanisms or the metabolic pathway (described below). The immunological mechanisms (as described in the introduction) may lead to rather whole brain than very specific effects since this mechanism is involved in various processes in the body. Additionally, gut-bacteria can produce metabolites (metabolic pathway) that can enter the bloodstream thereby affecting the local enteric nervous system as well as the central nervous system, which also suggests whole brain effects. Interactions of neurochemicals between the central nervous system and gut microbiota have been shown to be bidirectional (e.g., Lyte, 2014). The produced metabolites can be precursors of neurotransmitters or can affect those travelling through the bloodstream and cross the blood-brain barrier to affect neurotransmitter synthesis in the brain (Collins, Surette, & Bercik, 2012). Some bacterial strains have been indicated to produce, for instance gamma-Aminobutyric acid (GABA) or tryptophan, a precursor of serotonin (O'Mahony, Clarke, Borre, Dinan, & Cryan, 2015). Tryptophan, for example, can affect various brain regions as serotonergic projections can be found throughout the brain, indicating that in case this mechanism is one to play a role here, effects would concern whole brain effects rather than simply affecting specific brain regions (e.g., Charnay & Leger, 2010).

Additionally, gut microbiota can produce hormones and regulate their secretion (Neuman, Debelius, Knight, & Koren, 2015), which is important concerning stress regulation, which has been shown to be malfunctioning in germ-free mice (Sudo et al., 2004). The immunological mechanism is tightly coupled with other mechanisms and systems in the body such as the metabolic mechanism and endocrine system. Through the regulation of hormone levels the gut-bacteria can, for instance affect the immune system (Neuman et al., 2015). In order to find out which one of these mechanisms plays a significant role in the current study, or whether they interact, it is necessary to include blood samples of participants in future studies. These samples can be used to measure the amount of inflammation markers in the blood of subjects providing a more direct measurement of the mechanism via which the probiotic changes in gut microbiome might affect cognitive functions. In case results concerning the PFC disappear with the final sample and effects are specific to affective brain regions, we might have reason to conclude another gut-brain mechanism to underlie the findings. In a study by Cryan and O'Mahony (2011), anxiolytic and antidepressant effects of a bacterium in mice were prevented when animals underwent vagotomy, suggesting a role

of the vagus nerve in gut-brain communication. The vagus nerve transmits signals from the enteric nervous system to the central nervous system. Its efferent pathway is crucial for the regulation of a number of cytokines in response to stress signals in the gut, suggesting a certain degree of overlap between the vagus nerve and the immune system (Sherman, Zaghouani, & Niklas, 2015). The vagus nerve sends signals about sensations occurring in the gut back to the brain providing the nucleus tractus sollitarius with gut-related information. Subsequently the information is transmitted to the parabrachial nucleus, which is connected to various brain regions (including insular, hypothalamus and amygdala) (King, 2007; Mayer, 2011) amongst others involved in emotion processing and regulation and maintenance of bodily homeostasis. Thus, in case the vagus nerve may play a major role in gut-brain communication here, we would expect more specific, direct effects on brain areas involved in emotion processing and regulation (rather than whole brain effects, including the PFC). Nevertheless, it is likely that these different mechanisms work together to a certain degree rather than functioning completely independently.

Limitations, strengths, and future directions

For future studies, it might be interesting to include a vagus nerve intervention. Current research is intended to test effects of blocking the vagus nerve in patients with obesity (e.g., Shikora et al., 2015). In a future study the vagus nerve could be blocked for a specific period of time for one group of participants, thereby adding a direct measurement of vagal nerve contribution to gut-brain communication. However, so far there are only invasive techniques available to block or stimulate the vagus nerve, which is not ethically appropriate for the present study. However, it is also possible to make predictions about vagal activity by means of a person's heart rate variability (HRV), more specifically the respiratory sinus arrhythmia (RSA) – a measurement of heart rate change in response to inhalation and exhalation – which was first proposed to be associated with vagal activity by Hering (1910) (as stated by Berntson et al., 1997). If the vagus nerve plays a role in gut-brain communication, changes in vagal tone should be observable (Alcock, Maley, & Aktipis, 2014).

In this preliminary data set one of the biggest limitations is the small sample size per group, which limits the amount of meaningful analyses that can be applied and decreases the power of related results. The final study will include 60 participants, 30 per group, which will lead to more reliable results and presumably less noise in the data. Additionally, due to time limits in the present study trials in the classic colour-word Stroop paradigm were evenly split into congruent and incongruent ones. However, Stroop interference has been shown to be stronger when overall proportion congruency is higher, that is, when the paradigm consists of proportionately more congruent than incongruent trials, which was seen in several studies (e.g., Kane & Engle, 2003; Logan & Zbrodoff, 1979). In future studies, additional congruent trials could thus be added in order to increase the Stroop interference and related brain activations. Furthermore, we aim at performing additional analyses with the complete sample such as brain-behaviour correlation analyses, in order to receive a more detailed picture of the data. In order to investigate whether the probiotic intervention affected affective brain regions directly or rather indirectly via the PFC, connectivity analyses between mPFC regions and the amygdala will be carried out for this paradigm with the complete sample of 60 participants.

A future strength of this study concerns the collected feces samples from participants. With the final sample we will apply analyses to the feces, enabling us to look for different bacteria strains in the sample. By means of a database it is then possible to find out which metabolites these bacteria can produce.

Conclusion

Taken together, the present study investigated the effects of a multispecies probiotic in a first, small subsample of participants. We can conclude that our tasks functioned as expected, although they did not show complete congruency with findings of earlier studies. Additionally, we could observe that indications of probiotic effects extended beyond affective brain regions as was expected. These effects included *decreased* vmPFC activity during emotion regulation, and more importantly, lPFC activity enhancement during a 'pure' non-emotional cognitive control task.

Clinical relevance

This study is of high clinical relevance since it aimed at finding effects of a probiotic on brain regions such as the amygdala involved in a variety

of affective psychiatric disorders (e.g., Peluso et al., 2009). Our preliminary results indicate a reduction of amygdala activity due to the probiotics in one of the tasks, which is of specific relevance for depression as depression disorder is often associated with hyperactivation of the amygdala (Peluso et al., 2009). If this hyperactivation could be reduced by means of probiotics it might lead to a reduction in symptom severity. Thereby it might represent a potential new treatment for affective psychiatric disorders since gut microbiota have been suggested to play a crucial role in these kinds of disorders. Patients with depression disorder have been shown to have increased levels of antibodies in the blood that are secreted as immune response against LPS, suggesting a dysfunction of the epithelial barrier (Maes, Kubera, & Leunis, 2008). Restoring the epithelial barrier function might thus lead to fewer circulating inflammation markers in patients with depression disorder and might have beneficial effects on their symptom severity. Additionally, gut microbiota are involved in the production of metabolites of, for instance, precursors of serotonin. Due to evidence suggesting serotonin deficiency as a possible causal factor in a number of affective disorders, bacteria producing this precursor might be beneficial for this group of disorders (Dinan et al., 2013; Lakhan & Vieira, 2008). Future studies in these patient populations could be set up in order to test the effectiveness of a probiotic treatment for these specific disorders.

References

Alcock, J., Maley, C. C., & Aktipis, C. A. (2014). Is eating behavior manipulated by the gastrointestinal microbiota? Evolutionary pressures and potential mechanisms. *BioEssays, 36*(10), 940–949.

Allin, K. H., Nielsen, T., & Pedersen, O. (2015). Mechanisms in endocrinology: Gut microbiota in patients with type 2 diabetes mellitus. *European Journal of Endocrinology, 172*(4), R167–R177.

Bercik, P., Denou, E., Collins, J., Jackson, W., Lu, J., Jury, J., … & Collins, S. M. (2011). The Intestinal Microbiota Affect Central Levels of Brain-Derived Neurotropic Factor and Behavior in Mice. *Gastroenterology, 141*(2), 599–609.

Berntson, G. G., Thomas Bigger, J., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., … & Van Der Molen, M. W. (1997). Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology, 34*(6), 623–648.

Bishop, S., Duncan, J., Brett, M., & Lawrence, A. D. (2004). Prefrontal cortical function and anxiety: controlling attention to threat-related stimuli. *Nature Neuroscience, 7*(2), 184–188.

Capuron, L., & Miller, A. H. (2011). Immune system to brain signaling: Neuropsychopharmacological implications. *Pharmacology & Therapeutics, 130*(2), 226–238.

Charnay, Y., & Leger, L. (2010). Brain serotonergic circuitries. *Dialogues in Clinical Neuroscience, 12*(4), 471–487.

Chen, X., D'Souza, R., & Hong, S.T. (2013). The role of gut microbiota in the gut-brain axis: current challenges and perspectives. *Protein & Cell,* 4(6), 403–414.

Collins, S. M., Surette, M., & Bercik, P. (2012). The interplay between the intestinal microbiota and the brain. *Nature Reviews Microbiology, 10*(11), 735–742.

Courtney, S. M., Petit, L., Haxby, J. V., & Ungerleider, L. G. (1998). The role of prefrontal cortex in working memory: examining the contents of consciousness. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 353*(1377), 1819–1828.

Cryan, J. F., & Dinan, T. G. (2012). Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nature Reviews Neuroscience, 13*(10), 701–712.

Cryan, J. F., & O'Mahony, S. M. (2011). The microbiome-gut-brain axis: from bowel to behavior. *Neurogastroenterology & Motility, 23*(3), 187–192.

Desbonnet, L., Clarke, G., Shanahan, F., Dinan, T. G., & Cryan, J. F. (2014). Microbiota is essential for social development in the mouse. *Molecular Psychiatry, 19*(2), 146–148.

Dinan, T. G., Stanton, C., & Cryan, J. F. (2013). Psychobiotics: A Novel Class of Psychotropic. *Biological Psychiatry, 74*(10), 720–726.

Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences, 15*(2), 85–93.

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving Emotional Conflict: A Role for the Rostral Anterior Cingulate Cortex in Modulating Activity in the Amygdala. *Neuron, 51*(6), 871–882.

Fellows, L. K., & Farah, M. J. (2007). The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cerebral Cortex, 17*(11), 2669–2674.

Foster, J. A., & McVey Neufeld, K. A. (2013). Gut–brain axis: how the microbiome influences anxiety and depression. *Trends in Neurosciences, 36*(5), 305–312.

Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-Related fMRI: Characterizing Differential Responses. *NeuroImage, 7*(1), 30–40.

Ghosh, S., Van Heel, D., & Playford, R. J. (2004). Probiotics in inflammatory bowel disease: is it all gut flora modulation?. *Gut, 53*(5), 620–622.

Gimeno, D., Kivimäki, M., Brunner, E. J., Elovainio, M., De Vogli, R., Steptoe, A., … & Ferrie, J. E. (2009). Associations of C-reactive protein and interleukin-6 with cognitive symptoms of depression: 12-year follow-up of the Whitehall II study. *Psychological*

Medicine, 39(03), 413–423.

Gold, A. L., Morey, R. A., & McCarthy, G. (2015). Amygdala–Prefrontal Cortex Functional Connectivity During Threat-Induced Anxiety and Goal Distraction. *Biological Psychiatry, 77*(4), 394–403.

Hargreaves, K. M. (1990). Neuroendocrine markers of stress. *Anesthesia Progress, 37*(2–3), 99– 105.

Hariri, A. R., Bookheimer, S. Y., & Mazziotta, J. C. (2000). Modulating emotional responses: effects of a neocortical network on the limbic system. *Neuroreport, 11*(1), 43–48.

Haxby, J.V., Hoffman, E.A., & Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*(6), 223–233.

Joormann, J., & Gotlib, I. H. (2010). Emotion Regulation in Depression: Relation to Cognitive Inhibition. *Cognition and Emotion, 24*(2), 281–298.

Juruena, M. F. (2014). Early-life stress and HPA axis trigger recurrent adulthood depression. *Epilepsy & Behavior, 38*, 148–159.

Kajander, K., Hatakka, K., Poussa, T., Färkkilä, M., & Korpela, R. (2005). A probiotic mixture alleviates symptoms in irritable bowel syndrome patients: a controlled 6-month intervention. *Alimentary Pharmacology & Therapeutics, 22*(5), 387–394.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology. General, 132*(1), 47–70.

King, M. S. (2007). Anatomy of the Rostral Nucleus of the Solitary Tract. In R. M. Bradley (Ed.), *The Role of the Nucleus of the Solitary Tract in Gustatory Processing*. Boca Raton (FL): CRC Press/Taylor & Francis.

Kopp, M. V., Goldstein, M., Dietschek, A., Sofke, J., Heinzmann, A., & Urbanek, R. (2008). Lactobacillus GG has in vitro effects on enhanced interleukin-10 and interferon-gamma release of mononuclear cells but no in vivo effects in supplemented mothers and their neonates. *Clinical and Experimental Allergy: Journal of the British Society for Allergy and Clinical Immunology, 38*(4), 602–610.

Lakhan, S. E., & Vieira, K. F. (2008). Nutritional therapies for mental disorders. *Nutrition Journal, 7*, 2.

Logan, A. C., & Katzman, M. (2005). Major depressive disorder: probiotics may be an adjuvant therapy. *Medical Hypotheses, 64*(3), 533–538.

Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition, 7*(3), 166–174.

Lyte, M. (2014). Microbial endocrinology: Host-microbiota neuroendocrine interactions influencing brain and behavior. *Gut Microbes, 5*(3), 381–389.

Maes, M., Kubera, M., & Leunis, J.C. (2008). The gut-brain barrier in major depression: intestinal mucosal dysfunction with an increased translocation of LPS from gram negative enterobacteria (leaky gut) plays a role in the inflammatory pathophysiology of depression. *Neuroendocrinology Letters, 29*(1), 117–124.

Mayer, E. A. (2011). Gut feelings: the emerging biology of gut–brain communication. *Nature Reviews Neuroscience, 12*(8), 453–466.

Messaoudi, M., Lalonde, R., Violle, N., Javelot, H., Desor, D., Nejdi, A., … & Cazaubiel, J.M. (2011). Assessment of psychotropic-like properties of a probiotic formulation (Lactobacillus helveticus R0052 and Bifidobacterium longum R0175) in rats and human subjects. *The British Journal of Nutrition, 105*(05), 755–764.

Miller, A. H. (1998). Neuroendocrine and immune system interactions in stress and depression. *Psychiatric Clinics of North America, 21*(2), 443–463.

Moayyedi, P., Ford, A. C., Talley, N. J., Cremonini, F., Foxx-Orenstein, A. E., Brandt, L. J., & Quigley, E. M. M. (2010). The efficacy of probiotics in the treatment of irritable bowel syndrome: a systematic review. *Gut, 59*(3), 325–332.

Naseribafrouei, A., Hestad, K., Avershina, E., Sekelja, M., Linløkken, A., Wilson, R., & Rudi, K. (2014). Correlation between the human fecal microbiota and depression. *Neurogastroenterology and Motility: The Official Journal of the European Gastrointestinal Motility Society, 26*(8), 1155–1162.

Neufeld, K. M., Kang, N., Bienenstock, J., & Foster, J. A. (2011). Reduced anxiety-like behavior and central neurochemical change in germ-free mice. *Neurogastroenterology & Motility, 23*(3), 255–e119.

Neuman, H., Debelius, J. W., Knight, R., & Koren, O. (2015). Microbial endocrinology: the interplay between the microbiota and the endocrine system. *FEMS Microbiology Reviews, 39*(4), 509–521.

O'Mahony, S. M., Clarke, G., Borre, Y. E., Dinan, T. G., & Cryan, J. F. (2015). Serotonin, tryptophan metabolism and the brain-gut microbiome axis. *Behavioural Brain Research, 277*, 32–48.

O'Mahony, S. M., Marchesi, J. R., Scully, P., Codling, C., Ceolho, A.M., Quigley, E. M. M., … & Dinan, T. G. (2009). Early life stress alters behavior, immunity, and microbiota in rats: implications for irritable bowel syndrome and psychiatric illnesses. *Biological Psychiatry, 65*(3), 263–267.

Peluso, M. A. M., Glahn, D. C., Matsuo, K., Monkul, E. S., Najt, P., Zamarripa, F., … & Soares, J. C. (2009). Amygdala hyperactivation in untreated depressed individuals. *Psychiatry Research, 173*(2), 158–161.

Poser, B. A., Versluis, M. J., Hoogduin, J. M., & Norris, D. G. (2006). BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: Parallel-acquired inhomogeneity-desensitized fMRI. *Magnetic Resonance in Medicine, 55*(6), 1227–1235.

Roberts, K. L., & Hall, D. A. (2008). Examining a supramodal network for conflict processing: a systematic review and novel functional magnetic resonance imaging data for related visual and auditory Stroop tasks. *Journal of Cognitive Neuroscience, 20*(6), 1063–1078.

Rook, G. A. W., Lowry, C. A., & Raison, C. L. (2013).

Microbial "Old Friends", immunoregulation and stress resilience. *Evolution, Medicine, and Public Health, 2013*(1), 46–64.

Rothwell, N. J., & Hopkins, S. J. (1995). Cytokines and the nervous system II: Actions and mechanisms of action. *Trends in Neurosciences, 18*(3), 130–136.

Santos, J., Yang, P. C., Söderholm, J. D., Benjamin, M., & Perdue, M. H. (2001). Role of mast cells in chronic stress induced colonic epithelial barrier dysfunction in the rat. *Gut, 48*(5), 630-636.

Sherman, M. P., Zaghouani, H., & Niklas, V. (2015). Gut microbiota, the immune system, and diet influence the neonatal gut-brain axis. *Pediatric Research, 77*(1–2), 127–135.

Shikora, S. A., Toouli, J., Herrera, M. F., Kulseng, B., Brancatisano, R., Kow, L., … & Billingto, C. J. (2015). Intermittent Vagal Nerve Block for Improvements in Obesity, Cardiovascular Risk Factors, and Glycemic Control in Patients with Type 2 Diabetes Mellitus: 2-Year Results of the VBLOC DM2 Study. *Obesity Surgery, 26*(5), 1021–1028.

Shreiner, A. B., Kao, J. Y., & Young, V. B. (2015). The gut microbiome in health and in disease. *Current Opinion in Gastroenterology, 31*(1), 69–75.

Smith, K., McCoy, K. D., & Macpherson, A. J. (2007). Use of axenic animals in studying the adaptation of mammals to their commensal intestinal microbiota. *Seminars in Immunology, 19*(2), 59–69.

Söderholm, J. D., & Perdue, M. H. (2001). II. Stress and intestinal barrier function. *American Journal of Physiology-Gastrointestinal and Liver Physiology, 280*(1), G7-G13.

Sparkman, N. L., Buchanan, J. B., Heyen, J. R. R., Chen, J., Beverly, J. L., & Johnson, R. W. (2006). Interleukin-6 Facilitates Lipopolysaccharide-Induced Disruption in Working Memory and Expression of Other Proinflammatory Cytokines in Hippocampal Neuronal Cell Layers. *The Journal of Neuroscience, 26*(42), 10709–10716.

Steenbergen, L., Sellaro, R., Van Hemert, S., Bosch, J. A., & Colzato, L. S. (2015). A randomized controlled trial to test the effect of multispecies probiotics on cognitive reactivity to sad mood. *Brain, Behavior, and Immunity, 48,* 258–264.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662.

Sudo, N., Chida, Y., Aiba, Y., Sonoda, J., Oyama, N., Yu, X.N., … & Koga, Y. (2004). Postnatal microbial colonization programs the hypothalamic–pituitary–adrenal system for stress response in mice. *The Journal of Physiology, 558*(1), 263–275.

Swick, D., Ashley, V., & Turken, A. U. (2008). Left inferior frontal gyrus is critical for response inhibition. *BMC Neuroscience, 9*(1), 102.

Tillisch, K., Labus, J., Kilpatrick, L., Jiang, Z., Stains, J., Ebrat, B., … & Mayer, E. A. (2013). Consumption of Fermented Milk Product With Probiotic Modulates Brain Activity. *Gastroenterology, 144*(7), 1394–1401.

Van Hemert, S., & Ormel, G. (2014). Influence of the Multispecies Probiotic Ecologic®Barrier on Parameters of Intestinal Barrier Function. *Food and Nutrition Sciences, 5*(18), 1739–1745.

Winecoff, A., Clithero, J. A., Carter, R. M., Bergman, S. R., Wang, L., & Huettel, S. A. (2013). Ventromedial Prefrontal Cortex Encodes Emotional Value. *Journal of Neuroscience, 33*(27), 11032–11039.

Wu, H.J., & Wu, E. (2012). The role of gut microbiota in immune homeostasis and autoimmunity. *Gut Microbes, 3*(1), 4–14.

Zoccatelli, G., Beltramello, A., Alessandrini, F., Pizzini, F. B., & Tassinari, G. (2010). Word and position interference in Stroop tasks: a behavioral and fMRI study. *Experimental Brain Research, 207*(1–2), 139–147.

The Effects of Grammatical Gender on Reference Processing in German: an ERP Study

Julia Misersky[4]
Supervisors: Asifa Majid[1,3,4], Tineke M. Snijders[2,4], Pascal M. Gygax[5]

[1]Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands
[2]Radboud University Nijmegen, Donders Centre for Cognitive Neuroimaging, The Netherlands
[3]Radboud University Nijmegen, Centre for Language Studies, The Netherlands
[4]Max-Planck-Institute for Psycholinguistics, Nijmegen, The Netherlands
[5]Université de Fribourg, Fribourg, Switzerland

Grammatically masculine forms can refer to men and women, but may favour male-specific interpretations. Using a match-mismatch paradigm, the present event-related potential (ERP) study assessed how masculine grammatical gender in role nouns affects referent processing in German. Twenty participants read sentences in which a stereotypically neutral role noun (grammatically masculine or feminine) introduced a group of people. A sentence continuation specified the group as consisting in part of *men* or *women*, meaning continuations were either congruent (masculine–*men*, feminine–*women*) or incongruent (masculine–*women*, feminine–*men*) to the grammatical gender of the role noun. Incongruent continuations were expected to result in an N400-P600 complex. Between 300 and 500 ms, no N400-like effect was observed. Following masculine role nouns, all continuations were processed similarly (p = .891). Following feminine role nouns, incongruent continuations elicited more positive responses than congruent ones (p = .045). Between 500 and 800 ms, a P600-like effect was observed. For both masculine and feminine role nouns, incongruent continuations resulted in more positive responses than congruent ones (p = .039). The results are discussed in terms of a two-stage model: initially, the incongruency between the masculine and *women* continuations goes unnoticed, yet leads to processing difficulties later on, implying a male-specific interpretation.

Keywords: grammatical gender, reference processing, EEG, P600, N400

Corresponding author: Julia Misersky; **E-mail:** julia.misersky@mpi.nl

Gender is a social category we encounter on a daily basis, with gender equality having become an important socio-political issue. Cues to gender are present in language in many ways. For example, we find words describing gendered referents (e.g., *father* or *woman*). Language can also carry stereotype information, which is strongly linked to conceptual knowledge (e.g., *electrician* or *nurse*, associated with males and females, respectively). Additionally, grammatical gender may map onto social categories of male and female gender. Often, we find these mappings in so-called role nouns, which describe social importance or occupation, and constitute an important way of denoting human referents. Indeed, grammatical gender in role nouns has been considered increasingly relevant to the equality debate. It has been suggested that masculine grammatical gender in role nouns favours male-specific interpretations, yet the masculine is regularly used to generically talk about male and female referents. This raises the question of whether the masculine is adequate in representing society in a gender-neutral way. To assess this issue, the present study investigated how masculine grammatical gender affects our understanding of differently gendered referents.

Behavioural research on gender cues

As a linguistic category, gender has been described to be "almost universally present in language […]" (Irmen, Holt, & Weisbrod, 2010, p. 133). Indeed, the gender of a human referent can be indicated linguistically via diverse means. While lexical semantics (e.g., *woman*) cue gender explicitly, role nouns (e.g., *musician*) often only imply referent gender. Take the example of the role noun *musician(s)*. In English, the noun is not grammatically gendered in the singular or plural. Only via the use of pronouns, and only in the singular can the gender of the *musician* be revealed as she or he. In contrast, in many grammatically gendered languages, the same role noun is morpho-syntactically marked for gender. It seems, then, that role nouns are realised differently across languages, depending on the grammatical system.

Behavioural research across languages has investigated how role nouns are understood. In the absence of grammatical gender, referent gender is cued by stereotypicality (Gygax, Gabriel, Sarrasin, Oakhill, & Garnham, 2008; Sato, Gygax, & Gabriel, 2013). For example, in English, the noun *musicians* implies male referents (Misersky et al., 2014). In gendered languages, such as German, grammatical gender indicates referent gender, since specific forms are used to describe females and males. While feminine role nouns (e.g., $Musikerinnen_{\text{feminine}}$ [fem.]) cue specifically for females, the masculine (e.g., $Musiker_{\text{masculine}}$ [masc.]) can be used for males, but it can be used as a generic too. As such, $Musiker_{\text{masc.}}$ can be used for a group of men or a mixed-gender group consisting of both men and women. In short: the masculine may yield specific as well as generic interpretations.

Behavioural research has revealed using a specific form such as $Musiker_{\text{masc.}}$ for a mixed-gender group (i.e., including both males and females) results in an ambiguity in interpretation. For example, Gygax et al. (2008) assessed the gender interpretation of role nouns by comparing English, which has natural gender, to French and German, both of which have grammatically masculine and feminine gender categories. Participants were presented with a stereotypically male or female role noun (in the masculine for French and German) in one sentence, and with an anaphoric noun (*men* or *women*) in a second. With a Yes or No response, participants had to indicate whether the sentence containing the anaphor was a sensible continuation of the sentence with the role noun. In English, participants' responses and response times were linked to the stereotypicality of the role noun. There was no such effect of stereotypicality for French and German. Instead, in both these languages *men* continuations received more Yes responses (i.e., continuation is sensible) compared to *women* continuations. This was further reflected in the response times: participants were faster to judge *men* continuations as sensible compared to *women* continuations. Thus, despite being generically intended, the masculine was interpreted as specific. Recent studies have found similar effects in primary school children (Vervecken, Gygax, Gabriel, Guillod, & Hannover, 2015; Vervecken & Hannover, 2015). Presenting role nouns in the masculine or a gender-neutral form in German and Dutch, grammatical form was found to affect ratings of job status, difficulty, and accessibility. Specifically, the children reported lower self-efficacy –that is, whether they thought they could succeed in a job– when presented with the masculine. These results highlight the societal relevance of grammatical gender cues.

In sum, this research suggests grammatical gender to be highly relevant in guiding our understanding of human referents, which may even override stereotype information (Gygax et al., 2008; Irmen & Roßberg, 2004). In particular, grammatically masculine forms

seem to lead to male-specific interpretations, despite their use as a generic for all genders.

ERP research on gender cues

Electroencephalography (EEG), in particular event-related-potentials (ERPs), has been used to assess the effects of gender cues on comprehension during online processing. ERP studies on gender cues regularly find N400 and P600 effects. A negative deflection around 300 to 500 ms after stimulus presentation, the N400 is generally linked to lexical-semantic processing (Kutas & Federmeier, 2011). The P600 is sensitive to syntactic mismatch and integration difficulties (Osterhout, McLaughlin, & Bersick, 1997b), and is characterised by a positive deflection around 500 to 800 ms after stimulus presentation.

To a large extent, ERP research on the effects of grammatical gender has studied the processing of non-human referents, such as object nouns. Recent work investigated the processing of gender-to-ending consistency in Italian (Caffarra, Siyanova-Chanturia, Pesciarelli, Vespignani, & Cacciari, 2015). The materials consisted of sentences containing determiner-noun pairs, which either matched in gender (*il*$_{masc.}$ *cucchiaio*$_{masc.}$ [the spoon]) or did not (*la*$_{fem.}$ *cucchiaio*$_{masc.}$ [the spoon]). Gender mismatches resulted in processing difficulties reflected by a P600 effect. Similar results were obtained with noun-adjective gender mismatches (e.g., *faro*$_{masc.}$*–alta*$_{fem.}$ [lighthouse–high]) in Spanish sentences (Barber & Carreiras, 2005). Interestingly, the authors found the same mismatches presented as word-pairs outside of sentences elicited an N400 effect. According to Barber and Carreiras (2005), the N400 component can be driven by conceptual, lexical or morphological feature integration depending on task and stimuli. The authors thus concluded the N400 for the word-pairs reflected simple feature integration. By contrast, full sentences required more complex syntactic structure building, leading to the P600.

ERP research on human referent processing has predominantly focused on stereotype information embedded in role nouns. White, Crites, Taylor, and Corral (2009) used a match-mismatch paradigm and found word-pairs, which mismatched in stereotypicality (e.g., *secretary–aggressive*) elicited a larger N400 effect than matching word-pairs (e.g., *secretary–caring*). Research by Osterhout, Bersick, and McLaughlin (1997a) assessed the processing of reflexive pronouns (*himself/herself*) following a role noun in English. Pronouns either did or did not match an antecedent role noun's gender definition (e.g., *bachelor*) or stereotype (e.g., *doctor*). The results revealed a P600 effect for gender mismatches. Interestingly, the amplitude of the P600 varied for stereotypical compared to definitional nouns. Mismatches between reflexives and definitional role nouns were considered more anomalous as they leave little room for interpretation, thus leading to larger P600 amplitude. Interestingly, an N400 was also found for the last word of each sentence for the definition-violating reflexives. The authors explain that in the stereotype-violating sentences, an acceptable interpretation had been derived. For the definition-violating sentences, this was not possible, leading to the N400. Testing German speakers, Irmen et al. (2010) investigated how stereotypicality of role nouns affects the processing of a subsequent referent. They presented their participants with sentences consisting of a stereotypically male or female role noun (e.g., *computer scientist*; stereotypically male), and a co-referential continuation, which was either neutral (e.g., *these people*), matching (e.g., *these men*) or mismatching (e.g., *these women*) with regards to gender. An N400 effect across all continuations following a stereotypically male role was observed, suggesting participants may have anticipated a mismatch. A later P600 effect shows a clear interaction between the stereotypicality of the role noun and the continuation, and was taken as the integration of the two nouns. Irmen et al. (2010) linked this to a two-stage model of reference resolution (Garrod & Terras, 2000): initial linking (bonding) relies on lexical-semantic information, whereas resolution takes place once additional information has been taken into account.

In sum, two ERP components are most regularly observed in the research on gender cues: the N400 and P600. Roughly speaking, the N400 reflects lexical-semantic processes, whereas the P600 is associated with systematic syntactic processing. Additionally, an N400-P600 complex has been described (e.g., Irmen et al., 2010), which might be especially relevant to the processing of role nouns. Both conceptual and syntactic information are relevant to successfully building co-reference (Schmitt, Lamers, & Münte, 2002), and this might be especially true for languages where semantics are increasingly subject to syntactic constraints as a result of grammatical gender. In German, for example, the semantic and grammatical gender of words describing human referents (e.g., *Frauen*$_{fem.}$ [women]) tend to agree (Irmen et al., 2010). This means lexical-semantic and syntactic processing difficulties may co-occur during the processing of grammatical gender cues for human referents. In line with previous interpretations of

the two ERP components, initial reference might be established by the bonding of conceptual and/or morphological features (N400), with a systematic analysis of grammatical gender taking place later on (P600).

The present study

Behavioural research shows how the grammatical gender of role nouns can affect understanding of human referents. Following a feminine form, only female referents are pragmatically and grammatically acceptable. By contrast, it is pragmatically acceptable to use a masculine form for male and female referents. Grammatically, however, only male referents are an acceptable match for a masculine role noun. The masculine has thus been suggested to lead to male-specific interpretations (Gygax et al., 2008), at least when explicit decision-making is required. ERPs provide the ideal measure to assess referent processing as separate from decision-making. Thus far, ERP research on grammatical gender has predominantly studied processing of nonhuman referents (e.g., Caffarra et al., 2015). In the case of human referent processing, ERP studies using role nouns have focused on stereotype processing (e.g., White et al., 2009). Motivated by Gygax et al.'s (2008) findings, the present study investigated how masculine grammatical gender affects the processing of differently gendered human referents. Specifically, the study aimed to assess whether the masculine is understood as gender-neutral (i.e., encompassing both females and males) or as specific to males. Focusing on the N400 and P600, which relate to semantic and syntactic processing respectively, allowed for a comprehensive investigation of this question.

Grammatical gender may override stereotype information in decision-making, but the two are considered to interact during processing (Gygax et al., 2008; Irmen et al., 2010). In a match-mismatch paradigm, the present study thus used stereotypically neutral roles to systematically focus solely on the effects of grammatical gender. The experimental sentences introduced a group of people via a role noun (manipulated as grammatically masculine or feminine, e.g., *Studenten*$_{masc.}$ or *Studentinnen*$_{fem.}$), and a sentence continuation specified the group as consisting in part, yet not exclusively, of men (*Männer*$_{masc.}$) or women (*Frauen*$_{fem.}$). This meant role noun and continuation either matched (masculine–*men*; feminine–*women*) or mismatched (masculine–*women*; feminine–*men*) in grammatical gender. Table 1 gives an example of the sentence stimuli.

Both the N400 and P600 have been observed in ERP studies on role nouns, and both conceptual and syntactic processes have been considered important for reference building. Thus, in this study, both components resulting from grammatical gender incongruencies between role nouns and continuations were expected. The analyses specifically focused on ERPs following the onset of the continuations (*men, women*) to assess how these are integrated with the preceding role noun. In line with the main research question, the following findings were hypothesised:

Firstly, it was anticipated that a role noun in the masculine would favour a male-specific interpretation of the referent. As a result, masculine role nouns

Table 1.

Example of the sentence stimuli and conditions.

Grammatical gender	Continuation	
Masculine	congruent (*men*)	Die **Studenten**$_{masculine}$ gingen zur Mensa, weil manche der **Männer** Hunger hatten. [The **students** went to the canteen, because some of the **men** were hungry.]
	incongruent (*women*)	Die **Studenten**$_{masculine}$ gingen zur Mensa, weil manche der **Frauen** Hunger hatten. [The **students** went to the canteen, because some of the **women** were hungry.]
Feminine	congruent (*women*)	Die **Studentinnen**$_{feminine}$ gingen zur Mensa, weil manche der **Frauen** Hunger hatten. [The **students** went to the canteen, because some of the **women** were hungry.]
	incongruent (*men*)	Die **Studentinnen**$_{feminine}$ gingen zur Mensa, weil manche der **Männer** Hunger hatten. [The **students** went to the canteen, because some of the **men** were hungry.]

followed by *women* continuations were hypothesised to lead to processing difficulties. In particular, both an N400 and a P600 effect were expected to reflect these difficulties, with incongruencies leading to a relatively more negative amplitude of the N400, and a relatively more positive amplitude of the P600.

Secondly, the masculine is used as a default for both male and female referents, whereas the feminine is only used for female referents. Theoretically, this allows for a more flexible gender interpretation following masculine role nouns. Processing difficulties (as reflected by the N400 and P600) were thus expected to be reduced for the masculine (masculine–*women*), compared to the feminine role nouns (feminine–*men*).

Gygax and Gabriel (2008) suggest using materials with both feminine and masculine forms may lead participants to consider them in direct contrast. This may amplify the interpretation of the masculine as specific to males, since the feminine is always specific to females. However, the motivation for including the feminine in this study was twofold. Firstly, since role nouns in the feminine always denote female referents, a pairing with male referents would be highly incongruent. The high specificity of the feminine would establish a benchmark of incongruency between role noun and continuation. This would allow for a more comprehensive interpretation of the results concerning continuations after a masculine form, which has a flexible use for all genders. Secondly, we often encounter both forms in a variety of contexts. As such, including both forms in this study provided an appropriate representation of a real-life situation.

Methods

Participants

Twenty native speakers of German (13 female, age range 19 – 29 years, M = 22.3, SD = 2.68) recruited from Radboud University's SONA system participated in this study. Participants provided written informed consent, and received course credit or payment for their participation. All had normal or corrected to normal vision and were right-handed. Four participants were excluded from further analysis; two due to a low number of trials after pre-processing, and another two due to accuracy on the content questions (described below) that was at or below chance. The study was approved by the local ethical committee (Commissie Mensgebonden Onderzoek, Regio Arnhem-Nijmegen).

Materials and design

A total of 156 role nouns were selected from a recent norming study (Misersky et al., 2014) on the basis of their stereotypicality rating. Role nouns rated as stereotypically neutral (M = .47, SD = .08, ranging between 1 = stereotypically female and 0 = stereotypically male) were included in the materials. Using these role nouns at the beginning of the sentence, coordinate clauses were created inspired by previous research (Gygax et al., 2008; Irmen et al., 2010). The sentence-initial role noun introduced a group of people, with the role nouns grammatical gender manipulated as masculine (e.g., *Studenten*$_{masc.}$) or feminine (e.g., *Studentinnen*$_{fem.}$). Later in the sentence, this group was specified as being partially consisting of men (*Männer*$_{masc.}$) or women (*Frauen*$_{fem.}$). Inspired by previous research (Gygax et al., 2008), pseudo-randomised quantifiers were selected to highlight the group was not exclusively made up of men or women; *einige* (a few/some), *mehrere* (several), *manche* (some), *einzelne* (single ones) or *viele* (many); refer to Table 1 for an example. In addition, 80 filler items were created, half of which followed a similar format, and half of which were structurally different, resulting in a total of 236 experimental sentences. Sentences were pseudo-randomised for presentation.

Procedure

Participants were seated in a dimly illuminated sound-attenuating testing booth. They were instructed to attentively read the sentences, since they would receive content questions throughout the experiment. These instructions were both explained by the experimenter, and presented visually on the testing PC. Since eye movements distort the EEG recording, participants were also asked to only blink between sentences and during breaks. Participants were able to remain in contact with the experimenter via microphone. All interactions and instructions were in German.

Sentences were presented using Presentation software (Neurobehavioral Systems, www.neurobs.com). Each sentence was presented in a word-by-word serial visual presentation mode at the centre of a 24-inch PC monitor. The background was a dark grey with the words presented in white letters (Helvetica, font size 26). The beginning of each sentence was preceded by a fixation cross (+). Each word was flashed for 380 ms with an inter-

word-interval (IWI) of 145 ms. The second and fourth word of each sentence were flashed slightly longer, for 480 ms, due to word length. Sentence-final words were followed by a full stop, then a 1000 ms blank. Every ten sentences, a content question would appear on-screen, requiring a self-paced Yes or No response via button press with the left or right index finger, respectively. The question related to the activity carried out in the sentence; and there was no repetition of the role noun. The inter-trial-interval (ITI) was 2000 ms during which the fixation cross re-appeared. First, participants received nine practice sentences, after which remaining uncertainties about the task could be resolved. The experiment was split into four blocks of 59 trials each. There were self-paced pauses between blocks where a drink of water was offered to the participants. Figure 1 gives an example of the procedure.

EEG set-up and apparatus

Continuous EEG was recorded from 32 active electrodes (10-20 system) attached to an elastic cap (actiCAP), with a BrainAmp DC amplifier (Brain Products, Gilching, Germany). The signal was sampled at 500 Hz. One electrode in the cap provided an active ground. Electrooculogram (EOG) was recorded from electrodes above and below the eye, and at the outer canthi of the eyes. Electrode impedances were kept below 20 kΩ.

The data was pre-processed using the FieldTrip toolbox for EEG/MEG-analysis (www.fieldtriptoolbox.org; Oostenveld, Fries, Maris, & Schoffelen, 2011) in MATLAB. For each continuation noun, segments were chosen in the range from 200 ms before to 1000 ms after word onset. Offline-filtering included a low-pass filter at 35 Hz and a high-pass filter at 0.1 Hz. The

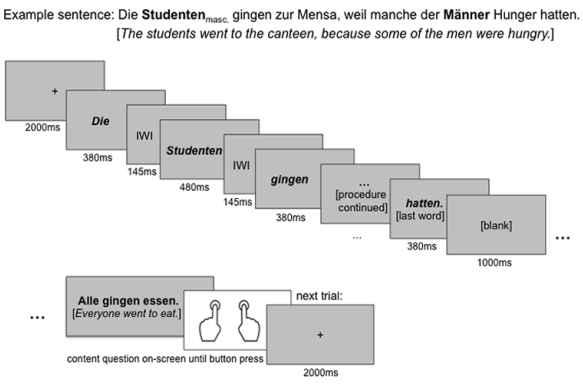


Fig. 1. Example of the word-by-word stimulus presentation followed by a content question.

data were then inspected visually, and trials showing electrode jumps and/or drifting were removed in preparation for an independent component analysis (ICA). ICA was performed to remove any remaining EOG and/or electrocardiogram (ECG) artefacts from the data. All channels were then referenced via the average of the signal of both mastoids (Luck, 2005). A baseline correction was applied in which the signal was normalised relative to a 200 ms stimulus preceding window. Trials containing signal exceeding ±75 μV were removed, and mean ERP amplitudes for the time windows of interest were calculated. The datasets of two subjects were excluded from further analysis, since less than 29 trials per condition (< 25 percent) remained after pre-processing. The average amount of kept trials per condition for the included subjects was 34.65 (88.85%, ranging from an average of 34.43 to 34.75 trials across all conditions).

Analysis

In line with existing work (Irmen et al., 2010; Osterhout et al., 1997a), mean ERP amplitudes were statistically analysed in two main time windows after the onset of the continuation noun; 300 to 500 ms for the N400, and 500 to 800 ms for the P600, respectively. The mean ERP amplitudes were analysed in SPSS. As in Irmen et al. (2010), nine electrodes in anterior, central and posterior positions of the left and right hemisphere and the midline were used for the statistical analyses (F3/z/4, C3/z/4, P3/z/4).

Results

Responses to content questions

The participants were accurate in correctly answering the content questions ($M_{\text{present correct}}$ = 99.18, *SD* = 3.26), meaning they understood the task and were attentively reading the sentences throughout the experiment.

Event-related potentials

Following Irmen et al. (2010), the mean amplitudes of the ERPs for the time windows of interest were subjected to repeated-measures analyses of variance (ANOVA). The factors submitted to each ANOVA included Anteriority (three levels: anterior, central, posterior), Laterality (three levels: left, midline, right), Grammatical Gender of the role noun (two levels:

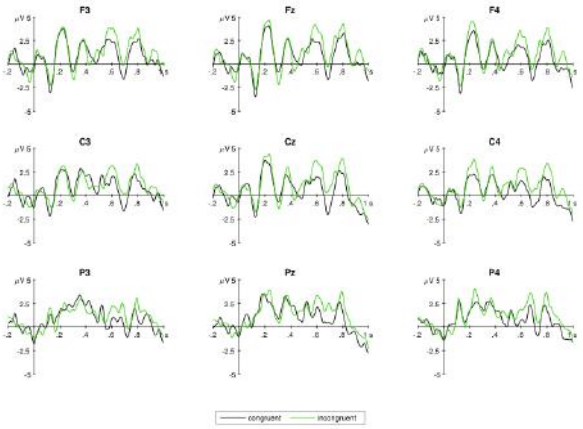


Fig. 2. ERPs for Continuations following a role noun with masculine grammatical gender.

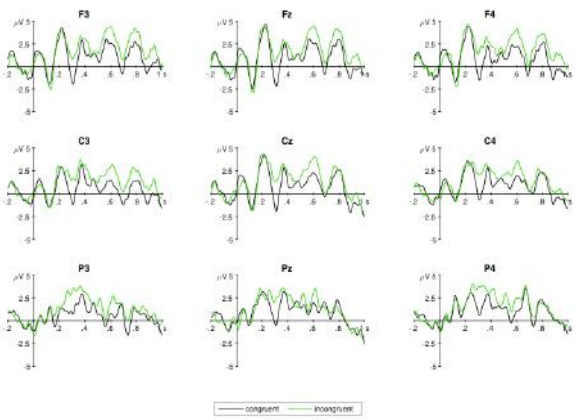


Fig. 3. ERPs for Continuations following a role noun with feminine grammatical gender.

masculine, feminine, see Table 1), and Continuation (two levels: congruent, incongruent, see Table 1). An alpha level of .05 was used for all statistical tests. Note that the effects for Grammatical Gender, Continuation, and the interaction between them are of prime relevance with regards to assessing the effects of grammatical gender on referent processing. Figures 2 and 3 represent the grand average ERPs for the nine electrodes separated by Grammatical Gender.

300 to 500 ms time window. The ANOVA revealed an interaction between Grammatical Gender and Continuation, $F(1, 15) = 5.82$, $p = .029$, $\eta^2 = .28$. Follow-up analyses were carried out for each Grammatical Gender separately (Fig. 4). For role nouns with masculine Grammatical Gender, there was no significant difference between congruent ($M = 1.23$ µV, $SEM = .42$) and incongruent ($M = 1.18$ µV, $SEM = .36$) continuations, $F(1, 15) = .02$, $p = .891$, $\eta^2 = .001$. For role nouns with feminine grammatical gender, however, responses to congruent Continuations differed significantly from those to incongruent ones, $F(1, 15) = 4.78$, $p = .045$, $\eta^2 = .24$. Incongruent Continuations elicited more positive responses ($M = 2.29$ µV, $SEM = .44$) compared to congruent Continuations ($M = 1.18$ µV, $SEM = .47$). There were no main effects of Grammatical Gender or Continuation, nor of Anteriority or Laterality.

500 to 800 ms time window. The ANOVA showed a main effect of Continuation, $F(1, 15) = 5.13$, $p = .039$, $\eta^2 = .25$. Regardless of the Grammatical Gender of the role noun, incongruent Continuations elicited significantly more positive responses ($M = 2.09$ µV, $SEM = .46$), compared to congruent Continuations ($M = 1.07$ µV, $SEM = .32$, see Fig. 5). In this time window, there were no main effects of Anteriority or Laterality, nor any interactions between any of the factors.

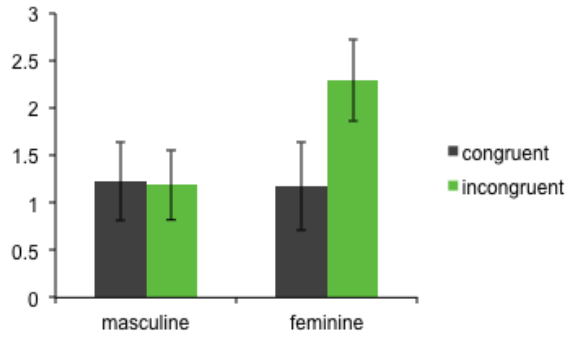


Fig. 4. Mean amplitudes in the 300–500 ms time window as a factor of Grammatical Gender (masculine vs. feminine) and Continuation (men vs. women).

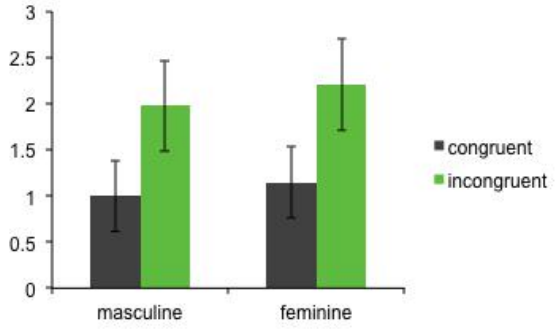


Fig. 5. Mean amplitudes in the 500–800 ms time window as a factor of Grammatical Gender (masculine vs. feminine) and Continuation (men vs. women).

Discussion

The present study assessed how grammatical gender affects referent processing. More specifically, the study focused on how grammatically masculine forms affect the processing of differently gendered referents. In line with previous research, masculine role nouns followed by *women* continuations were expected to lead to processing difficulties reflected by a complex of N400 (300 to 500 ms) and P600 (500 to 800 ms). Furthermore, effects of incongruency were expected to be reduced for continuations following a grammatically masculine role noun (masculine–*women*) compared to a grammatically feminine role noun (feminine–*men*).

In the 300 to 500 ms time window, an interaction between grammatical gender and continuation was found. Put differently, how the continuation was processed depended on the preceding grammatical gender. When participants read a role noun in the masculine, they processed both congruent and incongruent continuations similarly. By contrast, when they read a role noun in the feminine, the continuations were processed differently. Interestingly, the incongruent continuations elicited more positive responses compared to the congruent ones. This was unexpected, since for the N400, incongruencies tend to elicit relatively more negative responses compared to congruencies. Siyanova-Chanturia, Pesciarelli and Cacciari (2012) have looked into the processing of pronouns (*lei*[she]/*lui*[he]) following stereotypically (e.g., *insegnante*$_{\text{masc./fem}}$ [teacher]), and grammatically gendered (e.g., *pensionato*$_{\text{masc.}}$ [pensioner]) role nouns in Italian. For female participants, they observed positive responses to incongruencies between grammatically gendered roles and pronouns, and interpreted this effect as P300-like. The P300 has been linked to stimuli evaluation (Kutas & Hillyard, 1980a; Kutas & Hillyard, 1980b), and is sensitive to this evaluation being task-relevant (Holcomb, 1988). In line with this, Siyanova-Chanturia et al. (2012) attribute their finding to participants having made decisions on the word-pairs during the experiment. Unlike Siyanova-Chanturia et al. (2012), the present study used sentence stimuli. Additionally, participants did not need to explicitly evaluate the continuations, yet answering content questions may have increased the likelihood of participants evaluating the words more closely. While this opens up the possibility for a P300-like effect, additional exploration is needed to sufficiently examine this unexpected effect.

In the 500 to 800 ms time window, a relatively more positive response to incongruent compared to congruent continuations was observed, regardless of the grammatical gender of the role noun preceding it. This result reflects a P600-like effect, meaning participants encountered processing difficulties upon reading *women* when preceded by a grammatically masculine role noun, and upon reading *men* when preceded by a grammatically feminine role noun.

In line with Irmen et al. (2010), the results can be linked to Garrod & Terras' (2000) two-stage reference processing model. During an initial bonding stage, the Continuation is linked to the role noun by an automatic process. Note that complete congruency required the grammatical gender of the role noun to match both the grammatical and lexical-semantic gender of the continuation. Despite the unexpected positive ERP amplitude of the incongruent continuations following a feminine role noun, it was clear bonding was different for the congruent compared to the incongruent continuations. This could be due to the feminine form being highly specific, effectively constraining processing early on. For role nouns in the masculine, however, all continuations were processed similarly. The absence of differential processing in this initial stage could be a result of the masculine form being pragmatically used to describe mixed-gender groups. Only in the later time window, during the reference resolution stage, incongruent continuations lead to processing difficulties with no difference between the two grammatical genders.

To sum up, the observed pattern of effects suggests the grammatical gender of the role noun guided the processing of the continuation. Firstly, a role noun in the feminine leads to differential processing for *men* and *women* continuations early on. Secondly, the mismatch between *women* continuations following a masculine goes unnoticed during initial processing. Thirdly, and importantly, this same mismatch leads to difficulties later on. Linking back to the research question, the results suggest people struggled integrating female referents (*women*) with the grammatically masculine form during late systematic processing. Thus, the grammatically masculine form seems to create a bias towards males, despite being used to encompass females and males. This finding implies the masculine form to be inadequate in representing the whole of society. A grammatical form that is truly gender-neutral and/or encompasses males and females specifically could be key to reducing the effects observed in behavioural research and in this study. A follow-up including such grammatically gender-neutral forms can assess the possibilities of reducing the bias encountered.

Conclusion

The present study has shown that difficulties during referent processing are subject to grammatical constraints. In line with the behavioural research in this field, the data indicate grammatical gender to be an important cue to the understanding of human referents. This has relevant implications for the regular use of the masculine to talk about mixed-gender groups. Since its interpretation favours male referents, future research should aim to assess which grammatical forms can most effectively reduce the processing difficulties observed in this study.

References

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience, 17*(1), 137-153.

Caffarra, S., Siyanova-Chanturia, A., Pesciarelli, F., Vespignani, F., & Cacciari, C. (2015). Is the noun ending a cue to grammatical gender processing? An ERP study on sentences in Italian. *Psychophysiology, 52*(8), 1019-1030.

Garrod, S., & Terras, M. (2000). The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory and Language, 42*(4), 526-544.

Gygax, P., & Gabriel, U. (2008). Can a group of musicians be composed of women? Generic interpretation of French masculine role names in the absence and presence of feminine forms. *Swiss Journal of Psychology, 67*(3), 143-151.

Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2008). Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes, 23*(3), 464-485.

Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language, 35*(1), 66-85.

Irmen, L., Holt, D. V., & Weisbrod, M. (2010). Effects of role typicality on processing person information in German: Evidence from an ERP study. *Brain Research, 1353*, 133-144.

Irmen, L., & Roßberg, N. (2004). Gender markedness of language: The impact of grammatical and nonlinguistic information on the mental representation of person information. *Journal of Language and Social Psychology, 23*(3), 272-307.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual review of Psychology, 62*, 621.

Kutas, M., & Hillyard, S. A. (1980a). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology, 11*(2), 99-116.

Kutas, M., & Hillyard, S. A. (1980b). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, US: The MIT press.

Misersky, J., Gygax, P. M., Canal, P., Gabriel, U., Garnham, A., Braun, F., ... & Sczesny, S. (2014). Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behavior Research Methods, 46*(3), 841-871.

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*, 9 pages.

Osterhout, L., Bersick, M., & McLaughlin, J. (1997a). Brain potentials reflect incongruencies of gender stereotypes. *Memory & Cognition, 25*(3), 273-285.

Osterhout, L., McLaughlin, J., & Bersick, M. (1997b). Event-related brain potentials and human language. *Trends in Cognitive Sciences, 1*(6), 203-209.

Sato, S., Gygax, P. M., & Gabriel, U. (2013). Gender inferences: Grammatical features and their impact on the representation of gender in bilinguals. *Bilingualism: Language and Cognition, 16*(4), 792-807.

Schmitt, B. M., Lamers, M., & Münte, T. F. (2002). Electrophysiological estimates of biological and syntactic gender violation during pronoun processing. *Cognitive Brain Research, 14*(3), 333-346.

Siyanova-Chanturia, A., Pesciarelli, F., & Cacciari, C. (2012). The electrophysiological underpinnings of processing gender stereotypes in language. *PloS One, 7*(12), e48712.

Vervecken, D., Gygax, P. M., Gabriel, U., Guillod, M., & Hannover, B. (2015). Warm-hearted businessmen, competitive housewives? Effects of gender-fair language on adolescents' perceptions of occupations. *Frontiers in Psychology, 6*, 1-10.

Vervecken, D., & Hannover, B. (2015). Yes I can! Effects of gender fair job descriptions on children's perceptions of job status, job difficulty, and vocational self-efficacy. *Social Psychology, 46*(2), 76-92.

White, K. R., Crites, S. L., Taylor, J. H., & Corral, G. (2009). Wait, what? Assessing stereotype incongruities using the N400 ERP component. *Social Cognitive and Affective Neuroscience, 4*, 191-198.

Abstracts

Proceedings of the Master's Programme Cognitive Neuroscience is a platform for CNS students to publish their Master thesis. 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.

The Role of Entrained Oscillations in Segmenting Rhythmic Sentences During Foreign Language Listening

Sophie Arana, Anne Koesem, Tineke M. Snijders

When hearing a foreign language, listeners often have difficulties segmenting the continuous speech signal into individual words The mechanisms by which word segmentation occurs are largely unknown. Recently, neural oscillatory entrainment to the speech envelope has been proposed as a possible mechanism underlying speech processing. In the current study, we investigated whether the rhythm structure of foreign speech provides cues for word segmentation and whether neural oscillatory entrainment is involved during parsing of the signal. English natives listened to highly rhythmic and highly non-rhythmic Dutch stories. As an index of neuronal entrainment, we computed coherence between the neural activity measured with magnetoencephalography (MEG) and the speech envelope. In order to quantify word segmentation ability, participants performed a forced-choice lexical decision task after each story. They were prompted to recognise a word that had been repeated within each story and distinguish it from a novel word. Further, as a concurrent measure of successful word segmentation, we analysed event-related fields (ERFs) in response to familiarised versus novel words. High response accuracies as well as a suppression of the ERFs in response to familiarised versus novel words indicate that participants were able to segment words in the foreign speech stream. Moreover, we observed a significant increase in speech-brain coherence for rhythmic versus non-rhythmic speech input. Contrary to our predictions, we found no effect of rhythmicity on word segmentation. Thus, our results suggest that word segmentation of foreign speech does not solely rely on rhythmic regularities in the signal. Further research is required to elucidate the underlying mechanisms of processing both rhythmic and non-rhythmic speech and to integrate the role of neuronal entrainment.

Utilising FORCE Learning to Model Adaptive Behaviour

Jaap Buurman, Marcel van Gerven

Humans are able to learn from the environment and show a wide range of adaptive behaviours to solve the task at hand. They learn by trial and error. While reinforcement learning allows for artificial agents to learn via trial and error, they do so with algorithms that might not be the most biologically plausible. Recently, a new algorithm to train recurrent neural networks called FORCE learning has been proposed and this way of learning might be a lot more biologically plausible. We would like to investigate whether we can utilise this new algorithm to model adaptive behaviour. Performance on a set of three toy problems was evaluated and it was shown that these agents were indeed able to learn to perform these tasks. Interestingly, this way of learning showed phenomena that are comparable to phenomena found in biological brains.

The Role of the Orbitofrontal Cortex in the Cognitive Flexibility of Rats Assessed Using Novel Touchscreen-Based Tasks

Sarita Dam, Adam Mar, Johan Alsiö, Trevor Robbins

The orbitofrontal cortex (OFC) is critical for cognitive flexibility that aids organisms to adapt their behaviour within the ever-changing environments. Its role in flexible behaviour is typically assessed using tasks in which subjects must change an established behavioural response in order to adapt to new contingencies. In the present study, two tasks were performed to investigate the role of the OFC in rats. This was assessed using novel touchscreen-based tasks. One of these tasks is serial reversal learning, a test that was performed to explore the role of the OFC when the contingencies of the visual stimuli are reversed, specifically, when the rats receive different doses of the 5-HT2c receptor antagonist SB242084. This task used two visual cues, one positive (CS+) and one negative (CS-). Touching the screen was required for reward delivery. After training, rats received intra-OFC infusions via cannulae during the reversals. In the rats, reversal learning was significantly enhanced when receiving the 1.0 μg/ml dose, but not when receiving the vehicle or 3.0 μg/ml dose. This effect was shown only during the perseverative phase and thus seems phase-specific. The second experiment that was performed is the intra-dimensional/ extra-dimensional (IDED) set-shifting task. This is a test composed of different stages to measure rule acquisition and reversal. Again the task started with two visual cues during training. After that, the rats received exocitotic lesions in the medial orbitofrontal cortex (mOFC), the lateral orbitofrontal cortex (lOFC) or sham lesions. The rats were able to perform intra-dimensional set-shifting stages and therefore were able to form a dimensional set. These set shifting results also yielded differences between the lateral and medial OFC lesioned animals. Lesions in the mOFC showed enhanced intra-dimensional set shifting compared to the lateral and sham lesion groups. The reversal results, on the other hand, showed no clear differences between any of the groups. These results provide the first direct evidence from touchscreen-based tasks for the involvement of the 5-HT2c receptor in enhancing the perseverative phase of reversal learning. Moreover, they provide evidence for dissociable functional roles of the mOFC and lOFC during the performance of intra-dimensional set shifting.

The Role of Spatial Representations in MTL Regions in Generalising Fear Responses

Anouk van der Heide, Erno J. Hermans

Not only exaggerated reactions to threats, but more importantly deficits in suppressing fear in safe contexts can lead to fear-related disorders. It is therefore important to study the link between threat and context encoding. Animal studies demonstrated the importance of hippocampal place cells in representing contextual information in fear learning. Human functional magnetic resonance imaging (fMRI) virtual reality (VR) studies conducted so far showed that spatial locations within VR environments could be decoded from the (para)hippocampus and that this region also showed different activity patterns for threat and safe contexts, suggesting a role of medial temporal lobe (MTL) regions in contextual encoding. The relation between contextual representations and fear generalisation has not been investigated. We hypothesised that inadequate context representations in MTL regions play a major role in overgeneralisation of fear. To assess how the strength of contextual representations influences fear generalisation, we designed a differential contextual fear conditioning VR paradigm for fMRI and tested it in 24 healthy individuals. We investigated whether we could decode the contexts within the VR from patterns in four MTL regions. Fear generalisation was calculated from skin conductance responses (SCR) and compared the differential conditioning response between the safe and threat context. Multi voxel pattern analysis (MVPA) was used to decode activity patterns from MTL regions in the two contexts, where accurate decoding meant that the area contained information to dissociate the contexts. Our results demonstrate that spatial location could be decoded from patterns in the right entorhinal cortex, but not in other MTL regions. No correlation was found between individual SCR generalisation scores and decoding accuracies in any MTL region. In conclusion, contextual information seems to be encoded in MTL regions but it remains unclear how this relates to fear generalisation.

Evaluation of Phenotypic Abnormalities Among Patients with Intellectual Disability

Claudia J.M. Hellebrekers, Sandra Jansen, Jayne Y. Hehir-Kwa, Bert B.A. de Vries

The development of more advanced genetic techniques like the single nucleotide polymorphism (SNP)-array and genome wide sequencing techniques contributed largely to the identifying of genetic causes of intellectual disability (ID). Interpretation of the genetic variations found with these techniques, however, remains difficult. Multiple patients with the same phenotype and genotype are needed to decide about a variant's pathogenicity. It is although difficult to determine how many patients are needed for a statistically significant conclusion. More information about the specificity of phenotypic features will greatly improve these considerations, as genetic variants in combination with rare features can be declared pathogenic when fewer patients are found than variants occurring in combination with common features. However, at the moment, no comprehensive overview of the phenotype of patients with ID is available. This article provides an overview of the specificity of phenotypic features by an extensive evaluation of the phenotype of a large cohort of 7407 well-defined patients with ID. From 1320 of these patients very detailed information was available, these patients formed a best-defined group used for the analyses, the other 6087 a well-defined replication cohort. An overview of the frequency of all Human Phenotype Ontology (HPO) features in these patients is presented. In addition, we found 1534 combinations of positively associated features. These associations consisted of expected associations like Cleft lip (HP:0000204) and Cleft palate (HP:0000175; $p < .001$) but also unexpected associations like Short stature (HP:0004322) and Abnormality of the forebrain (HP:0100547; $p < .001$). Most of these 1534 combinations were also associated in the replication-cohort. We also proved that half of the major anomalies registered by Eurocat occur more in patients with intellectual disability than in the normal population and that patients with more congenital anomalies have a higher chance of having facial dysmorphism. This extensive description of phenotypes of intellectual disability patients provides an overview for clinicians that can contribute to the deliberation about an intellectual disability patient's diagnosis, which is valuable for the patients and their parents.

The Effect of HIV Infection on Frontostriatal Inhibitory Control in the Era of Combined Antiretroviral Therapy

Birgit Hermsen, Matthijs Vink

Studies on brain activity and connectivity in Human Immunodeficiency Virus (HIV) suggest that frontostriatal dysfunction plays a role in HIV-associated cognitive deficits. However, almost no research has been done to investigate frontostriatal functioning in a HIV population that is virally suppressed with combined antiretroviral therapy (cART). In this study, 27 HIV-positive (HIV+) participants, stable on cART and with an undetectable viral load, performed a stop signal anticipation task (SSAT) while being scanned with functional magnetic resonance imaging (fMRI). This task activates the striatum during reactive inhibition (i.e., outright stopping of a response) and engages the frontostriatal network during proactive inhibition (i.e., anticipation of stopping). Results showed normal striatal activity in HIV+ participants during both reactive and proactive inhibition. During proactive inhibition, HIV+ participants showed less reaction time slowing during anticipation of a stop signal. This was paralleled in the brain by decreased dynamic activation in the right inferior frontal gyrus (IFG) with increased stop-signal probability, as well as decreased functional connectivity between the right striatum and the right IFG. These results suggest normal striatal functioning in a virally suppressed HIV population, but dysfunctional frontal activity and decreased frontostriatal connectivity. These findings serve as a stepping stone to research that distinguishes between the effects of HIV-infection and cART on frontostriatal functioning.

Can I Predict Your Future?: Behavioural and Neurophysiological Evidence for Prediction of Observed Actions

Margreeth Hidding, Ricarda Braukmann, Edita Poljac, Sabine Hunnius

To be able to engage in social interaction with peers, we need to predict their movements and their action goals. Several studies have already shown that, when observing someone else move to act on an object, the observers show anticipatory eye movements towards the goal object before the hand of the actor arrives. Moreover, it has been shown that these anticipatory fixations become faster as the action sequence evolves. Furthermore, it has been found that one's own motor cortex becomes active not only when performing, but also when observing someone else perform an action. For the current study we combined eye tracking (ET) and electroencephalography (EEG) to study the behavioural and neurophysiological basis of how one predicts complex multi-step actions of others. In contrast to previous studies, the eye tracking data did not show a significant pattern of anticipatory fixations. There was no significant difference between the action steps and thus no indication of accumulation of evidence visible in the data. However, the EEG data demonstrated a significant increase of suppression over the motor cortex in both the mu (8 - 13 Hz) and beta (16 - 25 Hz) band and a significant difference between action steps. This suggests the presence of a neural predictive mechanism, aiding understanding in complex actions, whereas the behavioural evidence for anticipation of action end-goal seems not so well established yet.

Can Stress-Induced Amygdala – Frontal Connectivity Changes Predict Perceived Stress?

Iris Hulzink, Wei Zhang, Floris Klumpers, Karin Roelofs

Stress-related disorders are one of the most common mental disorders, and clarifying the influence stress exerts over our physiological systems could help screen vulnerable individuals and aid in early detection of disorders. To elucidate how stress affects our brain, we explicate an amygdala – frontal cortex connectivity model, distinguishing between their subregions. Participants (N = 71, mean age 24.5, 19 females) underwent two resting-state scans with a well-established stress induction protocol in between, to be able to detect changes in connectivity due to stress. These connectivity changes were then related to perceived stress later in life. Connectivity between the basolateral amygdala (BLA) and the dorsal and ventral frontal networks increased significantly due to stress induction. Stress-induced connectivity increases of the BLA and centromedial amygdala (CMA) with the dorsal network were marginally predictive of perceived stress, however, only for a subsample of our participant pool ($p = .07$ and $p = .06$ respectively). This indicates that stress-induced amygdala – frontal connectivity changes could be predictive of later perceived stress. Future research could focus on increasing sample size and predictive validity of amygdala – frontal connectivity in relation to other stress-related symptomatology.

Decoding Spatial Representations From Functional Magnetic Resonance Imaging Data

Steffen Kaiser, Tobias Navarro Schröder, Jacob Bellmund, Sander Bosch, Christian Doeller, Marcel van Gerven

The hippocampus is essential to our ability to navigate our environment. Rodent research on the underlying hippocampal code has revealed specialised place, grid, and head direction cells. Recent studies showed that humans employ a similar memory network to encode spatial representations in hippocampal regions. In this study we used recurrent neural networks to decode the trajectories of participants navigating a virtual environment from functional magnetic resonance imaging (fMRI) data. Encoding models were implemented trying to encode evoked blood-oxygenation-level-dependent (BOLD) responses from location feature sequences. Subsequently, decoding models were trained to predict the location sequence from BOLD responses. Although the models fell short in achieving sufficient performance, the study provides useful insights to improve future approaches to decode trajectories from fMRI data.

Sensitivity to Amphetamine Depends on a Lack of Serotonin Transporters

Janita Kelderhuis, Michel Verheij

Amphetamine abuse is a worldwide problem. Serotonin and dopamine are important neurotransmitters that are implicated in the acquisition, maintenance and relapse phase of amphetamine addiction. By blocking and reversing the action of dopamine reuptake transporters and reversing the action of serotonin reuptake transporters, amphetamine increases extracellular dopamine and serotonin levels. The present study aims to give insights in the role of the serotonin transporter and accumbal serotonin in amphetamine reward and addiction by combining microdialysis and behavioural testing in serotonin transporter knock out rats (SERT KO) and their wildtype (SERT WT) counterparts. In this study, SERT KO and SERT WT were subjected to an acute amphetamine challenge (2.5 mg/kg) while locomotor activity and ultrasonic vocalisations were measured as well as extracellular serotonin and dopamine levels were measured in the nucleus accumbens shell. A separate cohort of SERT KO and SERT WT was subjected to amphetamine self-administration (0.03 mg/kg/infusion) followed by a progressive ratio test to measure motivation. Microdialysis revealed an increase of both serotonin and dopamine levels after amphetamine (2.5 mg/kg) administration, but no significant genotype differences were found. However, the increase in extracellular accumbal serotonin, but not dopamine levels of SERT KO animals was more prolonged. SERT KO rats emitted more 50 kHz ultrasonic vocalisations and had higher locomotor activity compared to SERT WT rats after amphetamine. Self-administration revealed that SERT KO rats self-administered more amphetamine (0.03 mg/kg/infusion) compared to SERT WT rats. SERT KO animals showed a quadratic increase in their intake, whereas SERT WT showed a linear increase in their intake. Furthermore, SERT KO scored higher on the progressive ratio test.

All our results linked together support the hypothesis that SERT KO rats are more sensitive to amphetamine, which may lead to an increased intake compared to SERT WT rats. Accumbal dopamine levels did not differ between the genotypes, which may imply an important role for accumbal serotonin in individual differences in drug reward and sensitivity. While being careful generalising to humans, our findings could have implications for individualised treatments for patients with amphetamine addiction.

Sense of Agency in Infancy: Testing for Causal Model Building

Falma Kemalasari, Sabine Hunnius, Lorijn Zaadnoordijk

The sense of agency has been described as the experience that is sensed when we perform our own actions, such that we control events in the outside world through them. The development of sense of agency is crucial for children to be able to learn from their experience and interaction with the surrounding world. To date, evidence of a sense of agency in infancy is limited to a behavioural data pattern. However, a simulation with no capability to build internal model was able to replicate a similar data pattern. Therefore, additional investigation on different phenomena is necessary. It is assumed that to experience a sense of agency, an internal model is required. Having an internal model would enable one to make predictions upon it. When there is a mismatch between the predicted events with the actual events, a mismatch response in terms of the brain's event-related potential (ERP) is elicited. In this study, we proposed to investigate the presence of sense of agency in infants based on how the brain would react to a violation of its internal model. In this study, we investigated 3-to-4-month old infants' movement response along with the ERP response to movement in different experimental phases. There were three phases, namely baseline, connect, and disconnect. One of the infants' limb movement would trigger stimulus effect during the connect phase, whereas baseline corresponds to the phase before and disconnect corresponds to the phase after. The omission of the movement's effect was applied in the disconnect phase and aimed to violate infants' expectations. The results indicated infants experience violation of expectation in response to the omission of their movement effect, as shown in the ERP mismatch response. It indicates that infants build an internal model linking their movement with its effect, suggesting the presence of sense of agency.

Temporal but not Spatial Expectation Modulates Bottom-Up Attention

Felix Klaassen, Erik Meijs, Simon van Gaal, Floris de Lange

In the attentional literature there has been considerable debate about whether bottom-up attentional processes are purely stimulus driven or dependent on top-down sets and goals. One possible top-down factor that might influence bottom-up attentional processing is expectation. In this study we investigate the relationship between bottom-up attention and expectation, specifically asking whether bottom-up attention can be explained by a prediction-error. We hypothesise that unexpected stimuli generate a larger prediction error than expected stimuli, consequently leading to more attentional capture. In two experiments we used an exogenous cueing paradigm to investigate if spatial and temporal expectations about a distracting cue modulate the amount of attentional capture. In the first experiment we investigated the effect of spatial expectation on bottom-up attention, but found no evidence for an interactive relationship. In the second experiment we focused on the temporal predictability of the cue, and found a modulation of bottom-up attention by cue predictability. Specifically, unpredictable cues lead to more attentional capture compared to predictable cues, though only for a specific range of cue-target stimulus onset asynchronies. With these findings, we provide a first direct indication that expectation influences bottom-up attention, although the exact mechanism underlying the modulation is not yet clear. We propose a modulatory account of the interaction between expectation and bottom-up attention, and suggest that bottom-up attention is in principle stimulus-driven but can be modulated by expectation if the temporal relationship between the cue and the target is optimal.

Redox Imbalance, Myelin Deficits and Impaired Cognitive Function in a Rat Model for Schizophrenia

Kimberly M. A. de Kleijn, Astrid Vallès, Gerard Martens

Current treatments of the neurodevelopmental disorder schizophrenia are mainly targeted against the positive symptoms of the disorder. The treatment of cognitive symptoms, however, is still an unmet need. Moreover, the underlying mechanisms of how these symptoms develop are not fully understood. Recent literature suggests an involvement of redox imbalance, myelin deficits and oligodendrocyte abnormalities in this process. Our hypothesis poses that a redox imbalance might lead to stress in oligodendrocytes and myelin deficits, with subsequent cognitive symptoms. We used the APO-SUS rat line as a model for schizophrenia, with the APO-UNSUS counterpart animals as controls. We investigated whether deficits in redox-related gene, myelin-associated protein and oligodendrocyte transcription factor mRNA expression are present in several brain regions of female APO-SUS rats of pre- and post-adolescent ages. In medial prefrontal cortex and the dorsal striatum of female APO-SUS, myelin-associated protein mRNAs were downregulated at several ages. At these ages, mRNAs of two redox-related genes, Gstm4 and Prdx6, also showed dysregulated expression. Furthermore, the presence of cognitive deficits in the executive functioning domain was evaluated by a perceptual discrimination paradigm and the possibility for consequent remyelination was explored at the transcript level in APO-SUS and APO-UNSUS males. We found that male APO-SUS showed lower performance and shorter inter-trial intervals after an intra-dimensional shift in the perceptual discrimination paradigm. However, no increase in mRNA expression of myelin-associated proteins in the medial prefrontal cortex was observed after extensive neuronal activation. Lastly, myelin deficits at different developmental time points were investigated at the structural level in male APO-SUS by Third Harmonics Generation Microscopy validated by immunohistochemistry. Third Harmonics Generation microscopy showed differences in third harmonic signal contrast between male APO-SUS and their APO-UNSUS counterparts, and was able to visualise the developmental event of myelination. Summarising, the present study clearly suggests a link between redox imbalance and oligodendrocyte and myelin deficits. Further studies into this hypothesis are necessary to aid in the development of remyelination strategies that could ultimately alleviate cognitive symptoms in patients with schizophrenia.

The Interaction Effect of MAOA and Maltreatment on Aggression Subtypes, and Their Neural Correlates

Niels Las, Marjolein van Donkelaar, Barbara Franke, Jan Buitelaar

Aggression poses a major problem for society, through both financial costs and emotional distress. Although the genetics and etiology of human aggression are poorly understood, a polymorphism in the gene encoding monoamine oxidase A (MAOA) has been implicated through its effect on MAOA-expression. In an influential example of a gene-environment interaction, males with low MAOA-expression show a larger increase in aggressive behaviour in response to childhood maltreatment than those with high MAOA-expression. It has not yet been studied how, and through which biological mechanisms this interaction affects the reactive and proactive subtypes of aggression described in literature. Several subcortical brain regions have been implicated in aggression, most notably the amygdala, hippocampus, caudate nucleus and nucleus accumbens. We hypothesised that MAOA and maltreatment exert their effect on behaviour through influencing volume of these brain structures. We aimed to test the interaction effect of MAOA and maltreatment on subtypes of aggression, as well as the volumes of subcortical regions of interest. We also investigated whether the volumes of these regions were associated with aggression measures in our sample. Additionally, we tested whether functional connectivity in aggression-related brain circuits was associated with aggression subtypes. The 30bp MAOA uVNTR was genotyped in a sample of healthy adults with available structural and resting-state magnetic resonance imaging (MRI) data. For partially overlapping subsets of the sample, information on maltreatment and on aggression was available. Maltreatment was assesed by the List of Threatening Events and aggression was measured by the Reactive Proacive Questionnaire, and the Inventory of Callous-

Unemotional traits. We performed confirmatory factor analysis in order to test the two-factor and three-factor models of aggression proposed in literature (N = 661). General linear model and logistic regression were used to assess the effect of MAOA genotype and maltreatment on aggression measures (n = 82) and the subcortical volumes of interest (n = 258). General linear model was also used to examine the association between aggression measures and the volume of these subcortical regions (n = 574). Permutation testing was used to test the association between resting-state connectivity and aggression subtypes (n = 124). Confirmatory factor analysis showed a satisfactory fit for both the two-factor and the three-factor model, with a superior fit for the latter. No interaction between MAOA genotype and maltreatment was found for any of the aggression measures. However, this interaction did affect the left nucleus accumbens, where maltreatment was associated with a volume increase in low- but not in high-MAOA-expressing subjects. Aggression measures were not significantly associated with the volume of subcortical regions of interest or with resting-state functional connectivity. The current findings show that the reactive-proactive aggression distinction is valid, but that aggression is better described by a three-factor model in the current sample. A novel finding is that maltreated subjects with low MAOA-expression show a volume increase in the left nucleus accumbens, which may, in light of earlier findings, mediate increased aggression. No subcortical regions of interest were associated with aggression subtypes in the current sample. Resting-state functional connectivity was also not significantly associated with reactive or proactive aggression. Future research is needed to identify how, and through which biological mechanism the interaction of MAOA genotype and maltreatment affects aggression subtypes. Additionally, more research is needed to characterise the neural correlates of aggression subtypes.

Effects of Dopaminergic Medication on Reward and Punishment Sensitivity in Risky Decision Making

Karita E. Ojala, Guillaume Sescousse, Roshan Cools

Pathological gambling (PG) is a behavioural addiction similar in many aspects to substance use disorder. PG involves excessive risk taking and (monetary) reward. The neurotransmitter dopamine is of interest in relation to risk taking and PG due to its central role in learning from reward and punishment. Moreover, some Parkinson's disease patients develop PG following dopamine replacement therapy. The aim of this study was to examine the effects of dopaminergic modulation on risk-taking behaviour in healthy and PG individuals. Dopamine D2/D3 receptor antagonist sulpiride and a placebo drug were administered in order to transiently alter dopamine transmission during an economic decision-making task. Participants chose between sure choices of winning (or losing) a certain amount of money and gambles with different probabilities to win (or lose) money. A prospect theory modelling approach was used to estimate parameters reflecting sensitivity to outcomes and probabilities and optimism about risk, based on the varying amounts of money and probabilities in the task. We found that sulpiride decreased distortion in weighting the probabilities of potential gains. That is, participants overweighted low and underweighted moderate to high winning probabilities less in the sulpiride condition compared with placebo. However, the drug effect did not differ between the groups and was not found in the loss domain. In conclusion, we found evidence for a relationship between dopamine and risky decision making in the distortion of probability weighting.

The Segmentation Problem: Neuronal Entrainment as Underlying Mechanism for Rhythm-Based Word Segmentation in 9-Month-old Infants

Lisa Y. A. M. Rommers, Tineke M. Snijders

Infants need to start recognising words from continuous speech in order to learn their native language. As continuous speech does not have an equivalent to spaces in written text that clearly mark word boundaries, infants can exploit other cues, such as prosody cues, to segment the speech stream. So far, little is known about why infants can rely on prosodic cues in a very early language acquisition phase. Recent work proposed that neuronal oscillations that entrain to slow temporal prosody-specific (1 - 2 Hz) modulations in the speech stream can facilitate language processing. In this study, we investigate whether infants' segmentation ability is modulated by the prosodic structure of speech. We presented 9-month-old infants with rhythmic and non-rhythmic speech and investigated whether their electroencephalography (EEG) word recognition effect was modulated by the different prosodic presentations. Although we did not find significant evidence, our results suggest that infants' word segmentation ability is enhanced when speech is presented rhythmically. We speculate that neuronal entrainment in the delta frequency band (1 - 2 Hz) could possibly be the underlying mechanism of rhythm-based word segmentation in infants.

Exploring Oscillatory Signatures of Prediction and Prediction Error in Language Processing

Ksenija Slivac, Sybrine Bultena, Egbert Hartstra, Irina Simanova, Harold Bekkering

Language comprehension is a quick and dynamic process. The predictive coding theory suggests that we regularly make use of contextual clues in order to predict the upcoming content. In this study, we have used magnetoencephalography (MEG) to investigate oscillatory changes connected to predictions and prediction errors in discourse comprehension. While participants read short, semantically manipulated stories, we measured beta and gamma oscillatory powers as respective indicators of adaptation and prediction strengthening throughout the stories, and the one-off surprise, error-triggering effect, at the very end of the story. As hypothesised, we found evidence of adaptation and prediction strengthening, reflected in the significant beta band power increase as the reader becomes familiar with the content of the story. Furthermore, contrary to our expectations, we found a significant increase in both low and high gamma power in the same direction as beta band finding. This finding, while failing to confirm gamma frequencies as indicators of surprise in language comprehension, was in line with previous sentence comprehension studies that suggested gamma band to be indicative of the match between our expectations and input. Finally, we failed to find any difference in gamma power at the occurrence of the one-off surprise factor incongruent with the rest of the story. Our findings confirm a partial, yet consistent discrepancy between the predictive coding hypothesis on prediction and prediction errors oscillatory signatures, and evidence from the language comprehension research.

Decoding Noncommutative 3D Perceptual Consequences of Commutative 2D Swiping Movements

Katrin Sutter, Sara Fabbri, Pieter Medendorp

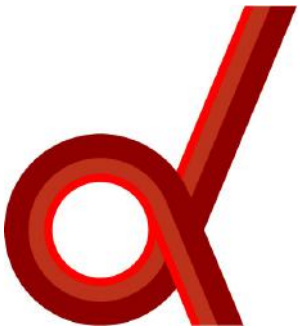
Interaction with objects is a skill that humans master. Rapidly developing touchscreen technology adds a new computational challenge to this expertise by allowing us to interact with virtual three-dimensional (3D) objects. On a touchscreen, reversing the order of swiping movements does not change the end point of the finger on the screen (commutative property of swiping movements). However, when interacting with a virtual 3D object, augmented by the screen, reversing the order of swiping movements leads to a different end orientation of the object (noncommutative property of rotations). The aim of our study was to unravel neural circuitry that transforms commutative actions into noncommutative visual consequences. Using functional magnetic resonance imaging (fMRI), we measured brain activity while participants performed swiping movements to rotate a virtual 3D object on a touchscreen. By manipulating the order of the swipes and the starting orientation of the object, we studied commutative and noncommutative processes. We used representational similarity analysis to find brain regions with similar activation patterns to our modeled prediction. Whilst our main results did not reach significance, they provide a tentative insight into brain regions that could be involved in coding perceptual commutative and noncommutative properties. They suggest that right inferior parietal lobule may be involved in noncommutative processes. Clusters in the right precunus, cuneus, lingual gyrus, and superior frontal gyrus respond following the commutative model.

Can Neural Activity During Encoding Under Stress Predict Memory Specificity over Time?

Hanjian Xu, Linda de Voogd, Erno Hermans

Events under stress are in general better remembered, but the effect of stress specifically on detail and gist information of the events is still inconsistent among studies. The mixed findings together may be explained by the fact that the quality of memories changes over time. Declarative memory depends less on the hippocampus over time, and this region is regarded essential for detailed memory suggested by multiple studies on pattern separation. Taking into account that stress-induced noradrenaline increase can enhance synaptic plasticity in the hippocampus, we hypothesised that details might initially be better remembered if encoded under stress, and a stronger gist may remain over time. Correspondingly, we predicted that encoding neural activity in the hippocampus should reflect such modulation, specifically be able to better predict subsequent detailed memories encoded under stress. The present study was designed to test the hypotheses in humans using event-related functional magnetic resonance imaging (fMRI). We investigated neural activity during memory formation in 49 healthy participants in an incidental encoding task embedded in either a neutral or stressful context. A subsequent recognition test containing identical targets, similar lures and novel foils, was performed after 24 h and 1 week by asking participants to respond 'new', 'similar', or 'old' per stimulus. Functional MRI and heart rate were acquired throughout procedures. Blood pressure and a set of questionnaires were obtained additionally at specific times. Our data show that heart rate frequency and the negative-affect scores of Positive and Negative Affect Schedule (PANAS) were higher in the stress group compared to the neutral group, but we did not find group differences in blood pressure and heart rate variance. These together indicated a moderate effect of stress induction. We used recog index and patsep index, calculated from response proportions in the recognition test, to respectively represent gist and detailed memory. No difference of two indices were detected between the neutral and stress groups at both delays. As for the encoding neural activity, we did not find different hippocampal activation in target trials that were later remembered than in trials later forgotten, a contrast for gist memory subsequent memory effect (SME). Critically, there was no difference of hippocampal activation in response to lures that were later responded as 'similar' than 'old', defined as an SME representing pattern separation process. Stress, moreover, showed no effects on both SMEs in the hippocampus. To conclude, during-encoding stress in this study neither showed improvement for gist memory, nor enhanced effect of detailed memory. Hippocampal activity is not predictive for the subsequent pattern separation, and is not further influenced by stress induction.

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
www.ru.nl/donders/

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 Research Center
Montessorilaan 3
6525 HR Nijmegen

P.O. Box 9101
6500 HB Nijmegen
http://babyresearchcenter.nl