Transcranial direct current stimulation to boost working memory

Talsma, L.J.

Publication date
2018
Document Version
Final published version
License
Other

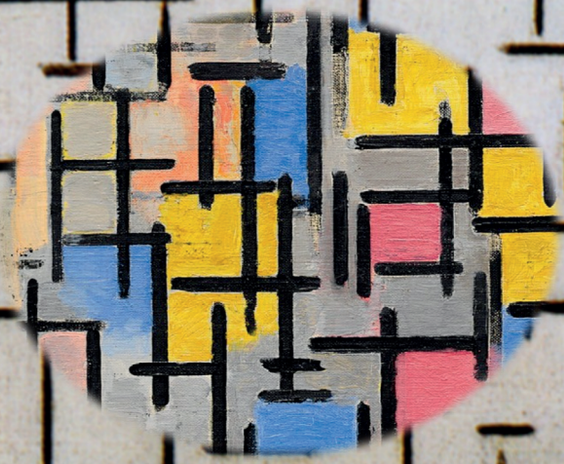
tDCS to boost working memory

tDCS to

boost

working

memory



lotte talsma

lotte talsma

Transcranial Direct Current Stimulation to Boost Working Memory

Lotte Talsma

Cover images:

Piet Mondriaan, 1915, *Pier en oceaan (Composition No. 10)*

Piet Mondriaan, 1916, *Composition*

Transcranial Direct Current Stimulation to Boost Working Memory

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

Prof. Dr. Ir. K.I.J. Maex

ten overstaan van een door het College voor Promoties ingestelde commissie,

in het openbaar te verdedigen in de Agnietenkapel

op dinsdag 12 juni 2018, te 12.00 uur

door

Lotte Joanne Talsma

geboren te Rotterdam

Promotiecommissie

Promotores:	Dr. H.A. Slagter	Universiteit van Amsterdam
	Prof. dr. K.R. Ridderinkhof	Universiteit van Amsterdam
Overige leden:	Prof. dr. A. T. Sack	Universiteit van Maastricht
	Prof. dr. L.S. Colzato	Ruhr University Bochum
	Prof. dr. J.M.J. Murre	Universiteit van Amsterdam
	Prof. dr. R.W.H.J. Wiers	Universiteit van Amsterdam
	Dr. J. O'Shea	University of Oxford
	Dr. I.G. Sligte	Universiteit van Amsterdam
Faculteit:	Faculteit der Maatschappij- en Gedragswetenschappen	

Have no fear of perfection – You'll never reach it.

(Salvador Dali)

Als het brein simpeler te begrijpen zou zijn, zouden wij te simpel zijn om het te begrijpen.

(Onze Taal)

Contents

Chapter 1

Introduction

Background

Working memory in an everyday situation

Imagine that you enter Amsterdam Central Station. For me, since I eternally seem to be running late, at this point usually the main question in my head is: ‘Can I still catch that train?’. To decide on the answer to a question like this, a first step may be to determine the status of the train that you want to catch. Often, you can find this information somewhere on a computer screen within 2 minutes walking distance, but if you are really in a rush, checking the NS-app perhaps provides the fastest way to obtain the desired information: ‘13.38, platform 8: Intercity to Groningen’. While searching for this information one way or the other (the board or the NS-app), chances are that at some point stumble upon the current time (‘13.32’) and combining these two may lead you to conclude: ‘You still have 6 minutes to get to platform 8. Go!’

In doing the math there you probably already used your working memory; a kind of mental place to store information to work from. Namely, putting these two times next to each other on a temporary mental whiteboard will have allowed you to calculate the difference. Now assume that after you solved this first problem, something else pops up into your head: ‘Oh! Actually.. I could very much use some coffee...’. You look around. Remembering that you have 6 minutes to get to the platform, you search the station for a place that may serve you coffee in time. Now, here is where the real working memory madness starts, with a multitude of options to consider and simultaneously keep in mind.

Perhaps you look to your right and about 50 meters away, you see a Starbucks with a short queue of neat looking businessmen (‘Starbucks: 1 minute walking. 2 minutes waiting time. 3 more minutes to Platform 8. Oh, that will be tight! But I love Starbucks cappuccino…’). To your left at about the same distance, you see an Albert Heijn packed with middle-school boys (‘AH: 1 minute walking (but wrong way!), 5 minutes waiting time. 4 minutes to Platform 8. Not an option.’). Also, while your eyes are scanning the area, your mind remembers that you have been on Platform 8 before and that you probably saw a Kiosk there at that time (‘Kiosk: 4 minutes walking. Waiting time unknown. No extra walking afterwards...’). Whilst you weigh the different coffee-options and lay out your possibilities within the grander scheme of catching your train, your brain will temporarily hold these

pieces of information for you on a mental storage board (i.e. the imaginary whiteboard) where they remain easily accessible for you to shuffle around with until you have successfully solved your puzzle. It may do so very efficiently, allowing you not just to add new information ('Option 1: Starbuck… Option 2: AH…), but also to remove information that has become irrelevant ('Option 2 will cause you to miss your train. You don't want that. Therefore, DELETE Option 2') and update your working memory with relevant new information again at any time in the process ('New Option 2: Kiosk at Platform 8').

Moreover, in allowing you to derive at a final decision, your brain will not only store the information that you gathered from the outside world at that moment ('The Starbucks and an Albert Heijn'), but relate this to information already inside your memory ('The Kiosk at Platform 8') as well as to your particular internal goals and desires at that moment ('Ok, I will have to let go of the idea of my favourite Grande Soy Cappuccino and go to the Kiosk instead. Right now, my goal of catching the train is the most important. A regular Kiosk espresso is the safest option and will do just fine.').

Whether involving big decisions or not, our days are almost continuously filled with situations like this coffee-and-train example, with circumstances that require us to briefly maintain and work with particular pieces of information in our minds. Thus, even though most of the time you may not be consciously aware of it, throughout a typical day your working memory is actually working very hard for you.

Working memory in the brain

In the scientific field of Cognitive Psychology, we use the term working memory (WM) to refer to the multimodal capacity to retain and monitor information over brief periods of time (Baddeley, 1996). As such, WM is considered a core faculty of cognition, that is of fundamental importance to many other higher-order cognitive functions (Engle, 2002; Kane & Engle, 2002). Moreover, as the example above illustrates, as WM provides an interface between perception, long-term memory and action, it is an aspect of cognitive functioning that we critically rely on in many everyday life situations.

In the traditional model of WM, WM is subdivided into two storage units: a verbal and a visuospatial one, and a control structure termed the 'central executive' (Baddeley, 2003; Baddeley, Sala, Robbins, & Baddeley, 1996). In contrast to the maintenance compartments

(i.e. the whiteboards themselves), the central executive is considered to be responsible for updating and manipulating the contents of our WM (i.e. the marker that writes things on the board and the sponge that wipes the board clean again). In the example above, the central executive function thus allows the maintenance of all new information into working memory ('Option 1: … Option 2: … '), reorganisation of existing information according to task demands and desires ('DELETE option 2'), as well as the use of this information to control automatic behaviour ('Out of habit I may be drawn towards Starbucks, but actually going to the Kiosk is the better option for me today').

With regard to the brain, the function of WM is thought to rely on a network of brain regions centred on the prefrontal and parietal cortices (Baddeley, 2003). Moreover, with regard to the process of executive control, the dorsal lateral prefrontal cortex (DLPFC) is considered to play a key role (Baddeley, 2003; Kane & Engle, 2002; Smith & Jonides, 1999). Functional neuroimaging studies in humans consistently report activity in the DLPFC for a range of WM tasks that involve a variety of WM content (for meta-analytical reviews see e.g., Owen, McMillan, Laird, & Bullmore, 2005; Wager & Smith, 2003). In line with this, lesion studies in monkeys indicate that the DLPFC may be specifically involved in the manipulation of information in WM (Petrides, 2000). Similarly, a recent large study with human lesion patients concluded that also in humans, the DLPFC is vital for the manipulation of knowledge in WM, both in the verbal and spatial domain (Barbey, Koenigs, & Grafman, 2013). The DLPFC may thus play a critical role in domain-general central executive processes (i.e. this may be a key area if we were to link the marker and sponges to the brain).

Individual differences in working memory

WM is pivotal to the majority of our day-to-day behaviours. Nevertheless, from our own experience it is already apparent that not all of us may always use WM to an optimal degree. Whereas some people always seem to be able to remember that grocery list after one glance, others notice that whenever they get to the fourth item, the first one has

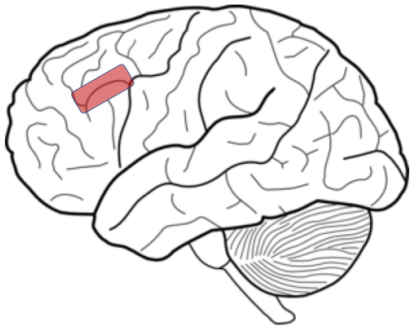


Figure 1. *The dorsolateral prefrontal cortex (DLPFC). The DLPFC plays a key role in the control and manipulation of information in WM (i.e., central executive processes), in interaction with other prefrontal as well as parietal areas.*

already slipped their mind again. Also, in the case of the train-and-coffee example, whereas some people may peacefully enjoy their train ride afterwards, daydreaming about that beautiful Italian Roman square where they enjoyed that 'most amazing espresso of their life', in others, thoughts about train times and coffee-options may keep popping up in their heads ('Option 2: AH! No good!! DELETE DELETE DELETE!').

In other words, substantial differences may exist between people with regard to both the amount of information that they can store at a given moment (i.e., the size of the whiteboard to work with, also termed Working Memory Capacity), as well as how well they can control the content of their WM (i.e. adding new information and wiping out old information by the central executive) (Baddeley, 2003). Although minor failures of our working memory may not make a crucial difference in most situations (It is ok to check your NS-app twice), adequate WM, however, may be of more importance in other cases (e.g., If getting distracted by the idea of Starbucks coffee would cause you to for a moment completely forget about catching your train, leading you to miss that important job interview…).

It has been speculated that impaired WM lies at the core of behavioural difficulties in many psychiatric conditions, such as schizophrenia, for which impaired WM functioning may be responsible for a lack of successful pursuit of long-term life goals (Barch & Ceaser, 2012). However, also in the healthy human population, proper WM functioning may yield many significant benefits. For example, correlational studies have shown that WM functioning may be related not only of performance in many other cognitive functions, like reasoning and math ability, but may even determine academic as well as professional life success (see Baddeley, 2003). Moreover, as WM performance is known to decrease in older age (Park et al., 2002), we are all bound to at some point experience that day where we really need to start our visit to the grocery store by taking out that written grocery list, because we cannot rely on our mental version of this anymore.

Thus, considering the implications of poor WM in virtually any situation in life, investigating ways to optimize WM functioning may be of direct societal and clinical relevance. Discovering ways in which WM can be improved may not only fundamentally help individuals that struggle daily with failures of their WM functioning, but possibly eventually also aid cognitive functioning in the healthy (aging) population.

Can we improve our working memory in adulthood?

For a long time, it was assumed that after our childhood the workings of our brains are relatively fixed. We may acquire new skills and memories, but the anatomical structures in our brain were thought to remain pretty much the same. However, two decades ago revolutionary research was done in monkeys that for the first time indicated that the plasticity of brain regions such as the sensorimotor cortex is not restricted to critical periods during childhood development, but that the capacity of brains to change in structure and function is in fact at least to some extent retained throughout adulthood (for an overview see e.g. Buonomano & Merzenich, 1998).

That this may also hold true in humans was first shown in a ground-breaking study that revealed that London taxi drivers, who are required to remember the complete complex roadmap of the city of London, have significantly larger hippocampi, a part of the brain that is critical for spatial memory (Maguire et al., 2000). Furthermore, that such

structural changes may actually be brought about relatively fast, was demonstrated in subsequent studies. For example, a study with novice jugglers showed that alterations could be found in both grey (neuronal cell bodies) and white matter (myelinated axons) in movement related visual areas (such as M1) after a mere 7 days of practice with three ball juggling (Bogdan Draganski et al., 2004; Driemeyer, Boyke, Gaser, Büchel, & May, 2008).

In other words, the adult brain maintains the capacity to learn and reorganize itself as a result of experience in the form of functional and structural neuroplasticity, albeit to a far lesser degree than in the initial years (for a deeper investigation of the underlying processes of neuroplasticity in adulthood, see e.g., Draganski & May, 2008). The majority of the initial research in this field primarily focused on experience or training-related changes in the motor and sensory domains, e.g. after new skill learning. Nevertheless, neuroplastic changes have also been observed in brain regions involved in higher-order cognitive functions, such as prefrontal and parietal regions, after short periods of intensive cognitive effort such as after studying for exams (Cecarrelli et al. 2009) or practicing mental calculations (Takeuchi, Taki, Sassa et al. 2011). Although preliminary, this research has paved the way for the possibility that also the brain regions involved in higher-order cognitive functions, like WM, still exhibit neuroplasticity.

Initial attempts with intensive computerized WM task training

To challenge the brain to change its fundamental circuity it is important to push it to its capacity limits and to promote neurocognitive plasticity, any type of training should thus be highly challenging (Lovden et al. 2010). The first attempts to improve WM have focused on designing effective intensive computerized WM task training regimes. In such regimes, subjects typically practice one or more WM tasks for several days a week for a number of weeks, preferably with a wide range of stimuli and task settings to promote domain-general learning.

Initial results with such WM trainings indicated that in healthy adults, behavioural WM training might lead to wide-spread cognitive improvements and even increased intelligence (Chein & Morrison, 2010; Jaeggi et al., 2008; Klingberg, 2010). However, more

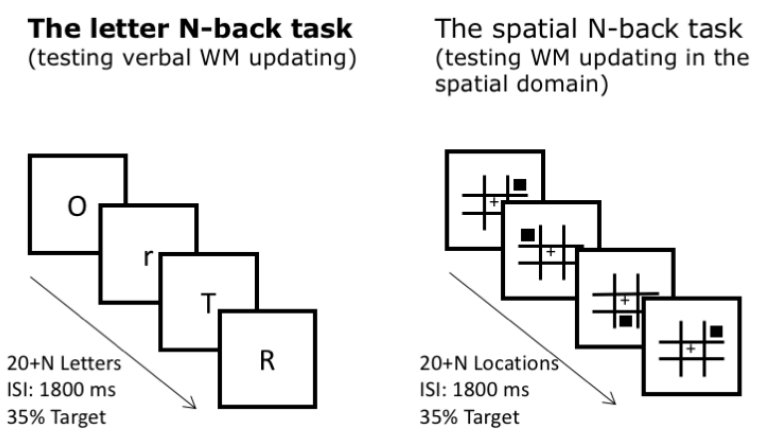


Figure 2 *– In the* ***Chapters 2, 3*** *and* ***4*** *of this thesis, we primarily used the letter N-back task to measure verbal WM functioning in our participants (left panel). In this task, the participant is presented with a stream of letters while seated in front of a computer screen and required to press a button when the current letter is the same as N stimuli before. By manipulating the level of N (i.e., number of items that need to be stored and updated), the difficulty of the task can be adjusted. In the N-back task, subjects are thus required to continuously update the content of their WM as well as make a comparison decision. As such, the N-back task is a challenging WM task, that is thought to rely heavily on the updating (and clearing) of information in WM. Seminal studies in both the field of WM training (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008) and single session tDCS research (Fregni et al., 2005) applied the letter N-back task, rendering it a key candidate to investigate and probe WM in the current thesis. Moreover, by changing the type of stimuli used from letters to e.g., locations in a grid, the task can neatly be adapted to test the generality of tDCS-induced changes in WM functioning, by investigating possible transfer to other domains (in this case spatial WM, see right panel).*

recent, well-controlled studies have shown that WM training may improve performance on the tasks practiced during the training as well as very similar WM tasks (Shipstead, Redick, & Engle, 2012), but that transfer of benefits to other cognitive domains is very limited (Harrison et al., 2013; Redick et al., 2013). Although the debate on the effects of behavioural WM training has not been resolved completely (see e.g., Au et al., 2015), WM training is currently generally considered to have negligible cognitive benefits. Central to this are observations that although WM training may lead to (short-term) benefits on similar WM tasks, these do not generalize to other tasks and settings (Melby-Lervåg, Redick, & Hulme, 2016; Shipstead et al., 2012).

Yet, transfer of training gains to non-trained stimuli, modalities and tasks is crucial for any cognitive enhancement method to be practically relevant for everyday life. Otherwise, behavioural advances do not reflect actual improvements in the working of the underlying cognitive faculty (Shipstead et al., 2012), but rather task-specific effects, such as simple stimulus familiarity practice effects, changes in stimulus-response mappings, or strategy learning (Lovden et al., 2010, Jonides, 2004).

Since transfer of the observed benefits of computerized WM task training to other situations is typically modest and takes relatively long training times (typically > 20 hrs) to establish, the practical value of computerized WM task training to effectively improve WM and the neural circuitries underlying this cognitive functioning is considered to be limited.

Transcranial Direct Current Stimulation: a fast lane to boost WM performance?

About a decade ago, an exciting new method started to receive attention as an alternative method to induce cognitive improvement: Transcranial Direct Current Stimulation (tDCS) (Kuo & Nitsche, 2012). This form of Non-Invasive Brain Stimulation (NIBS) allows us to directly influence brain functioning, thereby importantly, creating a possible short-cut compared to the traditional methods such as WM training that indirectly attempt to hone the brain for functional and structural alterations.

The method of tDCS can affect cortical excitability in specific parts of the human brain. In tDCS, a very small constant electrical current is run between (conventionally) two electrodes that are placed at specific locations on the scalp. Although the majority of this

current will be shunted through the skin and skull, a small part may actually reach the brain and affect the membrane potentials of neurons in the underlying cortical tissue. In this manner, tDCS can up- or down regulate the excitability of a particular brain region, making the neurons under the electrode more (anodal stimulation) or less (cathodal stimulation) prone to fire action potentials (Kuo & Nitsche, 2012; Michael A. Nitsche et al., 2008). Such changes in cortical excitability may influence performance on tasks that critically rely on the stimulated brain region. Furthermore, importantly in contrast to other NIBS, tDCS may as such provide a way to also directly improve cortical functioning, and not just interfere with it (in contrast to e.g. transcranial Random Noise Stimulation and Transcranial Magnetic Stimulation).

Most of the pioneering research with tDCS in humans has been done in the motor domain and involved the primary motor cortex. In line with earlier animal findings, these studies showed increased motor-evoked potentials after anodal tDCS to the motor cortex and decreased motor-evoked potentials after cathodal tDCS (e.g. Nitsche & Paulus, 2000), but see also the discussion between (Horvath, Carter, & Forte, 2014) and (Antal, Keeser, Priori, Padberg, & Nitsche, 2015). These findings support the notion that tDCS can affect local cortical excitability and hence directly manipulate human brain functioning.

By now, there are also numerous reports that tDCS may affect cortical functioning in other brain areas in a similar manner and thus can also be used to alter higher-order brain areas and corresponding functions, such as WM. Interestingly, a pioneering study in this regard is one by Fregni and colleagues (2005), which reported that a single session of anodal tDCS (vs. sham control stimulation) over the left dorsolateral PFC (lDLPFC) improved verbal WM in healthy adults. Since then, this finding has been replicated in a variety of populations (Andrews, Hoy, Enticott, Daskalakis, & Fitzgerald, 2011; Boggio et al., 2006; Ohn et al., 2008), providing support for the idea that anodal lDLPFC stimulation is a promising new tool for neurocognitive enhancement in healthy as well as clinical populations.

However, recent meta-analyses question the ability of anodal lDLPFC stimulation to robustly improve WM performance (see e.g. Bennabi et al., 2014; Brunoni & Vanderhasselt, 2014; Dedoncker, Brunoni, Baeken, & Vanderhasselt, 2016; Hill, Rogasch, Fitzgerald, & Hoy, 2017; Mancuso, Ilieva, Hamilton, & Farah, 2016). Although the large differences in experimental set-up and tDCS parameters in this new emerging field make a direct

comparison between studies difficult, these meta-analyses certainly show that the effect of anodal tDCS on WM functioning is less straight-forward than originally thought. The fact that these positive effects are not consistently replicated warrants care with drawing any firm conclusions about tDCS as a cognitive enhancement method. Moreover, they also raise questions about the factors that may influence the effects of anodal lDLPFC tDCS on WM.

Interestingly, in addition to temporarily affecting brain functioning, tDCS has also been associated with longer-lasting changes. For example, tDCS induced improvements in planning ability were still visible 6 months after stimulation (Dockery, Hueckel-Weng, Birbaumer, & Plewnia, 2009), indicating that tDCS may bring about lasting neuroplastic modifications. Moreover, findings from stroke revalidation studies indicate that anodal tDCS may stimulate learning and thereby speed up the effects of behavioural motor and visuomotor rehabilitation training (Hashemirad, Zoghi, Fitzgerald, & Jaberzadeh, 2016). For example, in one study, 3 months of visual field training combined with tDCS resulted in improvements typically observed only after 6 months of behavioural training (Plow, Obretenova, Fregni, Pascual-Leone, & Merabet, 2012). A similar role for tDCS as learning enhancer has been found in the cognitive domain and in healthy populations, where repeated tDCS has been shown to facilitate artificial number learning (Cohen Kadosh, Soskic, Iuculano, Kanai, & Walsh, 2010, with effects still apparent six months afterwards), as well as response inhibition training (Ditye, Jacobson, Walsh, & Lavidor, 2012). These effects may occur through NMDA-dependent mechanisms similar to the synaptic-use dependent mechanisms long-term potentiation and depression, which play a critical role in memory and learning (Liebetanz, Nitsche, Tergau, & Paulus, 2002; Monte-Silva et al., 2013; Nitsche et al., 2003).

Anodal lDLPFC tDCS may thus improve WM functioning in two ways. On the one hand, it may instantly enhance WM performance by directly improving functioning of the left prefrontal cortex. On the other hand, it may temporarily enhance plasticity, thereby facilitating learning and promoting lasting WM improvements when paired with intensive computerized WM training. Taking together, tDCS thus provides an interesting new method to enhance cognitive functioning. Especially, when concurrently applied with WM training, tDCS may be a promising method to speed up and/or strengthen the effects of WM training.

Combined tDCS and WM training: The best of both worlds?

In ***Chapter*** 2 of the current thesis I address a number of outstanding questions regarding the potential of combined tDCS and WM training as a method to lastingly and effectively enhance WM. Firstly, I investigated whether we can continue to boost WM if we stimulate not once but multiple times. In other words, are the tDCS-induced improvements a 'one time thing', or does repeated stimulation lead to additive effects? Secondly, I explored whether tDCS-induced WM improvements are specific to the task paired with stimulation or whether these in fact reflect enhancement of more domain-general WM processes.

Concurrent with the coming about of this thesis and following a similar line of reasoning, several fellow research groups also pursued the idea that combined tDCS and WM training may be a successful way to lastingly improve cognitive functioning. However, so far, initial findings provide little support for the notion that multiple sessions with anodal lDLPFC tDCS and verbal WM training may lead to larger, more persistent, and/or better transferable WM improvements (Lally, Nord, Walsh, & Roiser, 2013; Martin et al., 2013; Richmond, Wolk, Chein, & Olson, 2014). However, particular design choices may have been at the root of these null-findings, which I will discuss in greater depth in the Introduction of ***Chapter 2***. In short, in tDCS research, several parameters may be pivotal in determining the effects of tDCS, including electrode location, stimulation intensity and duration, and the task paired with stimulation, which may not have been optimal in these studies. Taking into account findings of the current thesis, furthermore, in ***Chapter 5***, I discuss these issues further as well as delineate important issues for future studies, that aim to use the concurrent application of tDCS to speed up the effects of behavioural WM training.

Exploring individual differences in tDCS response

Interestingly and unexpectedly, in ***Chapter 2***, we observed considerable individual differences in the effects of anodal tDCS on WM performance. In fact, whereas most participants showed the expected improvements in WM, WM performance in some participants actually seemed to deteriorate as a result of the stimulation. These findings are in line with a growing body of recent research reporting relatively large individual variability in the effects of tDCS on WM performance (Berryhill & Jones, 2012; London & Slagter, 2015). This exposes a fundamental problem with the way we currently apply tDCS; namely,

with one and the same stimulation protocol for everyone. Yet, for a successful practical implementation of combined tDCS and WM training as a method to boost cognition, it is crucial that stimulation protocols benefit all individuals (and not actually worsen some), and if possible, to an optimal degree in everyone. It is therefore vital that we advance our understanding of the causes that may underlie individual differences in the effects of tDCS observed with regard to WM performance.

In recent years, several hypotheses have been proposed to account for individual variability in tDCS responsiveness. Most of these revolve around the notion that although the amount of current at the scalp is the same, the percentage of this current that may actually affect the target brain area differs to a great extent between individuals. Recent modelling studies have indicated that the current flow in conventional tDCS set-ups (such as used in the current thesis) is strongly influenced by individual differences in anatomy, skull thickness and precise cortical folding (Opitz, Paulus, Will, Antunes, & Thielscher, 2015). Such factors may thus lead to substantial differences in how much of and where the admitted current may actually reach the brain. In other words, with our current standard tDCS set-ups we may simply be more or less successful in affecting the desired cortex in some individuals compared to others.

A role for baseline cortical excitability?

Another hypothesis (not mutually exclusive with the previous one) is that differences in baseline cortical functioning between individuals (i.e., the brain state before the stimulation is applied) can critically determine the effects of tDCS on behaviour. Since with tDCS, we are modulating the excitability of neurons in a particular cortical area, it is plausible that the effect of tDCS crucially depend on/ interact with the baseline activity state of the stimulated brain region. More precisely, if the brain region is not yet fully engaged in the task, enhancing its excitability with anodal tDCS may benefit the system,

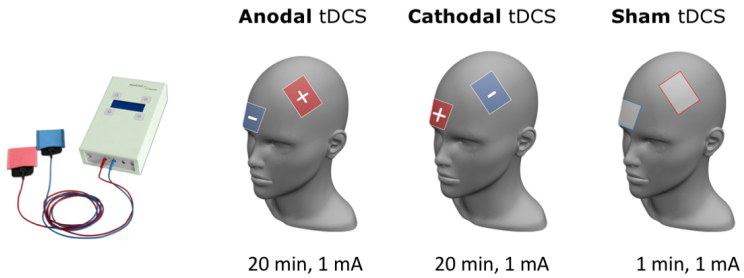


Figure 3 – *In **Chapters 2** and **4**, I used Transcranial Direct Current Stimulation over the left dorsolateral prefrontal cortex to manipulate WM functioning. For this, I used a battery-driven Eldith-DC stimulator (Neuroconn GmbH, Germany) and a conventional set-up with two 7x 5cm electrodes in saline soaked sponges, which were held in place using rubber bands. In tDCS research, there are many stimulation parameters that can be chosen and may critically determine the effect of tDCS. Therefore, in the current thesis we used a set-up that has most consistently been found to improve WM performance in single session tDCS studies (e.g. Fregni et al., 2005). The active electrode was always placed over the left dorsolateral prefrontal cortex (localized as F3 in the 10/20 system using an EEG cap) and the reference over the right supra-orbital cortex (i.e., centered above the right eye). There are three types of prefrontal stimulation used in this thesis. In the anodal stimulation condition (**Chapter 2** and **4**) the anode (i.e., the positive electrode) was the active electrode, while in the cathodal condition (**Chapter 4**) the cathode (i.e. the negative electrode) was the active electrode. In both cases, 1 mA of current was always applied for 20 minutes. The sham condition (**Chapter 2**) is a control condition in which tDCS was admitted for 1 minute. All types were tolerated well by all participants.*

resulting in improved WM performance observed at the level of behaviour. However, ifneurons of the cortical area are already optimally engaged in the task, tDCS may actually negatively interfere with the functioning of that specific brain area, thereby impairing behavioural performance.

The manner in which baseline cortical functioning may be represented at the cellular level in a particular region is in the form of cortical excitability; the excitation/inhibition balance. This balance is thought to be determined by two key neurotransmitters: Glutamate, which has an excitatory effect and GABA, the brain's main inhibitory neurotransmitter (Petroff, 2002).

Cortical excitability levels critically determine neuronal firing rates, thereby playing a pivotal role in regulating cortical functioning. In an optimal situation, the cortex is active enough for functional firing to effectively take place, but at the same time inhibited enough to reduce noise and unwanted firing (Turrigiano & Nelson, 2000; Turrigiano & Nelson, 2004). In this regard, both too high and too low cortical excitability can be detrimental for functional performance, albeit for different reasons.

In 2013, Krause and colleagues first suggested that baseline excitation/inhibition balances may determine the effect of tDCS and proposed a theoretical model to explain observed individual differences in tDCS response (Krause, Márquez-Ruiz, & Kadosh, 2013). More specifically, they proposed that the effect of tDCS, beneficial or unfavourable for brain function, may critically depend on a subject's original position on the cortical excitability spectrum; namely, tDCS may push or pull the cortical region towards or away from its optimal excitability level (Krause et al., 2013).

Interestingly, studies that have investigated the effects of tDCS on neurotransmitter concentrations have related tDCS to both changes in GABA and Glutamate. Anodal stimulation over the motor cortex was found to reduce GABA levels (Stagg et al., 2009), while cathodal stimulation in contrast reduced Glutamate levels (Clark, Coffman, Trumbo, & Gasparovic, 2011). Although probably through different mechanisms, this indicates that both types of stimulation may critically affect neuronal functioning by changing local excitation/inhibition balances.

To further pursue the hypothesis that individual differences in tDCS response may stem from individual differences in baseline cortical excitability levels, in ***Chapter 4*** of this

thesis, I investigated whether prefrontal cortical excitability predicts individual differences in the effect of left prefrontal tDCS on WM performance. To do this, I used the neuroimaging method of Magnetic Resonance Spectroscopy (see below and Figure 5) to quantify Glutamate and GABA neurotransmitter levels in the left DLPFC. These levels were used to calculate individual Glutamate/GABA ratio as a proxy for local excitation/inhibition balances, i.e. cortical excitability. To specifically test the hypothesis that the effects of tDCS on WM performance depend on cortical excitability, I investigated whether these ratios predicted observed WM improvements or deteriorations after anodal and cathodal tDCS across individuals.

Measuring GABA and Glutamate in-vivo with Magnetic Resonance Spectroscopy

Magnetic Resonance Spectroscopy (MRS) is a non-iodizing and relatively novel method that allows for the non-invasive in-vivo quantification of neurotransmitter levels such as GABA and Glutamate in a particular brain area or voxel in the human brain (Jansen, Backes, Nicolay, & Kooi, 2006) (see also Figure 5). MRS thus for the first time allows investigation of the relationship between neurotransmitter concentrations, such as GABA and Glutamate, in the human brain and behaviour (Duncan, Wiebking, & Northoff, 2014). Moreover, since with MRS we can measure brain levels of the key players in maintaining local excitation/inhibition balances, this method can also provide local Glutamate/GABA ratios as a proxy of cortical excitability, which, as operationalised in ***Chapter 4***, can be used to predict the strength and/or direction of individual tDCS responses.

However, while designing the study of Chapter 4, we realized that it was still unclear to what extent MRS-measured concentrations reflect individual 'trait'-levels of neurotransmitters, or are also affected by activity 'state' at the moment of measurement.

3T-MRS: Measuring 'State' or 'trait'?

So far, the vast majority of studies that have used MRS to investigate the relationship between neurotransmitter levels and cognitive performance have measured these levels during a resting-state condition (see e.g., the studies discussed in (Duncan et al., 2014)). This assumes that MRS-measured GABA and Glutamate levels reflect stable individual 'trait'-like levels. However, other studies have found that these concentrations may not be

so static, but can in fact change over relatively short time windows (Floyer-Lea, Wylezinska, Kincses, & Matthews, 2006; Michels et al., 2012; Shibata et al., 2017). Since brain functioning may be very different during a task compared to a rest situation, and research with other neuroimaging measures such as EEG and fMRI have often found better brain-behaviour relations in an on-task situation, MRS-measured neurotransmitter levels at rest may not be the strongest or most reliable index of function-specific cortical excitability to relate to behaviour.

Therefore, before determining to what extent the effect of tDCS depends on cortical excitability (**Chapter 4**), in **Chapter 3** I first addressed the following two outstanding methodological questions with regard to the use of MRS to in-vivo measure local Glutamate and GABA in the human brain. Firstly, do cortical Glutamate and GABA levels vary as a function of task demand, i.e., do they reflect 'trait' or 'state'? Secondly, if indeed (partly) reflecting activity state, in which activity state (e.g., rest, on-task) do these neurotransmitter concentrations best predict performance on brain region-related cognitive tasks? To answer these questions, I scanned both a primary sensory region (the medial occipital cortex) and a higher order cognitive region (the left dorsolateral prefrontal cortex) under different conditions, and related the obtained neurotransmitter levels to performance on region-related cognitive tasks.

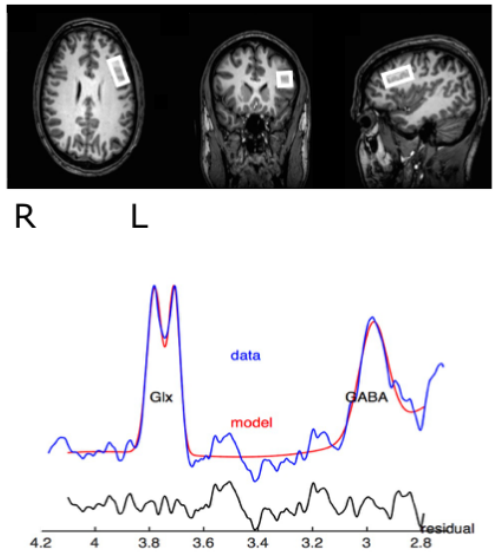


Figure 4 *– In* ***Chapters 3*** *and* ***4****, I used Magnetic Resonance Spectroscopy at a 3T-MRI (Philips Achieva TX MRI scanner, Philips Healthcare) scanner to non-invasively measure neurotransmitter concentrations of Glutamate and GABA in the healthy human brain. Shown here is a structural MRI scan with an exemplary 3D MRS-voxel placed over the target region of interest; in this case, the left dorsolateral prefrontal cortex (size of the voxel: 30x20x25 mm). During an MRS-scan (typically taking about 6 minutes), a spectrum is obtained from this volume, from which the relative concentrations of different biochemicals can be computed (using analyses packages such as GANNET (www.gabamrs.com) as these will show up as different peaks at specific known frequencies in the spectrum (bottom panel). At 3T, however, the peak of Glutamate is known to overlap with the one of Glutamine, resulting in the compound measure Glx as best measure for Glutamate. By calculating Glutamate/GABA ratios, MRS can furthermore be used to arrive at a measure for local excitation/inhibition balances, i.e., cortical excitability. This measure was used in* ***Chapter 4*** *to determine to what extent individual differences in the effect of tDCS on WM performance can be explained by individual differences in baseline cortical excitability.*

Overview current thesis

In the current thesis, I set out to (i) determine the potential of tDCS (combined with practice on a WM task) to induce lasting and transferable enhancements in WM functioning, and (ii) investigate the possible role of cortical excitability in influencing the strength and direction of the tDCS-induced effect on WM performance in a given individual.

To this end, in ***Chapter 2***, I investigated the effects of three daily sessions of anodal prefrontal tDCS (see Figure 2) combined with practice on a verbal WM task (see Figure 3) on WM functioning. This chapter is focused on two core questions. Firstly, if we stimulate more than once, can we continue to boost WM? Secondly, are tDCS-induced WM improvements restricted to the task paired with stimulation or do they reflect domain-general WM improvements? We used a between-subject design and submitted thirty subjects (fifteen in each group) to three sessions of either active or sham anodal stimulation. Moreover, we added a behavioural session before and another two days after the last day of stimulation to investigate transfer of tDCS-induced benefits to three WM tasks; namely, the same verbal WM task with a different stimulus-set, a spatial version of this task (see Figure 2) as well as a complex WM span task (the O-span (Unsworth, Heitz, Schrock, & Engle, 2005)).

To this end, in ***Chapter 2***, I investigated the effects of three daily sessions of anodal prefrontal tDCS (see Figure 2) combined with practice on a verbal WM task (see Figure 3) on WM functioning. This chapter is focused on two core questions. Firstly, if we stimulate more than once, can we continue to boost WM? Secondly, are tDCS-induced WM improvements restricted to the task paired with stimulation or do they reflect domain-general WM improvements? To answer these questions, 3T-MRS data was acquired from thirty healthy men in two brain areas and under different task demands: from the medial occipital cortex (OC), at rest (eyes closed) and while subjects watched a movie (on-task); and the left DLPFC, at rest, during an easy working memory (WM) task (letter 2-back), and during a challenging WM task (adaptive letter N-back). Moreover, in a separate behavioural session, we determined subjects' performance on a discrimination task (using oblique grating patches) to relate to neurotransmitter levels in the occipital cortex, and on two working memory tasks (letter N-back and the Sternberg task) to relate to neurotransmitter levels in the left DLPFC.

In ***Chapter 4,*** I investigated the possibility that baseline prefrontal cortical excitability may predict observed individual differences in response to prefrontal tDCS. For this, twenty of the participants that participated in the study of Chapter 3 came back to the lab for two separate stimulation sessions (spaced one week apart): one in which they received anodal and one in which they received cathodal tDCS over the left DLPFC (see Figure 2) while performing a verbal WM task (see Figure 3 for details). We used their prefrontal Glutamate and GABA concentrations as obtained in the study of Chapter 3 to individually calculate excitation/inhibition balance ratios (Glutamate/GABA) as a proxy for cortical excitability and related these to the behavioural effects of both types of tDCS.

Lastly, ***Chapter 5*** presents a summary of the main findings of the different studies presented in this thesis and puts these in perspective of the current literature in the field. In this chapter, I will also address some outstanding methodological issues and discuss directions for future tDCS studies. I will end this chapter with some concluding remarks concerning the possible potential application of combined tDCS and WM training as a method to lastingly enhance WM in all individuals.

Chapter 2

Boosting cognition: Effects of multiple session tDCS on working memory

Chapter published as

Talsma, L.J., Kroese, H.A., Slagter, H.A. (2016) Boosting cognition: Effects of multiple session tDCS on working memory. *Journal of Cognitive Neuroscience, 29(4),* 755-768

Abstract

Transcranial direct current stimulation (tDCS) is a promising tool for neurocognitive enhancement. Several studies have shown that just a single session of tDCS over the left dorsolateral prefrontal cortex (lDLPFC) can improve the core cognitive function of working memory (WM) in healthy adults. Yet, recent studies combining multiple sessions of anodal tDCS over lDLPFC with verbal WM training did not observe additional benefits of tDCS in subsequent stimulation sessions, nor transfer of benefits to novel WM tasks post-training. Using an enhanced stimulation protocol, as well as a design that included a baseline measure each day, the current study aimed to further investigate the effects of multiple sessions of tDCS on WM. Specifically, we investigated the effects of three subsequent days of stimulation with anodal (20 min, 1 mA) vs. sham tDCS (1 min, 1 mA) over lDLPFC (with a right supraorbital reference) paired with a challenging verbal WM task. WM performance was measured with a verbal WM updating task (the letter N-back) in the stimulation sessions, and several WM transfer tasks (different letter set N-back, spatial N-back, Operation span) before and two days after stimulation. Anodal tDCS over lDLPFC enhanced WM performance in the first stimulation session, an effect that remained visible 24 hours later. However, no further gains of anodal tDCS were observed in the second and third stimulation sessions, nor did benefits transfer to other WM tasks at the group level. Yet, interestingly, post-hoc individual difference analyses revealed that in the anodal stimulation group the extent of change in WM performance on the first day of stimulation predicted pre- to post changes on both the verbal and the spatial transfer task. Notably, this relationship was not observed in the sham group. Performance of two individuals worsened during anodal stimulation and on the transfer tasks. Together, these findings suggest that repeated anodal tDCS over lDLPFC combined with a challenging WM task may be an effective method to enhance domain-independent WM functioning in some individuals, but not others, or can even impair WM. They thus call for a thorough investigation into individual differences in tDCS respondence as well as further research into the design of multi-session tDCS protocols that may be optimal for boosting cognition across a wide range of individuals.

Introduction

Transcranial Direct Current Stimulation (tDCS) is a safe and noninvasive brain stimulation method in which a low-voltage electric current (<= 2 mA) is run between two scalp electrodes; the anode (the positive electrode) and cathode (negative electrode). By modulating the membrane potential of underlying cortical neurons, tDCS may alter brain functioning. More specifically, stimulation with tDCS may temporarily make neurons more (anodal; facilitating) or less (cathodal; inhibiting) prone to fire action potentials (Kuo & Nitsche, 2012; Nitsche et al., 2008).

Working memory (WM) is considered a core cognitive function underlying performance in many everyday life situations as it allows us to retain and monitor information over brief periods of time (Baddeley, Sala, Robbins, & Baddeley, 1996). As WM may be disturbed in psychiatric conditions such as schizophrenia (Barch & Ceaser, 2012) and decrease in older age, there is a growing interest in methods to enhance WM functioning, e.g. with intensive computerized task training. Although initial results of WM training studies suggested wide-spread cognitive benefits (Chein & Morrison, 2010; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Klingberg, 2010), more recent, well-controlled studies found only limited transfer of improvements after WM training (Harrison et al., 2013; Redick et al., 2013). Together with long training times (typically > 20 hours), this substantially limits the practical value of WM training as method to improve cognitive functioning.

Interestingly, a decade ago a pioneering study by Fregni and colleagues (2005) reported that a single session of anodal tDCS (vs. sham stimulation) over the left Dorsolateral Prefrontal Cortex (lDLPFC) could improve verbal WM in healthy adults. This finding has been replicated and extended to a variety of populations (Bennabi et al., 2014; Hill, Fitzgerald, & Hoy, 2016; Mancuso, Ilieva, Hamilton, & Farah, 2016; but also see Brunoni & Vanderhasselt, 2014), providing substantial support for the claim that directly modulating the brain with anodal lDLPFC stimulation may be a promising new tool for neurocognitive enhancement in healthy as well as clinical populations.

Moreover, the effects of tDCS on behaviour do not seem to be limited to temporary changes in excitability only, but may involve actual longer-lasting neuroplastic changes. This may make tDCS a specifically useful method for enhancing learning. Indeed, anodal tDCS

over relevant areas may speed up the effects of behavioural motor and visuo-motor revalidation training after stroke (Hashemirad et al., 2015). For example, in one study, three months of visual field training combined with tDCS resulted in improvements typically observed after 6 months of behavioural training only (Plow, Obretenova, Fregni, Pascual-Leone, & Merabet, 2012). Similar effects have been found in healthy individuals and in the cognitive domain, where repeated tDCS has been shown to facilitate artificial number learning (Cohen Kadosh, Soskic, Iuculano, Kanai, & Walsh, 2010) (with effects still apparent 6 months later) and response inhibition training (Ditye, Jacobson, Walsh, & Lavidor, 2012). Together, these findings raise the premise that anodal tDCS over lDPLFC paired with WM training may speed-up and/or strengthen WM training effects.

While many studies have reported effects of single session tDCS, so far only three studies have examined the effects of multiple sessions of anodal lDLPFC stimulation and verbal WM training on WM in healthy adults. Firstly, Lally et al. (2013) found no additional improvement on a verbal WM task during anodal vs sham stimulation over the course of two sessions (although a post-hoc analysis did show larger enhancements in the anodal group on the first day). Secondly, Martin et al. (2013) also found no differences between an anodal and sham group in a study with ten sessions of combined tDCS and verbal WM training, not on the trained task itself (when group baseline performance differences were taken into account), nor on other cognitive tasks administered in a separate session one day after stimulation to assess possible transfer of training benefits. Thirdly, with a similar design, Richmond et al. (2014) did find a larger increase in verbal WM performance over ten training sessions in the anodal compared to the sham tDCS group. Yet, they too failed to observe larger post-training improvements on additional cognitive transfer tasks. Collectively, these initial findings thus provide little support for the notion that multiple sessions with anodal lDLPFC tDCS and verbal WM training may lead to larger persistent and transferable WM improvements than WM training alone.

However, these null findings may be a consequence of particular design choices in these studies. In tDCS research, several parameters are pivotal in determining its effects, including electrode location, stimulation intensity and duration, and the task paired with stimulation.

In the three studies described above, we believe that some of these parameters may not have been optimal for inducing verbal WM enhancements.

First and perhaps most importantly, electrode location critically determines current flow through the brain and thereby the precise cortical regions that are affected (see Nitsche et al., 2008). Notably, all above studies used a different set-up than the single session studies that found WM improvements; with the anode (i.e., the active electrode) over lDLPFC (electrode site F3) and the cathode (i.e., the reference) over the right orbitofrontal cortex (rOFC) (i.e. the contralateral forehead). All three studies placed the anode over lDLPFC, but the cathode was placed differently. Both Lally et al. (2013) and Martin et al. (2013) chose extraencephalic references with the cathode on the contra-lateral cheek and shoulder respectively. Although common in e.g. the motor domain, it is conceivable that in with this set-up more medial parts of lDLPFC that are also important for WM are missed. Moreover, Richmond et al. (2014) placed the cathode over the right DLPFC (F4), a region known to be involved in WM (Au et al., 2016; Berryhill & Jones, 2012; Owen, McMillan, Laird, & Bullmore, 2005). The possible inhibitory effect of the cathode over this region may make this electrode set-up suboptimal for inducing WM improvements.

Two other parameters that play an important role in the effect of tDCS on behaviour are stimulation intensity and duration. Most effective single session tDCS studies used a stimulation strength of 1 mA to boost WM in healthy adults (Andrews, Hoy, Enticott, Daskalakis, & Fitzgerald, 2011; Fregni et al., 2005; Mulquiney, Hoy, Daskalakis, & Fitzgerald, 2011; Ohn et al., 2008). Notably, in contrast to the intuitive notion that higher intensities lead to stronger effects, a recent study showed that 1 mA and not 2 mA stimulation resulted in the most pronounced WM improvements (Hoy et al., 2013; but see also Teo, Hoy, Daskalakis, & Fitzgerald, 2011). The current strengths of 1.5 mA and 2.0 mA used by Richmond et al. (2014) and Martin et al. (2013) may thus have been suboptimal. Similarly, longer stimulation durations may not always result in larger effects either. In fact, in the motor domain, longer stimulation times have shown to diminish and sometimes actually result in opposite effects in behaviour (see Nitsche et al., 2008). In particular, the 30 minutes stimulation by Martin et al. (2013) is relatively long compared to the 10 to 20 minutes

typically used in the single session literature (and the 10 and 15 minutes used by Lally et al. (2013) and Richmond et al. (2014)), which may have reduced its effectiveness.

Finally, the 'state' of the stimulated area (i.e., what a subject is doing) may be critical in determining stimulation effects on behaviour. Anodal tDCS admitted concurrent with a task (online stimulation) has been shown to be more effective in boosting WM than tDCS admitted during rest (offline stimulation) (Andrews et al., 2011; Mancuso et al., 2016; Martin, Liu, Alonzo, Green, & Loo, 2014). Possibly, this is because the targeted brain networks are already engaged in the to-be-modulated cognitive activity. This likely also applies to repeated stimulation. Both Lally et al. (2013) and Martin et al. (2013) used on-line stimulation. However, in Richmond et al. (2014), the task was only paired with tDCS during the last 5 minutes of stimulation, and the remainder of the task was done without stimulation, conceivably reducing tDCS effectiveness in enhancing WM.

Adding to the literature in this field, the current study aimed to evaluate the effects of multiple-session lDLPFC stimulation on WM using a set-up that may maximize tDCS effectiveness. Similar to single session verbal WM enhancement studies, the anode was placed over lDLPFC (F3) and the cathode over rOFC (contralateral above the right eye) and stimulation was applied at 1 mA intensity for 20 minutes. Furthermore, stimulation was paired with a highly demanding verbal WM task (3- and 4 letter N-back task), as this may be critical for enhancing cognitive functioning (Gill, Shah-Basak, & Hamilton, 2015). The study was conducted using a randomized double blind design in which subjects underwent either active tDCS or sham stimulation (1 min of stimulation) on three consecutive days. The verbal WM task on these days was split in four equal blocks and stimulation was always applied during the second block. This design allowed us to look at the effects of tDCS at different time windows during and after stimulation. Moreover, for each session the effects of tDCS on behaviour could be contrasted to the first, baseline block of that day, which also permitted us to separate within-session effects of tDCS from between-session carryover effects of previous stimulation.

Furthermore, to determine if our combined tDCS and WM protocol could induce more general WM enhancements we assessed transfer of potential benefits to different stimuli and task contexts. To this end, prior to the first and two days after the last stimulation

session, subjects performed three other WM tasks, namely; the same verbal WM task with a different letter set, a spatial version of this task (spatial N-back) and a complex span task (the automated Operation span task (Unsworth, Heitz, Schrock, & Engle, 2005)).

We predicted that using our optimized stimulation protocol, three daily sessions with anodal vs. sham tDCS would first of all result in greater cumulative improvements in verbal WM in the stimulation sessions. Second, we expected tDCS effects to outlast the stimulation and remain apparent 24 hours later, i.e. in the baseline blocks of the next day. Third, we expected anodal (vs. sham) tDCS combined with WM practice to induce general WM improvements, as reflected in larger performance improvements on the WM transfer tasks post-training.

Recently, a growing number of studies have reported that the effect of tDCS may vary substantially across individuals. This may be due to differences in e.g., brain anatomy (Kim et al., 2014; Opitz, Paulus, Will, Antunes, & Thielscher, 2015), baseline performance (Berryhill & Jones, 2012; Learmonth, Thut, Benwell, & Harvey, 2015; London & Slagter, 2015; Meiron & Lavidor, 2013), and/or differences in cortical excitability (Krause & Cohen Kadosh, 2014). Consequently, in standardized tDCS protocols, some individuals may benefit more than others. Therefore, in addition to group level analyses, we post-hoc also explored if across subjects, protocol effectiveness (i.e. the extent to which WM was improved in the stimulation sessions) could predict transfer to the WM tasks post-training. Our final, fourth prediction was that such transfer of benefits should be most apparent in those individuals whose performance increased most in the tDCS combined with WM training sessions.

Methods

Participants

50 subjects were recruited via the University of Amsterdam and were compensated for their participation with money or research credits. Subjects gave written informed consent before the experiment, as approved by the local ethics committee. All reported no history of psychiatric conditions and had normal or corrected-to-normal vision. They were checked for tDCS contra-indications, such as metal implants and sensitive skin (see Nitsche et al.

2008). Furthermore, pilot analyses showed that subjects that already started out with high WM accuracy scores in the pre-stimulation session tended to reach almost perfect performance in the second or beginning of the third stimulation session. To ensure sensitivity to improvements throughout all three stimulation sessions for all subjects, we excluded high performers in the pre-stimulation session from further participation. To determine this, we calculated accuracy scores (in the form of A', see below) over the verbal WM task (level 3 and 4 only) during this initial session and excluded subjects who showed A' values above 0.90 (n=15). This threshold corresponded to an average hit rate of 0.85 in the excluded subjects. Four participants could not complete the study because of personal or health reasons unrelated to the study. One last subject was excluded because of very poor performance on the verbal WM task throughout the whole experiment (> 3 SDs below the mean). As a result, 30 participants were left for analysis (Anodal group: 4 male, 11 female, mean age 21.9 years, SD 2.8; Sham group: 5 male, 10 female, mean age 22.1 years, SD 2.3).

Design and Procedure

Participants came to the lab at the same time each day for a total of five sessions: a first behavioural session (pre-session), three consecutive days of tDCS stimulation combined with a verbal WM task (stimulation sessions) and a second behavioural session (post-session). Subjects were pseudo-randomly divided over the two stimulation groups; active vs sham (double-blind between-subject design), matching the groups on gender, age and WM performance in the pre-session. In the stimulation sessions, subjects received either active (1 mA, 20 min) or sham (1 mA, 1 min) anodal tDCS over the left DLPFC while performing a verbal WM updating task. Moreover, in the separate behavioural pre- and post-session (48 hours after), subjects performed three additional tasks: the same verbal WM task with a different stimulus set, a spatial version of the same WM task and a complex WM span task. As research in animals has indicated that cellular changes induced directly after tDCS have fully returned to baseline levels after 48 hours (see Nitsche, 2008), a two-day gap was implemented between the last stimulation session and the post-session to

ensure that the temporary effects of tDCS would have worn out. Please see figure 1A for a graphical rendering of the different experimental sessions.

Transcranial Direct Current Stimulation

Stimulation was delivered with a battery driven Eldith DC-stimulator (NeuroConn GmbH, Germany) using two 7 x 5 cm conductive electrodes. Electrodes were placed in saline-soaked sponges and held in place with rubber bands. The anodal electrode was always placed over the left DLPFC (F3 in the 10/20 system), while the cathodal electrode was placed over the right supra-orbitofrontal region (centered above the right eye pupil), see Figure 1B. This electrode arrangement is known to result in significant WM enhancements in single stimulation sessions (Fregni et al. 2005, Andrews et al. 2011, Mulquiney et al. 2011, Teo et al. 2011). In each subject, in the first session the position of F3 was localized using an EEG cap (64 channels Biosemi) and marked on the scalp to ensure the same electrode placement on the subsequent days. Participants in the active stimulation group received 20 min of 1 mA anodal stimulation, while those in the sham group received only 1 min of 1 mA anodal stimulation each stimulation session. To reduce discomfort and improve our shamming procedure, in both conditions, the current was ramped up over 30 seconds and down over 60 sec. Both participant and experimenter were blind to experimental group and thus which type of stimulation was applied.

Stimulation sessions: tDCS + verbal WM task

We investigated the immediate effects of anodal versus sham tDCS over left DLPFC on verbal WM across three daily sessions of stimulation. Similar to previous single session studies (e.g., Fregni et al. 2005), subjects performed a letter version of the N-back task to probe verbal WM. In this task subjects are presented with a stream of letters and have to indicate if the currently presented stimulus is the same as the one presented N stimuli back. N is an integer and the value of N determines the difficulty level of the task, with higher levels of N corresponding to higher WM loads as more stimuli have to be held in WM in sequential order. A recent study found that anodal (vs. sham) tDCS to lDPLFC only improved

post-stimulation performance on an attention task when combined with a challenging 3-back but not an easy 1-back verbal WM task (Gill et al., 2015). Therefore, to ensure a challenging task during all stimulation sessions, the level of N used alternated between 3 and 4.

While seated in a comfortable chair behind a computer screen (approximately 90 cm distance), subjects first practiced the task before actual data collection started. Stimuli were presented using Presentation software (Neurobehavioural Systems, Inc.) The task was divided into four blocks of about 15 minutes (non-stimulation blocks) or 20 minutes (stimulation blocks) each, and stimulation was always applied concurrent with the second block of the task (see Figure 1).

Each day, the first block of the task thus served as a baseline, which allowed us to investigate possible carry-over effects of tDCS to the next day as well as provided a more accurate measure of within session effects in each session. After this first 15-min block of the task, tDCS was administered for 20 minutes. To allow itching sensations in the first minutes of tDCS stimulation to wear off, the task was started two minutes after the onset of stimulation. The last 5 minutes of behavioural data during the stimulation block was discarded in the analyses to ensure comparison of blocks of equal length. Our design thus allowed us to compare subjects' verbal WM performance before, during and in two blocks after either active or sham stimulation over left DLPFC.

Each 15-min block of the task consisted of 24 so-called runs, in which level of N alternated every 3 runs between 3 and 4. Runs consisted of a stream of 20 + N stimuli each and were self-paced to allow the subject to take small breaks in between runs. Letters were presented in black (Arial, font size 72) at the centreof a white screen for 300 ms each, followed by a 1500 ms inter-stimulus interval in which a fixation cross (Arial, font size 20) was displayed centrally (see Figure 1C). Of the presented letters, 35% were so-called targets, i.e., the letter that was the same as the letter presented N letters back. If presented with a target letter, subjects were required to respond by pressing the space bar on the keyboard. Two letter sets were used in the experiment, namely: [a, b, c, d, e, f, g, h, j, k] and [k, m, n, o, p, r, s, t, u, w]. One of these was always used in the verbal WM in the stimulation sessions, while the other letter set was used in the verbal WM transfer task in the pre- and

post-session. Letter set assignment was counterbalanced across subjects. Furthermore, to prevent the use of a simple visual feature matching strategy by subjects, letters could be presented in upper or lower case and still would classify as the same letter (i.e., a target).

Each stimulation session started and ended with filling out questionnaires to assess possible side effects of stimulation on mood and arousal, and physical sensations. To assess mood and arousal levels, a Dutch translation of the short version of the Activation Deactivation Adjective Checklist (AD ACL) was used (Thayer, 1978), that asked subjects to respond to 20 items using a 4-point rating scale (namely; "definitely feel", "feel slightly," "do not really feel" and "definitely do not feel"). Answers are scored on four subscales; energy (general activation), tiredness (general deactivation), tension (high preparatory arousal) and calmness (low preparatory arousal). The AD ACL has proven reliable and valid, showing high test-retest reliability for each of its subscales (all > .79) (Thayer, 1978). The AD ACL was filled out twice each session and changes in mood and arousal were calculated. In addition, to investigate possible physical side effects of the tDCS stimulation, at the end of the session participants were asked to rate their experience on a five items scale (namely: "not", "a little", "somewhat", "strongly" and "very strongly") with each of eight following sensations: itching, prickling, burning, pain, headache, fatigue, dizziness and nausea.

Pre- and post-session: WM Transfer tasks

To investigate whether possible verbal WM enhancements after tDCS may reflect more general WM learning, before and two days after the three stimulation sessions, subjects participated in a behavioural session in which they performed three WM transfer tasks: the same verbal WM task but with a different letter set (i.e. a new stimulus set), a spatial WM task (i.e. a different domain), and a complex WM span (i.e. a different task).

The pre- and post-session was identical except that the pre-session ended with brief trial tDCS (30 seconds of stimulation) to familiarize the subject with the sensation of tDCS, while the post-session started with a block of the verbal WM task used the stimulation sessions. Adding this last baseline block allowed us to determine if any tDCS effects observed in the stimulation sessions were still measurable two days later. Before the actual

task started subjects received instructions and performed a series of practice trials with feedback. Order of the verbal and spatial WM task was counter-balanced between subjects. The complex WM span was always performed last as this task includes feedback about performance and may thus possibly lead to motivational differences between subjects.

The verbal WM task in the pre- and post-session was very similar to the verbal WM task of the stimulation sessions. However, to investigate possible transfer to a different stimulus set (i.e. stimulus independent learning), the other letter set was used. Also, level of N in the pre- and post-session ranged from 2 to 5 to index a broader range of participants' abilities. The task started with N level 2 and progressed to N level 5 twice, leading to 48 runs of the task in total.

The spatial WM task was a spatial version of the letter N-back task, with the same task structure and stimulus timing. This task was administered to determine possible transfer of tDCS-induced learning effects to a different domain; namely spatial WM. The stimuli in this task were blue squares (80 by 80 pixels) that could be presented in one of eight outer locations of a 3x3 grid (200 by 200 pixels, on a 23-in. LCD monitor with the screen set to 1280 by 1024). Please see figure 1C for a graphical rendering of the task. As in the pre-and post-session verbal WM task, level of N ranged from 2 to 5 in two sequences, again leading to a total of 48 runs of the task.

The complex span task that we administered was the automated version of the operation span (Ospan) task (Unsworth et al., 2005) using Eprime (Psychology Software Tools, Inc.). In this task, subjects are also required to remember sequences of letters, but now in between the presentation of each of these letters, they have to evaluate mathematical equations (75 in total). After 3 to 7 letters and math equations, participants are required to report the letters in the order in which they were presented. To account for individual differences in mathematical solving speed, a maximal response time is determined based on subjects' performance on 10 practice operations. To prevent

problematic short maximal response times in the second time the task was administered (i.e., in the post-session) because of familiarity with these practice operations, we composed 10 novel operations of similar difficulty level. Order of the two practice sets was

counterbalanced across subjects. Please see Unsworth et al. (2005) for further details of the task and stimulus structure of the Ospan.

Data analysis

Questionnaires

We first examined using the debriefing questionnaires if there were differences between the active stimulation and sham stimulation group in the number of subjects who believed to belong to the active stimulation group using a χ^2 test. To examine whether there were systematical differences in physical sensations between groups, repeated measures ANOVAs were conducted for each of the eight items on the tDCS side-effects questionnaire with Stimulation Session as a within-subject factor and Group as a between-subject factor. To determine whether there was a difference in the effects of anodal vs. sham stimulation on arousal states, scores on each of the four subscales of the AD ACL questionnaire were calculated before and after stimulation for each stimulation session separately and subsequently subtracted from each other to obtain a measure of the effect of electrical stimulation. For each subscale separately, a repeated measures ANOVA was then conducted comparing changes in the resulting difference scores across Sessions between the groups. A Bonferroni correction was applied to account for multiple comparisons for both questionnaires separately, resulting in an alpha of .05/8 = .0063 for the tDCS side effects questionnaire and an alpha of .05/4 = .0125 for the Short Form AD ACL questionnaire.

WM Performance

For the verbal and spatial N-back tasks, accuracy was operationalized using A' (A prime). A' is the non-parametric variant of signal detection theory's d' and takes into account both

Figure 1. *Procedure and tasks. (A) We investigated the effects of 3 sessions of anodal (vs. sham) lDLPFC tDCS combined with a verbal WM task (a letter N-back) on performance on this task, as well as on three WM transfer tasks in a post- versus presession; namely the same verbal WM task with a different stimulus set, a spatial WM task and a complex WM span task (the automated Ospan (Unsworth et al., 2005)). Order of the verbal and spatial WM transfer tasks was counter-balanced across participants, while the complex WM task was always performed last. (B) The active stimulation group received 20 minutes of 1 mA tDCS, while the sham group received only 1 minute of stimulation. The anode was always placed on the lDLPFC (F3), and the cathode on the rSOF (above the right eye). (C) In the verbal WM task, a stream of letters was presented and subjects were required to press a button if the current letter was the same as N stimuli before. In the spatial version of the task, the stimulus to respond to was a blue square that was presented in one of the eight outer locations of a 3x3 grid. In the stimulation sessions, level of N in the verbal WM task alternated between 3 and 4 to ensure a challenging task over all three sessions. In the transfer verbal and spatial WM tasks, level of N ranged from 2 to 5 to index a broader range of participants' abilities.*

hits (correct responses) and false alarms (incorrect responses). We reverted to A' because in our data we encountered blocks of the task in which participants did not have any false alarms, and d' cannot account for these situations. A' can be calculated from hit rate (H) and false alarm rate (F) with the following formula ((Zhang & Mueller, 2005)):

$$A' = \begin{cases} \frac{3}{4} + \frac{H-F}{4} - F(1-H) \; if \; F \leq 0.5 \leq H; \\ \frac{3}{4} + \frac{H-F}{4} - \frac{F}{4H} \qquad if \; F \leq H < 0.5; \\ \frac{3}{4} + \frac{H-F}{4} - \frac{1-H}{4(1-F)} \quad if \; 0.5 < F \leq H. \end{cases}$$

A' scores range from 0 to 1, in which 0 indicates chance performance and 1 perfect accuracy. To determine effects of stimulation on response speed and check whether any changes in accuracy may be explained by altered speed-accuracy trade-offs, we also computed average reaction times (RTs) using correct response trials only.

For the Ospan task, we used the so-called Total Score as our primary measure of WM functioning. Total Score is calculated as the sum of all the letters that were recalled in the correct order. Also, we looked at mathematical operations errors (Math errors) to check for possible trade-offs between letter memory and math performance.

Analytical Approach

To test our first prediction that anodal tDCS would produce larger gains on the WM task than sham stimulation in the three stimulation sessions, we ran a mixed 3 x 4 x 2 repeated measures analysis of variance (ANOVAs) for A' or RT separately with Session (day1, day2, day3) and Block (before, during tDCS, after(1), after(2)) as within-subject factors and Group (active vs. sham) as a between subject factor. As we did not have hypotheses for differential effects of stimulation on difficulty levels, accuracy and RT data was collapsed over levels of N.

Furthermore, to test our second prediction that potential effects of tDCS on verbal WM performance would remain apparent 24 (or 48) hours after stimulation, we conducted a 4 x 2 ANOVA on A' of the baseline blocks with Session (day1, day2, day3 and day5) as a within-subject factor and Group as a between-subject factor. Whenever an interaction with

Group was observed, follow-up tests were run to determine if effects could be ascribed to active or sham stimulation. Similarly, whenever an interaction with Session was observed, we ran additional analyses to investigate the exact time course of stimulation effects.

Additionally, to examine if in the stimulation sessions, physical sensations differed between the active and sham tDCS condition, we ran a 3 x 2 mixed ANOVA for each of the eight sensations on the tDCS side-effects questionnaire, with Session as a within-subject and Group as a between factor. To account for multiple comparisons a Bonferroni correction was applied, leading to an alpha of .05/8 = 0.0063. Similarly, to assess whether mood and/or arousal were differentially affected by stimulation, we ran a 3 x 2 mixed ANOVA separately for each of the four subscales of the short form AD ACL, with Session as within-subject factor and Group as a between-subject factor. Bonferroni correction led to an alpha of 0.05/4 = .0125 for the AD ACL questionnaire.

Our third prediction was that anodal stimulation combined with WM training would induce general improvements in verbal WM performance, i.e. that are not specific to the particular stimuli, domain and task paired with stimulation. To this end we ran mixed 2 x 4 x 2 repeated measures ANOVA's separately on the A' and RT data from the verbal and spatial WM transfer tasks with Session (day0 and day5) and level of N (2, 3, 4, 5) as within-subject factors, and Group as a between subject factor. Additionally, we analysed the pre- and post-session performance scores for the verbal and spatial WM task for each level of N separately, as we hypothesized that learning effects may be more pronounced at higher levels of difficulty, where there may be more room for improvement. Finally, to investigate transfer of possible learning effects to performance on the Ospan task, we ran a 2 x 2 ANOVA on Total score and Math errors with the within-subject factor Session and between-subject factor Stimulation Group.

To test our fourth, final prediction that transfer of benefits may be more pronounced in individuals who displayed the largest improvements in the stimulation sessions, we ran cross-subject Spearman correlations (2-tailed) between tDCS-induced changes in WM performance in the stimulation sessions and changes in performance on the transfer WM tasks, separately for the sham and anodal stimulation groups.

All statistical analyses were conducted using the Statistical Package for the Social Sciences for Mac OS, Version 20 (SPSS Inc, USA). In case of significant main or interaction effects, post-hoc analyses were performed to further clarify the results when suitable. Whenever appropriate, Greenhouse Geisser-corrected results are reported.

Results

Questionnaires

All participants tolerated the tDCS well. Moreover, debriefing questionnaires showed that the majority of subjects in both groups believed to belong to the active stimulation group (active group – 57.1% and sham – 64.3%; $\chi^2(1) = 0.15$, $p = .70$), indicating that our sham control procedure was successful.

On the tDCS side effects questionnaire, the active stimulation group reported slightly higher feelings of itching ($F(1,27) = 4.27$, $p = .049$) and prickling ($F(1,27) = 5.63$, $p = .025$) than the sham group. Also, over the sessions, both groups reported higher levels of headache (main effect Session; $F(2,54) = 3.45$, $p = .039$) and fatigue (main effect Session; $F(2,54) = 4.09$, $p = .022$). However, these effects did not remain significant after correction for multiple comparisons (all other p's > .061). Thus, the stimulation groups did not significantly differ in reported levels of physical sensations in the stimulation sessions.

The AD ACL questionnaire revealed a main effect of Session on the Energy subscale ($F(2,52) = 6.01$, $p = .008$), albeit no Group * Session $F(2,52) = .67$, $p = .49$), reflecting that both groups of subjects felt less energetic at the end of the first and second stimulation session compared to the third. For the subscale Tiredness, the main effect of Session almost reached significance after correction for multiple comparisons ($F(2,50) = 4.58$, $p = .015$), but reported tiredness also did not differ between groups (Group * Session $F(2,50) = .67$, $p = .49$). Although subjects in the sham group showed a small drop on the subscale Tension (Mean = -.615, StE = .274) while the active group did not (Mean = .143, StE = .226), this difference ($F(1,25) = 4.61$, $p = .042$) did not survive the Bonferroni correction. None of the other main effects or interactions reached significance (all p > .058), thus indicating that our active and sham stimulation did not exert differential effects on mood and arousal in our subjects.

Stimulation sessions: Immediate effects of multiple-session tDCS on verbal WM

We tested our first prediction, that multiple sessions with prefrontal anodal stimulation would lead to cumulative verbal WM enhancements by examining the effects of anodal vs sham tDCS on accuracy and RTs on the verbal WM task in the stimulation sessions.

Accuracy. Accuracy on the verbal WM task improved over the three stimulation sessions in both the active and the sham stimulation group (main effect Session - $F(2,56) = 25.89$, $p < .001$). However, whereas the active group shows a specific rapid improvement in the first stimulation session, the sham group displays a more gradual improvement in A' over all three sessions (see Figure 2). This pattern was captured by a significant three-way interaction between Session * Blocks * Stimulation condition ($F(6,168) = 3.53$, $p = .017$). The overall ANOVA furthermore showed a significant interaction between Session * Blocks ($F(6,168) = 4.33$, $p = .006$). The main effect of Stimulation Group ($p = .67$) and all other interactions were not significant (all p > .48).

Additional post-hoc repeated measures ANOVA's for each session separately confirmed that tDCS improved WM only in the initial stimulation session (Interaction between Group * Block ($F(3,84) = 3.20$, $p = .047$), but not in the subsequent second ($p = .427$) or third session ($p = .409$). Notably, performance was not yet at ceiling-level in these sessions (Mean A' Second session for the active group was .89 and sham was .89; Mean A' third session active group was .90 and for sham was .91) indicating that this effect cannot simply be explained by lack of room for further improvement (Mean A' first session active is .86 and sham is .87).

To determine if the difference between groups in change in WM performance in the first session was driven by changes in the active stimulation group, as one would expect, we conducted further follow-up analyses separately per stimulation group. These confirmed that subjects in the active group significantly improved over blocks in the first session ($F(3,42) = 6.75$, $p = .005$), while the sham stimulation group did not ($F(3,42) = .68$, $p = .50$). Moreover, planned contrasts in the active group showed that this main effect of block in particular reflected a significance increase in accuracy after stimulation ended, as indicated by significant higher accuracy in the first ($t(14) = 2.96$, $p = .011$) and second block ($t(14) = 2.96$, $p = .010$) after stimulation compared to baseline. Performance in the block during

stimulation differed from baseline at trend level (t(14) = 1.80, p = .094). An independent t-test between the active and sham group in the baseline block in the first stimulation session showed no significant difference in WM performance (t(28) = 1.14, p = .26), indicating that the groups did not differ in performance before tDCS was applied.

To investigate our second prediction, that the immediate effects of tDCS should still be present 24 to 48 hours after stimulation, we compared accuracy in the baseline blocks of the three stimulation sessions and the post-session. The active and sham group showed similar improvements over sessions (main effect Session (F(3,84) = 31.90, p < .001), main effect Stimulation Group (p = .53), Session * Group (p = .13)). However, as the effect of active tDCS seemed to be specific to the first stimulation session, we post-hoc also compared baseline performance between groups in the first and second stimulation session only. This revealed a larger improvement in the active than in the sham stimulation group (Session * Group F(1,28) = 4.99, p = .034), indicating that the stimulation effects observed in the first session may have carried over to the next day.

RTs. Next to accuracy, we also examined if anodal stimulation combined with WM training speeded up RTs on the WM task. Participants in both groups became faster both within each (main effect Blocks (F(3,84) = 14.20, p < .001)) and across the three stimulation sessions (main effect Session (F(2,56) = 15.24, p < .001). However, this reduction was the same in the active and sham stimulation group (Session * Group; F(2,56) = 0.20, p = .76). The main effect of Stimulation Group (p = .67) and all other interaction effects were not significant (all p > .21). Thus, anodal tDCS did not affect response speed on the verbal WM task in the stimulation sessions.

To summarize, partially in line with our first prediction, anodal tDCS improved verbal WM accuracy in the first, but not in the second and third stimulation session. As tDCS did not affect RTs these effects cannot simply be explained as a speed-accuracy trade-off, nor did we observe differential levels on physical, mood and arousal scales. Furthermore, partially in line with our second prediction, the effects of tDCS were visible in both blocks after stimulation and remained apparent 24 hours after the first stimulation session.

Pre- and post-session: transfer effects of combined tDCS and verbal WM practice

Next, to test our third prediction, we examined if the observed improvements in verbal WM performance by anodal tDCS in the stimulation sessions may reflect more general WM

learning effects. To this end, we investigated differences between the active and sham stimulation group in performance on the three transfer tasks administered in the pre-and post-session.

Verbal WM with a different stimulus set

Accuracy. In contrast to our expectations, multiple sessions with WM practice paired with active stimulation did not enhance verbal WM transfer performance more than sham stimulation (Group * Session (F(1,28) = .96, p = .34) and Group * Session * Level N (F(3,84) = .61, p =.58). Yet, a typical practice effect was observed with subjects performing significantly better on the transfer letter N-back task in the post- (Day5 mean = .901, StE = .013) compared to the pre-session (Day0 mean = .804, StE = .017) (main effect Session (F(1,28) = 35.39, p < .001) and at lower compared to higher levels of N (main effect Level N (F(3,84) = 71.91, p < .001). Furthermore, a significant interaction between Session and Level N (F(3,84) = 9.41, p < .001) likely reflects that the largest transfer gains were found for level 3 and 4 (see Figure 3). This is conceivable because subjects practiced at these levels in the stimulation sessions.

RTs. Contrary to our prediction, analyses of the RT data also did not reveal enhanced performance on the transfer verbal WM tasks after active vs. sham stimulation (F(3,84) = .17, p = .84). All subjects were faster in the post- (Day5 mean = 567 ms; StE = 24 ms) compared to the pre-session (Day0 mean = 654 ms; StE = 21 ms) (main effect Session F(1,28) = 14.69, p = .001), but this did not differ between the stimulation groups (Session * Group F(1,28) = .35, p = .56), thus indicating a general practice effect. Furthermore, a trend was observed towards faster reaction times on the lower compared to the higher levels of N (F(3,84)= 2,69 , p = .079).

Spatial WM

Accuracy. A similar pattern was observed on the spatial WM transfer task. The amount of pre- to post improvement in accuracy did not differ between the experimental groups (Group * Session F(1,28) = .49, p = .49; Group * LevelN F(3,84) = 1.01, p = .39; Group * Session * LevelN F(3,84) = .43, p =.62), indicating that anodal tDCS did not improve

accuracy on the spatial WM task more than sham. Participants displayed significantly higher accuracy scores in the post- (Day 5 mean = .855, StE = .015) compared to the pre-session (Day 0 mean = .803, StE = .014) (main effect Session: $F(1,28) = 15.66$, $p < .001$) and performed better at lower compared to higher levels of N (main effect Level N: $F(3,84)= 119.88$, $p < .001$). The Session * Level N interaction was significant at trend level ($F(3,84) = 2.88$, $p = .075$).

RTs. Again similar to the verbal WM transfer task, no differences in change in RT over time were found between the stimulation groups on the spatial WM transfer task (Group * Session $F(1,28) = .34$, $p = .57$; Group * LevelN $F(3,84) = .70$, $p = .51$; Group * Session * LevelN $F(3,84) = .49$, $p = .66$), indicating that anodal stimulation did not affect RT on the spatial WM transfer task differently from sham stimulation. Again, a practice effect was observed: subjects were faster on the spatial WM task in the post- (Day5 mean = 535 ms; StE = 22 ms) compared to the pre-session (Day0 mean = 594 ms; StE = 19 ms) (main effect Session: $F(1,28) = 9.18$, $p = .005$). At trend level, they were also faster for lower compared to higher levels of N (main effect Level N: $F(1,28) = 2.60$, $p = .079$).

A complex WM task

Active stimulation was also not associated with greater improvements in performance on the complex WM transfer task. Total scores of the Ospan showed no difference between stimulation groups in the number of letters recalled between the pre- and post-session (Session * Stimulation Group - $F(1,28) = .03$, $p = .86$) or between sessions (main effect Session $F(1,28)=1.46$, $p = .24$). Also, no differences were observed in number of errors made in the mathematical operations on the post-compared to the pre-session (main effect Session – $F(1,28) = .26$, $p = .62$; Session * Group – $F(1,28) = .48$, $p = .50$). However, despite our two versions of practice operations, post-hoc analyses revealed that the time set to solve the math operations significantly differed between the pre- (Day0 mean = 6771 ms; StE = 836) and post-test (Day5 mean = 5469 ms; StE = 483) (main effect Session $F(1,28) = 4.92$, $p = .035$). However, this did not differ between the stimulation groups (Group * Session – $F(1,28) = .00$, $p = .99$) and is thus not likely to affect our observed lack of tDCS effects on Ospan task performance.

In summary, our analyses of the transfer task data showed that subjects were more accurate and faster in the post- compared to the pre-session on both the verbal and spatial WM transfer tasks, but their performance did not increase on the complex WM span task. Moreover, the extent of improvement did not differ between the experimental groups. Thus, contrary to our prediction, combined anodal tDCS and verbal WM practice was not associated with larger general WM benefits at the group-level, as measured by our transfer tasks.

Individual difference in tDCS respondence and transfer

Recent research shows that the effects of tDCS can vary greatly across individuals (Berryhill & Jones, 2012; London & Slagter, 2015). We therefore explored if the extent to which a subject benefitted from combined stimulation and WM practice (i.e. tDCS respondence) could predict pre- to post enhancement on the transfer tasks. As no changes were found on the Ospan, analyses were done for the verbal and spatial WM transfer tasks only. As the effects of tDCS on behaviour were only found in the two blocks after stimulation in the first session, tDCS respondence was computed as the difference in accuracy between the before (baseline) block and the average of both blocks after stimulation. Pre- to post-session improvements were calculated per transfer task as the difference in accuracy, collapsed over level of N. One subject (from the sham group) showed exceptionally large improvements pre- to post on the letter WM task (>4 STD from the mean) and was therefore excluded from the analyses.

Interestingly, for the verbal WM task pre- to post improvement significantly correlated with tDCS respondence in the active (Spearman's rho = .550, $p = .034$) but not stimulation group (Spearman's rho = -.042, $p = .89$). A Fisher transformation showed a trend-level difference between these correlations ($z = 1.58$, $p = .05$, (1-tailed) calculated

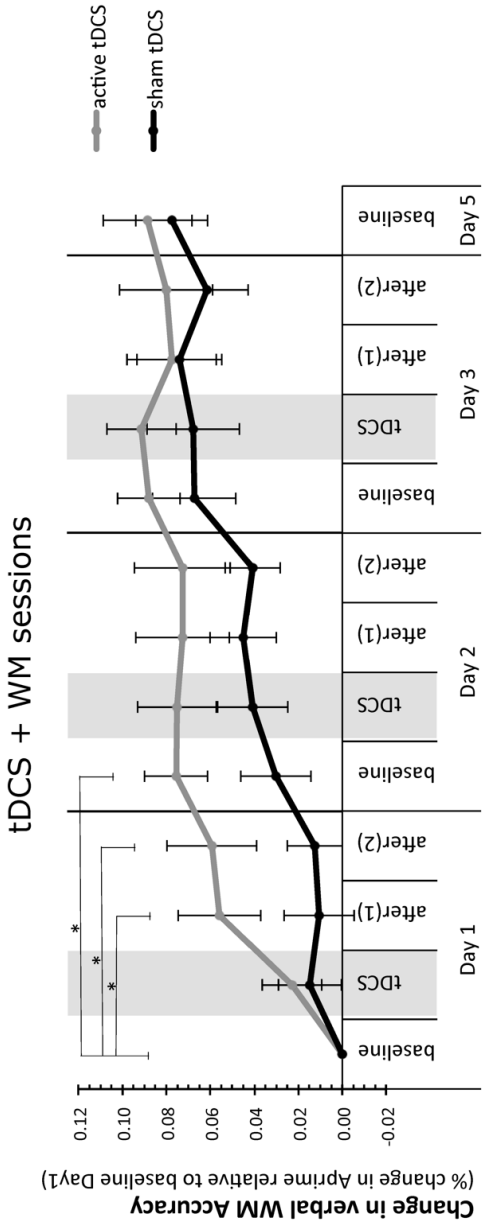


***Figure 2**. Immediate effects of active vs. sham stimulation on the verbal WM task in the combined tDCS and WM sessions. Anodal tDCS over lDLPFC increased accuracy only on the first day of stimulation. Improvements were significant for the two blocks after stimulation ended and remained apparent 24 hours after. No effect of active vs. sham tDCS on WM was found in the second and third day of stimulation. (Note, results are displayed here as percent change from baseline, as groups did not differ in baseline performance. Statistical analyses were done on the actual A' values.)*

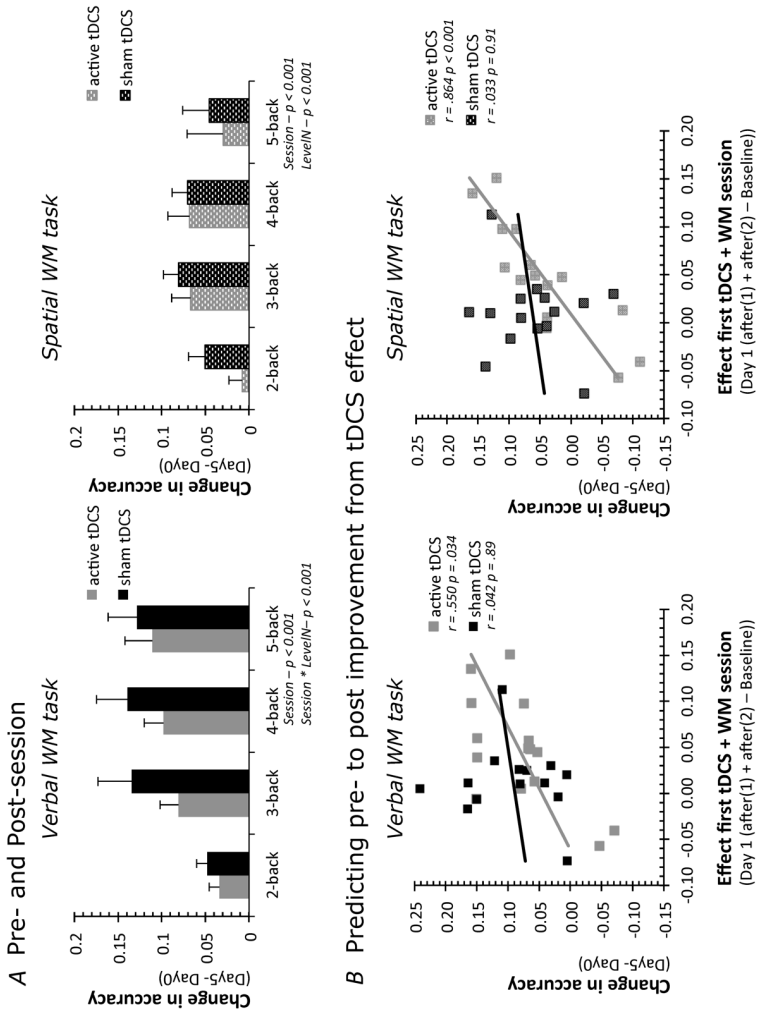


Figure 3. *Transfer effects of active vs. sham stimulation paired with WM training on the different WM tasks. (A) Both groups showed improvements on the verbal and spatial WM transfer task in the post- compared to the pre-session, indicative of a general practice effect. However, at the group-level, active tDCS did not result in greater improvements than sham stimulation. (B) Post-hoc individual difference analyses revealed that in the active stimulation group, subjects that showed larger WM improvements in the first stimulation session also showed the largest pre- to post improvements on both the verbal and spatial WM transfer task. Notably, this relationship was not observed in the sham stimulation group. Thus, tDCS may have enhanced WM functioning specifically in subjects for which the tDCS was most effective. However, performance of two subjects actually worsened after active stimulation in session 1 and on both transfer tasks in the post session. This raises the possibility that repeated stimulation paired with WM training can also impair WM functioning in some subjects.*

with vassarstats.net) indicating that only in the active group, subjects with the largest verbal WM improvements in the stimulation session also showed the largest pre- to post increases on the verbal WM transfer task.

What's more, for the spatial WM transfer task tDCS respondence in the first session also predicted pre- to post improvements in the active (Spearman's rho = .864. $p < .001$), but not the sham stimulation group (Spearman's rho = .033, p = .91). Again Fisher's transformation showed a significant difference between these correlation coefficients ($z = 3.06$, $p = .001$), showing that only in the anodal group, subjects with the largest verbal WM enhancements showed the largest improvements on the spatial WM transfer task (see Figure 3B).

Notably, closer inspection of the data showed that two subjects in the active stimulation group showed WM decrements both in the first stimulation session and between the pre- and posttest sessions. This may indicate that active stimulation may have actually impaired WM function in some individuals. To determine to what extent these subjects contributed to the observed correlation between change in performance during active stimulation and change in performance on the transfer tasks, we ran a control analysis in the active group in which we excluded these subjects. For the verbal WM transfer task this resulted in a non-significant correlation (Spearman's rho = .313, p = .297). However, for the spatial WM transfer task the correlation remained highly significant (Spearman's rho = .819, p = .001), and was furthermore still significantly different from the correlation observed for the sham stimulation group ($z = 2.57$, $p = .0052$).

Care should be taken in interpreting these data because of the small sample size of the current study, but at the very least these results stress the importance of looking at individual differences, as tDCS may improve WM in some individuals, but impair WM in others.

Discussion

The current study aimed to investigate the effects of multiday tDCS stimulation over lDLPFC on verbal WM performance. More specifically, we examined if three sessions with anodal stimulation (1 mA for 20 min) over lDLPFC (anode F3, cathode rOFC) combined with a challenging verbal WM task may result in cumulative as well as general WM improvements. There were five main findings. First, stimulation improved verbal WM in the first stimulation session, replicating findings from previous single-session studies (Bennabi et al., 2014). Notably, these effects were only apparent after but not during stimulation. Furthermore, the greater WM enhancements observed in the first stimulation session in the stimulation (but not sham) group were still present 24 hours later, indicating that the effects of tDCS on WM may not simply reflect short-lived changes in neuronal excitability. However, third, no additional enhancements in verbal WM performance were observed in the subsequent two stimulation sessions. This is in contrast to our expectations, but corroborates previous reports with different stimulation set-ups in which also no additional effects of tDCS were observed in multiple daily stimulation sessions (Lally et al., 2013; Richmond et al., 2014). Fourth, subjects improved on both the verbal and spatial WM transfer tasks, but not the complex WM span, but not significantly more so after they had received anodal stimulation compared to sham. Thus, in line with previous findings (Martin et al., 2013; Richmond et al., 2014), we found no evidence that anodal stimulation might lead to enhanced transfer of WM benefits at the group level. However, fifth and finally, individual difference analyses revealed that within the group that had received anodal stimulation, gains in verbal WM in the first stimulation session predicted pre- to post- training improvements on both the verbal and spatial WM transfer tasks. This relationship was not found for participants that received sham stimulation. Although this cross-subject relationship should be interpreted with caution due to our small sample size, it may provide support for the idea that when effective, anodal tDCS over lDLPFC paired with WM practice may induce WM improvements that outlast the temporary effects of stimulation as well as transfer to a different modality, and thus may reflect true changes in WM functioning. Yet, two subjects in the active stimulation group actually displayed worse WM performance after stimulation and on the verbal and spatial WM transfer tasks. This observation highlights the importance of taking individual differences into account and the need for future studies to determine the factors

that may underlie such individual differences in tDCS respondence. These studies should also examine the extent of potential negative effects of tDCS in some individuals.

Contrary to our expectations, on the second and third day of stimulation, anodal tDCS did not further boost verbal WM in our subjects. Hence, also with an electrode set-up identical to the one that has repeatedly shown effective verbal WM improvements in single-session studies (with the reference over rOFC) and more optimal parameters in the form of current strength (1 mA), duration (20 min) and task paired with stimulation (verbal WM on a challenging level), anodal tDCS only significantly enhanced WM compared to sham on the first of three consecutive days of stimulation. Several reasons may account for this.

Firstly, in previous research with multi-session tDCS and WM it has been proposed that tDCS may be only effective in boosting the "early" phases of learning (Lally et al., 2013; Richmond et al., 2014). Indeed, for example in the different domain of threat detection learning, Bullard et al. (2011) found that subjects who received tDCS during the first of two hours of training showed greater improvements than those that received stimulation during the second hour. Importantly, in the current study, subjects had already performed the task for over 60 minutes (in the pre-session and the first baseline block of the task) before stimulation was admitted. Therefore, they supposedly were already beyond these very first stages of learning in the first stimulation session, making it unlikely that this explanation accounts for the current results.

Secondly, it has been speculated that rather than improving actual WM functioning, tDCS-induced enhancements may be the result of strategy learning which presumably takes effect rapidly, but also reaches ceiling level quickly. Among other functions, the lDLPFC has been related to strategic processes (Bor, Duncan, Wiseman, & Owen, 2003; MacDonald, Cohen, Stenger, & Carter, 2000). Yet, if anodal tDCS would facilitate strategy learning only, one would not expect such benefits to transfer to tasks that rely on different strategies. Our subjects consistently reported the use of a verbal strategy in the verbal WM task and non-verbal strategies in the spatial WM task. Still, we found that in those subjects that benefited most from the stimulation paired with practice, improvements in verbal WM transferred to a spatial version of the task. This makes it unlikely that the here observed effects of tDCS reflect verbal WM strategy learning solely.

Third and finally, the absence of tDCS effects in the second and third day of stimulation may be related to the time implemented between stimulation sessions. tDCS effects on behaviour that are caused by neuroexcitability changes are generally assumed to have worn out after minutes (with very short stimulation durations) or a few hours (>10 min of stimulation) (Nitsche et al., 2008; Nitsche & Paulus, 2000). Spacing the sessions with 24 hours in between has therefore generally been considered safe in ensuring that neuroexcitability effects from previous sessions have worn out before new stimulation is admitted. Interestingly, in the current study WM performance was significantly higher in the baseline block in the second stimulation session in the anodal compared to the sham group. This may reflect that anodal tDCS induced longer lasting effects or it could indicate that neuroexcitability levels in the stimulated areas had in fact not yet returned to baseline. Interestingly, a new study with seven sessions of anodal tDCS over DLPFC paired with a spatial WM training found significantly larger WM gains between the third and fourth session in subjects in which these sessions were separated by a weekend (i.e., 72 hours) than for those that received them on consecutive days (i.e., 24 hours) (Au et al., 2016). This implies that to achieve cumulative tDCS effects, multiple stimulation sessions may in fact need to be spaced more than 24 hours apart.

However, as the study by Au and colleagues (2016) included no daily baseline measure, it remains unclear whether their results should be interpreted as a larger within-session effect of tDCS, or stems from better learning consolidation after 72 hours of 'rest' time. Similar to muscles after physical exercise, it is conceivable that for optimal learning to take place in brain regions a minimum of rest time is required to consolidate effects. Unfortunately, little is currently known about optimal intervals for enhancing cognitive functioning with training or tDCS (Goldsworthy, Pitcher, & Ridding, 2015). Furthermore, although the current study did not find behavioural effects of tDCS in subsequent sessions, we do not know whether the same learning effects would have been found with only one of the three daily stimulation sessions. Research that systematically investigates the effects of spacing of stimulation sessions on WM performance is necessary to determine which multiple session protocol(s) combining tDCS and WM training may be most effective to cumulatively and lastingly enhance WM functioning. In addition to such stimulation parameters, these studies should also take other variables into account that may facilitate

transfer of WM learning, such as stimulus- and task-variability and optimal levels of arousal (e.g. see Slagter, 2012).

In line with previous studies (Martin et al., 2013; Richmond et al., 2014), at the group-level anodal tDCS over lDLPFC was not associated with greater pre- to post WM improvements than sham stimulation. This may seem contrary to the conclusion of a recent meta-analysis study by Mancuso et al. (2016), that reported a small, but significant effect of left DLPFC anodal stimulation coupled with WM training (). However, this analysis, based on 10 studies in total, included 6 single-session studies. It is hence possible that the reported effect is completely driven by effects of anodal stimulation in the first session. This would be quite in line with our finding of an effect of anodal tDCS on WM performance in the first session only.

However, additional individual differences analyses showed that the degree to which tDCS combined with WM practice was effective in boosting WM in the first session, predicted gains on the verbal and spatial WM transfer task post-training. Our results thereby add to the growing number of studies that report that the effects of tDCS may vary substantially across individuals (e.g. London & Slagter, 2015). Notably, two subjects in the active stimulation group actually displayed decrements in WM performance, suggesting that anodal lDLPFC stimulation may also impair WM performance in some individuals. Without these subjects, the relationship between individual tDCS respondence and individual change in performance on the spatial WM transfer task remained significant, but this was no longer the case for the verbal WM transfer task. The latter could reflect reduced statistical power, but may also indicate that this relationship was spurious.

Future studies with larger samples sizes are necessary to determine whether tDCS respondence during WM training determines the strength of transfer effects. This research should also include a stimulation-only (i.e., without WM training) group, so that effects of tDCS and WM training can be better separated. Lastly, future studies should determine why some individuals may and why others may not respond to stimulation, or even in a negative manner.

Several explanations have been proposed for individual differences in tDCS respondence. Firstly, a recent modelling study has indicated that current flow may be strongly affected by individual differences in anatomy, skull thickness and folding of the

cortex (Opitz et al., 2015). As a result of this, standard tDCS set-ups may be more or less effective in affecting cognitive functioning, simply because they are more or less successful in delivering current to the target brain area. Interestingly, a recent study indeed observed a direct relationship between individually simulated current density values in the DLPFC and behavioural effects of prefrontal anodal tDCS on verbal WM (Kim et al., 2014). Secondly, the effect of tDCS may be dependent on the 'baseline' activation level of an area, needing some activity to 'grasp' on (Berryhill & Jones, 2012), but also interacting with a delicate optimal balance of activity levels for the best cognitive performance (London & Slagter, 2015). Thirdly, and likely related to this, it has been suggested that baseline excitability/inhibitory balances in the stimulated cortex (reflected in GABA/Glutamate concentration ratios), may predict the effectiveness of tDCS (Krause, Márquez-Ruiz, & Cohen Kadosh, 2013). Future studies that combine stimulation with neuroimaging and current flow modelling are needed to shed more light on individual differences in tDCS respondence and help predict which subjects may benefit most from tDCS and why, but also how possible negative effects of stimulation can be prevented.

As no neuroimaging was included, we can only speculate about the underlying mechanisms through which anodal tDCS over lDLPFC modulated WM functioning in our study. We expect the effects to stem primarily from changes in functioning in the left DLPFC itself, an area that is known to play a key role in WM. Furthermore, as activation in this region has been related both to verbal and spatial WM (Owen et al., 2005), it provides a logical neural basis for the transfer of stimulation effects we found to the spatial domain. Nonetheless, studies that combined tDCS with neuroimaging have reported more widespread changes after lDLPFC stimulation (Keeser et al., 2011; Stagg et al., 2013), making it likely that the tDCS effects on WM in the current study are not confined to lDLPFC alone, but may also include other regions.

Conclusion

To conclude, repeated anodal tDCS over lDLPFC concurrent with a challenging verbal WM task improved verbal WM performance only in the first of three daily stimulation sessions. Furthermore, individual differences in respondence to stimulation paired with WM practice predicted the extent and direction of WM improvement on a verbal and spatial transfer task (2 days after stimulation). More research is needed to determine which individuals may benefit the most from stimulation and why some individuals may be negatively affected, as well as to determine the optimal spacing of sessions for WM learning to take place. With a growing aging population and WM that is known to decrease over the lifespan, future research in this direction may help delineate the optimal parameters to use tDCS most effectively to enhance WM functioning in a range of individuals.

Chapter 3

State or trait? MRS-measured GABA and Glutamate concentrations are not modulated by task demand and do not robustly predict task performance

Talsma, L.J., Van Loon, A., Scholte, H.S., Slagter, H.A., State or trait? MRS-measured GABA and Glutamate concentrations are not modulated by task demand and do not robustly predict task performance. ***Submitted Manuscript***.

Abstract

In recent years, Magnetic Resonance Spectroscopy (MRS) has become a popular method to non-invasively study the relationship between in-vivo concentrations of neurotransmitters such as GABA and Glutamate and cognitive functioning in humans. However, currently, it is unclear to what extent MRS measures reflect stable trait-like neurotransmitter levels, or may be sensitive to the brain's activity state as well. Therefore, this study investigated if cortical GABA (GABA+/Cr) and Glutamate (Glx/Cr) levels differ as a function of task demand, and if so, in which activity state these measures may best predict behavioural performance. We acquired 3T-MRS data from thirty healthy men in two brain areas during different task demands: the medial occipital cortex (OC), at rest (eyes closed) and while subjects watched a movie (on-task); and the left dorsolateral prefrontal cortex (lDLPFC), at rest, during an easy working memory (WM) task, and during a challenging WM task. Task demand had no effect on the concentration of GABA or Glutamate in either brain region. Moreover, we observed no correlations between GABA and Glutamate concentrations and behavioural performance; occipital neurotransmitter concentrations did not predict visual discrimination nor did those in lDLPFC predict WM updating accuracy, capacity or maintenance. These null findings were supported by Bayesian statistics. In conclusion, these results suggest that with 3T-MRS we measure relatively stable trait-like neurotransmitter concentrations, but at the same time question the validity of 3T-MRS as a method to relate GABA and Glutamate concentrations to behaviour.

Introduction

Magnetic Resonance Spectroscopy (MRS) is a non-ionizing technique that can be used to non-invasively determine in-vivo neurotransmitter concentrations, such as GABA and Glutamate, in the human brain. Being the primary inhibitory and excitatory neurotransmitters, GABA and Glutamate play a key role in regulating neuronal excitability and hence determining cortical functioning. As such, MRS is a promising neuroimaging technique to investigate the relationship between neurotransmitter concentrations and brain functioning and behaviour has opened new avenues for investigating the relationship between neurotransmitter concentrations and brain functioning and behaviour (Isaacson & Scanziani, 2011).

In recent years, MRS-measured cortical GABA in specific brain areas has been linked to inter-individual differences in related functions in a range of cognitive domains. For example, GABA concentrations in the sensorimotor cortex have shown to be predictive motor performance, as well as tactile discrimination (Puts, Edden, Evans, McGlone, & McGonigle, 2011), while GABA in the occipital cortex (OC) has been related to visual performance (Edden, Muthukumaraswamy, Freeman, & Singh, 2009; Sandberg et al., 2014; Song, Sandberg, Andersen, Blicher, & Rees, 2017; van Loon et al., 2013). Similarly, GABA levels in prefrontal areas have been found to relate to higher-order cognitive functions, such as working memory (Yoon, Grandelis, & Maddock, 2016) and attention (Kihara, Kondo, & Kawahara, 2016). Yet, in many of these studies, sample sizes were relatively small and their findings hence warrant replication.

Moreover, so far, the vast majority of studies linking cortical neurotransmitter levels to behaviour have quantified these in rest only, thus assuming that MRS-measured GABA and Glutamate levels reflect stable individual 'trait' differences. Yet, other studies that have looked at differences in GABA and Glutamate concentrations as a function of experimental manipulation have found these concentrations may in fact not be so static but can change over relatively short time windows, e.g. as a function of time on task (Michels et al., 2012) or after learning (Floyer-Lea, Wylezinska, Kincses, & Matthews, 2006; Shibata et al., 2017).

At present, it thus remains unclear to what extent MRS-measured neurotransmitter levels reflect stable and consistent, trait-like neurotransmitter concentrations, or in fact are

sensitive to changes in metabolic activity as a function of task demand. However, this knowledge is important for our theoretical understanding of what it is that we measure with MRS ('trait' or 'state' concentrations) and also has pivotal implications for the design of future studies that aim to experimentally manipulate neurotransmitter concentrations. Moreover, it is still unclear whether neurotransmitter levels that are measured during performance of a task that activates the brain area of interest, are more indicative of performance than neurotransmitter levels measured at rest. Namely, similar to what has been observed with other neuroimaging methods such as EEG and fMRI, neurotransmitter activity investigated in an active state may in fact be a better predictor of behaviour than the measures acquired at rest that we currently use in MRS research.

To address these outstanding issues, the current study investigated whether 3T-MRS GABA and Glutamate concentrations vary as a function of task demand, and if so, in which brain state (rest or on-task), these concentrations may best predict cognitive performance. To this end, we scanned both a primary sensory (medial occipital cortex, OC) and higher-order cognitive brain region (left dorsolateral prefrontal cortex lDLPFC).

The OC, key for visual processing was scanned once at rest (eyes closed) and once while subjects watched a movie (on-task). The lDLPFC, which has consistently been shown to be active with temporarily holding and manipulating information in working memory (WM) (Owen, McMillan, Laird, & Bullmore, 2005), was measured three times: at rest, during an easy WM task (letter 2-back), and during a challenging WM task (adaptive letter N-back).

In a separate behavioural session participants performed a visual discrimination task (with oblique grating patches) and two WM tasks (letter N-back updating and Sternberg task). This design critically permitted us to determine, first of all, if MRS-measured GABA and Glutamate levels reflect stable trait-like indicators of brain neurotransmitter concentrations or whether they are influenced by the cognitive state of the subject. Secondly, it allowed us to examine, if concentrations fluctuate with state and task, which state may best predict individual differences in performance outside the scanner.

We expected to replicate previous findings which associated higher occipital GABA levels with better visual discrimination performance (Edden et al., 2009) and higher lateral prefrontal GABA levels with better WM performance (Yoon et al., 2016). We expected no such correlations with Glutamate levels. Moreover, it has recently been proposed that not so much the concentration of each neurotransmitter individually, but the relative concentrations of GABA and Glutamate (i.e. the cortical excitation/inhibition balance) may provide a more accurate reflection of cortical functioning and hence be a better predictor of cognitive performance (Krause & Cohen Kadosh, 2014). In line with this, we expect that combining information from both measures into Glutamate/GABA ratios may better predict individual differences in performance than GABA levels only.

Methods

Participants

Thirty healthy volunteers (mean age: 21,2 years, StD: 2,5; all men) were recruited via the university subject pool and participated in return for a monetary reward or course credit. As cortical GABA concentrations have shown to vary with the menstrual cycle (De Bondt, De Belder, Vanhevel, Jacquemyn, & Parizel, 2015; Harada, Kubo, Nose, Nishitani, & Matsuda, 2011), only male participants were included. Subjects gave written informed consent before the experiment and the experiment was approved by the University of Amsterdam ethical committee. All reported no history of psychiatric conditions, complied to the rules for MRI safety, and had normal or corrected-to-normal vision.

Procedure

Subjects came to the lab for two sessions (see Figure 1A), a behavioural and an MRS session, planned at the same time of day with maximally 11 days (mean: 4,2 StD: 3,2) in between. In the first behavioural session, they were seated in a comfortable chair in front of a computer screen (at approximately 90 cm distance) and performed three WM tasks and a visual discrimination task. Order of the tasks was counter-balanced across subjects, and they first practiced each task before data collection started. At the end of the behavioural session, subjects also performed an attentional blink task, but these data were not analysed for the current paper.

The visual discrimination task was an orientation task with oblique gratings, similar to the one used by Edden et al. (2009), as described in more detail below. WM performance

was measured with two versions of the letter N-back task; one with level N fixed (WM updating accuracy) and one with level N adapted to performance (WM updating capacity). Also, to be able to examine the extent to which metabolite levels in lDLPFC could predict WM performance in a more generalized manner (i.e., on a different WM task than administered during scanning), we furthermore administered a Sternberg task to determine WM maintenance more specifically. Importantly, the Sternberg task has consistently been related to functioning in the lDLPFC in particular, both in functional neuroimaging (Altamura et al., 2007) and non-invasive brain stimulation (Jansma et al., 2013) studies. All three WM tasks are described in more detail below.

In the second MRS session, five MRS scans and an anatomical scan were acquired. In two of the MRS scans, the voxel was placed over the medial occipital cortex (OC) (primary visual cortex, see figure 1B) and in the other three over the left dorsolateral prefrontal cortex lDLPFC (see figure 1C). The OC voxel was scanned twice, once when subjects had their eyes closed (rest condition) and once when they watched a movie (active condition). The lDLPFC voxel was scanned three times: once when subjects had their eyes closed (rest condition), and once while they performed an easy (letter 2-back) and once a challenging (adaptive letter N-back) WM task (see for more info about the task below). By manipulating WM task difficulty for the lDLPFC voxel, we aimed to also investigate possible differences in neurotransmitter levels depending on the extent of cortical engagement. To prevent carry-over effects of task activity in the MRS signal between the different activity states, the lDLPFC and occipital voxels were scanned in an interleaved manner. Also, order of the tasks (and thus activity states) was counter-balanced across subjects. Due to a shortage of time, for one subject, the lDLPFC rest condition scan could not be acquired.

MRS data acquisition and analysis

Scanning was performed on a 3T Philips Achieva TX MRI scanner (Philips Healthcare) with an eight-channel head coil. Spectroscopy voxel localization was performed by the experimenter according to the individual's anatomical landmarks as visible from an initial anatomical scan. The occipital voxel (30 x 25 x 20 mm) was placed bilaterally over the calcarine sulcus (see Figure 1B) (cf. van Loon et al., 2013). For the left dorsolateral prefrontal cortex lDLPFC voxel, the centre of the voxel (30 × 20 × 25 mm) was placed on the left middle

frontal gyrus, with the posterior border of the voxel positioned anterior to the precentral sulcus (see Figure 1C). Both voxels were placed with care to exclude cerebral spinal fluid (CSF) from the ventricles or the cortical surface.

Edited ^{1}H J-difference spectra were acquired for each voxel using a GABA-specific sequence of the Mescher-Garwood point-resolved spectroscopy (MEGA-PRESS) method (Waddell, Avison, Joers, & Gore, 2007). Scanning took approximately 12 minutes per acquisition, during which 384 transients were collected (TE = 73 ms; TR = 2,000 ms). On the odd transients, a 15,64 ms sinc-centreediting pulse (64 Hz full width at half maximum) was applied in an interleaved manner at 1,9 ppm and 4,6 ppm to excite GABA and suppress water respectively.

Spectral data were analysed with the MATLAB-based package GANNET v2.1 (Edden et al. 2014, www.gabamrs.com). Using the in-build options of the GannetLoad-function, the following processing steps were performed: time-domain frequency-and-phase correction using spectral correction, line broadening with an exponential apodization function, Fast Fourier Transform (FFT), time averaging, frequency and phase correction based upon fitting of the Choline and Creatine signals, pairwise rejection of the data for which fitting parameters are greater than 3 SDs from the mean, and finally, subtraction of the even from the odd transients to generate the edited difference spectrum. Notably, in this edited difference spectrum, the GABA signal is contaminated by the macromolecule homocarnosine (Edden, Puts, & Barker, 2012), a GABA derivative, and thus often referred to as GABA+. Also, as the spectra of Glutamate and Glutamine are known to overlap at 3T, the combined measure of Glx was used as the best measure for Glutamate.

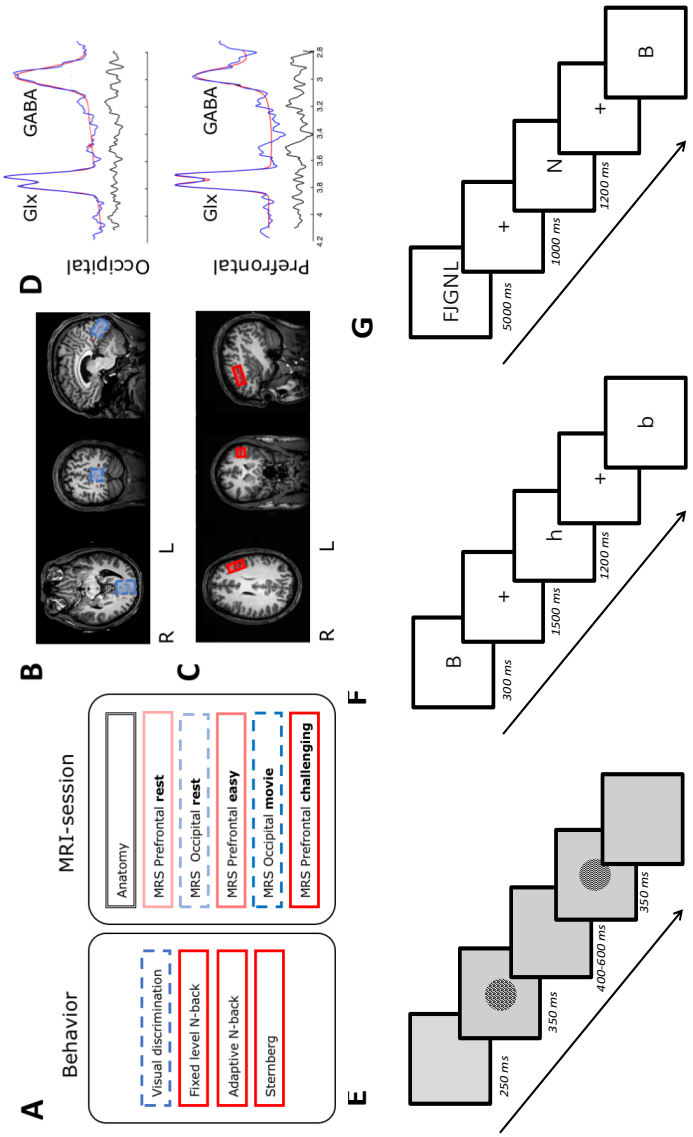


Figure 1. *Schematic illustration of the research design and methods. Subjects came to the lab for two sessions, an MRI and behavioural session. (**A**). In the MRI session, 3T-MRS (MEGA-PRESS) was used to measure GABA (GABA+/Cr) and Glutamate (Glx/Cr) levels in an occipital (OC) and a prefrontal voxel (lDLPFC) under different activity conditions. The occipital voxel (**B**) was scanned twice: once when subjects had their eyes closed (rest) and once while they watched a movie (on-task). The prefrontal voxel (**C**) was scanned three times: once with eyes closed (rest), once while subjects performed an easy WM updating task (letter 2-back) and once while they performed a challenging WM updating task (an adaptive letter N-back). Order of the activity conditions was counter-balanced between subjects, but the occipital voxel was always scanned in between the prefrontal voxels. (**D**) Outcome of the modelling of the GABA and Glx signal in the occipital and prefrontal voxel for a typical subject (output from the Gannet analysis toolbox (Edden et al. 2014, www.gabamrs.com). In blue the edited spectrum is shown, overlaid in red is the model of best fit (using a simple Gaussian model) and the residual of these is shown in black. (**E-F**): In a separate behavioural session, we administered four tasks to determine cognitive performance. An oblique visual discrimination task (**E**) was performed to relate to neurotransmitter levels in the occipital voxel. Furthermore, three WM tasks (**F**) were administered to relate to neurotransmitter levels in the prefrontal voxel: two versions of the letter N-back WM task (**F**) to determine both WM updating accuracy (level N ranged 2-5) and Capacity (level N on-line adapted to performance) as well as a Sternberg WM maintenance task (**G**).*

Subsequently, using the GannetFit function of GANNET, GABA+ and Glx functions were modelled to the data together (see Figure 1D) and ratios relative to Creatine (Cr) were calculated (i.e. GABA+/Cr and Glx/Cr). Normalizing values to Creatine has been shown to be superior to normalizing to H2O with regard to intra-subject stability (Bogner et al., 2010; Greenhouse, Noah, Maddock, & Ivry, 2016) and is known to substantially reduce inter-subject variance as a result of differences in global signal strength, as well as those stemming from differences in tissue fractions in the scanned voxel (grey matter, white matter, and cerebrospinal). Calculating GABA and Glutamate levels relative to Creatine thus makes coregistration, segmentation and the calculation of CSF corrected values superfluous.

Scans were excluded when no Creatine peak was visible in the data (N=3; corresponding to Creatine Signal to Noise ration (SNR) < 50), model fit turned out to be poor (N=2; GABAGlxModelfiterror > 15), or the GABA+ or Glx peak could not be confidently be determined (N=1; GABA SNR < 3). Furthermore, we used the Statistical Package for the Social Sciences for Mac OS, Version 24 (IBM, Armonk, NY) to identify outliers as a result two GABA+/Cr values in the lDLPFC rest condition (0.359, 0.238) and three in the easy WM condition (0.347, 0.246, 0.201) as extreme outliers. As these outlier values were also much higher than previously reported in the prefrontal cortex (De Bondt et al., 2015; Greenhouse et al., 2016), these were excluded from further analyses. The remaining GABA+/Cr and Glx/Cr ratios in the OC and lDLPFC voxels fell all in agreement with previous studies that measured similar regions during rest (De Bondt et al., 2015; Edden et al., 2009; Greenhouse et al., 2016; Iwabuchi et al., 2017; Michels et al., 2012; Yoon et al., 2016).

Visual discrimination task

In the behavioural session only, participants performed a visual discrimination task that was based on the one used by Edden and colleagues (2009). In this task, subjects were sequentially shown two circles with oblique grating patterns and asked to indicate if the second of the two was rotated clockwise (left mouse button) or counter clockwise (right mouse button) with respect to the first one (see Figure 1E). The circular gratings (diameter: 4 degrees visual angle, Spatial frequency: 3 cycles/degree, contrast: 80%, mean luminance: 44,5 cd/m2) were displayed for 350 ms each, with an inter stimulus interval chosen

randomly between 400 and 600 ms. During the task, the difference in orientation was adjusted logarithmically, using two interleaved staircases that applied the principle of one up, two down. Mean orientation of both gratings was always 45 degrees, as Edden et al. (2009) observed the highest correlation between GABA and orientation discrimination threshold in an oblique compared to a vertical average condition. An auditory tone provided feedback on each trial, and one run of the task continued until both staircases completed 12 reversals. Subjects completed two runs of the task, but only the second run was used for analysis due to expected early task training effects (as reported by Edden et al., 2009). Of this second run, the first two reversals were discarded and visual discrimination thresholds were subsequently computed for each participant by averaging the angle difference between the two stimuli over the last 10 reversals and both staircases (cf. Edden et al., 2009).

Working memory tasks

The primary working memory (WM) task that subjects performed in both the behavioural and MRS session was a letter N-back task (see Figure 1F). In this WM updating task, subjects are presented with a stream of letters and asked to indicate if the currently presented letter is the same as the one presented N stimuli back. Hereby, N is an integer and the value of N determines the difficulty level of the task. With higher levels of N, more stimuli have to be held in WM in sequential order, increasing WM load. As WM content has to be continuously updated, the letter N-back task is considered to be a demanding WM task. Therefore it is a standard task to investigate WM updating performance (Jaeggi, Buschkuehl, Perrig, & Meier, 2010) and importantly, has consistently been related to processing in the lDLPFC (e.g., see meta-analysis by Owen et al., 2005).

In our letter N-back task letters (Arial, font size 72, letterset ["A", "B", "C", "D", "E", "F", "G", "H", "J", "K"]) were presented for 300 ms at the centre of a screen, followed by a 1500 ms inter-stimulus interval in which a fixation cross was displayed (Arial, font size 20). In the behavioural session, we presented black letters on a white screen, while in the MRS session we showed subjects white letters on a black background because of the dimly lit nature of the scanning room. Of the presented letters, approximately 35% were so-called targets, i.e., the letter in the current trial matched the letter presented N letters back.

Letters could be presented in upper or lower case and would still classify as the same letter (i.e., a target). If presented with a target, subjects were required to press the space bar on the keyboard in front of them in the behavioural session, or one of the buttons on the button-box in the scanner. Runs consisted of a stream of 20 + N stimuli each and were self-paced in the behavioural session to allow the subject to take small breaks in between and enhance focus during the runs, but they started automatically in the MRI-session to ensure the task was performed for the entire time of the scan.

Subjects performed both a fixed-level and an adaptive version of the letter N-back task in the behavioural and in the MRS session. First of all, in the behavioural session, subjects performed 24 runs of a fixed-level version of the task in which level N sequentially increased from 2 to 5, with steps of 1. We used this version of the task to calculate WM updating accuracy, which was operationalized using A' (A prime). A' is the non-parametric variant of signal detection theory's d' and takes into account both hits (correct responses) and false alarms (incorrect responses). In contrast to d', A' can account for situations in which participants do not show any false alarms, which sometimes occurred on lower N levels of our task. A' scores range from 0 to 1, with 0 indicating chance performance and 1 perfect accuracy. A' can be calculated from hit rate (H) and false alarm rate (F) with the following formula (Zhang & Mueller, 2005):

$$A' = \begin{cases} \frac{3}{4} + \frac{H-F}{4} - F(1-H) & if\ F \leq 0.5 \leq H; \\ \frac{3}{4} + \frac{H-F}{4} - \frac{F}{4H} & if\ F \leq H < 0.5; \\ \frac{3}{4} + \frac{H-F}{4} - \frac{1-H}{4(1-F)} & if\ 0.5 < F \leq H. \end{cases}$$

Secondly, in the adaptive version of the task, level of N always started with N = 2 (set as the lowest possible level N) and subsequently adapted to performance by going up one step (current N+1) if subjects made fewer than three errors, and down one step (current N-1) if they may more than five errors (similar to Jaeggi, Buschkuehl, Jonides, & Perrig, 2008). The adaptive version of the task in the behavioural session also consisted of 24 runs, of which we calculated average level N over the last 21 runs only (disregarding the first 3 runs to allow each individual some ramp-up time to their average level), to use as our measure for WM updating capacity. Additionally, in the MRS-session, subjects once performed the task with N fixed to level 2 (the *easy* WM condition) and once with N adapted to performance in the same way as in the behavioural session (the *challenging* WM condition). Hereby, the amount of runs of the task was determined to ensure it covered the whole MRS-scan. We used Presentation software (Neurobehavioural Systems, Inc.) to administer the letter N-back task.

In the Sternberg task (see Figure 1G), subjects were presented with a string of five letters that they were required to remember (5000 ms). Consequently, one letter was shown on the screen at a time (1200 ms per letter, 1000 ms fixation cross in between) for which subjects had to indicate whether that letter was in the currently remembered string (press 'N' key) or not (press 'Z' key). All letters were presented in uppercase and came from a predetermined letterset (["B", "D", "F", "G", "H", "J", "K", "L", "M", "N"], Arial, Font size 60). Per run, 10 letters were presented of which 50% was a target. Subjects completed 10 runs of the task. We determined WM maintenance accuracy by calculating the number of correct trials divided over the total. The data of one subject was discarded because of extreme low below chance performance (accuracy = 30% correct).

Specific hypotheses and statistical approach

To investigate whether GABA and Glutamate levels measured with MRS differed between different levels of task demand, repeated measures ANOVA's were run with task demand as the within-subjects factor and GABA+/Cr or Glx/Cr as the dependent variable for the occipital and lDLPFC voxel separately. Additionally, to investigate within- subject stability of MRS measured GABA and Glutamate levels across activity states, after testing for normality, individual Pearson correlations were run between the GABA+/Cr and Glx/Cr levels measured under the different task conditions, again separately for the prefrontal and occipital brain region. Namely, we reasoned that if these neurotransmitter measures reflect stable, trait-like neurotransmitter concentrations, they should correlate across activity states across subjects. Furthermore, we wanted to replicate previous reports of within-subject regional specificity for GABA-levels (Bogner et al., 2010; Greenhouse et al., 2016), therefore we also correlated occipital and lDLPFC neurotransmitter measures in the resting state.

Depending on the outcomes of our first ANOVA's, we followed one of two approaches. In case of no specific effect of activity state, GABA+/Cr and Glx/Cr concentrations were averaged over all conditions separately for the OC and lDLPFC and these average measures were related to individual task performance; OC: rest and movie, DLPFC: rest, easy (letter 2-back) and challenging (adaptive N-back). In case of systematic differences in MRS measures between activity states, multiple regression analyses including all activity states as predictors were run to determine which activity state best predicted behavioural performance. Analyses were conducted separately for GABA and Glutamate and included the relevant brain area and corresponding task only. In addition, we also ran control analyses in which neurotransmitter concentrations of the task-unrelated brain region were related to the behavioural measures, for which we did not expect to find significant correlations.

To test the hypothesis that not so much GABA and Glutamate individually, but in fact their relative concentrations (i.e. the excitation/inhibition balance) may provide the best predictor of cognitive functioning, for both brain areas we also calculated glutamate/GABA ratios from our data and investigated the extent to which these ratios predicted behavioural performance.

The correlation with our behavioural measures were investigated with 2-tailed Pearson correlations. To account for the fact that GABA and the Glutamate/GABA ratio are highly related and investigate a similar research question, we divided alpha over two to determine significant levels and correct for multiple comparisons.

All statistical analyses were conducted using the Statistical Package for the Social Sciences for Mac OS, Version 24 (IBM, Armonk, NY). Furthermore, we additionally repeated our analyses with Bayesian statistics using the open-software package JASP (http://www.jasp-stats.org, Wagenmakers, Marsman, et al., 2017)). The resulting Bayes factors, which grade the intensity of evidence for the null (H0) and alternative hypothesis (H1), and values were interpreted according to the corresponding classification scheme (see for elaboration Wagenmakers, Love, et al., 2017): 1/30 < Bf < 1/10, Strong evidence for H0; 1/10 < Bf < 1/3, Moderate evidence for H0; 1/3 < Bf < 1, Anecdotal evidence for H0; Bf = 1, No evidence; 1 < Bf < 3, Anecdotal evidence for H1; 3 < Bf < 10, Moderate evidence for H1; 10 < Bf < 30, Strong evidence for H1.

Results

Descriptives cognitive performance

Subjects performed in line with expectations on all tasks in both the behavioural and the MRS session. Visual discrimination angle thresholds ranged between 0.845 and 3.873 (Mean: 2.347, StD: 0.868), similar to the range reported by e.g., Edden et al. (2009). For the WM tasks, accuracy was well above chance for all participants on both the fixed WM updating letter N-back task (range A': 0.623 to 0.920, mean: 0.827, StD: 0.075), and the Sternberg maintenance task (range Accuracy: 0.850 to 0.950, mean: 0.921, StD: 0.038). Moreover, on the adaptive N-back task, subjects showed a relatively wide spread in WM updating capacity (range mean level N: 1.53 to 6.86, mean: 4.17, StD: 1.28), i.e., inter-individual differences in WM updating capacity were relatively large.

In the MRS session, accuracy on the 2-letter N-back task ranged between 0.79 and 1.00 (A' Mean: 0.95, StD: 0.04), indicating ceiling level or close to ceiling level performance in all subjects. Also, WM capacity levels on the adaptive N-back task were similar to those observed in the behavioural session (range mean level N: 2.20 to 7.80, Mean: 4.54, StD: 1.50) and correlated well within subjects ($r = .826$, $p < .001$; Bf = 743996). These findings show that our task manipulation was effective, with the adaptive WM task (i.e. challenging WM condition) placing greater demands on WM processes than the 2-letter N-back task (i.e. easy WM condition).

Resting-state versus on-task occipital and prefrontal GABA and Glutamate levels

To address our first research question, we assessed whether MRS-measured GABA and Glutamate levels differed as a function of task demand (i.e. reflect activity state). To this end, we compared neurotransmitter levels measured in rest with those measured during stimulus- or task-induced activity, separately for the OC and lDLPFC voxels.

In the occipital voxel, we observed no difference in GABA or Glutamate concentrations between the rest and active (movie watching) condition (GABA: ($F(1,29) = .904$, $p = .350$, Bf = 0.37) and Glutamate: ($F(1,29) < .001$ $p = .981$, Bf = 0.26). Similarly, GABA levels in lDLPFC did not show significant differences between the three activity conditions (Rest; Easy WM and Challenging WM) ($F(2,38) = .210$, $p = .811$, Bf = 0.16) and neither did

Glutamate ($F(2,38) = .210$, $p = .811$, $Bf = 0.30$). Thus, both in the occipital and in the prefrontal cortex we found no effect of task demand on GABA and Glutamate concentrations. Importantly, in all these cases our Bayesian statistics showed moderate evidence for the null-hypothesis. Together, these findings thus indicate that our MRS measure was insensitive to possible stimulus- or task-induced changes in GABA and Glutamate levels.

Additionally, within subjects GABA levels correlated well between the Rest and Movie conditions in OC ($r(29) = .568$ $p < .001$, $Bf = 37.8$) as well as between all task demand conditions in lDLPFC (Rest and Easy WM: $r(21) = .400$ $p = .032$, $Bf = 1.3$, Rest and Challenging WM: $r(25) = .544$, $p = .002$, $Bf = 10.4$, Easy and Challenging: $r(20) = .410$, $p = .032$, $Bf = 1.4$). Our Bayesian statistics thus indicate strong overall intra-subject stability of GABA levels in the OC, but only anecdotal evidence for intra-subject stability of GABA levels in the lDLPFC. Similarly, Glutamate levels significantly correlated between the two conditions in OC ($r(29) = .531$, $p = .003$, $Bf = 17.5$. However, in lDLPFC, Glutamate correlated well between the Rest and Easy WM ($r(25) = .476$, $p = .014$, $Bf = 4.287$), but not between the Rest and the Challenging WM ($r(26) = .157$, $p = .434$, $Bf = 0.3$) and the easy and the Challenging WM ($r(23) = -.120$, $p = .575$, $Bf = 0.3$) conditions. In this case, our Bayesian statistics produce a similar picture, providing strong evidence for the within-region intra-subject stability of our Glutamate measure in the occipital cortex and anecdotal to moderate evidence for within-region intra-subject stability of our Glutamate measure in lDLPFC.

Replicating previous findings of regional specificity of neurotransmitter levels (Bogner et al., 2010; Greenhouse et al., 2016), resting-state GABA levels did not correlate between the OC and lDLPFC voxel ($r(26) = -.256$, $p = .198$, $Bf = 0.5$), nor did Glutamate levels ($r(28) = .124$, $p = .521$, $Bf = 0.3$).

In sum, GABA and Glutamate levels did not systematically change depending on the activity state of the brain region but were relatively stable over the different task demand conditions within subjects. This indicates that although current 3T-MRS neurotransmitter concentrations do not capture possible differences in neurotransmitter activity between activity states (rest versus on-task), they do reliably capture stable trait-like measures of individual neurotransmitter levels in the human brain, especially with regard to GABA and in the occipital cortex.

Linking occipital and prefrontal GABA and Glutamate to region-related cognitive performance

Our second main aim was to determine, if MRS-measured neurotransmitter concentrations were sensitive to the activity state of the brain region, in which activity state the concentrations would best predict individual differences in behavioural performance. However, given that there were no significant differences between activity

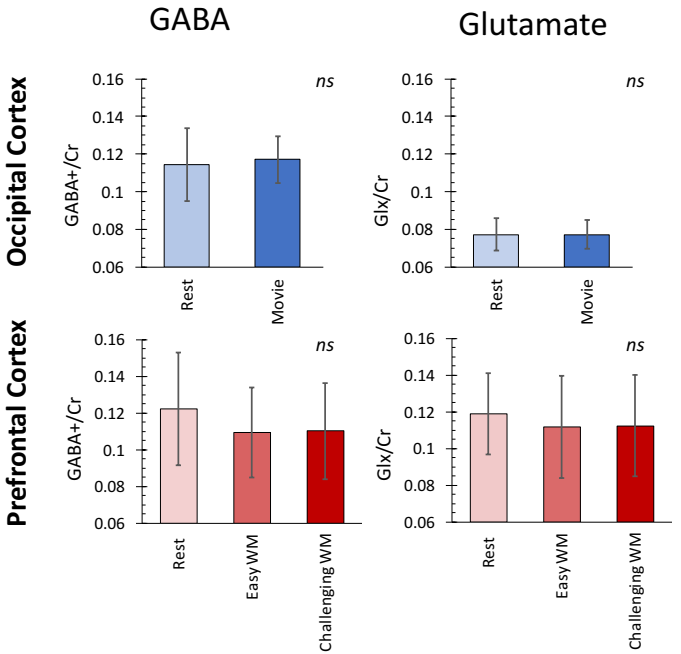


Figure 2. *GABA and Glutamate levels did not vary as a function of activity state. Group-level and of GABA (GABA+/Cr) and Glutamate (Glx/Cr) levels per brain area (Occipital cortex: top panel; Prefrontal cortex (lDLPFC): bottom panel) and activity condition.*

states (on task, rest), we continued by averaging GABA and Glutamate per brain region over the different task demand conditions.

Occipital GABA and Glutamate and visual discrimination performance

First, we related occipital GABA and Glutamate levels to visual discrimination performance. We expected to replicate the negative correlation between resting-state occipital GABA levels and visual discrimination performance previously reported by Edden et al. (2009), but no correlation between Glutamate and visual discrimination performance. However, in contrast to our expectations and the findings by Edden et al. (2009), participants' average GABA levels in OC did not predict their performance on the visual discrimination task ($r(29) = .287$, $p = .124$, $Bf = 0.7$). In line with our expectations, average Glutamate levels in OC did not either ($r(29) = -.298$, $p = .110$; $Bf = 0.8$). In both cases, Bayesian statistics reported anecdotal support for the null hypothesis of no relationship. When using the Glutamate/GABA ratios as an index of cortical excitability (Krause, Márquez-Ruiz, & Kadosh, 2013), we observed a correlation, namely higher Glutamate/GABA ratios correlated with lower discrimination thresholds ($r(29) = -.384$, $p = .036$; $Bf = 1.8$), but this correlation does not survive our multiple comparison correction and is backed up with only anecdotal evidence according to Bayesian statistics.

A post-hoc analysis revealed that even when we correlated GABA levels only at rest like Edden et al. (2009) (i.e., not averaged across conditions (Rest and Movie)), we observed a trend-level correlation also in the opposite direction ($r(29) = .328$, $p = .077$, $Bf = 1.0$). However, this correlation would again not survive a multiple comparison correction and moreover was supported with zero to no evidence according to our Bayesian results.

Thus, contrary to our expectations, we conclude that occipital cortex GABA, Glutamate and cortical excitability levels were not related to visual discrimination performance in our study.

Prefrontal GABA and Glutamate and WM performance

Next, we examined the relationship between GABA and Glutamate levels in the lDLPFC and behavioural performance on the three WM tasks performed in the separate behavioural session. In line with Yoon et al. (2016), we predicted that higher lDLPFC GABA levels would predict better WM performance. As Yoon et al. (2016) specifically found a correlation between prefrontal resting-state GABA and performance degradation as a result of increased WM load, but not increased maintenance time or as a function of distractor

presence, we furthermore expected that this relation would be specifically apparent for our WM updating capacity measure, as this measure may be the most sensitive to inter-individual differences in WM load. We did not expect lDLPFC Glutamate levels to significantly predict WM performance, nor did we expect occipital GABA levels to predict WM performance (both also similar to Yoon et al., 2016).

In contrast to expectations, but mirroring the OC results, average GABA levels in the lDLPFC did not predict accuracy on the fixed level Letter N-back task ($r(29) = .052$, $p = .785$, $Bf = 0.2$), mean level N on the adapted N-back ($r(29) = .044$, $p = .816$, $Bf = 0.2$), or accuracy on the Sternberg maintenance task ($r(27) = -.052$, $p = .797$, $Bf = 0.2$). In line with our expectations, lDLPFC Glutamate levels did not either (WM updating accuracy: $r(29) = -.015$, $p = .936$, $Bf = 0.3$; WM capacity: $r(29) = -.148$, $p = .436$, $Bf = 0.3$; WM maintenance (Sternberg): $r(27) = .098$, $p = .626$, $Bf = 0.3$). Furthermore, we looked at the Glutamate/GABA ratio as a possibly more sensitive index of cortical excitability, but this measure also did not significantly correlate with WM updating accuracy ($r(29) = -.024$, $p = .902$, $Bf = 0.3$), WM capacity ($r(29) = -.102$, $p = .592$, $Bf = 0.3$), nor WM maintenance ($r(27) = .052$, $p = .797$, $Bf = 0.3$). In all cases, Bayesian analyses indicated moderate evidence for the null hypothesis that the lDLPFC neurotransmitter measures do not relate WM performance. Even when we reran all analyses, but looked at resting-state only to stay closest to current literature (Yoon et al., 2016), no significant relationship between neurotransmitter levels or cortical excitability and WM performance was observed for any of our three WM tasks (all p's $> .139$, all Bf's < 0.7).

We reasoned that perhaps the delay between the behavioural session and the MRS session could have decreased our sensitivity to neurotransmitter concentration and brain-behaviour correlations. Therefore, post-hoc, we also explored these correlations with WM performance measured during the MRS scanning procedure. This, however, produced qualitatively the same pattern of null findings: nor GABA, nor Glutamate, nor cortical excitability measured during the Easy WM task could predict simultaneously acquired accuracy scores on this 2-letter N-back task (all p's $> .142$, Bf's < 0.7).

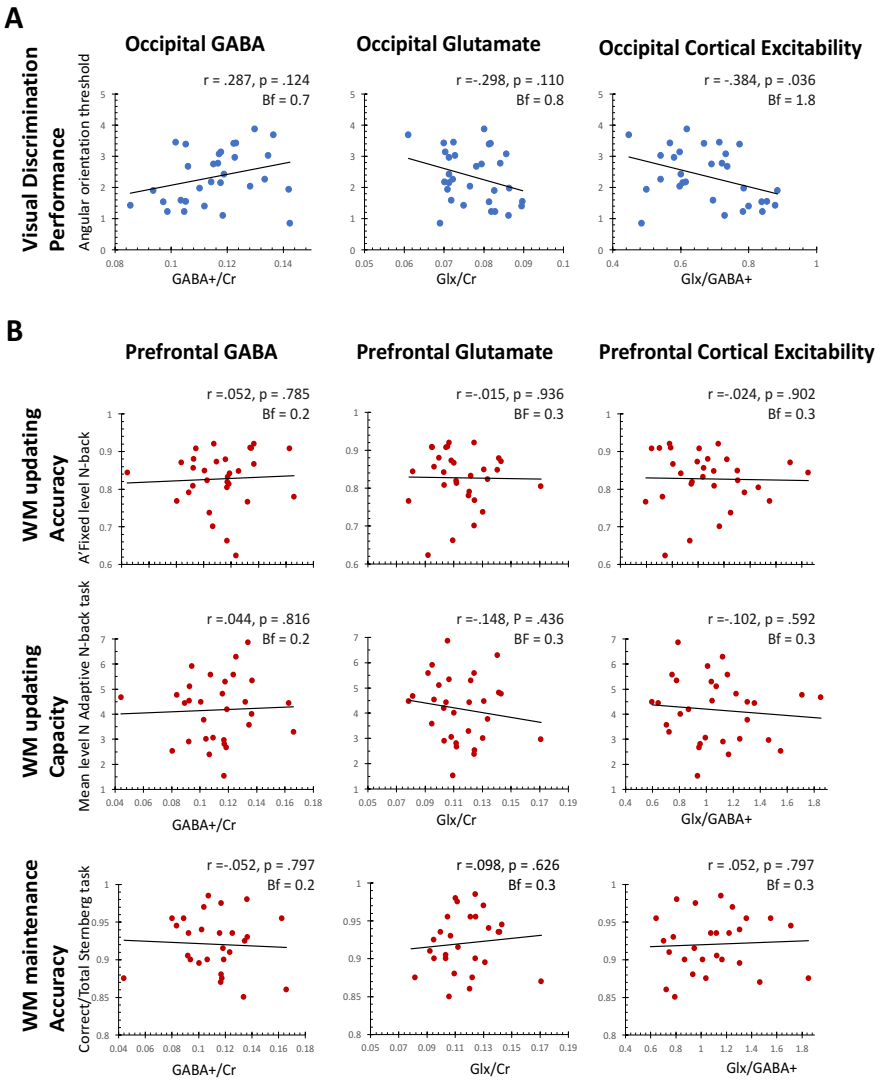


***Figure 3**. Scatter plots displaying the relationship between GABA and Glutamate concentrations as well as cortical excitability (Glutamate/GABA ratio) (collapsed across activity state) and performance on the brain-region related tasks. None of these metabolite-behaviour relationships was significant, indicating that our 3T-MRS measures of GABA and Glutamate did not robustly related to performance outside the scanner. More specifically, occipital cortex neurotransmitter levels, nor cortical excitability predicted visual discrimination performance (**A**), neither did these measures in the lateral prefrontal cortex predict performance on any of the three WM tasks (**B**). Pearson correlation coefficients and two-tailed p statistics are reported (alpha = .025; adjusted for multiple comparisons) as well as Bayes factors.*

Similarly, neurotransmitter levels acquired during the Challenging WM task did not predict WM updating capacity as measured during scanning either (all p's > .279, Bf < 0.5).

In summary, in contrast to our predictions, the neurotransmitter concentrations we measured in the occipital cortex and the left dorsolateral prefrontal cortex did not correlate with visual discrimination and WM performance, respectively. Thus, while our first set of findings suggested that GABA and Glutamate levels measured with 3T-MRS may reflect relatively stable measures of individual neurotransmitter concentrations, they seem to fail to predict individual differences in behavioural performance on brain region-relevant tasks. Remarkably, we thus did not replicate previous reports of such relationships, even though our sample size was substantially larger than in both of these studies (30 versus 15 and 23 respectively (Edden et al., 2009; Yoon et al., 2016)).

Discussion

The current study set out to investigate to what extent 3T-MRS measured GABA and Glutamate levels capture changes in cognitive activity state, as well as to determine under which activity state (rest vs. on task) these concentrations may best predict behavioural performance. We observed no differences in GABA or Glutamate levels during resting state compared to active, on-task conditions, neither in the primary visual cortex (the occipital cortex) nor in a higher-order prefrontal area (left DLPFC). Importantly, in general, we did observe strong within-subject correlations between the GABA and Glutamate levels for the different conditions within each brain area, showing that the measurements themselves where reliable. Furthermore, in contrast to previous findings, in this study levels of GABA and Glutamate, or their ratio (averaged over activity states), did not predict inter-individual differences in behavioural performance on brain region-related cognitive tasks.

Together, these findings therefore suggest that 3T-MRS may provide relatively stable 'trait'-like measures of GABA and Glutamate at the neurochemical level which are insensitive to subtle functionally-related changes as a function of cortical activation. At the same time, however, they question a robust relation between these trait-like neurotransmitter concentrations and behavioural individual differences in brain region-related cognitive performance.

Our finding that current 3T-MRS measures of GABA and Glutamate are insensitive to task demand and reflect stable 'trait' rather than 'state' levels has important implications for the interpretation of previous studies that did observe changes in GABA over relatively short time-windows. More specifically, these studies have consistently reported decreases in GABA concentrations over time; in the sensorimotor cortex after thirty minutes of performance on a motor task (Floyer-Lea et al., 2006), in the occipital cortex after twenty minutes of performance on a visual perceptual learning task (Shibata et al., 2017), and in prefrontal regions after forty minutes of performance on working memory task (Michels et al., 2012). The fact that we did not observe any activity-related changes in GABA in two of these three brain regions suggests that these earlier findings cannot simply be explained by transient modulations in activation because of longer time spent on the task and thus likely reflect learning-related structural changes in GABA activity. Also, indirectly, this implies that 3T-MRS may be a useful method to investigate the role of GABA in such learning-related cortical plasticity, as these changes seem substantial enough to be picked up with this measure.

In both the occipital and prefrontal brain region, GABA and Glutamate levels correlated strongly within subjects between the rest and task activity conditions (except for prefrontal Glutamate in the Challenging WM condition). This indicates that our measures are reliable and relatively stable within subjects. Yet, the obtained correlation coefficients are somewhat lower than previously reported in studies that looked at resting state blocks only (e.g. Bogner et al., 2010; Greenhouse et al., 2016). This could suggest that very subtle differential effects of cognitive activity on GABA and Glutamate across individuals may be picked up by our measure. In line with this, a recent 7T-MRS study did not observe any changes in GABA or glutamate as a function of acute psychosocial stress (Hoetepen et al., 2017). In this study, GABA and Glutamate levels were significantly correlated over time in the control condition, but were not correlated in the stress condition. These observations support the notion that activity state (in this case, stress) could indeed have very small, differential effects on GABA and glutamate across individuals. Because of these inter-individual differences, in current MRS practices, these 'state'-related fluctuations may fail to become visible at the group-level.

Although the measured GABA and Glutamate levels were found to reflect stable, 'trait'-like neurotransmitter concentrations, we observed no relationship between these levels and individual differences in behavioural performance on region-related tasks. More specifically, in contrast to a previous study by Edden et al. (2009), in our study occipital GABA (both when averaged over conditions and when only looked at rest) did not predict visual discrimination performance. This was unexpected, as we used the same task, observed a similar spread in subject's performance, and included an MRS voxel that covered a highly similar area of the visual cortex as Edden et al. Considering the relatively small sample size of the previous study (N=15) and only moderate sample size of the current study (N=30), future replication studies with larger sample sizes and thus greater statistical power may be necessary to further investigate the possible absence or presence of a relation between occipital GABA and visual discrimination performance. Indeed, our Bayesian correlation analyses suggested that even with our relatively large sample size, evidence for the null hypothesis of no relationship was only anecdotal.

Mirroring the occipital cortex findings, lateral dorsolateral prefrontal GABA did not predict performance on any of the three WM tasks; measuring WM updating, accuracy, WM capacity as well as WM maintenance. In this case, our Bayesian correlation analyses suggested moderate evidence in our data for the absence of such relationships. Here too, we thus failed to replicate findings by a previous study (Yoon et al. 2016, N=23) in which resting-state lateral prefrontal GABA levels correlated with individual differences on a face WM maintenance task. More specifically, in this study, Yoon et al. found that prefrontal GABA correlated positively with the extent to which subjects' performance decreased as WM load increased (one versus two to be remembered faces) Yet, no correlations with GABA were found when Yoon and colleagues looked at WM performance differences as a result of increased maintenance time or the absence or presence of distractors, indicating that this relation may hold only for a rather specific aspect of WM. Although the three WM tasks included in the current study are different than the tasks used by Yoon et al., both WM updating and maintenance have been robustly associated with activation of the lDLPFC and may thus be considered region-related WM functions (Altamura et al., 2007; Owen et al., 2005). As the current study included a larger subject sample (N=30), and applied a more extended range of WM tasks, our null results at the very least suggest that the previously

reported relationship between WM performance and GABA concentrations in lDLPFC is not very robust. Furthermore, together with the lack of a neurotransmitter-behaviour relation in the occipital cortex, they cast doubt on the claim that with current 3T-MRS practices we can detect relationships between neurotransmitters levels and region-related behavioural performance.

One important direction for future studies may therefore be to examine the role of neurotransmitters in cognitive functions using 7T- instead of 3T-MRS. Although less widely available, 7T has two important advantages over 3T with regard to MRS. Firstly, increased spectral resolution at the 7T-MRS enables better discrimination and quantification of neurotransmitter concentrations of both Glutamate (independent from Glutamine (An et al. 2014)) and GABA (uncontaminated by macromolecules (Ganji et al. 2014)). Secondly, at higher field strengths, better signal-to-noise ratios may be obtained (Choi et al., 2010), which enables the use of smaller sized MRS voxels, thereby increasing sensitivity to study a precise target region.

Namely, an important limitation of current 3T-MRS is the relatively large MRS voxel size that is necessary to acquire sufficient signal strength. Placing this relatively large voxel common over an actually much smaller region of interest may substantially 'delute' the signal, as small differences in the relevant cortical region (i.e. the desired signal) may drown in a sea of irrelevant fluctuations in the surrounding cortical regions (i.e. noise) that are also included in the voxel and thus together create the average that we measure. In other words, measuring GABA and Glutamate concentrations in a voxel that is much larger than the relevant brain region may substantially reduce the sensitivity of the method to investigate small-scale relevant regional specific neurotransmitter concentrations to relate to behaviour. Future studies should therefore investigate if the higher spectral and spatial resolution of 7T-MRS may create a method that is more sensitive to detect changes in neurotransmitter activity induced by task demand, as well as investigate relationships between neurotransmitter function in a specific brain region and related cognitive and behavioural performance.

Another direction which may aid in increasing sensitivity of MRS to detect local neurotransmitter concentrations may be to combine MRS with functional neuroimaging. More specifically, localizing individual peak activations for the region of interest may

significantly help to increase spatial acuity in the placement of the MRS voxel over the relevant area. This may be particularly helpful for higher order cortical areas, including the prefrontal cortex, where variability in functional neuroanatomy is particularly high. For example, peak activations on a WM maintenance task are known to be spread along the middle frontal gyrus across individual subjects (Jansma et al., 2013), and thus, a one-fits-all approach here may be less effective. The (relatively large) voxel used in the current study ensures peak activation was covered for all subjects, but conceivably also included surrounding cortical regions not engaged by our tasks. Eventually, therefore, smaller voxels such as may be enabled by higher magnetic field strengths, that are placed individually after functional localization, may enhance spatial acuity substantially and thereby result in higher sensitivity and more accurate measures of neurotransmitter concentrations for a specific functional region of interest.

A last explanation for the lack of brain-behaviour correlations in the current study is that the hypothesized relation between neurotransmitter levels and functional performance is actually more complex than the standard simple linear correlation we generally apply to investigate such inter-individual correlations. In fact, with regard to the excitation/inhibition balance, it has been proposed that an inverted U-curve may best describe the relation with performance, with performance being highest when the cortex is active enough for functional firing to effectively take place, but at the same time inhibited enough to reduce noise and unwanted firing (Krause & Cohen Kadosh, 2014). However, to adequately investigate this, many more data points are needed than are currently generally available in neuroimaging studies. This, again, calls for the use of larger sample sizes in studies that attempt to link neurotransmitter levels to behaviour.

Conclusion

To conclude, the current study found that 3T-MRS measures of GABA and Glutamate generally reflect stable and reliable 'trait'-like neurotransmitter levels and do not capture task demand-induced changes. However, in contrast to previous findings, we did not observe correlations of neurotransmitter concentrations with behavioural performance on region-related tasks. This questions to what extent GABA and Glutamate concentrations measured with current 3T-MRS practices reflect neurotransmitter activity that is relevant for behaviour. The use of higher magnetic field strengths (e.g., 7T), and/or individually localized voxel placement in future studies may improve the sensitivity to subtle task-induced changes in GABA and Glutamate levels, allowing further investigation of in-vivo measured neurotransmitter levels in the human brain as well as their relationship with behaviour.

Chapter 4

No evidence that baseline prefrontal cortical excitability (3T-MRS) predicts effects of prefrontal tDCS on working memory performance

Talsma, L.J., Broekhuizen, J.A., Huisman, J., Slagter, H.A., No evidence that baseline prefrontal cortical excitability (3T-MRS) predicts the effects of prefrontal tDCS on WM performance. ***In revision.***

Abstract

Transcranial Direct Current Stimulation (tDCS) over the left dorsolateral prefrontal cortex (lDLPFC) is a promising tool to enhance working memory (WM) in clinical as well as healthy populations. Yet, tDCS does not affect everyone similarly: whereas tDCS improves WM in most individuals, some individuals do not, or actually show detriments in WM performance after stimulation. One hypothesis that has been put forward to account for individual differences in tDCS response is that baseline cortical excitability levels in the stimulated cortex may determine the strength and the direction of the effects of tDCS. Specifically, by locally affecting neuronal excitability, tDCS may interact with baseline cortical excitability levels, thereby pushing or pulling individuals towards or away from an optimal level of cortical functioning. In the current study, we put this hypothesis to the test with regard to prefrontal cortex stimulation and WM. In 20 healthy male participants, using Magnetic Resonance Spectroscopy (MRS) at 3T, we measured concentrations of Glutamate and GABA in the lDLPFC and calculated individual Glutamate/GABA ratios as a measure for cortical excitability. Subsequently, in two stimulation sessions, we once applied anodal and once cathodal tDCS over the lDLPFC (20 min, 1 mA). Stimulation was always applied in the second block of three blocks of a WM updating task. Surprisingly, at the group-level, we found no effects of anodal or cathodal stimulation on WM performance. Yet, in line with previous studies, large individual variability was observed in the strength and direction of tDCS effects; whereas about half of the participants improved, the other half showed lower accuracy after stimulation. This was true for both anodal and cathodal tDCS. Nevertheless, contrary to our expectations, individual baseline prefrontal cortical excitability did not predict these individual differences in the effect of anodal or cathodal stimulation on WM accuracy. Future studies with larger sample sizes, which use higher magnetic field strengths (e.g., 7T) to measure cortical excitability and/or apply individualized stimulation protocols, are necessary to shed more light on the influence of baseline cortical excitability on effects of anodal and cathodal tDCS over lDLPFC on WM performance.

Introduction

Transcranial direct current stimulation (tDCS) is a non-invasive brain stimulation technique that has rapidly gained scientific interest as a promising tool to enhance cognitive functions, such as working memory. In tDCS, a low-voltage electrical current (typically < 2 mA) is run between two or more electrodes placed over specific brain areas at the scalp. A small portion of this current reaches the brain and influences the membrane potentials of neurons such that they are more (under the anode) or less (under the cathode) prone to fire action potentials (Kuo & Nitsche, 2012; Nitsche et al., 2008). Thus, tDCS can directly modulate neuronal excitability in particular brain regions, thereby affecting brain and cognitive functioning. As such, tDCS may be used as a tool to enhance brain function and cognitive abilities such as working memory.

Our working memory (WM) allows us to maintain and monitor information over brief periods of time (Baddeley, Sala, Robbins, & Baddeley, 1996) and thus plays a core role in many daily-life situations. One brain region that is critically involved in WM is the left dorsolateral prefrontal cortex (lDLPFC) (Owen, McMillan, Laird, & Bullmore, 2005). Initial studies with tDCS found that anodal tDCS stimulation over the lDLPFC could improve verbal working memory in healthy (Andrews, Hoy, Enticott, Daskalakis, & Fitzgerald, 2011; Fregni et al., 2005; Ohn et al., 2008) and clinical populations (e.g. Boggio et al., 2006), making it a promising method for enhancing working memory functioning. However, since those pioneering studies, the effects of tDCS on cognition have been less conclusive (Jacobson, Koslowsky, & Lavidor, 2011), with several studies questioning the ability of anodal lDLPFC stimulation to robustly improve WM performance (see meta-analyses by Bennabi et al., 2014; Brunoni & Vanderhasselt, 2014; Dedoncker, Brunoni, Baeken, & Vanderhasselt, 2016; Hill, Fitzgerald, & Hoy, 2016; Mancuso, Ilieva, Hamilton, & Farah, 2016).

Recently, we examined if multiple sessions of anodal tDCS over the lDLPFC during a verbal WM task (a letter N-back updating task) would lead to increasing gains in WM performance across training sessions in healthy adults (Talsma, Kroese, & Slagter, 2016). Replicating previous single session studies, we found that anodal compared to sham tDCS led to an increase in WM performance, but only in the first session. Moreover, we observed that the effects of anodal tDCS were quite variable across individuals in both strength and direction, and that these individual differences in the effect of anodal tDCS during WM

training predicted the extent to which individuals performed better on subsequent WM transfer tasks. Looking at the individual subject data, we found that 2 of our 15 subjects in fact showed worse performance after anodal stimulation both on the trained and transfer WM tasks.

Large variation in the effect of tDCS on WM between subjects is problematic with regard to the practical use of tDCS as method to enhance WM function in everybody. Moreover, a better understanding of individual differences in tDCS responsiveness may help resolve current inconsistencies in the literature, as it may explain why overall tDCS effects are found in some groups of subjects, but not in others. Therefore, for the current progression of the tDCS field, investigating the determinants of individual differences in tDCS response is a pivotal scientific direction to explore.

In recent years, several possible explanations have been proposed for the relative large variability in tDCS response. Currently, many of these are directed at the question whether the admitted current may in fact reach the target brain area in all subjects. Modelling studies have indicated that tDCS current flow with conventional standard tDCS set-ups can be strongly affected by individual differences in anatomy, skull thickness and folding of the cortex (Opitz, Paulus, Will, Antunes, & Thielscher, 2015). Another proposal that has been put forward to account for inter-individual differences in tDCS response is that tDCS effects may depend on baseline functioning of the stimulated area. Specifically, it has been proposed that prefrontal tDCS may enhance WM performance only in subjects in which the prefrontal cortex is in fact engaged in the task, assuming that tDCS needs some baseline activation to 'grasp' onto (Berryhill & Jones, 2012). Postulated more broadly, the effect of tDCS may depend on the extent to which a stimulated brain region is already activated by the task. Namely, when brain region engagement is already optimal, tDCS may cause overstimulation, resulting in worse performance, whereas when brain region engagement is suboptimal, tDCS may optimize brain function, resulting in improved performance.

Recently, Krause and colleagues proposed a theoretical model to explain these baseline and tDCS effects interactions at the cellular level (Krause, Márquez-Ruiz, & Kadosh, 2013). More specifically, they suggested that since the cortical excitability level of a particular brain area critically determines neuronal firing rates, it plays a pivotal role in

cortical functioning. In an optimal situation, the cortex is active enough for functional firing to effectively take place, but at the same time inhibited enough to reduce noise and unwanted firing (Turrigiano & Nelson, 2000; Turrigiano & Nelson, 2004). However, both too high and too low excitability may be detrimental for functional performance, and the relation between cortical excitability and performance can thus be described as an inverted U-curve. Depending on an individual's initial position on the curve, Krause et al. suggested that a specific type of stimulation may be either beneficial or unfavourable for local brain functioning, depending on whether it pushes or pulls the brain region towards or away from its optimal excitability level (Krause et al., 2013).

Cortical excitability can be quantified by the excitation/inhibition balance in a particular cortical region. This balance is thought to be determined by two key neurotransmitters: GABA, which has an inhibitory effect, and Glutamate, the brain's main excitatory neurotransmitter (Petroff, 2002). Magnetic Resonance Spectroscopy (MRS) is a relatively novel method that allows for non-invasive, in-vivo quantification of neurotransmitter levels such as GABA and Glutamate in a particular voxel in the human brain. Interestingly, MRS can thus be used to acquire individual Glutamate/GABA ratios that can be taken as a proxy for local cortical excitability in a specific target brain area of interest.

In line with the cortical excitability hypothesis, studies that combined MRS with tDCS in humans have related tDCS stimulation with both changes in GABA and Glutamate. More specifically, anodal stimulation over the motor cortex was shown to reduce resting-state GABA levels (Stagg et al., 2009), while cathodal stimulation in contrast reduced Glutamate levels (Clark, Coffman, Trumbo, & Gasparovic, 2011). Although through different mechanisms, both types of stimulation may thus change the local excitation/inhibition balance and thereby critically alter neuronal functioning within the underlying cortex.

So far, most of the research applying both MRS and tDCS has been done in the motor domain and focused on the motor cortex. However, as effects of tDCS at the cellular level are not expected to be different for different parts the cortex, effects of tDCS on GABA and Glutamate should be similar for brain regions involved in higher-order cognitive functions, such as WM. In the current study, we aimed to investigate possible interactions between tDCS response and baseline cortical excitability further with regard to prefrontal tDCS and working memory. More specifically, we examined if prefrontal cortical excitability levels

(Glutamate/GABA ratios) determine behavioural effects of left DLPFC tDCS on verbal WM performance across individuals.

In an initial MRS session, we used 3T-MRS to measure GABA and Glutamate levels in the lDLPFC to determine baseline cortical excitability in this region in 20 healthy male subjects. Subsequently, in two stimulation sessions (separated by one week), we admitted once anodal and once cathodal tDCS over the lDLPFC (reference supraorbital in both cases, cf. Talsma et al., 2016). In both tDCS sessions, before, during and after stimulation, subjects performed a verbal WM task (the letter N-back task) to determine WM performance. The difficulty of this task was tailored to subjects' individual WM updating capacity to allow for enough room for tDCS to increase or decrease WM performance, as well as to make the task equally challenging in all subjects.

Based on previous findings (Andrews et al., 2011; Fregni et al., 2005; Lally, Nord, Walsh, & Roiser, 2013; Ohn et al., 2008; Talsma et al., 2016), we expected that anodal tDCS over the lDLPFC would improve WM accuracy in the majority of our subjects, resulting in a general improvement in WM performance compared to cathodal tDCS. Yet, in line with earlier reports, we also expected the effects of anodal stimulation to vary across subjects, with some subjects showing larger improvements after anodal prefrontal tDCS than others, and some perhaps showing decrements in WM performance. As the effects of cathodal stimulation in the cognitive domain are less conclusive (Jacobson et al., 2011), we expected no group-level effect of cathodal stimulation or a general decrement in performance.

We made two predictions with regard to the relation between baseline cortical excitability and the effect of stimulation. First, as anodal tDCS is associated with reducing GABA (Stagg et al., 2009), we expected a negative relationship between baseline cortical excitability and anodal tDCS-induced WM improvements. That is, we expected that subjects with lower baseline Glutamate/GABA ratios in lDLPFC (i.e., relatively higher baseline GABA concentrations) would show the biggest enhancements, as here anodal tDCS may help increase initial lower than optimal activation in this area. In contrast, as cathodal tDCS may specifically lower Glutamate levels (Clark et al., 2011), for cathodal tDCS, secondly, we expected a positive relationship, with cathodal stimulation being most beneficial in subjects with high baseline cortical excitability levels.

Methods

Participants

20 healthy, right-handed male participants participated in the study (Age range: 18 to 26, Mean 21.8, StD Age: 2.6). Female participants were excluded because cortical GABA concentrations have been reported to vary over the menstrual cycle (De Bondt, De Belder, Vanhevel, Jacquemyn, & Parizel, 2015; Harada, Kubo, Nose, Nishitani, & Matsuda, 2011). Subjects were recruited from a previous study sample for which we had already acquired 3T-MRS data (Talsma et al., submitted). They were screened for tDCS contra-indications (see Nitsche et al., 2008) and were paid for their participation in the form of course credit or with a monetary compensation. One subject did not complete the study because of excessive itching during stimulation. All procedures in this experiment were approved by the University of Amsterdam's Ethical Committee.

Procedure

Participants came to the lab for a total of four sessions (the first two sessions were part of a previous study (Talsma et al., submitted)). In a first behavioural session, we determined working memory updating capacity (WMC) for each participant using an adaptive version of a verbal WM updating task (the letter N-back). In a second MRS-session, we used 3T-MRS to measure individual GABA and Glutamate concentrations in the left DLPFC (note: this data has previously been reported in Talsma et al. (submitted)).

In the third and fourth session, subjects came to the lab for two stimulation sessions at the same time of the day and spaced exactly one week apart. In one of the two stimulation sessions, subjects received anodal tDCS over the lDLPFC, while in the other session, they received cathodal stimulation (both 1 mA, 20 min). Order of stimulation type was counter-balanced between subjects and the first stimulation session took place on average 40 days after the MRS session (StD: 8,5, range: 29-67). As high intra-subject stability of neurotransmitter levels has previously been reported over the course of four

Figure 1. *Schematic illustration of the research design and methods. Subjects came to the lab for a total of four sessions (A). In a first behavioural Pre session, working memory updating capacity (individual N) was determined for each participant using an adaptive version of the verbal WM updating task. In a second MRS-session, 3T-MRS was used to measure individual GABA and Glutamate concentrations in the left DLPFC under three conditions: rest, an easy and a challenging WM task. Notably, because we observed no differences between these three conditions (as reported in Talsma et al. (submitted)), for this study we averaged over all conditions and calculated Glutamate/GABA ratios to use as a measure for prefrontal cortical excitability. In two subsequent stimulation sessions, participants performed three blocks of a verbal WM task. During the second block, they received either anodal or cathodal stimulation over the lDLPFC (active electrode - F3, reference - above the right eye). (B) In the verbal WM task (letter N-back), a stream of letters was presented and participants were required to press a button if the current letter was the same as N stimuli before. In the adaptive version in the Pre session, the level of N was on-line adjusted to performance and gave us a measure for individual updating capacity for each subject (individual N). For the stimulation sessions, level of N consequently ranged between -1 and +2 around this individual updating capacity level to ensure a challenging task level for all subjects. (C) MRS-voxel location over the left Dorsolateral Prefrontal Cortex (size: 30 × 20 × 25 mm). (D) Modelling of the GABA and Glx (Glutamate + Glutamine) signal for a typical subject (output from Gannet). In blue the edited spectrum is shown, overlaid in red is the model of best fit (using a simple gaussian model) and the residual of these is shown in black.*

weeks (Bogner et al., 2010), as well as over different activity 'states' (Talsma et al. (submitted)), the GABA and Glutamate concentrations that we measure with MRS likely reflect relatively stable 'traits' that can assumed to show consistency over time.

In both stimulation sessions, subjects performed a total of three blocks of a verbal WM updating task (the letter N-back): one before, one during and one after stimulation. See figure 1A for a schematic overview of the study design.

Measuring WM performance: The letter N-back task

A letter N-back task was used to measure WM performance (see also Figure 1B). In this task, a stream of letters is presented and subjects are asked to indicate if the currently presented letter is the same as the one presented N stimuli back. N is an integer and the value of N hence determines the difficulty level of the task: with higher levels of N, more stimuli have to be held in WM in sequential order, increasing WM load.

Because WM content has to be continuously updated, the letter N-back task is well suited to investigate WM updating performance (Jaeggi, Buschkuehl, Perrig, & Meier, 2010). Moreover, performance on this task has consistently been related to processing in the lDLPFC (e.g., see meta-analysis by (Owen et al., 2005). Furthermore, although recent meta-analyses raise questions with regard to the reliability of anodal tDCS to lDLPFC to enhance WM performance (e.g., (Dedoncker et al., 2016)), we previously found that anodal lDLPFC tDCS enhanced accuracy on a very similar version of this letter N-back task, although as noted in the introduction, the stimulation effects varied both in strength and direction across individuals (Talsma et al., 2016).

To ensure a challenging task level for all subjects, but also leave enough room for tDCS to improve or impair WM performance, we individually determined the level of N in the two tDCS sessions for each subject based on their average WM updating capacity score, using their performance on an adaptive version of the letter N-back task in the first behavioural session and in the MRS-session. In this adaptive version, level of N always started at 2, but was adjusted per run according to performance, with N incrementing one level after fewer than three errors (false alarms + misses) and decrementing one level after more than five errors (similar to (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008)). To determine WM updating capacity, we took the mean level N that subjects achieved in the last 21 runs

of this task in the behavioural session and 12 runs in the MRS session (in both sessions disregarding the first 3 runs to allow some ramp-up time), and averaged scores obtained in the behavioural and MRS-session. As expected, in our sample, we observed a relatively large spread in the resulting individual WM updating capacity scores, with level N ranging between 2,7 and 6,1 (Mean: 4,4 StD: 1,2).

The individual capacity scores were used to choose the levels of N for each subject separately in the stimulation sessions to ensure similar task difficulty for all subjects. Specifically, we first determined *individual N*'s by rounding off WM updating capacity score to the nearest integer. Across subjects, this resulted in individual N's ranging between 3 and 7 (number of subjects per level – 3:5, 4:3, 5:6, 6:5). Then, in the stimulation sessions, level N's ranged between *individual N* – 1 and *individual N* + 2. Thus, task level on the letter N-back ranged over four levels, which allowed for enough room to observe tDCS-related improvements as well as possible decrements in performance in every subject. In the stimulation sessions, each block of the task consisted of 24 runs, where N incremented twice over the different levels (individual N -1, individual N, individual N +1, individual N +2).

Presentation software (Neurobehavioural Systems, Inc.) was used to administer the letter N-back task. Black letters were presented (Arial, font size 72, letterset ["A", "B", "C", "D", "E", "F", "G", "H", "J", "K"]) for 300 ms at the centreof a white screen, followed by a 1500 ms inter-stimulus interval in which a fixation cross was displayed (Arial, font size 20). Of the presented letters, approximately 37,5% were so-called targets, i.e., the letter that was the same as the letter presented N trials back. Letters could be presented in upper or lower case and still classified as the same letter (i.e., a target). When presented with a target, subjects were required to press the space bar on the keyboard in front of them. Runs consisted of a stream of 20 + N stimuli each and were self-paced to allow the subject to take small breaks in between and enhance focus during the runs.

For each stimulation session separately, working memory performance accuracy on the letter N-back task was operationalized by calculating A' scores for each of the three blocks of the task (before, during and after stimulation), averaged over the levels of N (cf. Talsma et al., 2016). A' is the non-parametric variant of signal detection theory's d' and takes into account both hits (correct responses) and false alarms (incorrect responses). In

contrast to d', A' can account for situations in which participants do not show any false alarms, which may occur on easy task levels. A' scores range from 0 to 1, in which 0 indicates chance performance and 1 perfect accuracy. A' can be calculated from hit rate (H) and false alarm rate (F) with the following formula (Zhang & Mueller, 2005):

$$A' = \begin{cases} \frac{3}{4} + \frac{H-F}{4} - F(1-H) & if\ F \leq 0.5 \leq H; \\ \frac{3}{4} + \frac{H-F}{4} - \frac{F}{4H} & if\ F \leq H < 0.5; \\ \frac{3}{4} + \frac{H-F}{4} - \frac{1-H}{4(1-F)} & if\ 0.5 < F \leq H. \end{cases}$$

Additionally, to allow the investigation of possible speed-accuracy trade-offs, as well as to investigate possible stimulation effects on WM response speed, we calculated average reaction times over the correct responses for each block of the letter N-back task and each stimulation session separately.

Measuring prefrontal cortical excitability: 3T-MRS data acquisition and analysis

In the MRS-session, for each subject we measured GABA and Glutamate levels in the left Dorsolateral Prefrontal Cortex (see also Talsma et al., submitted). Scanning was performed on a 3T Philips Achieva TX MRI scanner (Philips Healthcare) using an eight-channel head coil. According to individual anatomical landmarks as visible on an initial anatomical scan, the experimenter positioned the MRS voxel (30 × 20 × 25 mm) on the middle frontal gyrus and with the rear of the voxel anterior to the precentral sulcus (see also Figure 1C). Care was taken not to include cerebral spinal fluid (CSF) from the ventricles or the cortical surface.

We used a GABA-specific sequence of the Mescher-Garwood point-resolved spectroscopy (MEGA-PRESS) method (Waddell, Avison, Joers, & Gore, 2007) to acquire Edited ^{1}H J-difference spectra. The acquisition of this scan took approximately 12 min, during which 384 transients were collected (TE = 73 ms; TR = 2,000 ms). On the odd transients, a 15,64 ms sinc-centreediting pulse (64 Hz full width at half maximum) was applied in an interleaved manner at 1,9 ppm and 4,6 ppm to excite GABA and suppress water respectively.

Neurotransmitter levels in the lDLPFC were measured three times in every subject: once at rest (eyes closed), once while they performed an easy WM task (letter 2-back) and once during a challenging WM task (adaptive letter N-back). Due to time constraints, the

rest scan of one subject is missing. In our previous study, we found that GABA and Glutamate levels did not differ between activity states (i.e., at rest vs. on-task (Talsma et al., submitted). Therefore, for the current study we averaged GABA and Glutamate concentrations across the different activity conditions.

Spectral data were analysed with the MATLAB-based package GANNET v2.1 (Edden et al. 2014, www.gabamrs.com) as also described in Talsma et al. (submitted). Using the in-build options of the GannetLoad-function, the following processing steps were performed (in this order): time-domain frequency-and-phase correction using spectral correction, line broadening with an exponential apodization function, FFT, time averaging, frequency and phase correction based upon fitting of the Cho and Creatine signals, pairwise rejection of the data for which fitting parameters are greater than 3 SDs from the mean, and finally, subtraction of the even from the odd transients to generate the edited difference spectrum. Note that in this edited difference spectrum, the GABA signal is contaminated by the macromolecule homocarnosine (Edden, Puts, & Barker, 2012), a GABA derivative, and thus often referred to as GABA+.

Subsequently, using the GannetFit function of GANNET, GABA and Glx (the combined signal for Glutamate and Glutamine) functions were modelled to the data together (see also Figure 1D) and ratios relative to Creatine (Cr) were calculated (i.e. GABA+/Cr and Glx/Cr). Normalizing values to Creatine has been shown to reduce inter-subject variance as a result of differences in global signal strength, as well as differences stemming from tissue fractions in the scanned voxel (grey matter, white matter, and cerebrospinal), thus making coregistration, segmentation and the calculation of CSF corrected values superfluous. Moreover, normalizing to Creatine has shown superior to normalizing to H2O with regard to intra-subject stability and therefore can be considered the most reliable measure for concentration estimates (Bogner et al., 2010).

Data of scans was excluded when the modelfit was poor (N=2; corresponding to FitError >15), and when the GABA or Glx-peak could not be confidently be determined (N=1; GABA SNR<3). Furthermore, in SPSS we identified outliers and excluded these from the data (N=4, all values for GABA). Because previous analyses did not reveal differences between the three activity conditions, for the current study we averaged GABA (GABA+/Cr) and Glutamate (Glx/Cr) concentrations over the remaining scans per subject. Subsequently, a

measure for cortical excitability was calculated by dividing Glutamate over GABA, resulting in a prefrontal Glutamate/GABA ratio for each subject.

Prefrontal Anodal and Cathodal Transcranial Direct Current Stimulation

In each of the two stimulation sessions, participants were seated comfortably behind a computer screen (at approximately 90 cm distance). Before the WM task started, rubber straps were put into place and the lDLPFC was localized in each participant (see below; cf. (Talsma et al., 2016). This allowed for a fast placement of the electrodes right before the stimulation block, but prevented the sponges from drying out. After a brief practice session with feedback, subjects performed three blocks (+/-20 minutes each) of the letter N-back updating task (see for details above) (cf. Talsma et al., 2016). The first block of the task was administered before stimulation started and thus served as a baseline condition. The second block began 90 seconds after stimulation was started and ran throughout the entire stimulation time. The third block of the task was started after stimulation had ended.

Transcranial Direct Current Stimulation was delivered with a battery-driven Eldith DC-stimulator (NeuroConn GmbH, Germany) using two 7 × 5 cm conductive electrodes. Electrodes were placed in saline-soaked sponges and held in place with rubber bands. In both sessions, after the baseline task block, one electrode was placed over the left DLPFC (F3 in the 10/20 system) and the other was placed over the right supra-orbitofrontal region (centered above the right eye pupil) (cf. e.g., Talsma et al., 2016), see Figure 1A). In the first stimulation session, the position of F3 was localized in each participant using an EEG cap (64 channels, Biosemi, Amsterdam, The Netherlands). This position was marked on the scalp as well as measured relative to landmarks like the tip of the nose, the inion, and ears, to ensure identical electrode positioning in the second session. In both tDCS conditions, stimulation was applied for 20 min on 1 mA, once with the anodal electrode over the lDLPFC (i.e. anodal tDCS condition) and once the cathode (i.e. the cathodal tDCS condition). To reduce discomfort, in both conditions, the current was ramped up over 90 sec and down over 90 sec. Both participant and experimenter were blind to the type of stimulation that was applied in each session.

Additionally, at the beginning and the end of each stimulation session, subjects filled out a questionnaire to assess physical sensations and determine possible side effects of tDCS on mood and arousal levels. To investigate possible physical side effects of tDCS, at the end of each tDCS session, participants were asked to rate their experience on a 5-item scale (namely “not,” “a little,” “somewhat,” “strongly,” and “very strongly”) of each of eight following sensations: itching, prickling, burning, pain, headache, fatigue, dizziness, and nausea. In addition, to assess mood and arousal levels, a Dutch translation of the short version of the Activation Deactivation Adjective Checklist (AD ACL) was used (Thayer, 1978), which requires participants to rate 20 items using a 4-point scale (namely “definitely feel,” “feel slightly,” “do not really feel,” and “definitely do not feel”). Answers are scored on four subscales: energy (general activation), tiredness (general deactivation), tension (high preparatory arousal), and calmness (low preparatory arousal). The AD ACL has proven reliable and valid, showing high test–retest reliability for each of its subscales (all >.79; Thayer, 1978). The AD ACL was filled out pre- and post-stimulation in each tDCS session, and changes in mood and arousal were calculated for each session separately.

Analytical Approach and Data analysis

Firstly, we investigated the group-level effects of anodal and cathodal stimulation on WM performance. For this, we first conducted a 2x3 repeated measures ANOVA on accuracy scores, with Stimulation type (Anodal vs Cathodal) and the three blocks of the task (before, during and after stimulation) as within-subject variables. We repeated this analysis, but with RT as the dependent variable. In case of significant effects, post-hoc analyses were performed to further investigate findings and whenever appropriate Greenhouse Geisser corrected values are reported. Additionally, to investigate whether order of the stimulation sessions or individual differences in WM updating capacity may have affected the effects of stimulation, we also reran the repeated measures ANOVA’s for both Accuracy and RTs, adding session order and individual WM capacity (individual N) separately as a covariate.

Secondly, next to determining group effects of tDCS, we tested our hypothesis that baseline lDLPFC cortical excitability levels may predict individual differences in the effect of anodal and cathodal stimulation on WM performance using correlation analyses. We previously observed the largest effects of tDCS not during, but after stimulation and on WM accuracy specifically (Talsma et al., 2016). Therefore, we quantified the tDCS effect on WM accuracy (A’) by subtracting baseline performance (i.e. in the first block of the task, before

stimulation was applied) from performance in the block after stimulation, and divided this over baseline again to get a measure of relative improvement after tDCS per subject (i.e. (after-before)/before). This was done separately for the anodal and cathodal stimulation session. Because of our relatively small sample, we ran Spearman rank correlations to determine the relationship between the MRS-measured prefrontal Glutamate/GABA ratios and these individual effects of anodal (cathodal) tDCS on WM performance.

Lastly, to examine possible non-specific physical or arousal effects of anodal and cathodal stimulation, we ran repeated-measures ANOVAs for each of the eight items on the tDCS side-effects questionnaire with Stimulation Type (anodal, cathodal) as a within-subject factor and Session Order as a covariate. To determine whether there was a difference in the effects of the two types of stimulation on arousal states or mood, scores on each of the four subscales of the AD ACL questionnaire were calculated before and after stimulation for each stimulation session separately and subsequently subtracted from each other to obtain a measure of the effect of each type of stimulation on arousal and mood. For each subscale separately, we then conducted a repeated-measures ANOVA with Stimulation Type (anodal, cathodal) as a within-subject factor and Session Order as a covariate, thus comparing changes in the resulting difference scores between the tDCS conditions. A Bonferroni correction was applied to account for multiple comparisons for both questionnaires separately, resulting in an alpha of .05/8 = .0063 for the physical side effects questionnaire and an alpha of .05/4 = .0125 for the Short Form AD ACL questionnaire.

All statistical analyses were conducted using the Statistical Package for the Social Sciences for Mac OS, Version 24 (IBM, Armonk, NY). Furthermore, because of significant advantages over conventional statistics (Wagenmakers, Marsman, et al., 2017), we additionally ran Bayesian analyses using the open-software package JASP (http://www.jasp-stats.org, see also Wagenmakers, Love, et al., 2017). Bayes factors will be reported, grading the intensity of evidence for the alternative hypothesis (Bf10), and values will be interpreted according to the corresponding classification scheme (see for elaboration Wagenmakers, Love, et al., 2017): 1/30 < Bf < 1/10, Strong evidence for H0; 1/10 < Bf < 1/3, Moderate evidence for H0; 1/3 < Bf < 1, Anecdotal evidence for H0; Bf = 1, No evidence; 1 < Bf < 3, Anecdotal evidence for H1; 3 < Bf < 10, Moderate evidence for H1; 10 < Bf < 30, Strong evidence for H1.

Results

General WM performance

In both stimulation sessions, all subjects showed good, but not ceiling level overall WM performance (Mean A': 0.84, StD: 0.038, range: 0.74 – 0.91; Mean RT: 785, StD = 141, range: 590 – 1146). This indicates that our method to adapt task-levels according to individual WM updating capacity worked well.

Importantly, baseline performance did not differ between the two stimulation sessions, not in accuracy ($t(18) = .504$, $p = .621$, $Bf = 0.266$) or in reaction times ($t(18) = .892$, $p = .384$, $Bf = 0.338$). Moreover, accuracy scores in this first block of the task ranged between 0.78 and 0.92 in the anodal (Mean: 0.84, StD: 0.04) and between 0.79 and 0.87 in the cathodal stimulation session (Mean: 0.84, StD: 0.03), indicating enough room to improve (as well as possibly deteriorate) as a function of tDCS stimulation in both stimulation sessions. Please see Table 1 for the Mean and StD of both accuracy and RTs over the different task blocks for both stimulation conditions.

Group-level effects of anodal and cathodal tDCS on WM performance

We first investigated the effect of anodal and cathodal stimulation on WM performance at the group level by running a 2 (Stimulation Type) x 3 (Block) repeated measure ANOVA for Accuracy and RTs separately. See Table 1 for an overview of mean accuracy and RT per Stimulation type and Block.

Overall, accuracy did not change over the different blocks of the task (Main effect Block $F(2,36) = .575$, $p = .499$, $Bf = 0.09$), nor did it significantly differ between anodal and cathodal stimulation (Main effect Stimulation Type: $F(1,18) = .007$, $p = .933$, $Bf = 0.47$). Moreover, the critical interaction effect between Stimulation Type and Block was non-significant indicating that anodal and cathodal did not differentially affect WM accuracy

	Anodal tDCS		**Cathodal tDCS**	
	Accuracy	RTs	Accuracy	RTs
	Mean (*StD*)	Mean (*StD*)	Mean (*StD*)	Mean (*StD*)
Before	0.84 (*0.04*)	674 (*113*)	0.84 (*0.03*)	663 (*114*)
tDCS	0.85 (*0.05*)	843 (*169*)	0.84 (*0.05*)	828 (*130*)
After	0.85 (*0.05*)	856 (*193*)	0.84 (*0.08*)	845 (*170*)

Table 1. *Mean and Standard Deviations shown separately for accuracy (A') and RTs on the verbal WM updating task in the two stimulation sessions (N=19), split out for the three different blocks of the task and the anodal and cathodal tDCS stimulation condition.*

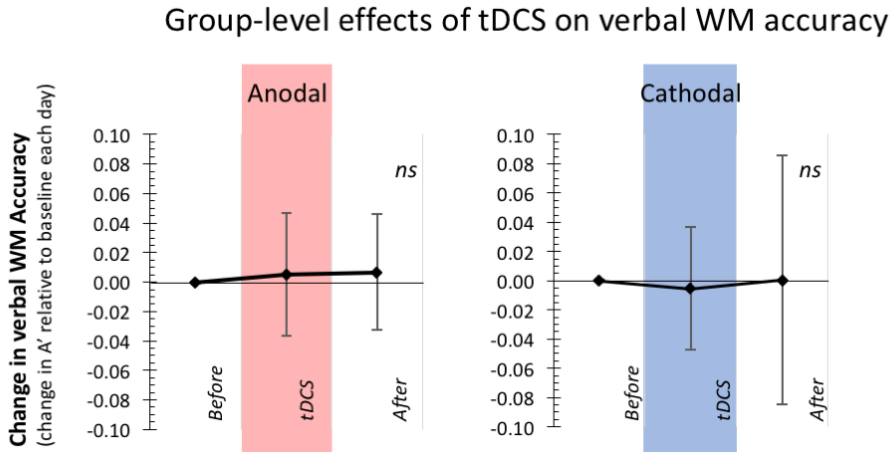


Figure 2. *Group-level analyses showed that anodal and cathodal prefrontal tDCS stimulation did not consistently affect WM performance. Displayed here is the change in Accuracy (A') for the blocks of the task during and after tDCS stimulation relative to the baseline block of that day (error bars represent Standard Deviations from the mean).*

(Interaction Type * Block: F(2,36) = .560, p = .523). Furthermore, a Bayesian model including the two main effects (Bf = 0.05) and one which additionally included the interaction (Bf = 0.01), both showed more evidence for the null-hypothesis. Thus, anodal and cathodal stimulation did not (differentially) affect verbal WM accuracy.

As to Reaction Times, response times did not significantly differ between the anodal and cathodal session (Main effect Stimulation Type: F(1,18) = .996, p = .331, Bf = 0.231). Although in both sessions, subjects' responses became slower over time (Main effect Block F(2,36) = 92.053, p = .000, Bf = 6.01 * 10^22), the extent to which responses became slower over time did not differ between the stimulation conditions (Interaction Block * Stimulation Type (F(2,36) = .025, p = .975). These findings indicate that anodal and cathodal stimulation did not have a differential effect on WM response times. Indeed, our Bayesian analyses showed extreme evidence for the alternative hypothesis of no effect, both in a model that included both main effects (Bf = 2.13*10^22) and in one which additionally included the interaction between the two (Bf = 3.17*10^21). Moreover, a direct comparison between these two models critically shows moderate evidence in favour of a model in which the interaction is not included (Bf = 0.15). Thus, we found no effect of stimulation type on RT either. The observed slowing in RT likely reflects a general fatigue effect in both conditions.

To control for possible confounding effects of session order and individual differences in WM updating capacity, we ran the analyses on both accuracy and RT again adding these as covariates. This did not change the pattern of findings.

Together, these findings indicate that at the group-level, neither type of tDCS stimulation over the lDLPFC (anodal nor cathodal) consistently altered accuracy (see also figure 2) or reaction times on the verbal WM updating task. Thus, in contrast to our expectations, in the current study we do not replicate our previous findings (Talsma et al., 2016) that anodal stimulation over the lDLPFC concurrent with a verbal WM updating task improves WM accuracy.

Does lDLPFC cortical excitability levels predict the effect of tDCS on WM?

To answer our main research question, we next examined if individual differences in the effects of anodal and cathodal tDCS on WM performance across subjects can be predicted by baseline prefrontal cortical excitability levels, as measured with 3T-MRS. For this, we first quantified the effect of anodal and cathodal stimulation on WM accuracy for every subject as a relative change to baseline per session (After - Before/Before). Eyeballing our data, we found that about half of our subjects improved in the block after anodal tDCS (n=11), while the other half showed decreased WM accuracy after stimulation compared to before (n=8). Similarly, in the cathodal stimulation session, accuracy improved after stimulation in approximately half of our subjects (n=10), while it deteriorated in the other subjects (n=9).

To test our main hypothesis, we subsequently correlated prefrontal cortical excitability levels (Glutamate/GABA ratios) with these behavioural effects across subjects. In contrast to our expectations, cortical excitability levels in lDLPFC did not predict the effect of anodal tDCS on WM performance (r(18) = .182, p = .453, Bf = 0.32). Similarly, prefrontal cortical excitability also did not predict the effect of cathodal prefrontal tDCS stimulation on verbal WM (r(18) = .058, p = .815, Bf = 0.29) Removing one subject that showed extreme deterioration in the cathodal stimulation condition (>3 StD from the mean) did not change this result (r(17) = 0.091, p = 0.720, Bf = 0.31). In both cases, Bayesian statistics indicated moderate evidence for the lack of a relation between baseline cortical excitability and individual differences in the effect of tDCS on WM accuracy (see also Figure 3). Furthermore, post-hoc additional analyses that related GABA (GABA+/Cr) and Glutamate (Glx/Cr) separately to the effects of anodal and cathodal tDCS were not significant either (all p's > 0.375; all Bf's < 0.42).

In conclusion, even though we observed large variability in both the extent and direction of the effects of anodal and cathodal on WM performance, prefrontal cortical excitability did not predict the effect of anodal or cathodal lDLPFC stimulation on WM performance across subjects.

Figure 3. *In contrast to our expectations, baseline prefrontal cortical excitability (Glutamate/GABA ratios) did not predict individual differences in the extent and direction of the effects of anodal and cathodal tDCS on WM. As can be seen in these scatterplots, in both stimulation conditions about half the subjects showed improved verbal WM after stimulation, while the other half showed worsened performance. Pearson correlation coefficients and two-tailed p statistics are reported, as well as Bayes factors (Bf10).*

Questionnaires

On the tDCS side effects questionnaire, no differences were reported between the anodal and cathodal stimulation condition for any of the possible physical sensations (all uncorrected p's > 0.25, Bf's < 0.6), and no significant interactions were found with session order (all uncorrected p's > 0.12, Bf's < 0.6). Overall, subjects reported to have felt 'somewhat' of an itching (Mean: 1.87, StD: 1.16) and prickling (Mean: 1.84, StD: 1.16) sensation, and experienced 'a little' of a burning (Mean: 1,25, StD: 1,12) and a feeling of tiredness (Mean: 1,43, StD: 1,21). However, importantly, they did not report general feelings of pain (Mean: 0,24, StD: 0,49), headaches (Mean: 0,37, StD: 0,65), dizziness (mean: 0,22, StD: 0,47) or nausea (Mean: 0,03, StD: 0,12).

Similarly, for the mood and arousal questionnaire, traditional statistics revealed that subjects reported equal changes for the anodal and cathodal stimulation condition on all subscales (all uncorrected p's > 0.17), independent of the order in which they received each type of stimulation (all uncorrected p's > 0.24). However, Bayesian statistics indicated that there is strong evidence for a difference in change in the subscale energy between the stimulation conditions (Bf = 11; unrelated to session order (Bf = 0.5)), but not any of the other subscales (Main effects Stimulation Type: Bf's < 0.5, Interactions Stimulation Type * Session Order: Bf's < 0.24). The change in the level of energy was on average 0,68 (StD: 2,52) in the anodal and -2,26 (StD: 2,82) in the cathodal stimulation condition. Overall, subjects reported lower levels of energy (Mean: 1,79, StD: 2,67), higher levels of tiredness (Mean: -2,21, StD: 2,38) but no substantial differences in feelings of tension (Mean: 0,45, StD: 1,97) or calmness (Mean: -0,21, StD: 1,39) at the end compared to the beginning of the stimulation session.

Discussion

In the current study, we aimed to investigate if baseline cortical excitability can explain individual differences in how tDCS affects cognitive functioning. Specifically, we tested if prefrontal cortical excitability levels, as indexed by 3T-MRS measured Glutamate/GABA ratios, can predict the extent to which anodal or cathodal prefrontal tDCS stimulation improves or impairs verbal WM performance across subjects.

Replicating previous observations of large individual variability in tDCS effects on WM performance (Berryhill & Jones, 2012; London & Slagter, 2015; Talsma et al., 2016), for both types of stimulation, about half of the participants showed improved verbal WM updating accuracy after stimulation, while the other half showed detriments. Yet, in contrast to our main expectations, baseline prefrontal cortex excitability did not predict the effects of anodal or cathodal tDCS on WM functioning across subjects. Moreover, in contrast to earlier studies (e.g. Andrews et al., 2011; Fregni et al., 2005; Lally, Nord, Walsh, & Roiser, 2013; Ohn et al., 2008; Talsma et al., 2016), at the group-level, neither anodal nor cathodal stimulation affected WM performance.

Since we used the exact same stimulation parameters and a very similar verbal WM updating task as in our previous study (Talsma et al., 2016), this latter finding is surprising. Nevertheless, at the same time, these null findings add to the growing number of reports that the relation between anodal prefrontal stimulation and WM improvements is not as consistent as initially assumed (e.g., see the meta-analyses by Bennabi et al., 2014; Brunoni & Vanderhasselt, 2014; Dedoncker, Brunoni, Baeken, & Vanderhasselt, 2016; Hill, Fitzgerald, & Hoy, 2016; Mancuso, Ilieva, Hamilton, & Farah, 2016). Given the also noted variability in individual tDCS response, especially in smaller subject samples, group-level conclusions may be substantially affected by the specific selection of subjects within the sample, thereby creating inconsistencies in conclusions across the field.

At the same time, however, our findings point out the importance of looking at why the effects of tDCS vary so substantially across subjects. In the current study, we tested the hypothesis that baseline cortical excitability levels (partly) determine the effect of tDCS on WM performance. However, in contrast to this notion, we found no evidence that baseline prefrontal cortical excitability levels predicted individual differences in the effects of anodal and cathodal prefrontal stimulation on WM. Being a pioneering report in this regard, below we will discuss some limitations of the current study and suggest directions for future research necessary to confidently determine the absence of presence of this relationship between baseline cortical excitability and tDCS effect.

A first important limitation of the current study is the use of 3T-MRS to quantify Glutamate and GABA concentrations to calculate prefrontal cortical excitability levels. In the past decade, 3T-MRS has repeatedly been used to in-vivo measure concentrations of

the main inhibitory (GABA) and excitatory (Glutamate) neurotransmitters in the human brain, and to relate these neurotransmitter concentrations to individual differences in behaviour (e.g., Yoon, Grandelis, & Maddock, 2016; see for an overview also Duncan, 2013). However, in a recent study (Talsma et al. submitted) with a larger subject sample than earlier studies, we failed to replicate two of these previously reported relations (Edden, Muthukumaraswamy, Freeman, & Singh, 2009; Yoon et al., 2016). Specifically, prefrontal GABA did not predict WM updating, updating capacity or WM maintenance, nor did occipital GABA predict visual discrimination performance. These findings challenge the idea that 3T-MRS provides a measure that is sensitive enough to adequately quantify Glutamate and GABA concentrations predictive of cortical functioning and behaviour. If so, 3T-MRS may also fail to provide the sensitivity that is necessary to successfully investigate the role of baseline cortical excitability in the effects of tDCS on behavioural performance.

A pivotal direction for future studies that aim to investigate the relationship between cortical excitability and tDCS effects may be to use 7T-MRS, which has two important advantages over 3T-MRS. Firstly, increased spectral resolution at higher magnetic field strengths enables better discrimination, and thereby quantification of the two neurotransmitters critical for the cortical excitation/inhibition balance: Glutamate (independent from Glutamine (An et al. 2014)) and GABA (uncontaminated by macromolecules (Ganji et al. 2014)). Secondly, because of the better signal-to-noise ratios (Choi et al., 2010), smaller sized MRS voxels can be used, which may substantially enhance the spatial precision of the brain area for which one aims to determine cortical excitability.

Indeed, another factor that may have obscured a relationship between prefrontal cortical excitability and tDCS effects is the fact that the region of lateral prefrontal cortex that is considered critical to WM functioning (Owen et al., 2005) is much smaller than the voxel area that we need with current 3T- MRS to obtain a good enough signal. Placing a relatively large voxel over an actually much tinier region of interest may ‘delute’ the measure significantly, thereby reducing its sensitivity for what we know are very regional specific concentrations. Functionally localizing the part of the lDLPFC that is involved in WM functioning in every subject individually, and subsequently placing a smaller (7-T) MRS voxel over this area, could hence also substantially improve sensitivity of the measure of cortical excitability, and be important to successfully study its relation to tDCS effects on behaviour.

Lastly, two additional factors may have played a role in our findings. Firstly, due to the lack of a no stimulation control condition, non-specific effects related to time on task, such as changes in fatigue level or learning, may have confounded our measure of tDCS-induced change in WM performance. Secondly, next to cortical excitability, in conventional non-individualized tDCS set-ups, other factors may determine the effectiveness of tDCS to an even larger extent, such as anatomical differences affecting current flow (Kim et al., 2014). Although the low spatial accuracy of conventional tDCS set-ups heightens the chance that the lDLPFC target area is affected at least to some extent in all participants, the exact amount of current that reaches the targeted cortical neurons in each subject likely greatly varies. These inaccuracies may mask contributions of more subtle factors such as delicate interactions of tDCS with the baseline cortical excitation/inhibition balance of the area. Developing more individualized stimulation protocols that allow for a more precise deliverance of a specific amount of current to the (individually localized) target brain area in every subject may therefore be a critical next step before we can further investigate the role of baseline cortical excitability in determining the effect of anodal and cathodal tDCS on cognitive performance.

Conclusion

To our knowledge, the current study is the first to test the hypothesis that baseline cortical excitability levels critically determine the effects of tDCS on cognitive functioning. Although we observed large individual differences in tDCS response, baseline prefrontal cortical excitability levels did not predict which subjects improved and which actually deteriorated after anodal or cathodal stimulation. However, being a pioneering study, these findings should be interpreted with care and should first and foremost serve to direct the design of future studies in this field. Hopefully, this will eventually lead to a better understanding of tDCS and how it may improve WM. This knowledge is not only essential to help resolve current inconsistencies in the field, but also to ensure the practical application of tDCS to enhance WM functioning not just in some, but in all individuals.

.

Chapter 5

Summary and discussion

Summary of aims and main findings

In the current thesis, I set out to (i) determine the potential use of tDCS combined with practice on a WM task to induce lasting and transferable enhancements in WM functioning, and (ii) investigate the possible role of cortical excitability in determining how tDCS affects WM performance in a given individual (positively or negatively).

Specifically, in ***Chapter 2***, I investigated the effects of multiple sessions of anodal tDCS over the left dorsolateral prefrontal cortex (lDLPFC) on WM performance to determine if it is possible to continue to boost verbal WM across three daily stimulation sessions. Since I was the first to include a baseline measure of WM performance each day, I could importantly discriminate within-session effects of tDCS from carry-over effects of previous stimulation sessions. Furthermore, by administering several WM transfer tasks in a behavioural pre- and post-session, I examined whether tDCS induced improvements may transfer to other WM tasks and domains and thus reflect domain-general enhancements.

In line with previous reports (see e.g. Bennabi et al., 2014), anodal tDCS (compared to sham), significantly improved verbal WM performance in the first stimulation session. This effect remained visible 24 hours after stimulation. However, in contrast to our expectations, no effect of anodal (compared to sham) tDCS on verbal WM was found in the second and third stimulation session. This leads us to conclude that multi-day stimulation may not necessarily extend the benefits of a single session with anodal prefrontal tDCS on WM.

What's more, at the group-level we found no differences between anodal and sham stimulation on any of our three transfer tasks: a verbal WM task with a different letter-set, a spatial WM task (see for both Figure 2 of the Introduction) or a complex WM span task (the Operation-span (Unsworth, Heitz, Schrock, & Engle, 2005)). Yet, post-hoc individual difference analyses revealed that in the anodal stimulation group, the extent of WM improvement in the first stimulation session predicted pre- to post changes on both the verbal and spatial WM transfer task. This relationship was not observed for the sham stimulation control group. Although this finding should be interpreted with care given the small sample size of the study (N=15), it does suggest that if effective, tDCS paired with practice on a WM task may induce domain-general WM improvements that transfer to other stimulus modalities (i.e., spatial) than the one paired with stimulation (i.e., verbal). At

the same time, it illustrates the importance of looking at individual differences in tDCS response. This was further highlighted by the fact that performance of two participants in the anodal stimulation group worsened after stimulation and showed similar pre- to post decrements on both transfer tasks. Thus, remarkably, the same tDCS stimulation that may benefit the majority of individuals may actually impair WM functioning in others.

One factor that may determine how tDCS affects individual performance is baseline cortical excitability. In ***Chapter 3*** and ***4***, I therefore explored the relation between individual differences in tDCS-induced effects on WM and GABA and Glutamate levels in the left dorsolateral prefrontal cortex. In ***Chapter 3***, I addressed two important methodological questions with regard to the use of Magnetic Resonance Spectroscopy (MRS) to in-vivo measure local Glutamate and GABA concentrations in the human brain. Specifically, I investigated whether concentrations measured with 3T-MRS reflect 'state' or 'trait'-like neurotransmitter levels, as well as looked at the optimal conditions under which these may be measured in relation to behaviour.

We found that neither in the medial occipital brain area (i.e. the visual cortex), nor in the left lateral prefrontal cortex, concentrations of Glutamate and GABA differed when we measured them at rest (eyes closed), or under a challenging task condition (watching a movie (OC) or performing a challenging WM task (lDLPFC)). This suggests that 3T-MRS does not capture small-scale demand-induced fluctuations in Glutamate or GABA levels (i.e., brain region activity 'state'), but rather provides a measure of relatively stable 'trait'-like neurotransmitter levels (that notably also correlated well over activity conditions within subjects).

However, in contrast to our expectations, individual differences in neurotransmitter concentrations did not correlate with individual differences in behavioural performance on region-related tasks. More specifically, medial occipital GABA and Glutamate did not predict visual discrimination performance (in contrast to Edden, Muthukumaraswamy, Freeman, & Singh, 2009), nor did left lateral prefrontal concentrations predict WM updating accuracy, capacity or maintenance performance (as may be expected from e.g. (Yoon, Grandelis, & Maddock, 2016)). Our findings thereby question to what extent Glutamate and GABA concentrations measured with conventional 3T-MRS practices reflect neurotransmitter activity that is relevant for behaviour.

Since cortical excitability (i.e., the excitation/inhibition balance) is considered to be primarily maintained by Glutamate (the brain's main excitatory neurotransmitter) and GABA (inhibitory), in ***Chapter 4***, I calculated Glutamate/GABA ratios to use as a proxy for cortical excitability and related these to the effects of anodal and cathodal prefrontal tDCS on WM. Hereby, I thus tested whether baseline prefrontal cortical excitability levels can predict which subjects may benefit and which deteriorate from anodal and/or cathodal tDCS.

In line with previous studies and our earlier findings (***Chapter 2***), in ***Chapter 4***, we observed large individual variability with regard to the strength and direction of the effects of both anodal and cathodal prefrontal tDCS on WM performance. After both anodal and cathodal stimulation session, about half the participants improved, whereas in the other half WM performance actually seemed to worsen after stimulation. Yet, at the group level we did not observe a differential effect of anodal compared to cathodal tDCS on WM, thereby adding to the general inconsistency of reports in the field (e.g. (Hill, Fitzgerald, & Hoy, 2016)). Moreover, baseline prefrontal cortical excitability levels did not predict which subjects benefited from tDCS and which did not. In this chapter, we thus did not find evidence for the hypothesis that baseline cortical excitability may critically determine the effects of prefrontal tDCS on WM performance in a given individual.

General discussion

In the current thesis, I aimed to take some pioneering steps in investigating the potential of tDCS to improve WM and the possible role of cortical excitability in determining the effects of tDCS. In this general discussion, I will relate our findings summarized above to the broader literature and address current methodological limitations of, and issues related to tDCS and 3T-MRS that may be important in interpreting our findings. In doing so, I will also delineate important directions for future studies.

Can anodal prefrontal tDCS improve WM functioning?

In line with several previous studies (e.g. Andrews, Hoy, Enticott, Daskalakis, & Fitzgerald, 2011; Fregni et al., 2005; Mulquiney, Hoy, Daskalakis, & Fitzgerald, 2011; Ohn et al., 2008), in ***Chapter 2*** we found that in a between-subject design, a first application of anodal tDCS

over the left DLPFC compared to sham tDCS significantly improved WM performance. Yet, in ***Chapter 4***, we did not observe a beneficial effect of anodal compared to cathodal prefrontal tDCS on WM performance in our subjects. These contradicting findings corroborate conclusions of recent meta-analyses that suggest that the effect of anodal prefrontal tDCS on WM is not reliable (Bennabi et al., 2014; Brunoni & Vanderhasselt, 2014; Dedoncker, Brunoni, Baeken, & Vanderhasselt, 2016; Hill et al., 2016; Mancuso, Ilieva, Hamilton, & Farah, 2016).

One explanation for the contradicting observations of the effects of anodal tDCS on WM in the current thesis may be the control condition against which effects were compared. In Chapter 2, we compared changes in WM performance before and after stimulation between an anodal and sham stimulation group. However, in Chapter 4, we compared changes in WM before and after stimulation within subjects, between an anodal and cathodal stimulation condition (spaced one week apart).

Although cathodal stimulation seems to consistently inhibit cortical activity in studies in the motor domain, this inhibitory effect seems less consistent in the cognitive domain (Jacobson, Koslowsky, & Lavidor, 2011). Possibly, this is the result of more complex networks playing a role in higher-order functioning in which other brain areas may, for example, start to compensate. So far, few studies have combined tDCS with functional neuroimaging or M/EEG to investigate this possibility. Nevertheless, one study that combined EEG to investigate prefrontal effects of tDCS on WM and brain oscillations found that whereas anodal tDCS enhanced oscillatory power (in the theta and alpha frequency bands) and WM performance, cathodal tDCS had the opposite effect and reduced oscillatory power and WM performance (Zaehle, Sandmann, Thorne, Jäncke, & Herrmann, 2011). Additionally, as addressed in more detail in ***Chapter 4***, the hypothesis that the effect of both anodal and cathodal tDCS on WM functioning critically depends on its interaction with baseline excitability levels, predicts that for individuals with relatively high baseline cortical excitability levels, cathodal tDCS may in fact be beneficial to cortical functioning and WM performance.

At present, however, at the neural level, the precise effects of cathodal tDCS on cortical activity remain largely elusive. Before this outstanding issue is addressed, future studies that aim to investigate effects of both anodal and cathodal tDCS on cognitive functioning should therefore make sure to also include a sham stimulation control group. This may help resolve these inconsistencies in the future.

Notably, in ***Chapter 2*** and ***4***, I used the exact same stimulation set-up and settings for anodal prefrontal stimulation (cathode contra-supraorbital, 1 mA, 20 minutes, see also Figure 3 in the Introduction). Therefore, parameter choices of electrode location, current intensity and duration (the importance of which is discussed in more depth below) that may explain inconsistencies in findings across research groups (Antal, Keeser, Priori, Padberg, & Nitsche, 2015), cannot explain the contradicting effects of tDCS observed in these chapters. Moreover, the task that we paired with stimulation was the same in both studies; a verbal WM updating task (the letter N-back, see Figure 2 of the Introduction). However, the levels of the WM task (level N) was different for both chapters, which may have affected anodal tDCS effects on behaviour.

To ensure a challenging task and enough room to improve in everyone, in ***Chapter 2*** I preselected a homogenous group of subjects based on their performance in the pre-session (thereby excluding ceiling level performers). However, for the study of ***Chapter 4,*** I recruited a subset of subjects for which we measured MRS in ***Chapter 3***, and to ensure a challenging performance level in all subjects, I adjusted WM task difficulty to individual WM capacities. Subjects in Chapter 4 turned out to have surprisingly high WM capacities (mean level of N = 4,4, StD 1,2). This caused task levels to commonly involve extraordinary levels of N, such as 7, 8 and 9 (i.e., the length of sequence of letters to be maintained in order and to be continuously updated). This may have tempted subjects to use strategies, such as chunking, which may have also affected the ability of tDCS to improve WM functioning per se. Moreover, including a group of relatively good WM performers may have diminished the chance of finding anodal tDCS-induced enhancements across the board; namely, WM functioning may have already been optimal in the majority of these subjects (i.e., at their individual ceiling levels), leaving little room for improvements at the group level. Thus, several factors may have contributed to the different pattern of findings between Chapters 2 and 4, including the control stimulation condition, differences in task difficulty and strategy use.

In contrast to our predictions, in our multi-session stimulation design in ***Chapter 2***, we did not observe any enhancements after anodal compared to sham prefrontal tDCS on

the second or third day of stimulation. Albeit unexpected, these findings corroborate other reports that the effects of anodal prefrontal tDCS in multiple daily sessions with stimulation, if present at all, are confined to the first session (e.g. Lally, Nord, Walsh, & Roiser, 2013). With regard to these previous studies, I reasoned that these null-effects may be the result of suboptimally chosen stimulation parameters (i.e., electrode positioning, stimulation intensity and duration), and the task paired with stimulation (see for elaboration the introduction of ***Chapter 2*** and the section on determining tDCS parameters below). However, also in our study of ***Chapter 2*** in which I attempted to resolve these issues, we failed to observe a beneficial effect of tDCS beyond the first stimulation session.

In recent years, findings in the motor domain have lead researchers to question the original assumption that the aftereffects of tDCS on neuro-excitability wear out within hours after stimulation is applied (Monte-Silva et al., 2013). Therefore, one likely explanation for the absence of tDCS effects in multiple sessions with tDCS is that these may result from unexpected and unwanted interactions between spill-over effects of previous stimulation and stimulation in subsequent sessions. This interaction may diminish the positive effects of tDCS or actually result in negative effects on neuronal functioning in the underlying cortex. In the section below on additional issues in multiple sessions of tDCS stimulation, I will discuss this issue in more detail.

Interestingly, in ***Chapter 2***, the (group-level) effect of anodal tDCS after the first application was still apparent in the baseline block of the second stimulation session (i.e. 24 hours afterwards). Though it is currently unknown whether this reflects prolonged changes in neuro-excitability, this finding may imply that tDCS can actually induce longer lasting WM improvements. Furthermore, taking an individual differences approach, in the anodal stimulation group, the extent to which tDCS improved WM in an individual significantly correlated with the extent to which he or she showed improvement on the verbal and spatial WM transfer tasks. Although preliminary, because of the small sample size and the concurrently observed lack of transfer effect at the group level, this might indicate that if effective, tDCS can induce longer lasting effects in some individuals, which importantly, transfer to other domains and therefore may reflect actual domain-general WM learning.

In general, the research field of prefrontal tDCS and WM desperately needs a better understanding of the factors that may determine and promote positive effects of prefrontal

tDCS on WM functioning (in a given individual). In the next section, I will therefore discuss a number of stimulation parameter choices that may critically influence the effects of tDCS on brain and cognitive functioning.

Determining tDCS stimulation parameters: Developing the best recipe for success

In tDCS research, there are a number of parameters that need to be set before stimulation is delivered, that may each critically determine the specific effect of the tDCS on cortical functioning (see e.g. Michael A. Nitsche et al., 2008). These parameters include the choice of electrodes (i.e., electrode number, size and shape) and the specific positioning of the electrodes on the head; stimulation intensity and duration; and the conditions under which the tDCS is applied (i.e., what the subjects are doing during stimulation). This multi-dimensional parameter space allows for a lot of freedom in the use of tDCS to affect brain functioning, and presently, the precise role of and optimal settings for each of these parameters are still unclear. Suboptimal stimulation settings may prevent tDCS to affect behaviour, and lead to an accumulation of reports of null findings, unjustly diminishing faith in the potential of tDCS to improve cognitive functions, such as WM.

One factor that is considered critical in determining the effects of tDCS is the stimulation electrode set-up. Especially in conventional two electrode tDCS set-ups, the positioning of the reference electrode crucially influences the part of the brain that is affected by the tDCS (directly under the electrodes) as well as current flow (between both electrodes). Illustrative of this, a recent study, albeit with a relatively small sample size per group, reported that 5 days of anodal prefrontal stimulation with the cathode placed over the contralateral supraorbital cortex (i.e. above the right eye, as used in this thesis), no tDCS effects in WM were observed (Möller, Nemmi, Karlsson, & Klingberg, 2017). However, when the cathode was placed over the visual cortex in the combined tDCS and WM training sessions, WM performance was actually impaired. Possibly, in this latter case, cathodal stimulation of the occipital lobe interfered with visual processing of the task stimuli, thereby (indirectly) impairing WM performance. Similarly, when placing the reference electrode over the dorsolateral prefrontal cortex in the right brain hemisphere (as e.g. done in Richmond, Wolk, Chein, & Olson, 2014), inhibition of the area under this cathodal electrode may interfere with potential beneficial effects of the anodal electrode over the left DLPFC,

as the right DLPFC is also known to be involved in WM functioning (Au et al., 2016; Berryhill & Jones, 2012; Owen, McMillan, Laird, & Bullmore, 2005).

However, even when the reference electrode is strategically placed to not affect cortical functioning, spatial accuracy of conventional tDCS set-ups that aim to stimulate the lDLPFC is currently poor. With traditionally sized electrodes, the brain area that is affected is at least as large as the size of the active electrode with which the current is delivered, e.g., 5 x 7 = 35 cm2. Moreover, recent modelling studies suggest that in reality, spatial accuracy may be even lower and lDLPFC tDCS may stimulate almost the entire left prefrontal cortex and possibly part of the right counterpart, too (see e.g. Truong, Magerowski, Blackburn, Bikson, & Alonso-Alonso, 2013). The human prefrontal cortex houses many functions. Whereas parts of the dorsolateral prefrontal cortex play a key role of in executive functions, such as WM, more anterior parts of the frontal cortex are important for e.g., goal-directedness and motivation (Miller & Cohen, 2003). If lDLPFC tDCS affects many different prefrontal brain areas simultaneously, this may render it difficult to optimize the neural network for a specific function, such as WM, because a delicate balance may exist between these brain networks (Wokke, Talsma, & Vissers, 2015). Moreover, it makes pinpointing the precise source of the tDCS effects observed in behaviour difficult.

To address the issue of spatial precision in tDCS, in the last five years, more advanced tDCS set-ups have been developed, including the so-called High Density tDCS set-ups (HD-tDCS). By surrounding the active electrode with a multitude of smaller reference electrodes (typically four), the electrical current is contained in between these electrodes, significantly enhancing the spatial precision of the tDCS stimulation (Alam, Truong, Khadka, & Bikson, 2016; Kuo et al., 2013; Truong et al., 2013). Such set-ups also automatically solve the issue of choosing an optimal reference location. With considerably better control over the brain area that is affected by brain stimulation, advancing HD-tDCS set-ups is therefore an important avenue for future research to further explore. Nevertheless, due to the nature of the method (i.e. current delivered via scalp-based electrodes), tDCS will always have a much poorer spatial resolution compared to other non-invasive brain stimulation methods such as Transcranial Magnetic Stimulation. Yet, there may also be an upside to lower spatial accuracies, as this renders the method less susceptible to individual differences in the precise location of the part of the lDLPFC involved in WM. This increases the potential of a

(to some extent) one-fits-all approach. Enhancing spatial precision to optimize brain region-specific functioning may at some point thus inevitably also require more individualized approaches to functional localization (e.g., by using fMRI) and electrode placements.

Next to electrode montage and size, another important parameter that can affect the effectiveness of tDCS is stimulation intensity. Findings from a recent study, which examined the parameters necessary to affect brain activity in vivo in humans and rats, indicate that stimulation with higher intensity than in conventional protocols, like ours, is necessary to actually affect neuronal circuits (Vöröslakos et al., 2018). Nonetheless, other studies suggest that 'more' stimulation is not always 'better'. For example, a study with a within-subject design showed that 1 mA and not 2 mA anodal lDLPFC tDCS resulted in the most pronounced WM improvements (Hoy et al., 2013). Modelling studies have also shown that individual differences in anatomy, skull thickness and cortical folding may to a large degree determine the amount of current that actually reaches the brain (Opitz, Paulus, Will, Antunes, & Thielscher, 2015). Rather than searching for a group-level 'optimal' current setting, it may thus be more valuable to develop individually tailored stimulation protocols. Using anatomically (and possibly functionally) based modelling, in the future, we may be able to individually determine optimal current densities and electrode placements (as well as current density per electrode). This step will be critical to ensure a better control of the amount of tDCS stimulation that actually reaches the brain area of interest in each and every individual.

Yet another parameter that can critically influence the effects of tDCS on cortical functioning is stimulation duration. Little research has so far been done on optimal settings for tDCS durations either for inducing immediate or lasting WM improvements. At the beginning of the century, a pioneering study in the motor domain found that by increasing stimulation duration from 5 to 13 minutes, the after-effects of anodal stimulation on motor cortex excitability could be increased from 1 to 2 hours, indicating that longer stimulation may induce longer-lasting effects (Nitsche & Paulus, 2000). However, a later study showed that extending stimulation durations from 13 to 26 minutes actually resulted in reduced excitability (Monte-Silva et al., 2013). Possibly, this opposing effect of tDCS is caused by homeostatic metaplastic mechanisms (Siebner et al., 2004) by which the brain attempts to counter the induced change. This may eventually lead to the opposite effect than originally

anticipated (see for an elaboration on this mechanism at a cellular level (Karabanov et al., 2015)) and thus possibly a negative effect on brain functioning and behaviour. Stimulation duration is thus another parameter that requires more investigation.

A final critical factor that should be taking into account when predicting the effects of tDCS is the condition under which tDCS is applied, i.e., what do subjects during stimulation? In single-session tDCS studies, it has been shown that anodal tDCS admitted concurrent with a task (i.e., online stimulation) is more effective in boosting WM than tDCS admitted during rest (i.e., offline stimulation) (Mancuso et al., 2016; Martin, Liu, Alonzo, Green, & Loo, 2014; Andrews et al., 2011). Likely, this is because tDCS needs some baseline level of activity in the stimulated cortex to 'grasp' onto. By ensuring that the WM brain network is active by involving it in a challenging WM task while tDCS current is applied, this may lead to an optimal situation of tDCS to do so. Future research is necessary to advance our understanding of the interaction of tDCS with ongoing activity in the stimulated cortex.

Next to determining the 'optimal' parameters to consistently induce direct effects of tDCS on cortical functioning and performance, it is also crucial that research determines the optimal conditions under which tDCS can promote synaptic-dependent neural plasticity and longer lasting effects on cognitive performance and learning. In the current thesis, I have discussed the role of cortical excitability predominantly in relation to ensure optimal cortical functioning at that moment. However, if the aim is to use tDCS to induce longer lasting improvements in WM, that outlast the stimulation for more than e.g., 24 hours, it is also critical that future research investigates how tDCS may best promote neural plasticity. Currently, it is not known whether the optimal parameter settings to bring about immediate effects on behaviour are the same parameters that optimally induce longer lasting (and transferable) learning effects. Further complicating matters, we also know little about whether such potential optimal learning conditions (i.e., for neural plasticity to take place) are the same for everyone and all regions of the cortex, or if these vary as a function of brain region, function and/or individual. In ***Chapter 2***, we found that the extent to which anodal tDCS improved WM functioning in the first session predicted how much subjects would improve on the verbal and spatial WM transfer tasks. This may be an indication for a causal relation between the direct effects of tDCS and eventual learning effects. However,

it could also merely reflect individual differences in our success to deliver the current to the targeted cortex across subjects.

Additional issues in multiple session tDCS

In addition to determining the most optimal stimulation parameters for inducing WM enhancement with a single tDCS session, researchers that aim to use longitudinal tDCS to speed up and/or enhance the effects of cognitive training are introduced to a whole new issue: how should we space multiple sessions with tDCS? Since tDCS has effects on cortical excitability that outlast the actual stimulation period (Nitsche et al., 2008), the time that is left in between sessions for these effects to establish or wear out may critically define both the immediate effects of tDCS on cortical excitability in subsequent sessions, as well as on neuroplasticity mechanisms related to learning and consolidation.

At the time that we designed the study in ***Chapter 2***, it was considered safe to assume that the immediate effects of tDCS on neuro-excitability would have worn out if stimulation sessions were spaced 24 hours apart (Nitsche et al., 2008). However, more recent evidence casts doubt on this initial assumption. For example, research in the motor domain has shown that if two tDCS sessions (of 13 minutes each) are admitted with a short interval (3 to 20 minutes) in between, the increase in motor cortex excitability observed directly after the first stimulation period disappears after the second stimulation (Monte-Silva et al., 2013). Moreover, with even longer intervals (12 or 24 hours) between stimulation sessions, the effects of anodal tDCS can actually reverse, surprisingly leading to inhibited excitability after motor cortex stimulation (Monte-Silva et al., 2013). This pattern of findings may suggest that if the second instance of stimulation is applied too soon after the first, it can interact with the after-effects of the first stimulation, and this can actually induce the opposite effect on cortical functioning than intended.

Assuming that the same neuronal homeoplastic mechanisms are at play in the motor and prefrontal cortex, interactions between stimulation sessions may have played a role in our multiple-session tDCS study of ***Chapter 2***. Interestingly, in this Chapter, I observed that the beneficial effects of the initial application of tDCS in the first session was still apparent in the baseline block in the second day of stimulation. This could indicate that the combined anodal tDCS and WM practice induced longer-lasting learning effects, that

remained apparent in behaviour 24 hours later. However, another explanation for this finding could be that neuro-excitability levels in the lDLPFC had in fact not yet returned to baseline and this enhanced prefrontal excitability state still facilitated WM performance in our subjects. In any case, tDCS effects of earlier sessions may have interacted with subsequent stimulation, possibly diminishing or even masking its effects. This could explain the observed lack of a within-session effect of tDCS on WM performance on the second and third day of stimulation.

Following up on this line of reasoning, perhaps then, if I would have spaced the stimulation sessions in ***Chapter 2*** apart further in time, I would have observed a boost in WM performance after every instance of tDCS, leading to the desired accumulation of WM improvements with multiple-session tDCS. Interestingly, in line with this idea, a recent study with seven sessions of anodal prefrontal tDCS paired with WM found significantly larger gains between the third and fourth stimulation session in subjects in which these sessions were separated by a weekend (i.e., 72 hours) compared to those that received them on consecutive days (i.e., 24 hours) (Au et al. 2016). Unfortunately, as this study did not include a daily baseline measure, it is unclear whether these findings should be interpreted as a larger effect of tDCS in this fourth session, or stem from better learning consolidation after longer 'rest' time relative to the previous (third) session. Therefore, future studies with designs that include a baseline measurement each day and systematically vary the spacing of sessions are necessary to elucidate the role of time in between multiple stimulation sessions on the effect of tDCS on WM.

To summarize, there are two ways in which prefrontal anodal tDCS may enhance WM performance. Firstly, because of its direct effects on cortical excitability, tDCS may instantly enhance prefrontal functioning and thereby improve WM. Secondly, tDCS may enhance neuroplasticity, thereby promoting (hopefully domain-general) WM learning. Currently, much is still unknown about both of these potential mechanisms, and many avenues are left unexplored. Future research that advances our understanding of the influence of different parameter settings in determining the effects of stimulation is pivotal for tDCS to develop into a method that can be used safely and robustly to non-invasively improve brain functioning also in clinical populations in the future.

Importance of individual differences: Interactions of tDCS with baseline cortical excitability

When investigating the optimal settings for tDCS to affect WM functioning, it is crucial that a close eye is kept on individual differences in tDCS response. By its very nature (being a neuromodulatory technique), tDCS stimulation will always critically interact with the baseline situation. Therefore, rather than focusing on finding the holy grail in the form of a one-fit-to-all kind of stimulation, it is of fundamental importance that future research also determines how tDCS stimulation can be tailored such that it is most beneficial for everyone.

In both ***Chapters 2*** and ***4***, we observed large individual differences in extent and direction of the effect of tDCS on WM performance. These observations add to a growing body of research that indicates large variability in how tDCS affects cognitive performance in a given individual (Berryhill & Jones, 2012; Learmonth, Thut, Benwell, & Harvey, 2015; London & Slagter, 2015). A substantial part of this variability may result from anatomical differences that critically affect the amount of stimulation that reaches the targeted brain area and/or the precise region that is affected by the current. Modelling studies and tailored stimulation protocols (using HD-tDCS set-ups) in which current densities and precise electrode placements are determined individually are hence crucial to ensure better control of the amount of current that actually reaches the target brain area in the anticipated manner in the future. However, even in an 'ideal' situation in which we have full control over the current density in the brain area of interest, there are several additional factors which role will need to be explored before we can decide on individually tailored 'best recipes for success'.

Especially critical in this regard is to establish how tDCS interacts with baseline cortical excitability (as I investigated in ***Chapter 4***). As described in the general introduction of this thesis in more detail, ideally the type and intensity of stimulation should be individually tailored to help push or pull the excitability in the relevant brain region towards the level for optimal functioning in each individual (Krause, Márquez-Ruiz, & Kadosh, 2013).

In ***Chapter 4***, we found no evidence that baseline prefrontal cortical excitability determines the effects of tDCS on WM. However, in ***Chapter 3***, we also failed to find a general correlation between behavioural WM performance and baseline GABA and

Glutamate levels in the prefrontal cortex. These findings question the validity of our measure to pick up small-scale alterations in GABA and Glutamate to relate to behaviour. Currently, the spatial resolution of 3T-MRS is very poor. The typical size of a prefrontal cortex voxel at 3T for which sufficient signal (i.e., a good signal-to-noise ratio) is obtained is typically 3 x 2 x 2,5 cm = 15 cm3. With current 3T-MRS methods we are thus limited to quantifying average GABA and Glutamate concentrations over a relatively large portion of brain tissue.

Variability in functional neuroanatomy between individuals is known to be particularly high in higher-order cortical areas, such as the prefrontal cortex. For example, with regard to WM specifically, functional neuroimaging research has observed peak lDLPFC activations during a challenging WM task to be spread across subjects along the middle frontal gyrus (Jansma et al., 2013). Although the brain area that we usually refer to as the left dorsolateral prefrontal cortex may be relatively large (see Figure 1 of the Introduction), the precise portion of this region that is involved in WM in each individual specifically may be much smaller. Placing a relatively large voxel over an actually much smaller region of interest may cause the signal of interest to drown in a sea of irrelevant signals from the surrounding area. Given the size of our 3T-MRS voxel, this may have significantly 'deluted' our measure. This may be a key reason why we failed to find a correlation between GABA and Glutamate neurotransmitter levels in left "DLPFC" with individual differences in WM performance in ***Chapter 3*** and between cortical excitability ratio and tDCS response in ***Chapter 4***.

A pivotal next step for future studies is thus to increase the spatial accuracy of their MRS measurement. One option is to measure neurotransmitter concentrations at higher magnetic field strengths than at 3T as done in this thesis. Because of better signal-to-noise ratios at 7T (Choi et al., 2010), significantly smaller sized MRS voxels can be used, hence increasing sensitivity to study neurochemical concentrations in a precise target region (e.g., lDLPFC). Yet, as also discussed in relation to potential future tDCS set-ups, enhanced spatial accuracy also calls for better individual function localization, as measuring from a smaller, yet irrelevant part of cortex would still miss the point.

Another reason why despite the likely theoretical framework, we may have failed to find evidence for a predictive role of baseline prefrontal cortical excitability in ***Chapter 4*** is

that due to our study design, other factors may have played a more dominant role in determining the effect of tDCS. For practical reasons, I applied tDCS with a conventional set-up and a one-fits-all approach. However, although the actual amount of current that reaches the target cortex in each subject is strongly affected by individual neuroanatomical differences (see the section on this above), we did not account for these in the current design. Unfortunately, such individual differences may have obscured the effects of subtle factors such as baseline cortical excitability in our study. Therefore, I believe that future research in which these methodological issues are addressed, is necessary before we can make firm conclusions about the role of baseline cortical excitability in determining how anodal and cathodal tDCS may affect WM performance.

Perhaps the most pressing concern at the moment with regard to our lack of knowledge about tDCS is that with conventional tDCS set-ups, we may not only not optimally succeed in benefitting everyone's WM, we can actually worsen WM functioning in some people. Although the implications of this may be acceptable within the controlled surroundings of scientific research, it imposes clear ethical problems in a widespread application of tDCS as cognitive enhancer in clinical as well as healthy populations.

Standardization of tDCS protocols across the research field will allow for a better comparison between studies and (hopefully) eventually a more cohesive body of findings with regard to the (lack of) effects of tDCS on cortical functioning and neuroplasticity. However, at the same time, determining the factors that influence the effect of tDCS within each individual is crucial. Only by incorporating this knowledge in the design of future stimulation protocols may we potentially eventually reach the goal of optimizing WM in clinical populations for which WM failures critically affect everyday life functioning.

In order to achieve this, fundamental research should furthermore be conducted in close interaction with studies in populations characterized by WM problems, such as in aging adults or patients with schizophrenia. In this, a special focus should be placed on baseline cortical excitability levels as these may particularly be important to consider when findings from healthy subjects are extrapolated to clinical patient groups, such as schizophrenia (e.g.(Crabtree et al., 2017). In the clinical domain, effects of tDCS also seem variable. For example, one study that combined tDCS with WM training suggest that tDCS can enhance WM training benefits in aging adults (Stephens & Berryhill, 2016). Yet, another

recent study (and meta-analysis) found no benefits in older adults (Nilsson, Lebedev, Rydström, & Lövdén, 2017, but see also Passow, Thurm, & Li, 2017). Similar observations have been made in the domain of addiction research (e.g., den Uyl, Gladwin, Rinck, Lindenmeyer, & Wiers, 2017). Fundamental research in healthy populations is pivotal in creating a better understanding of how tDCS can enhance brain and cognitive functioning, but to eventually allow application in clinical settings, it is important to also specifically determine if and how tDCS can best be used to reduce WM problems in clinical populations and the elderly.

Conclusion

In the current thesis, I set out to determine (i) the potential use of tDCS to induce lasting and transferable enhancements in WM functioning, and (ii) investigate the possible role of cortical excitability in determining tDCS-induced effects on WM. Excitingly, we found preliminary evidence that if effective, left prefrontal anodal tDCS combined with practice on a challenging WM updating task may induce WM improvements that appear to reflect domain-general learning improvements. Yet, we also observed great variability in tDCS response across individuals. In fact, tDCS may have impaired WM function in some individuals. Moreover, our findings also indicate that the effects of multiple sessions of tDCS are not linearly additive, and spill over effects from previous days may interact with the effects of additional tDCS in daily stimulation.

In this thesis, I also investigated the role of cortical excitability in determining individual tDCS response, and used MRS to quantify cortical excitability based on GABA and Glutamate levels in the prefrontal cortex. However, we found no evidence that baseline prefrontal cortical excitability levels (i.e., 3T-MRS measured Glutamate/GABA ratios) predicted the effects of anodal and cathodal prefrontal tDCS on WM performance. Yet, we also found that current 3T-MRS practices may be insensitive to successfully pick up individual differences in neurotransmitter concentrations that are relevant for behaviour, possibly due to the large brain area over which concentrations are averaged: in contrast to our expectations, prefrontal GABA and Glutamate concentrations did not predict individual differences in WM performance, nor did occipital neurotransmitter levels predict visual discrimination performance.

Increasing the spatial accuracy of both tDCS and MRS methods (e.g., by reverting to HD-tDCS and 7T-MRS) is a crucial step for future studies that aim to examine the neurochemical mechanisms that underlie the effects of prefrontal tDCS on WM functioning. Moreover, a systematic investigation of the optimal parameter setting for tDCS to induce lasting WM improvements is essential if tDCS is ever to be used to enhance WM functioning in everyday life or clinical settings in the future.

Furthermore, while developing this ‘best recipe for success’, it will be crucial to keep an eye on individual differences in tDCS response. In particular, potential interactions of tDCS stimulation with the baseline situation should be taken into account. For this, it is

important to concurrently advance our understanding of the mechanisms underlying the effects of tDCS on synaptic-dependent learning mechanisms and behaviour. Only if we succeed to develop tDCS protocols that optimize WM functioning in everyone, may tDCS transcend from a scientific research method to a safe and effective method to counter impaired cognitive functioning in clinical populations.

In the introduction of this thesis, I illustrated the importance of WM in our daily life by describing its role in a situation where you need to make a decision about where to buy a cup of coffee at the train station before you board your train. In making this decision, it is crucial to simultaneously keep all the available options in mind (both the ones that may be visible to you, as well as the ones that you remember) and relate this information to your internal goals and desires at that particular moment. Our WM plays a pivotal role in everyday situations like this, providing us with a mental whiteboard to temporarily store and manipulate information on until we have found the best and most satisfactory solution. The non-invasive brain stimulation method of tDCS may be a promising method to improve WM in people in which poor WM functioning may affect their everyday lives. However, at the moment, we are still very much in the initial stages of understanding how we may use tDCS to consistently and safely enhance WM functioning in everyone. Eventually, an ethical discussion will also be necessary to determine whether it is desirable to introduce methods such as tDCS into the world to optimize an already well functioning (i.e., healthy) brain, or whether these methods should be reserved for clinical populations in which brain functioning is impaired.

In any case, in the end, application should adhere to strong regulatory standards and presently care should be taken with uncontrolled commercial uses of tDCS devices (Steenbergen et al., 2015). Little is currently known about the potential negative side effects of improving functioning in one brain region or network. Considering the delicate balance of activation of different brain regions both within and between brain networks, tDCS-induced excitation in one network (improving its functioning) may actually be accompanied by inhibition of a different network (possibly resulting in decreased functioning (Wokke et al., 2015)). More research is necessary to gain a better idea of the possible costs or side effects of enhancing functioning of one brain area on functioning of other cortical networks and related functional capacities.

As of yet, for these reasons, I would not recommend widespread application of low-current tDCS stimulation to enhance WM outside of research contexts. However, ultimately, protocols in which prefrontal tDCS is paired with WM training may become a helpful tool to lastingly and generally enhance WM functioning in clinical populations with WM problems.

References

Alam, M., Truong, D. Q., Khadka, N., & Bikson, M. (2016). Spatial and polarity precision of concentric high-definition transcranial direct current stimulation (HD-tDCS). *Physics in Medicine & Biology*, *61*(12), 4506. https://doi.org/10.1088/0031-9155/61/12/4506

Altamura et al. (2007). Dissociating the effects of Sternberg working memory demands in prefrontal cortex. *Psychiatry Research: Neuroimaging*, *154*(2), 103–114. https://doi.org/10.1016/j.pscychresns.2006.08.002

Andrews, S. C., Hoy, K. E., Enticott, P. G., Daskalakis, Z. J., & Fitzgerald, P. B. (2011). Improving working memory: the effect of combining cognitive activity and anodal transcranial direct current stimulation to the left dorsolateral prefrontal cortex. *Brain Stimulation*, *4*(2), 84–89. https://doi.org/10.1016/j.brs.2010.06.004

Antal, A., Keeser, D., Priori, A., Padberg, F., & Nitsche, M. A. (2015). Conceptual and Procedural Shortcomings of the Systematic Review "Evidence That Transcranial Direct Current Stimulation (tDCS) Generates Little-to-no Reliable Neurophysiologic Effect Beyond MEP Amplitude Modulation in Healthy Human Subjects: A Systematic Review" by Horvath and Co-workers. *Brain Stimulation*, *8*(4), 846–849. https://doi.org/10.1016/j.brs.2015.05.010

Au, J., Katz, B., Buschkuehl, M., Bunarjo, K., Senger, T., Zabel, C., … Jonides, J. (2016). Enhancing Working Memory Training with Transcranial Direct Current Stimulation. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn_a_00979

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: a meta-analysis, *22*(2), 366–377. https://doi.org/10.3758/s13423-014-0699-x

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829–839. https://doi.org/10.1038/nrn1201

Baddeley, A., Sala, S. D., Robbins, T. W., & Baddeley, A. (1996). Working Memory and Executive Control [and Discussion]. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *351*(1346), 1397–1404. https://doi.org/10.1098/rstb.1996.0123

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral Prefrontal Contributions to Human Working Memory. *Cortex; a Journal Devoted to the Study of the Nervous System and Behaviour*, *49*(5), 1195–1205. https://doi.org/10.1016/j.cortex.2012.05.022

Barch, D. M., & Ceaser, A. (2012). Cognition in Schizophrenia: Core Psychological and Neural Mechanisms. *Trends in Cognitive Sciences*, *16*(1). https://doi.org/10.1016/j.tics.2011.11.015

Bennabi, D., Pedron, S., Haffen, E., Monnin, J., Peterschmitt, Y., & Waes, V. V. (2014). Transcranial direct current stimulation for memory enhancement: from clinical research to animal models. *Frontiers in Systems Neuroscience*, *8*. https://doi.org/10.3389/fnsys.2014.00159

Berryhill, M. E., & Jones, K. T. (2012). tDCS selectively improves working memory in older adults with more education. *Neuroscience Letters*, *521*(2), 148–151. https://doi.org/10.1016/j.neulet.2012.05.074

Boggio, P. S., Ferrucci, R., Rigonatti, S. P., Covre, P., Nitsche, M., Pascual-Leone, A., & Fregni, F. (2006). Effects of transcranial direct current stimulation on working memory in patients with Parkinson's disease. *Journal of the Neurological Sciences*, *249*(1), 31–38. https://doi.org/10.1016/j.jns.2006.05.062

Bogner, W., Gruber, S., Doelken, M., Stadlbauer, A., Ganslandt, O., Boettcher, U., … Hammen, T. (2010). In vivo quantification of intracerebral GABA by single-voxel 1H-MRS—How reproducible are the results? *European Journal of Radiology*, *73*(3), 526–531. https://doi.org/10.1016/j.ejrad.2009.01.014

Bor, D., Duncan, J., Wiseman, R. J., & Owen, A. M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron*, *37*(2), 361–367.

Brunoni, A. R., & Vanderhasselt, M.-A. (2014). Working memory improvement with non-invasive brain stimulation of the dorsolateral prefrontal cortex: A systematic review and meta-analysis. *Brain and Cognition*, *86*, 1–9. https://doi.org/10.1016/j.bandc.2014.01.008

Buonomano & Merzenich, 1998. (n.d.). CORTICAL PLASTICITY: From Synapses to Maps | Annual Review of Neuroscience. Retrieved February 7, 2018, from http://www.annualreviews.org/doi/full/10.1146/annurev.neuro.21.1.149?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub%3Dpubmed

Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin & Review*, *17*(2), 193–199. https://doi.org/10.3758/PBR.17.2.193

Choi et al. (n.d.). Improvement of resolution for brain coupled metabolites by optimized 1H MRS at 7 T [Wiley Online Library]. Retrieved November 29, 2017, from http://onlinelibrary.wiley.com/doi/10.1002/nbm.1529/full

Clark, V. P., Coffman, B. A., Trumbo, M. C., & Gasparovic, C. (2011). Transcranial direct current stimulation (tDCS) produces localized and specific alterations in neurochemistry: A 1H magnetic resonance spectroscopy study. *Neuroscience Letters*, *500*(1), 67–71. https://doi.org/10.1016/j.neulet.2011.05.244

Cohen Kadosh, R., Soskic, S., Iuculano, T., Kanai, R., & Walsh, V. (2010). Modulating Neuronal Activity Produces Specific and Long-Lasting Changes in Numerical Competence. *Current Biology*, *20*(22), 2016–2020. https://doi.org/10.1016/j.cub.2010.10.007

Crabtree, G. W., Sun, Z., Kvajo, M., Broek, J. A. C., Fénelon, K., McKellar, H., … Gogos, J. A. (2017). Alteration of Neuronal Excitability and Short-Term Synaptic Plasticity in the Prefrontal Cortex of a Mouse Model of Mental Illness. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(15), 4158–4180. https://doi.org/10.1523/JNEUROSCI.4345-15.2017

De Bondt, T., De Belder, F., Vanhevel, F., Jacquemyn, Y., & Parizel, P. M. (2015). Prefrontal GABA concentration changes in women—Influence of menstrual cycle phase, hormonal contraceptive use, and correlation with premenstrual symptoms. *Brain Research*, *1597*, 129–138. https://doi.org/10.1016/j.brainres.2014.11.051

Dedoncker, J., Brunoni, A. R., Baeken, C., & Vanderhasselt, M.-A. (2016). A Systematic Review and Meta-Analysis of the Effects of Transcranial Direct Current Stimulation (tDCS) Over the Dorsolateral Prefrontal Cortex in Healthy and Neuropsychiatric Samples: Influence of Stimulation Parameters. *Brain Stimulation*, *9*(4), 501–517. https://doi.org/10.1016/j.brs.2016.04.006

den Uyl, T. E., Gladwin, T. E., Rinck, M., Lindenmeyer, J., & Wiers, R. W. (2017). A clinical trial with combined transcranial direct current stimulation and alcohol approach bias retraining. *Addiction Biology*, *22*(6), 1632–1640. https://doi.org/10.1111/adb.12463

Ditye, T., Jacobson, L., Walsh, V., & Lavidor, M. (2012). Modulating behavioural inhibition by tDCS combined with cognitive training. *Experimental Brain Research*, *219*(3), 363–368. https://doi.org/10.1007/s00221-012-3098-4

Dockery, C. A., Hueckel-Weng, R., Birbaumer, N., & Plewnia, C. (2009). Enhancement of Planning Ability by Transcranial Direct Current Stimulation. *The Journal of Neuroscience*, *29*(22), 7271–7277. https://doi.org/10.1523/JNEUROSCI.0065-09.2009

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Neuroplasticity: changes in grey matter induced by training. *Nature*, *427*(6972), 311–312. https://doi.org/10.1038/427311a

Draganski, B., & May, A. (2008). Training-induced structural changes in the adult human brain. *Behavioural Brain Research*, *192*(1), 137–142. https://doi.org/10.1016/j.bbr.2008.02.015

Driemeyer, J., Boyke, J., Gaser, C., Büchel, C., & May, A. (2008). Changes in Gray Matter Induced by Learning—Revisited. *PLOS ONE*, *3*(7), e2669. https://doi.org/10.1371/journal.pone.0002669

Duncan, J. (2013). The Structure of Cognition: Attentional Episodes in Mind and Brain. *Neuron*, *80*(1), 35. https://doi.org/10.1016/j.neuron.2013.09.015

Duncan, N. W., Wiebking, C., & Northoff, G. (2014). Associations of regional GABA and glutamate with intrinsic and extrinsic neural activity in humans—A review of multimodal imaging studies.

Neuroscience & Biobehavioural Reviews, *47*, 36–52. https://doi.org/10.1016/j.neubiorev.2014.07.016

Edden, R. A. E., Muthukumaraswamy, S. D., Freeman, T. C. A., & Singh, K. D. (2009). Orientation Discrimination Performance Is Predicted by GABA Concentration and Gamma Oscillation Frequency in Human Primary Visual Cortex. *The Journal of Neuroscience*, *29*(50), 15721–15726. https://doi.org/10.1523/JNEUROSCI.4426-09.2009

Edden, R. A. E., Puts, N. A. J., & Barker, P. B. (2012). Macromolecule-suppressed GABA-edited magnetic resonance spectroscopy at 3T. *Magnetic Resonance in Medicine*, *68*(3), 657–661. https://doi.org/10.1002/mrm.24391

Engle, R. W. (2002). Working Memory Capacity as Executive Attention. *Current Directions in Psychological Science*, *11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Floyer-Lea, A., Wylezinska, M., Kincses, T., & Matthews, P. M. (2006). Rapid Modulation of GABA Concentration in Human Sensorimotor Cortex During Motor Learning. *Journal of Neurophysiology*, *95*(3), 1639–1644. https://doi.org/10.1152/jn.00346.2005

Fregni, F., Boggio, P. S., Nitsche, M., Bermpohl, F., Antal, A., Feredoes, E., … Pascual-Leone, A. (2005a). Anodal transcranial direct current stimulation of prefrontal cortex enhances working memory. *Experimental Brain Research*, *166*(1), 23–30. https://doi.org/10.1007/s00221-005-2334-6

Fregni, F., Boggio, P. S., Nitsche, M., Bermpohl, F., Antal, A., Feredoes, E., … Pascual-Leone, A. (2005b). Anodal transcranial direct current stimulation of prefrontal cortex enhances working memory. *Experimental Brain Research*, *166*(1), 23–30. https://doi.org/10.1007/s00221-005-2334-6

Gill, J., Shah-Basak, P. P., & Hamilton, R. (2015). It's the thought that counts: examining the task-dependent effects of transcranial direct current stimulation on executive function. *Brain Stimulation*, *8*(2), 253–259. https://doi.org/10.1016/j.brs.2014.10.018

Goldsworthy, M. R., Pitcher, J. B., & Ridding, M. C. (2015). Spaced Noninvasive Brain Stimulation: Prospects for Inducing Long-Lasting Human Cortical Plasticity. *Neurorehabilitation and Neural Repair*, *29*(8), 714–721. https://doi.org/10.1177/1545968314562649

Greenhouse, I., Noah, S., Maddock, R. J., & Ivry, R. B. (2016). Individual differences in GABA content are reliable but are not uniform across the human cortex. *NeuroImage*, *139*, 1–7. https://doi.org/10.1016/j.neuroimage.2016.06.007

Harada, M., Kubo, H., Nose, A., Nishitani, H., & Matsuda, T. (2011a). Measurement of variation in the human cerebral GABA level by in vivo MEGA-editing proton MR spectroscopy using a clinical 3 T instrument and its dependence on brain region and the female menstrual cycle. *Human Brain Mapping*, *32*(5), 828–833. https://doi.org/10.1002/hbm.21086

Harada, M., Kubo, H., Nose, A., Nishitani, H., & Matsuda, T. (2011b). Measurement of variation in the human cerebral GABA level by in vivo MEGA-editing proton MR spectroscopy using a clinical 3 T instrument and its dependence on brain region and the female menstrual cycle. *Human Brain Mapping*, *32*(5), 828–833. https://doi.org/10.1002/hbm.21086

Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working Memory Training May Increase Working Memory Capacity but Not Fluid Intelligence. *Psychological Science*, 0956797613492984. https://doi.org/10.1177/0956797613492984

Hashemirad, F., Zoghi, M., Fitzgerald, P. B., & Jaberzadeh, S. (2016). The effect of anodal transcranial direct current stimulation on motor sequence learning in healthy individuals: A systematic review and meta-analysis. *Brain and Cognition*, *102*, 1–12. https://doi.org/10.1016/j.bandc.2015.11.005

Hill, A. T., Fitzgerald, P. B., & Hoy, K. E. (2016). Effects of Anodal Transcranial Direct Current Stimulation on Working Memory: A Systematic Review and Meta-Analysis of Findings From Healthy and Neuropsychiatric Populations. *Brain Stimulation*, *9*(2), 197–208. https://doi.org/10.1016/j.brs.2015.10.006

Hill, A. T., Rogasch, N. C., Fitzgerald, P. B., & Hoy, K. E. (2017). Effects of prefrontal bipolar and high-definition transcranial direct current stimulation on cortical reactivity and working memory in healthy adults. *NeuroImage*, *152*, 142–157. https://doi.org/10.1016/j.neuroimage.2017.03.001

Horvath, J. C., Carter, O., & Forte, J. D. (n.d.). No Significant Effect of Transcranial Direct Current Stimulation (tDCS) found on Simple Motor Reaction Time comparing 15 Different Simulation Protocols. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2016.09.017

Hoy, K. E., Emonson, M. R. L., Arnold, S. L., Thomson, R. H., Daskalakis, Z. J., & Fitzgerald, P. B. (2013). Testing the limits: Investigating the effect of tDCS dose on working memory enhancement in healthy controls. *Neuropsychologia*, *51*(9), 1777–1784. https://doi.org/10.1016/j.neuropsychologia.2013.05.018

Isaacson, J. S., & Scanziani, M. (2011). How Inhibition Shapes Cortical Activity. *Neuron*, *72*(2), 231–243. https://doi.org/10.1016/j.neuron.2011.09.027

Iwabuchi, S. J., Raschke, F., Auer, D. P., Liddle, P. F., Lankappa, S. T., & Palaniyappan, L. (2017). Targeted transcranial theta-burst stimulation alters fronto-insular network and prefrontal GABA. *NeuroImage*, *146*, 395–403. https://doi.org/10.1016/j.neuroimage.2016.09.043

Jacobson, L., Koslowsky, M., & Lavidor, M. (2011). tDCS polarity effects in motor and cognitive domains: a meta-analytical review. *Experimental Brain Research*, *216*(1), 1–10. https://doi.org/10.1007/s00221-011-2891-9

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.0801268105

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18*(4), 394–412. https://doi.org/10.1080/09658211003702171

Jansen, J. F. A., Backes, W. H., Nicolay, K., & Kooi, M. E. (2006). 1H MR Spectroscopy of the Brain: Absolute Quantification of Metabolites. *Radiology*, *240*(2), 318–332. https://doi.org/10.1148/radiol.2402050314

Jansma, J. M., van Raalten, T. R., Boessen, R., Neggers, S. F. W., Jacobs, R. H. A. H., Kahn, R. S., & Ramsey, N. F. (2013). fMRI Guided rTMS Evidence for Reduced Left Prefrontal Involvement after Task Practice. *PLoS ONE*, *8*(12). https://doi.org/10.1371/journal.pone.0080256

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*(4), 637–671. https://doi.org/10.3758/BF03196323

Karabanov, A., Ziemann, U., Hamada, M., George, M. S., Quartarone, A., Classen, J., … Siebner, H. R. (2015). Consensus Paper: Probing Homeostatic Plasticity of Human Cortex With Non-invasive Transcranial Brain Stimulation. *Brain Stimulation*, *8*(3), 442–454. https://doi.org/10.1016/j.brs.2015.01.404

Keeser, D., Meindl, T., Bor, J., Palm, U., Pogarell, O., Mulert, C., … Padberg, F. (2011). Prefrontal transcranial direct current stimulation changes connectivity of resting-state networks during fMRI. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(43), 15284–15293. https://doi.org/10.1523/JNEUROSCI.0542-11.2011

Kihara, K., Kondo, H. M., & Kawahara, J. I. (2016). Differential Contributions of GABA Concentration in Frontal and Parietal Regions to Individual Differences in Attentional Blink. *Journal of Neuroscience*, *36*(34), 8895–8901. https://doi.org/10.1523/JNEUROSCI.0764-16.2016

Kim, J.-H., Kim, D.-W., Chang, W. H., Kim, Y.-H., Kim, K., & Im, C.-H. (2014). Inconsistent outcomes of transcranial direct current stimulation may originate from anatomical differences among individuals: Electric field simulation using individual MRI data. *Neuroscience Letters*, *564*, 6–10. https://doi.org/10.1016/j.neulet.2014.01.054

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, *14*(7), 317–324. https://doi.org/10.1016/j.tics.2010.05.002

Krause, B., & Cohen Kadosh, R. (2014a). Not all brains are created equal: the relevance of individual differences in responsiveness to transcranial electrical stimulation. *Frontiers in Systems Neuroscience*, *8*, 25. https://doi.org/10.3389/fnsys.2014.00025

Krause, B., & Cohen Kadosh, R. (2014b). Not all brains are created equal: the relevance of individual differences in responsiveness to transcranial electrical stimulation. *Frontiers in Systems Neuroscience*, *8*. https://doi.org/10.3389/fnsys.2014.00025

Krause, B., Márquez-Ruiz, J., & Cohen Kadosh, R. (2013). The effect of transcranial direct current stimulation: a role for cortical excitation/inhibition balance? *Frontiers in Human Neuroscience*, *7*, 602. https://doi.org/10.3389/fnhum.2013.00602

Krause, B., Márquez-Ruiz, J., & Kadosh, R. C. (2013). The effect of transcranial direct current stimulation: a role for cortical excitation/inhibition balance? *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00602

Kuo, H.-I., Bikson, M., Datta, A., Minhas, P., Paulus, W., Kuo, M.-F., & Nitsche, M. A. (2013). Comparing Cortical Plasticity Induced by Conventional and High-Definition 4 × 1 Ring tDCS: A Neurophysiological Study. *Brain Stimulation*, *6*(4), 644–648. https://doi.org/10.1016/j.brs.2012.09.010

Kuo, M.-F., & Nitsche, M. A. (2012). Effects of Transcranial Electrical Stimulation on Cognition. *Clinical EEG and Neuroscience*, *43*(3), 192–199. https://doi.org/10.1177/1550059412444975

Lally, N., Nord, C. L., Walsh, V., & Roiser, J. P. (2013). Does excitatory fronto-extracerebral tDCS lead to improved working memory performance? *F1000Research*, *2*. https://doi.org/10.12688/f1000research.2-219.v2

Learmonth, G., Thut, G., Benwell, C. S. Y., & Harvey, M. (2015). The implications of state-dependent tDCS effects in aging: Behavioural response is determined by baseline performance. *Neuropsychologia*, *74*, 108–119. https://doi.org/10.1016/j.neuropsychologia.2015.01.037

Liebetanz, D., Nitsche, M. A., Tergau, F., & Paulus, W. (2002). Pharmacological approach to the mechanisms of transcranial DC-stimulation-induced after-effects of human motor cortex excitability. *Brain: A Journal of Neurology*, *125*(Pt 10), 2238–2247.

London, R. E., & Slagter, H. A. (2015). Effects of Transcranial Direct Current Stimulation over Left Dorsolateral pFC on the Attentional Blink Depend on Individual Baseline Performance. *Journal of Cognitive Neuroscience*, *27*(12), 2382–2393. https://doi.org/10.1162/jocn_a_00867

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the Role of the Dorsolateral Prefrontal and Anterior Cingulate Cortex in Cognitive Control. *Science*, *288*(5472), 1835–1838. https://doi.org/10.1126/science.288.5472.1835

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(8), 4398–4403.

Mancuso, L. E., Ilieva, I. P., Hamilton, R. H., & Farah, M. J. (2016). Does Transcranial Direct Current Stimulation Improve Healthy Working Memory?: A Meta-analytic Review. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn_a_00956

Martin, D. M., Liu, R., Alonzo, A., Green, M., & Loo, C. K. (2014). Use of transcranial direct current stimulation (tDCS) to enhance cognitive training: effect of timing of stimulation. *Experimental Brain Research*, *232*(10), 3345–3351. https://doi.org/10.1007/s00221-014-4022-x

Martin, D. M., Liu, R., Alonzo, A., Green, M., Player, M. J., Sachdev, P., & Loo, C. K. (2013a). Can transcranial direct current stimulation enhance outcomes from cognitive training? A randomized controlled trial in healthy participants. *International Journal of Neuropsychopharmacology*, *16*(9), 1927–1936. https://doi.org/10.1017/S1461145713000539

Martin, D. M., Liu, R., Alonzo, A., Green, M., Player, M. J., Sachdev, P., & Loo, C. K. (2013b). Can transcranial direct current stimulation enhance outcomes from cognitive training? A randomized controlled trial in healthy participants. *The International Journal of Neuropsychopharmacology / Official Scientific Journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)*, *16*(9), 1927–1936. https://doi.org/10.1017/S1461145713000539

Meiron, O., & Lavidor, M. (2013). Unilateral Prefrontal Direct Current Stimulation Effects are Modulated by Working Memory Load and Gender. *Brain Stimulation*, *6*(3), 440–447. https://doi.org/10.1016/j.brs.2012.05.014

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working Memory Training Does Not Improve Performance on Measures of Intelligence or Other Measures of "Far Transfer": Evidence From a Meta-Analytic Review. *Perspectives on Psychological Science*, *11*(4), 512–534. https://doi.org/10.1177/1745691616635612

Michels, L., Martin, E., Klaver, P., Edden, R., Zelaya, F., Lythgoe, D. J., … O'Gorman, R. L. (2012). Frontal GABA Levels Change during Working Memory. *PLOS ONE*, *7*(4), e31933. https://doi.org/10.1371/journal.pone.0031933

Miller, E. K., & Cohen, J. D. (2003, November 28). AN INTEGRATIVE THEORY OF PREFRONTAL CORTEX FUNCTION [review-article]. Retrieved April 11, 2017, from http://www.annualreviews.org/doi/10.1146/annurev.neuro.24.1.167

Möller, A., Nemmi, F., Karlsson, K., & Klingberg, T. (2017). Transcranial Electric Stimulation Can Impair Gains during Working Memory Training and Affects the Resting State Connectivity. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00364

Monte-Silva, K., Kuo, M.-F., Hessenthaler, S., Fresnoza, S., Liebetanz, D., Paulus, W., & Nitsche, M. A. (2013). Induction of Late LTP-Like Plasticity in the Human Motor Cortex by Repeated Non-Invasive Brain Stimulation. *Brain Stimulation*, *6*(3), 424–432. https://doi.org/10.1016/j.brs.2012.04.011

Mulquiney, P. G., Hoy, K. E., Daskalakis, Z. J., & Fitzgerald, P. B. (2011). Improving working memory: Exploring the effect of transcranial random noise stimulation and transcranial direct current stimulation on the dorsolateral prefrontal cortex. *Clinical Neurophysiology*, *122*(12), 2384–2389. https://doi.org/10.1016/j.clinph.2011.05.009

Nilsson, J., Lebedev, A. V., Rydström, A., & Lövdén, M. (2017). Direct-Current Stimulation Does Little to Improve the Outcome of Working Memory Training in Older Adults. *Psychological Science*. https://doi.org/10.1177/0956797617698139

Nitsche, M. A., Cohen, L. G., Wassermann, E. M., Priori, A., Lang, N., Antal, A., … Pascual-Leone, A. (2008). Transcranial direct current stimulation: State of the art 2008. *Brain Stimulation*, *1*(3), 206–223. https://doi.org/10.1016/j.brs.2008.06.004

Nitsche, M. A., Cohen, L. G., Wassermann, E. M., Priori, A., Lang, N., Antal, A., … Pascual-Leone, A. (2008). Transcranial direct current stimulation: State of the art 2008. *Brain Stimulation*, *1*(3), 206–223. https://doi.org/10.1016/j.brs.2008.06.004

Nitsche, M. A., Fricke, K., Henschke, U., Schlitterlau, A., Liebetanz, D., Lang, N., … Paulus, W. (2003). Pharmacological Modulation of Cortical Excitability Shifts Induced by Transcranial Direct Current Stimulation in Humans. *The Journal of Physiology*, *553*(1), 293–301. https://doi.org/10.1113/jphysiol.2003.049916

Nitsche, M. A., & Paulus, W. (2000). Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. *The Journal of Physiology*, *527*(3), 633–639. https://doi.org/10.1111/j.1469-7793.2000.t01-1-00633.x

Ohn, S. H., Park, C.-I., Yoo, W.-K., Ko, M.-H., Choi, K. P., Kim, G.-M., … Kim, Y.-H. (2008). Time-dependent effect of transcranial direct current stimulation on the enhancement of working memory. *Neuroreport*, *19*(1), 43–47. https://doi.org/10.1097/WNR.0b013e3282f2adfd

Opitz, A., Paulus, W., Will, S., Antunes, A., & Thielscher, A. (2015). Determinants of the electric field during transcranial direct current stimulation. *NeuroImage*, *109*, 140–150. https://doi.org/10.1016/j.neuroimage.2015.01.033

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, *25*(1), 46–59. https://doi.org/10.1002/hbm.20131

Park, D. C., Hedden, T., Davidson, N. S., Lautenschlager, G., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the life span. *Psychology and Aging*, 299–320.

Passow, S., Thurm, F., & Li, S.-C. (2017). Activating Developmental Reserve Capacity Via Cognitive Training or Non-invasive Brain Stimulation: Potentials for Promoting Fronto-Parietal and Hippocampal-Striatal Network Functions in Old Age. *Frontiers in Aging Neuroscience*, *9*. https://doi.org/10.3389/fnagi.2017.00033

Petrides, M. (2000). Dissociable Roles of Mid-Dorsolateral Prefrontal and Anterior Inferotemporal Cortex in Visual Working Memory. *Journal of Neuroscience*, *20*(19), 7496–7503.

Petroff, O. A. C. (2002). Book Review: GABA and Glutamate in the Human Brain. *The Neuroscientist*, *8*(6), 562–573. https://doi.org/10.1177/1073858402238515

Plow, E. B., Obretenova, S. N., Fregni, F., Pascual-Leone, A., & Merabet, L. B. (2012). Comparison of Visual Field Training for Hemianopia With Active Versus Sham Transcranial Direct Cortical Stimulation. *Neurorehabilitation and Neural Repair*, *26*(6), 616–626. https://doi.org/10.1177/1545968311431963

Puts, N. A. J., Edden, R. A. E., Evans, C. J., McGlone, F., & McGonigle, D. J. (2011). Regionally Specific Human GABA Concentration Correlates with Tactile Discrimination Thresholds. *The Journal of Neuroscience*, *31*(46), 16556–16560. https://doi.org/10.1523/JNEUROSCI.4489-11.2011

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., … Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology: General*, *142*(2), 359. https://doi.org/10.1037/a0029082

Richmond, L. L., Wolk, D., Chein, J., & Olson, I. R. (2014a). Transcranial Direct Current Stimulation Enhances Verbal Working Memory Training Performance over Time and Near Transfer Outcomes. *Journal of Cognitive Neuroscience*, *26*(11), 2443–2454. https://doi.org/10.1162/jocn_a_00657

Richmond, L. L., Wolk, D., Chein, J., & Olson, I. R. (2014b). Transcranial direct current stimulation enhances verbal working memory training performance over time and near transfer outcomes. *Journal of Cognitive Neuroscience*, *26*(11), 2443–2454. https://doi.org/10.1162/jocn_a_00657

Sandberg, K., Blicher, J. U., Dong, M. Y., Rees, G., Near, J., & Kanai, R. (2014). Occipital GABA correlates with cognitive failures in daily life. *NeuroImage*, *87*, 55–60. https://doi.org/10.1016/j.neuroimage.2013.10.059

Shibata, K., Sasaki, Y., Bang, J. W., Walsh, E. G., Machizawa, M. G., Tamaki, M., … Watanabe, T. (2017). Overlearning hyperstabilizes a skill by rapidly making neurochemical processing inhibitory-dominant. *Nature Neuroscience*. https://doi.org/10.1038/nn.4490

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*(4), 628–654. https://doi.org/10.1037/a0027473

Siebner, H. R., Lang, N., Rizzo, V., Nitsche, M. A., Paulus, W., Lemon, R. N., & Rothwell, J. C. (2004). Preconditioning of low-frequency repetitive transcranial magnetic stimulation with transcranial direct current stimulation: evidence for homeostatic plasticity in the human motor cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *24*(13), 3379–3385. https://doi.org/10.1523/JNEUROSCI.5316-03.2004

Slagter, H. A. (2012). Conventional working memory training may not improve intelligence. *Trends in Cognitive Sciences*, *16*(12), 582–583. https://doi.org/10.1016/j.tics.2012.10.001

Smith, E. E., & Jonides, J. (1999). Storage and Executive Processes in the Frontal Lobes. *Science*, *283*(5408), 1657–1661. https://doi.org/10.1126/science.283.5408.1657

Song, C., Sandberg, K., Andersen, L. M., Blicher, J. U., & Rees, G. (2017). Human Occipital and Parietal GABA Selectively Influence Visual Perception of Orientation and Size. *Journal of Neuroscience*, *37*(37), 8929–8937. https://doi.org/10.1523/JNEUROSCI.3945-16.2017

Stagg, C. J., Best, J. G., Stephenson, M. C., O'Shea, J., Wylezinska, M., Kincses, Z. T., … Johansen-Berg, H. (2009). Polarity-Sensitive Modulation of Cortical Neurotransmitters by Transcranial Stimulation. *The Journal of Neuroscience*, *29*(16), 5202–5206. https://doi.org/10.1523/JNEUROSCI.4432-08.2009

Stagg, C. J., Lin, R. L., Mezue, M., Segerdahl, A., Kong, Y., Xie, J., & Tracey, I. (2013). Widespread modulation of cerebral perfusion induced during and after transcranial direct current stimulation applied to the left dorsolateral prefrontal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(28), 11425–11431. https://doi.org/10.1523/JNEUROSCI.3887-12.2013

Steenbergen, L., Sellaro, R., Hommel, B., Lindenberger, U., Kühn, S., & Colzato, L. S. (2015). "Unfocus" on foc.us: commercial tDCS headset impairs working memory. *Experimental Brain Research*, *234*(3), 637–643. https://doi.org/10.1007/s00221-015-4391-9

Stephens, J. A., & Berryhill, M. E. (2016). Older Adults Improve on Everyday Tasks after Working Memory Training and Neurostimulation. *Brain Stimulation*, *9*(4), 553–559. https://doi.org/10.1016/j.brs.2016.04.001

Talsma, L. J., Kroese, H. A., & Slagter, H. A. (2016). Boosting Cognition: Effects of Multiple Session Transcranial Direct Current Stimulation on Working Memory. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn_a_01077

Teo, F., Hoy, K. E., Daskalakis, Z. J., & Fitzgerald, P. B. (2011). Investigating the Role of Current Strength in tDCS Modulation of Working Memory Performance in Healthy Controls. *Frontiers in Psychiatry*, *2*. https://doi.org/10.3389/fpsyt.2011.00045

Thayer, R. E. (1978). Factor Analytic and Reliability Studies on the Activation-Deactivation Adjective Check List. *Psychological Reports*, *42*(3), 747–756. https://doi.org/10.2466/pr0.1978.42.3.747

Truong, D. Q., Magerowski, G., Blackburn, G. L., Bikson, M., & Alonso-Alonso, M. (2013). Computational modelling of transcranial direct current stimulation (tDCS) in obesity: Impact of head fat and dose guidelines. *NeuroImage : Clinical*, *2*, 759–766. https://doi.org/10.1016/j.nicl.2013.05.011

Turrigiano & Nelson, 2000. (n.d.). Hebb and homeostasis in neuronal plasticity - ScienceDirect. Retrieved February 6, 2018, from https://www.sciencedirect.com/science/article/pii/S095943880000091X?via%3Dihub

Turrigiano, G. G., & Nelson, S. B. (2004). Homeostatic plasticity in the developing nervous system. *Nature Reviews Neuroscience*, *5*(2), 97. https://doi.org/10.1038/nrn1327

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behaviour Research Methods*, *37*(3), 498–505.

van Loon, A. M., Knapen, T., Scholte, H. S., St. John-Saaltink, E., Donner, T. H., & Lamme, V. A. F. (2013). GABA Shapes the Dynamics of Bistable Perception. *Current Biology*, *23*(9), 823–827. https://doi.org/10.1016/j.cub.2013.03.067

Vöröslakos, M., Takeuchi, Y., Brinyiczki, K., Zombori, T., Oliva, A., Fernández-Ruiz, A., … Berényi, A. (2018). Direct effects of transcranial electric stimulation on brain circuits in rats and humans. *Nature Communications*, *9*. https://doi.org/10.1038/s41467-018-02928-3

Waddell, K. W., Avison, M. J., Joers, J. M., & Gore, J. C. (2007). A Practical Guide to Robust Detection of GABA in Human Brain by J-difference Spectroscopy at 3 Tesla Using a Standard Volume Coil. *Magnetic Resonance Imaging*, *25*(7), 1032. https://doi.org/10.1016/j.mri.2006.11.026

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., … Morey, R. D. (2017). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, 1–19. https://doi.org/10.3758/s13423-017-1323-7

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., … Morey, R. D. (2017). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, 1–23. https://doi.org/10.3758/s13423-017-1343-3

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cognitive, Affective & Behavioural Neuroscience*, *3*(4), 255–274.

Wokke, M. E., Talsma, L. J., & Vissers, M. E. (2015). Biasing neural network dynamics using non-invasive brain stimulation. *Frontiers in Systems Neuroscience*, *8*. https://doi.org/10.3389/fnsys.2014.00246

Yoon, J. H., Grandelis, A., & Maddock, R. J. (2016). Dorsolateral Prefrontal Cortex GABA Concentration in Humans Predicts Working Memory Load Processing Capacity. *Journal of Neuroscience*, *36*(46), 11788–11794. https://doi.org/10.1523/JNEUROSCI.1970-16.2016

Zaehle, T., Sandmann, P., Thorne, J. D., Jäncke, L., & Herrmann, C. S. (2011). Transcranial direct current stimulation of the prefrontal cortex modulates working memory performance: combined behavioural and electrophysiological evidence. *BMC Neuroscience*, *12*, 2. https://doi.org/10.1186/1471-2202-12-2

Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika*, *70*(1), 203–212. https://doi.org/10.1007/s11336-003-1119-8

Acknowledgements

Chapter 2

L.J.T. and H.A.S. designed the study. H.A.K. and L.J.T. collected the data. L.J.T. performed the data analysis and interpretation in consultation with H.A.S. L.J.T. drafted the manuscript and H.A.S. and H.A.K. provided critical revisions.

Chapter 3

All authors designed the study. L.J.T. collected the data. L.J.T. performed the data analysis and interpretation in consultation with A.vL., H.S.S. and H.A.S. L.J.T. drafted the manuscript and H.A.S., A.vL and H.S.S. provided critical revisions.

Chapter 4

L.J.T. and H.A.S. designed the study. J.A.B. and J.H. (in equal amount) and L.J.T collected the data. L.J.T. performed the data analysis and interpretation in consultation with H.A.S. L.J.T. drafted the manuscript and H.A.S., J.A.B. and J.H. provided critical revisions.

All studies reported in this thesis were supported by a VIDI grant from the Netherlands Organisation for Scientific Research (NOW), awarded to H.A. Slagter.

List of publications

Research presented in this thesis

Talsma, L.J., Kroese, H.A., Slagter, H.A. (2016) Boosting cognition: Effects of multiple session tDCS on working memory. *Journal of Cognitive Neuroscience, 29(4),* 755-768

Talsma, L.J., Van Loon, A., Scholte, H.S., Slagter, H.A., State or trait? MRS-measured GABA and Glutamate concentrations are not modulated by task demand and do not robustly predict task performance. Submitted Manuscript.

Talsma, L.J., Broekhuizen, J.A., Huisman, J., Slagter, H.A., No evidence that baseline prefrontal cortical excitability (3T-MRS) predicts the effects of prefrontal tDCS on WM performance. Submitted Manuscript.

Other publications

Reteig, L.C., Talsma, L.J., Van Schouwenburg, M.R., Slagter, H.A. (2017) Transcranial electrical stimulation as a tool to enhance attention. *J Cogn Enhanc, 1:10-25*.

Slagter, H.A., Vissers, M.E., Talsma, L.J., & Ridderinkhof, K.R. (2016). Cognitive enhancement: it's about time. *Cognitive Neuroscience, 8(2)*, 119-120.

Wokke, M.E., Talsma, L.J., & Vissers, M.E. (2014). Biasing neural network dynamics using non-invasive brain stimulation. *Frontiers in Systems Neuroscience, 8*, 246.

Nederlandse samenvatting

Ons werkgeheugen biedt een plek in ons brein waar we tijdelijk informatie kunnen opslaan, herorganiseren of simpelweg ‘actief’ houden bij het maken van complexe keuzes. Naast het toevoegen van nieuwe informatie, die we oppikken uit onze omgeving, kunnen we hierbij ook kennis uit het lange termijn geheugen aanroepen, alsmede onze persoonlijke doelen en motivaties voor de geest blijven houden.

Bijvoorbeeld wanneer we boodschappen doen en we tegelijkertijd zowel ons lijstje, de staat van onze portemonnee, maar ook bijvoorbeeld de persoonlijke voorkeuren van favoriete merken in ons hoofd moeten houden. Met behulp van ons werkgeheugen kunnen we deze stukjes informatie vervolgens combineren, wegen naar waarde, aanpassen en herorganiseren om uiteindelijk tot de ‘beste’ of meest bevredigende conclusie te komen. Een goed functionerend werkgeheugen is daarbij belangrijk voor veel alledaagse situaties en zorgt ervoor dat we doeltreffend onze weg kunnen vinden in het leven.

Echter, deze functie van ‘tijdelijke mentale opslagplaats’ werkt niet bij iedereen hetzelfde. Waar sommigen na een blik op hun boodschappenlijstje moeiteloos door de supermarkt navigeren, moeten anderen het papiertje na elk nieuw artikel in hun karretje opnieuw tevoorschijn toveren. Desondanks werkt het werkgeheugen over het algemeen in alle gezonde breinen goed genoeg dat iedereen uiteindelijk met de gewenste boodschappen de deur uit wandelt. Dit is alleen niet het geval in sommige klinische populaties, zoals schizofrenie. Mogelijk zorgt in deze groep een gebrekkig werkgeheugen ervoor dat informatie uit de externe omgeving niet effectief gebruikt kan worden om interne doelen te bereiken, wat belemmeringen veroorzaakt in het dagelijks functioneren.

In dit proefschrift heb ik onderzocht op welke manier een vorm van non-invasieve hersenstimulatie genaamd transcraniale Direct Current Stimulatie (tDCS), gebruikt zou kunnen worden om werkgeheugen functioneren te verbeteren in gezonde volwassenen. Bij tDCS worden twee elektrodes op specifieke plekken van het hoofd geplakt (een anode en een cathode). Hiertussen loopt een heel laag stroompje (vaak +/- 1 mA), waarvan het merendeel door de huid en schedel gaat en slechts een klein gedeelte de hersenen bereikt. Niettemin is uit onderzoek gebleken dat we met tDCS de werking van de hersenen onder de elektrodes kunnen beïnvloeden en een hersengebied tijdelijk meer (anodale) of minder (cathodale stimulatie) actief kunnen maken. Met tDCS hebben we dus een direct middel in handen om de werking van onze hersenen manipuleren en mogelijk het functioneren van een specifiek hersengebied verbeteren.

Traditioneel onderscheiden we twee vormen van werkgeheugen: het verbaal werkgeheugen dat we aanspreken bij het oproepen en onthouden van woorden, cijfers en letters, en het visueel-spatieel werkgeheugen voor figuren, vormen en locaties. Daarnaast is er een centraal executief orgaan dat deze beide beheert en verantwoordelijk wordt geacht voor het toevoegen, manipuleren en uitwissen van informatie in deze beide opslagplaatsen. Voor dit centrale orgaan lijkt de dorsolaterale prefrontale cortex (DLPFC) in het voorste gedeelte van de hersenen een belangrijke rol te spelen. Als we werkgeheugen in een bredere context willen verbeteren, waar we ook buiten het lab wat aan hebben, is het waarschijnlijk het effectiefst dit gedeelte van het werkgeheugen als doelwit te nemen.

Verscheidene eerdere onderzoeken hebben laten zien dat één sessie met anodale tDCS stimulatie over de linker DLPFC prestaties op een verbale werkgeheugentaak significant kan verbeteren. In navolging hiervan heb ik in *Hoofdstuk 2* van dit proefschrift gekeken wat er gebeurt als we niet één, maar drie keer stimuleren (steeds met een dag ertussen). Kunnen we met elke stimulatie sessie deze functie stapsgewijs blijven verbeteren? Ook heb ik in deze studie onderzocht of werkgeheugenprestatie alleen verbetert op de verbale werkgeheugentaak (met letters), die proefpersonen tijdens de stimulatie uitvoerden, of ook terug te zien is op een spatiële versie van dezelfde taak (waarbij locaties binnen een 3x3 raster dienden te worden te onthouden). Met andere woorden, kunnen we de gevonden verbeteringen ook gebruiken in andere situaties dan degene waarmee we geoefend hebben tijdens stimulatie?

Hoewel onze proefpersonen in *Hoofdstuk 2* in de eerste sessie een vergelijkbare verbetering lieten zien op de verbale werkgeheugentaak ten opzichte van een placebo stimulatie als eerdere studies, vonden we geen effect van anodale stimulatie in de tweede en derde sessie. Mogelijk is dit het gevolg van het feit dat de sessies te dicht op elkaar gepland waren en heeft het brein een bepaalde periode van rust nodig na elke stimulatie-sessie. Aan de andere kant is er op basis van recent onderzoek ook twijfel gerezen over de mogelijkheid om werkgeheugen functioneren te verbeteren na anodale tDCS over de linker DLPFC. Misschien werkt de stimulatie dus wel niet zo goed als we hoopten, of in elk geval niet voor iedereen. We vonden namelijk ook dat er grote variatie bestond tussen onze deelnemers in de verbetering die ze lieten zien na stimulatie in de eerste sessie. Bovendien bleek de prestatie van sommige van onze proefpersonen onverwacht juist te verslechteren.

Hoewel speculatief, gaven onze resultaten in dit hoofdstuk echter ook aanwijzingen dat indien effectief, tDCS verbeteringen tot stand kan brengen, die zich niet alleen tot het verbale werkgeheugen beperken, maar ook in de spatiële versie zichtbaar zijn. Dit impliceert deze verbeterde functie mogelijk in een verscheidenheid van situaties (ook buiten het lab) gebruikt kunnen worden. Uitgebreider onderzoek naar de effecten van tDCS op functionere in het dagelijks leven moet echter nog worden gedaan.

Ook heb ik in dit proefschrift verder onderzocht waar de individuele verschillen in het effect van tDCS vandaan zou kunnen komen. Hierbij heb ik specifiek gekeken naar de mogelijke rol van baseline corticale excitabiliteit (prikkelbaarheid). Hoe snel zenuwcellen vuren hangt grotendeels af van de balans tussen de neurotransmitters GABA (dat veelal een inhiberende werking heeft) en Glutamaat (veelal activerend). Met andere woorden, in een hersengebied met een hoge corticale excitabiliteit zijn de neuronen ‘actiever’; ze reageren sneller. Echter, *meer* is in dit geval niet altijd *beter*: een te hoge excitabiliteit in een bepaald hersengebied kan zorgen voor een verslechterde functie van het gedrag.

Onderzoek suggereert dat tDCS tijdelijk corticale excitabiliteit kan beïnvloeden en een effect heeft op de daarmee samenhangende concentraties van GABA en Glutamaat. Verschillen in de beginsituatie van corticale excitabiliteit kan daarom mogelijk verklaren waarom een en dezelfde stimulatie voor sommigen een positief effect, maar voor anderen juist een negatief effect kan hebben.

In *Hoofdstuk 4* heb ik de neuroimaging techniek Magnetic Resonance Spectroscopy (3T-MRS) gebruikt om de concentraties van GABA en Glutamaat te meten in de linker DLPFC. Daar heb ik vervolgens de ratio tussen berekend als maat van corticale excitabiliteit. In tegenstelling tot mijn hypothese, bleek ik hiermee de gedragseffecten van anodale (activerende) of cathodale (inhiberende) tDCS echter niet te kunnen voorspellen.

In het meer methodologische *Hoofdstuk 3* heb ik geprobeerd beter grip te krijgen op wat we precies meten met hedendaagse MRS-methodes. Weerspiegelen de neurotransmitter concentraties die we hiermee bepalen vooral stabiele eigenschappen van de individu (reflecteren ze 'trait') of veranderen ze als het gemeten gebied actief is of niet (meer 'state' of toestand gevoelig). Onze resultaten laten zien dat de gemeten concentraties onderling goed samenhangen en niet afhangen van of proefpersonen een relevante taak doen of niet; we meten dus vooral 'trait' en zijn niet gevoelig voor 'state'. In tegenstelling tot eerdere bevindingen bleken de gemeten neurotransmitter concentraties in onze studie echter niet te correleren met individuele verschillen in gerelateerde cognitieve taken. Prefrontale GABA concentraties waren niet voorspellend voor verbale werkgeheugen prestatie en occipitale GABA concentraties voorspelden niet hoe goed iemand het doet op een visuele perceptie taak.

Dit zaait twijfel over de validiteit van huidige (3T-)MRS methodes om de kleine verschillen in GABA en Glutamaat concentraties op te pikken die we terugzien in gedrag. Dit gebrek aan gevoeligheid in de metingen kan tevens de oorzaak zijn waarom in *Hoofdstuk 4* we geen correlatie vinden tussen baseline corticale excitabiliteit en individuele verschillen in het effect van tDCS op werkgeheugen.

De resultaten in dit proefschrift roepen meer nieuwe vragen op dan dat ze beantwoorden. Wat is de optimale manier om het beoogde hersengebied met tDCS te bereiken? Wat is de rol van rusttijd tussen stimulatiesessies op de effectiviteit van tDCS? Kunnen we als we langer wachten met een volgende stimulatiesessie verbeteringen meer stapsgewijs blijven opbouwen? Hoe beïnvloedt tDCS functioneren in het dagelijks leven? Als we de sensitiviteit van onze methode MRS verhogen, voorspelt corticale excitabiliteit dan wel het effect van stimulatie? Kan een stap naar een scanner met een hoger magnetisch veld en/of het gebruik van een nauwkeurigere localisatie van de functie in het brein hierbij helpen?

Momenteel staat het gebruik van tDCS als 'cognitieve booster' nog in de kinderschoenen. We zullen eerst een beter beeld moeten vormen van wat deze techniek precies doet, en hoe we kunnen zorgen dat iedereen uiteindelijk met een daadwerkelijk verbeterd brein de deur uitstapt, voordat deze op brede schaal gebruikt kan worden. Parallel hieraan zullen we ons bezig moeten houden met enkele ethische kwesties; moet tDCS voor iedereen beschikbaar worden of alleen voor klinische populaties met aanwijsbare problemen? Kunnen we het gezonde brein nog wel verder optimaliseren, of zullen ogenschijnlijke verbeteringen later een keerzijde blijken te hebben? Is het oneerlijk als een deel van de mensheid zijn brein kan verbeteren en de rest niet? En hoe werkt het bij kinderen?

Al het onderzoek in dit proefschrift heb ik gedaan in gezonde jongvolwassenen. Uiteindelijk denk ik echter dat het waarschijnlijk vooral klinische populaties zullen zijn waarvoor tDCS een uitkomst kan bieden. Daartoe moet eerst nog meer onderzoek worden gedaan. De menselijke hersenen vormen een ongelooflijk complex geheel van neuronen, neurotransmitters en activiteit. Lang leve de magie van het brein!

Dankwoord

Allereerst wil ik graag mijn beide promotoren bedanken. Heleen, ik heb de afgelopen jaren ontzettend veel geleerd van je kritische geest, nieuwsgierige blik en enthousiasme voor de wetenschap. Bedankt voor de vrijheid en het vertrouwen dat je me hebt gegeven bij het ontwikkelen van mijn ideeën en experimenten. Ook toen ik steeds verder afdwaalde van ons oorspronkelijke plan bleef je betrokken, dacht actief mee en ging indien nodig voor dag en dauw met mijn documenten aan de slag. Ik ben ontzettend dankbaar dat je altijd in mij bent blijven geloven. Als er een iemand is zonder wie dit proefschrift er niet was geweest dan ben jij het.

Richard, ook al was je niet altijd direct inhoudelijk bij mijn experimenten betrokken, je kennis, wijsheid en advies vanuit de coulissen hebben altijd als een grote steun gevoeld. Ik heb genoten van onze filosofische gesprekken en waardeer je kritische maar open en positieve blik op de wetenschap. Dank dat je mijn promotor wilde zijn.

Ook veel dank aan de directe betrokkenen van dit proefschrift: Alle deelnemers die ondanks de middeleeuwse stimulatiehelmpjes braaf mijn werkgeheugentaken bleven uitvoeren. De drie studenten Rik, Job en Julia die zorgvuldig en aandachtig deze rubberen kunstwerkjes op de hoofden fabriceerden. Mijn co-auteurs Steven en Anouk voor hun medewerking bij mijn MRS-project en het Spinoza-centrum voor extra scanuren. De Methodenwinkel voor hun onuitputtelijke statistische kennis. En tot slot natuurlijk mijn mede Brein&Cognitie members; voor het delen van hun kennis en ervaring op onderzoeksgebied maar ook wat betreft het hele promotiecircus hieromheen.

Ik ben uiteindelijk meer dan zes jaar met dit proefschrift bezig geweest. In deze periode heb ik een prachtige cirkel loyale, betrokken, mooie, lieve, slimme en creatieve

jonge mannen en vrouwen om me heen gehad. Ik ben ontzettend dankbaar voor de betrokkenheid en steun die ik de afgelopen jaren heb gevoeld van al mijn goede vrienden en vriendinnen, mijn huidige en oud-huisgenootjes, mijn collega's, mijn familie, mijn collega's, mijn jaarclub en last but not least, mijn akrobatiek-maatjes. Dank voor de manier waarop jullie elke kerst, festival, training, borrel of Franse zomer op zachte toon bleven informeren naar mij en mijn proefschrift vorderingen. Met jullie bemoedigende woorden, relativerende grapjes en warme knuffels hebben jullie mij een hart onder de riem gestoken op momenten dat ik het nodig had. I could not have done this without you!

Een aantal mensen wil ik wel graag nog extra persoonlijk bedanken. Mijn ouders, Joke en Fedde. Ik heb jullie liefde en onvoorwaardelijke steun de afgelopen jaren weer extra mogen voelen. Dank dat jullie er altijd voor mij zijn. Ik kan me geen betere of lievere ouders wensen dan jullie. Marlies, ik ben heel blij de highs en lows van deze promotie-weg met jou te hebben kunnen delen en kijk met veel plezier terug naar al onze avonturen onderweg. Je bent de beste aio-siss ever! En tot slot uiteraard mijn beide paranimfen. Fleur, met jou kan ik zowel het leven analyseren als kinderlijk schaterlachen. Je bent mijn beste en oudste vriendinnetje en dat voelt inmiddels verdomd veel als familie! Malu, met jou samen kan ik als met geen ander van live muziek, festivals en sfeer genieten. Ik waardeer het enorm hoe je daarnaast op de juiste momenten met praktisch levensadvies voor me klaar blijkt te staan. Duizendmaal dank aan jullie beiden voor het helpen bij de voorbereidingen voor 'mijn grote dag'. Ik heb heel veel zin deze dag samen met jullie te beleven!