Åbo Akademi University

Karolina M. Lukasik

Studies in Working Memory: Reliability of Measurement, Psychological Correlates, and Malleability to Brain Stimulation

Studies in Working Memory: Reliability of Measurement, Psychological Correlates, and Malleability to Brain Stimulation

Karolina M. Lukasik

Psychology

Faculty of Arts, Psychology, and Theology

Åbo Akademi University

Åbo, Finland, 2019

Supervised by
Professor Matti Laine, PhD
Department of Psychology
Åbo Akademi University
Finland

Associate Professor Minna Lehtonen, PhD
Center for Multilingualism in Society across the Lifespan
Department of Linguistics and Scandinavian Studies
University of Oslo
Norway

Dr. Anna Soveri, PhD
Department of Clinical Medicine
University of Turku
Finland

Reviewers
Professor Teresa Bajo, PhD
Faculty of Psychology
University of Granada
Spain

Director Martin Buschkuehl, PhD
MIND Research Institute
United States of America

Opponent
Professor Teresa Bajo, PhD
Faculty of Psychology
University of Granada
Spain

ISBN 978-952-12-3801-7 (printed)
ISBN 978-952-12-3802-4 (digital)

Painosalama Oy — Turku, Finland, 2019

TABLE OF CONTENTS

ACKNOWLEDGMENTS 4
LIST OF ORIGINAL PUBLICATIONS 5
SVENSK SAMMANFATNING 6
ABSTRACT 7
LIST OF ABBREVIATIONS 8
1 INTRODUCTION 9
1.1 Working memory and other executive functions: a dynamic system 9
1.2 Psychometric properties of WM tasks and other executive measures 11
1.3 Determinants of WM performance 12
1.4 Brain stimulation and its effects on WM performance 15
2 AIMS OF THE THESIS 17
3 MATERIALS AND METHODS 19
3.1 Participants 19
3.1.1 Study I 19
3.1.2 Studies II-IV 19
3.1.3 Study V 20
3.2 Methods 20
3.2.1 Study I 20
3.2.2 Studies II-IV 23
3.2.3 Study V 25
4 RESULTS 27
4.1 Study I 27
4.2 Study II 27
4.3 Study III 27
4.4 Study IV 28
4.5 Study V 28
5 DISCUSSION 29
5.1 Psychometric properties of WM tasks and other EF measures 29
5.2 Determinants of WM performance 30
5.3 Brain stimulation and its effects on WM performance 33
5.4 Conclusions 34
REFERENCES 37

ACKNOWLEDGMENTS

Most importantly, I would like to thank my supervisors: Professor Matti Laine, Dr Minna Lehtonen and Dr Anna Soveri. It was a great honour to be a mentee of such brilliant scientists. Thank you for your guidance and help, and the many discussions we had. You gave me the possibility to grow and thrive as a researcher, and inspired me to pursue my academic interests.

This dissertation would not be possible without the BrainTrain group. The project was an incredible learning experience, but also a great joy. Thank you for the seminars, conferences, and the truly impressive work we've done together.

The Department of Psychology at Åbo Akademi, and especially Kuvösen, provided an amazing, inspiring and fun atmosphere, and I have no doubts that this is thanks to the people who work there. I would like to thank you all for both challenging discussions and support, for your good humor and keen observations.

I would also like to thank my funders: the Åbo Akademi University Endowment (the BrainTrain project), Åbo Akademi Psychology and Logopedics Doctoral Network, and the Rector of Åbo Akademi.

My loved ones and friends had relentless faith in me, even when I myself thought that I could never complete this undertaking. They listened to me and cheered for each published article. Thank you, mom and dad, for helping me start this adventure. Zula, you are the best sister and an amazing scientist yourself. You inspire me. Elżbieta and Sara, thank you for always being there for me. Inari, Tero, Joanna, Tommi, Alexej - you made Turku my second home.

Research is not a solitary endeavor, even if it sometimes seems so: behind this dissertation is the support of all of the incredible, talented people I mentioned above. The best thing I can do is to give back to this communit of people who gave me so much: let this thesis be my contribution and my thanks.

LIST OF ORIGINAL PUBLICATIONS

I Soveri, A., Lehtonen, M., Karlsson, L. C., Lukasik, K., Antfolk, J., & Laine, M. (2018). Test–retest reliability of five frequently used executive tasks in healthy adults. *Applied Neuropsychology: Adult*, 25(2), 155-165. https://doi.org/10.1080/23279095.2016.1263795

II Waris, O., Soveri, A., Lukasik, K. M., Lehtonen, M., & Laine, M. (2018). *Working memory and the Big Five. Personality and Individual Differences*, 130, 26-35. https://doi.org/10.1016/j.paid.2018.03.027

III Lukasik, K. M., Lehtonen, M., Soveri, A., Waris, O., Jylkkä, J., & Laine, M. (2018). Bilingualism and working memory performance: Evidence from a large-scale online study. *PloS One*, 13(11). https://doi.org/10.1371/journal.pone.0205916

IV Lukasik, K.M., Waris, O., Soveri, A., Lehtonen, M., Laine, M. The relationship of anxiety and stress with working memory performance in a large non-depressed sample. *Frontiers in Psychology*, 10:4. https://doi.org/10.3389/fpsyg.2019.00004

V Lukasik, K. M., Lehtonen, M., Salmi, J., Meinzer, M., Joutsa, J., & Laine, M. (2018). No effects of stimulating the left ventrolateral prefrontal cortex with tDCS on verbal working memory updating. *Frontiers in Neuroscience*, 11:738. https://doi.org/10.3389/fnins.2017.00738

SVENSK SAMMANFATTNING

Arbetsminnet syftar på det temporära minnessystemet som möjliggör bevarandet och hanteringen av aktuell, relevant information. Arbetsminnet är en nyckelkomponent i mänsklig kognition och är associerad med flera andra förmågor och prestationer, såsom allmän intelligens, språkinlärning, känslomässig kontroll och akademisk prestation. Den föreliggande avhandlingen behandlade tre frågor om arbetsminnet: 1) reliabiliteten av arbetsminnesmått, 2) sambanden mellan arbetsminnesprestation och valda psykologiska faktorer, och 3) formbarhet genom hjärnstimulering.

Med hjälp av data från laboratorieundersökningar utvärderades retest-reliabiliteten på vanliga arbetsminnes- och andra exekutiva uppgifter. Uppgifternas reliabilitet har fått ringa uppmärksamhet inom forskningen, trots att den är kritisk med tanke på användningen av dessa mått. Den andra frågan utreddes med ett stort onlinesampel av vuxna i tre olika studier där sambanden mellan arbetsminnet och valda psykologiska faktorer (personlighetsdrag, psykologiskt välbefinnande, och tvåspråkighet) analyserades. För den tredje frågan undersöktes hur verbala arbetsminnesprestationer påverkas av transkraniell elektrisk stimulering (tDCS) av undre halvan av vänstra pannloben, ett hjärnområde som är relaterad till arbetsminnet.

Reliabilitetsanalyserna tydde på stor variation i retest-reliabilitet bland exekutiva uppgifter, med högsta retest-reliabilitet hos arbetsminnesuppgifterna. I onlinestudien mättes arbetsminnet med faktorbaserade summavariabler (verbalt arbetsminne, visuospatialt arbetsminne och uppdatering av arbetsminnets innehåll). Man fann endast svaga samband mellan de psykologiska faktorerna och arbetsminnesprestationer. Personlighetsdragen samvetsgrannhet och öppenhet hade små, men statistiskt signifikanta negativa samband med uppdatering av arbetsminnet. När det gäller stress och ångest observerades endast en tendens till ett negativt samband mellan övergående ångest och arbetsminnesprestationer. Gällande språklig bakgrund presterade sent tvåspråkiga personer bättre än tidiga tvåspråkiga och enspråkiga på uppdatering. Däremot observerades inga samband mellan prestationerna på arbetsminnesuppgifter och centrala drag i individuell tvåspråkig erfarenhet inom de tvåspråkiga grupperna. I tDCS-studien observerades inga signifikanta effekter av hjärnstimulering på arbetsminnesprestationer.

Sammanfattningsvis visar arbetsminnesmått (men inte vissa andra exekutiva mått) mestadels rimlig reliabilitet, vilket stöder deras fortsatta användning. I detta icke-kliniska sampel av vuxna verkar arbetsminnesprestationerna vara ganska robusta mot mätfel, normal inter-individuell variation i vissa relevanta psykologiska faktorer, och excitatorisk hjärnstimulering med tDCS.

ABSTRACT

Working memory (WM) refers to the temporary memory system that enables us to maintain and manipulate currently relevant information. As a key aspect of human cognition, it is associated with several other abilities and achievements, such as general intelligence, language learning, emotional control, and academic performance. The present thesis addressed three issues concerning WM, ranging from measurement to intervention: the reliability of its measures, psychological correlates, and malleability to brain stimulation. First, using data from laboratory studies, the test-retest reliability of commonly used WM and other executive tasks was examined. This issue has received scant research attention, even though it is crucial for the use of these measures. Second, three separate studies, employing a large online sample of adults, probed the relationships between WM and selected psychological factors (personality traits, psychological well-being, and bilingual experience). Third, the malleability of verbal WM performance to stimulation of WM-related left inferior frontal regions was probed with transcranial direct current stimulation (tDCS).

The reliability analyses indicated considerable variability of test–retest reliabilities in executive measures, with highest test-retest reliability values for the WM tasks. In the online study, WM was measured by factor-based composite scores (verbal WM, visuospatial WM, and updating of WM contents). Concerning the psychological factors and WM performance measures, only weak associations were found. Personality traits Conscientiousness and Openness had small but statistically significant negative associations with WM updating. Concerning stress and anxiety, there was only a trend towards a negative effect of transient anxiety on WM performance. With regard to linguistic background, late bilinguals outperformed early bilinguals and monolinguals on WM updating. However, within the bilingual groups, no associations were found between WM task performance and central features of bilingual experience. The tDCS study failed to find any significant effects of left prefrontal cortex stimulation on WM performance.

In conclusion, WM measures (but not some other executive measures) show mostly reasonable reliability that speaks for their continued use. In the present non-clinical samples of adults, WM performance appears to be rather robust against measurement error, normal inter-individual variability in some pertinent psychological factors, and excitatory brain stimulation as applied by tDCS.

LIST OF ABBREVIATIONS

WM — working memory

tDCS — transcranial direct current stimulation

EF — executive function

ADHD — attention deficit-hyperactivity disorder

PFC — prefrontal cortex

RT — reaction time

AoA — age of acquisition

L2 — second language

QIDS-SR16 — Quick Inventory of Depressive Symptomatology – Self-Report

ANCOVA — analysis of covariance

ANOVA — analysis of variance

1. INTRODUCTION

1.1 Working memory and other executive functions: a dynamic system

Whenever we face a change in our environment, stay focused in spite of distractors, or retain a piece of information while new stimuli appear, we rely on executive function (EF), a cluster of top-down mental processes enabling guided behavior (Miyake et al., 2000). EF has been considered to be at the same time a unitary and a diverse construct: while EF seems to tap some common underlying ability, it can be also broken down into separate, specific processes (Miyake et al., 2001). One of such core EF is working memory (WM): a system devoted to maintenance and manipulation of information relevant to the tasks at hand (Baddeley, 1998). WM is more than just a passive short-term memory system: Logie (2011) dubbed it the "mental workspace", highlighting the many processes involved in manipulating the contents of WM. These processes, which include updating (maintaining accurate representations of information changing over time; Ecker, Lewandowsky, & Oberauer, 2010), monitoring, rehearsal and integration of stimuli, are in constant interaction (Kessler & Meiran, 2008; Logie, 2011; Szmalec, Verbruggen, Vandierendonck, & Kemps, 2011).

The active and dynamic nature of WM enables higher-order cognitive processes such as planning, self-control, and creative thinking — and thereby flexible, goal-directed everyday functioning (Conway, MacNamara, & Engel de Abreu, 2013). There is also ample evidence that WM performance and general (fluid) intelligence are highly correlated (Conway, Kane, & Engle, 2003; Oberauer et al., 2005). It should be of little surprise, then, that WM is a predictor of language skills (Baddeley, 2003) and academic achievement (Alloway et al., 2005; Alloway & Alloway, 2010). WM is also sensitive to various neuropsychiatric disorders, such as depression, attention deficit-hyperactivity disorder (ADHD), and schizophrenia (Rose & Ebmeier, 2006; Gropper & Tannock, 2009; Park & Gooding, 2014). The thriving scientific interest in WM is proportional to its role in human behavior.

However, research into WM has proved challenging, as it comprises of numerous processes, and one task cannot be expected to assess them all[1]. Thus, a variety of task paradigms has been developed to measure different aspects of WM. The

1 WM performance is often conflated with WM capacity, but it should be emphasized that a person's WM task score is not a pure capacity measure. For example, the score can also be indicative of the ability to separate relevant and irrelevant information, and the accuracy of retrieval (Unsworth & Engle, 2007).

span tasks are amongst the most commonly used measures in cognitive psychology (Waters & Caplan, 2003; Conway et al., 2005; Schmiedek, Lövdén, & Lindenberger, 2014). This paradigm includes, among others, the simple span where the participant is asked to memorize and recall a sequence of stimuli; the running memory span where a stimulus sequence suddenly stops and one should recall the last *n* items (Broadway & Engle, 2010); and the complex span where a stimulus sequence should be maintained while intermittently performing another task, which increases the cognitive load. Span tasks tap the capacity of WM, which is limited and varies considerably between individuals (Cowan, 2010). Running memory task requires constant monitoring and updating of WM contents while also retaining the stimuli in proper order of appearance, and there is evidence that participants use both passive processing and active strategies (Botto et al., 2014). Another extensively used paradigm is the n-back (Kirchner, 1958; Jaeggi et al., 2009), where the participant is to decide whether the most recent stimulus is the same as the one that appeared *n* trials back. Hence, n-back taps into WM updating, but it also involves decision-making and interference resolution. The cognitive load in n-back can be systematically manipulated by increasing the value of *n*. All in all, the aforementioned paradigms operationalize key WM processes: storage, updating and manipulation of information.

There seem to be prominent individual differences in WM (Kane & Engle, 2002; Unsworth & Engle, 2007). WM performance has been linked to the organization and activity of the prefrontal cortex (PFC; Kane & Engle, 2002), and appears to carry a significant genetic component (Ando, Ono, & Wright, 2001). PFC is known to be one of the key brain areas involved in cognition and executive control (Koechlin, Ody, & Kouneiher, 2003). However, the limits of an individual's WM are shaped not only by the neurophysiological makeup, but also by environment, health status, and experience, to which the brain reacts – and changes (McEwen & Morrison, 2013). To take examples of these very diverse factors, Evans and Schamberg (2009) argued that childhood poverty and chronic stress are related to WM impairment that lasts throughout adulthood. On a different note, some studies suggest that childhood experience of bilingualism can be beneficial for WM development (Morales, Calvo, & Bialystok, 2013). Age is also an important factor: WM performance declines with ageing (Archer et al., 2018). The examples listed above concern long-lasting influences, but even temporary factors such as sleep deprivation and changes in motivation levels can modulate WM performance (Chee & Choo, 2004; Jaeggi et al., 2014). It has also been proposed that WM can be trained through intensive practice with WM tasks, but whether such training indeed affects WM capacity is controversial (Klingberg, 2010; von Bastian & Oberauer, 2014; Soveri et al., 2017). Some studies have shown success in pairing WM training with brain stimulation techniques, such as transcranial direct current stimulation (tDCS) (Brunoni & Vanderhasselt, 2014; Au et al., 2016), but meta-analytic evidence indicates that these effects are usually minor if observed at all (Mancuso et al., 2016; Horvath et al., 2015a, b). All in all, various aspects of WM[2] appear to be potentially modulated by diverse factors in ways that are not yet fully clear.

The present thesis had three aims. The first was to investigate the psychometric properties of commonly used WM tasks and other EF measures. The temporal stability of EF tasks, despite its crucial importance to researchers and clinical psychologists alike, has not been thoroughly studied. Here the test-retest reliability of commonly used EF tasks was examined. The second aim was to probe the associations between WM performance and some factors related to personality, psychological well-being, and linguistic background. Inter-individual variability in these factors can be related to the differences observed in cognitive performance, and charting these relationships for WM may shed light on the interplay of cognitive, emotional and linguistic domains. The third aim was to examine the malleability of verbal WM performance to externally provided excitatory neural stimulation by means of tDCS.

1.2 Psychometric properties of WM tasks and other executive measures

In 1996, Monsell lamented the "embarrassing zone of almost total ignorance" (p. 93) in cognitive psychologists' understanding of how EF engages and interacts in complex cognitive tasks. Since then, many advancements have been made, also by pairing cognitive tasks with neuroimaging (for example, see Owen et al., 2005). However, studies examining the validity and reliability of EF tasks are still remarkably scarce, and our incomplete understanding of psychometric properties of commonly used tasks calls for further research.

Reliability and validity of a task are intertwined properties: a task needs to be both valid and reliable in order to be useful both for research and practical applications (Bordens & Abbott, 2005). Task validity describes the extent to which a task measures what it is aimed to measure; how "pure" a task is and what else does it tap besides the targeted mental process. In cognitive psychology, this has been coined as the task impurity problem (Miyake et al., 2000), and it is a parti-

2 The internal structure of WM has been debated throughout the years. The early three-component model of WM proposed by Baddeley and Hitch (1974) highlighted the material-specificity of the subordinate systems (phonological loop, visuospatial sketchpad), whereas Oberauer and colleagues (2003) suggested a functional division between storage and processing. The present thesis does not address the mental architecture of WM as such, but employs a "hybrid" division by content and function based on the data-driven three-factor solution (verbal WM, visuospatial WM, updating in n-back) reported by Waris and colleagues (2017). Their analysis was based on the same large online sample that is employed in this thesis as well (Studies II-IV).

cularly relevant issue for EF tasks. Since all tasks use physical items (be it numbers, letters, shapes in different locations or auditory stimuli), every task will tap not only on the cognitive processes of interest, but also the perceptual and motor processes involved in task performance (Shah & Miyake, 1996). Hence, it is impossible to disentangle a "pure" EF score from task-specific and paradigm-specific variance (consider e.g. the uncertainty of stimulus list length in a running memory task which is not a feature of an n-back task, even though both paradigms require WM updating). Moreover, participants might approach a task differently and use different strategies to meet the task demands on separate occasions, or even within one session (Miyake et al., 2000). This could explain the usually low correlations between EF tasks that should tap the same executive processes, but use different materials and/or response features (Miyake et al., 2000; Kane et al., 2007), even when the tasks are isomorphic variants (Waris et al., 2017). Furthermore, the changing participant approaches could explain the low-to-moderate test-retest reliabilities of EF tasks (Wöstmann et al., 2013). However, factor analyses can reveal the hidden commonality in task variances (see for example, Schmiedek, 2009; 2014), and point to the existence of a shared underlying EF factor despite those drawbacks.

The question of temporal stability of EF tasks is crucial for psychological research, especially longitudinal and intervention studies. If a person's task score spontaneously varies greatly over time, it cannot be used as a screening tool or to mark the effectiveness of training or rehabilitation. Of course, no task is perfectly stable; random and systematic effects bring variability to a person's task performance. Surprisingly, only a limited number of studies report test-retest reliability of EF tasks. Moreover, the reliability coefficients for EF tasks in these studies are usually quite low. Waters and Caplan (2003) investigated the test-retest reliability for several WM span tasks, noting that the Pearson's r values ranged from .41 to .83. In their own subsequent study, only one of the span tasks met the value of $r >$ .7, considered as a liberal cutoff for adequate reliability (Nunnally, 1978; see also Anastasi, 1982, for more conservative criteria). For the n-back task, reliability measures across studies show a vast range from $r = .02$ to $r = .91$, and the reliability coefficients are usually the highest for the most demanding conditions in a task (Jaeggi et al., 2010). It should also be noted that reaction time (RT) measures usually have higher reliability than accuracy measures (Wöstmann et al., 2013; Gonçalves, Pinho, & Simões, 2016; Paap & Sawi, 2016). Nonetheless, even the adequate reliability of $r = .7$ indicates that the measurements at two time points share only half of the variance in task performance.

1.3 Determinants of WM performance

As WM performance varies significantly between individuals (e.g., Luck & Vogel, 2013), it opens the search for correlates of this variation. Knowledge of factors associated with higher or lower WM performance can help to protect this central cognitive function, and also to develop measurement tools when for example planning interventions for aging populations.

Individual differences in WM performance have neural underpinnings (Gevins & Smith, 2000; Luck & Vogel, 2013) and, for example, dopamine-related gene polymorphisms explain some portion of the variance (Bertolino et al., 2006). However, WM performance has been assumed to be affected also by diverse factors such as personality features, psychological well-being, and linguistic experience (Luo et al., 2013; Dick, Eccleston, & Crombez, 2002; Curtis et al., 2015). These particular factors and their associations with WM performance were examined in the present thesis.

Consistent inter-individual differences in human cognition and personality raise the issue as to whether they are related. Research conducted within the dominant Big Five framework (Goldberg, 1993) has indicated that personality traits are associated with, among others, occupational success and academic achievement (Paunonen & Ashton, 2001; Roberts et al., 2007). It has been suggested that personality has a long-term effect on cognition through the decisions each individual makes about the investment in their own cognitive abilities and environment (von Stumm, Chamorro-Premuzic, & Ackerman, 2011). As regards WM, a narrative review by Curtis and colleagues (2015) found a positive association between the personality feature Openness and cognitive ability, as well as a negative link between Neuroticism and WM. However, the authors state that the results are often inconclusive and many studies suffer from methodological drawbacks, prompting further research.

The demands posed by environment, impacting psychological well-being, can affect cognitive performance. It has been shown that chronic stress, trauma and disease can have adverse effects on cognitive functioning, including WM (Schoofs, Preuß, & Wolf, 2008). These influences can be long-lasting: Evans and Schamberg (2009) reported that psychological stress experienced in childhood has a detrimental effect on WM performance in adulthood. Chronic stress triggers repeated mobilizations of physiological systems, which in turn deplete one's physical and cognitive resources (Evans & Schamberg, 2009). These detrimental effects have been also observed in everyday and work-related stress (Sliwinski et al., 2006; Nguyen et al., 2012). In turn, the attentional control theory states that anxiety — a state of heightened vigilance due to uncertainty and conflict — will give more influence to bottom-up, stimulus-driven attentional processes, causing disruption in cognitive performance as it creates strain on higher-order, top-down functions (Eysenck et al., 2007). It has been observed that state anxiety impairs the efficiency of processing when the tasks demand executive control (Eysenck, Payne, & Derakshan, 2005; see also Christopher & MacDonald, 2005) and that it limits the ability to focus on relevant information (Vytal et al., 2012; 2013). However, while the effects of stress and anxiety on cognition are as such well-studied, this research mostly concerns clinical populations and, in the case of state anxiety, laboratory-induced affect. Much less is known as to how everyday experiences causing stress or transient anxiety influence WM, a topic taken up in the present work.

The idea that persistent, challenging activity, such as speaking two languages, serves as natural training for cognitive functions has garnered popularity in EF research (for example, see Bialystok & Craik, 2010; Prior & Macwhinney, 2010; Soveri et al., 2011). It is generally accepted that bilingual persons' both languages are active in their mind all the time (Macizo, Bajo, & Cruz Martin, 2010), and thus it is assumed that there is constantly potential interference to be controlled. Bilinguals' language selection, inhibition and switching are considered to engage various aspects of EF, and this could provide a general boost to EF (e.g., Colzato et al., 2008; Stocco & Prat, 2014). However, in the case of language, there appears to be no consensus regarding which aspects of bilingual experience are beneficial for cognition. Age of acquisition (AoA) is often brought up as a key factor; there is evidence that learning two languages in childhood affects neural networks related not just to linguistic processing, but also EF (Perani et al., 2003; Tao et al., 2011). Furthermore, people who have been using two languages since childhood have had more time to practice managing them efficiently (Luk, De Sa, & Bialystok, 2011). A related factor is L2 proficiency: suppressing a stronger language, to prevent its interfering influence, is more difficult than in the case of a weaker language. Therefore higher L2 proficiency can be assumed to have led to increased demands for cognitive control and thus higher training gains than lower L2 proficiency. It has been also suggested that language switching is a practice that puts greater strain on EF (e.g., Prior & Macwhinney, 2010; Prior & Gollan, 2011; Stocco & Prat, 2014). On the other hand, studies show that switching is highly dependent on sociolinguistic context, and can often be an effortless communication strategy in diglossic communities rather than ecological EF training (for example, see Costa et al., 2009; Hartanto & Yang, 2016).

Meta-analyses on studies that have compared bilingual and monolingual speakers on EF measures, including WM, have yielded inconclusive results. Adesope and colleagues (2010) reported a positive relationship between bilingualism and WM, in line with a later meta-analysis by Grundy and Timmer (2017) who found a small positive effect of bilingualism on WM in their study that included both adult and children samples. On the other hand, after correcting for publication bias, the recent large meta-analysis on adult studies by Lehtonen and colleagues (2018) found no effect of bilingual experience on WM performance or other aspects of EF in adults. Likewise, results from studies examining correlations between key features of bilingualism (AoA, level of proficiency, frequency of language switching) have yielded variable results (Luk, De Sa, & Bialystok, 2011; Paap & Greenberg, 2013; Soveri et al., 2011; Tao et al., 2011; Verreyt et al., 2016), albeit such correlations would be expected if bilingual experience provides a boost for WM and other aspects of EF. Thus, more research is clearly needed on the relationships between bilingualism and WM.

1.4. Brain stimulation and its effects on WM performance

tDCS represents a rather new brain stimulation method for influencing cognitive function. The use of tDCS in the enhancement of cognition, including WM, has garnered scientific interest thanks to its non-invasive character and relative ease of use (Nitsche et al., 2008). In tDCS, two electrodes are placed on the scalp, and a weak electrical current passes between them through brain tissue. One proposed mechanism of action is the modulation of neuronal membrane potentials: neurons located near the anode are hypopolarized, while cells near the cathode are hyperpolarized (Stagg & Nitsche, 2011). Early tDCS studies investigating the effects of stimulation on motor cortex have revealed that hypopolarization of neurons near the anode leads to increase of cortical excitability of those cells, whereas cathodal stimulation has the opposite, inhibitory effect (Nitsche & Paulus, 2004). The shift in resting membrane potential affects the threshold for neuronal discharge, either facilitating (anodal tDCS) or inhibiting (cathodal tDCS) neuronal firing (Nitsche et al., 2003). The second suggested mechanism is that more sustained stimulation changes the synaptic activity of the neurons under electrodes through long-term potentiation and depression (Horvath, Forte, & Carter, 2015a,b). Coupling anodal tDCS with cognitive or motor tasks is hypothesized to result in long-term potentiation (LTP) (Fritsch et al., 2010), facilitating the establishment of stable synaptic connections in the stimulated brain regions. In terms of behavior, this should facilitate long-lasting learning.

Successful performance on a cognitive task can be understood as efficient top-down control, which in turn is conceptualized as communication of key brain areas involved in completing the task (Fries, 2005). Although the idea of cognition conceived as brain regions "speaking" to each other is obviously simplistic, it is nonetheless a useful shorthand when designing relevant studies. Thanks to brain imaging and lesion studies, we have identified task-specific and task-general areas in the large brain networks; for example, there is ample evidence that PFC is involved in decision-making and cognitive control. If this is the case, enhancing neural communication within this crucial region by tDCS-induced increase in cortical excitability could result in better performance on cognitive tasks that call for cognitive control.

However, the evidence from experimental tDCS studies addressing WM is inconclusive, and recent meta-analyses have shown mixed results. Brunoni and

Vanderhasselt (2014) compared tDCS to transcranial magnetic stimulation (TMS) and found no effect of tDCS on WM in healthy participants. Similarly, Horvath and colleagues (2015a, b) claimed that not only does tDCS have no reliable cognitive effects, including WM performance, in healthy populations, but that it generates little-to-no reliable neurophysiological effects. Overall, the stable electrode-related excitation/inhibition patterns observed during motor cortex stimulation have not been replicated in studies of cognition (Horvath et al., 2015a, b). On the other hand, two more recent meta-analyses (Hill et al., 2016; Dedoncker et al., 2016) reported faster responses on WM tasks post-stimulation ("offline") in healthy cohorts, and greater accuracy during stimulation ("online") in clinical samples.

It has been argued that the heterogeneity of results in studies examining tDCS effects on cognition is rooted in the inherent complexity of cognitive tasks, which engage rich brain networks (Jacobson et al., 2012), and in the differences in experimental design (Fertonani & Miniussi, 2017). There is also evidence that participants who had low scores in WM tasks would benefit from the stimulation while high-performers would not (Wu et al., 2016; Juan et al., 2017; Katz et al., 2017), which suggests that individual differences also affect the outcome of stimulation. All in all, the scope and applicability of tDCS in WM enhancement remain unclear.

2. AIMS OF THE THESIS

This thesis had three main aims related to different aspects of WM performance, namely the temporal stability of EF task performance including WM, associations between WM and some key psychological factors, and the malleability of verbal WM performance to brain stimulation by tDCS. These three aims are described in more detail below. Due to the scarcity or inconclusiveness of previous research, all studies were exploratory in nature.

First, the present thesis investigated the test-retest reliability of commonly used executive tasks, including WM tasks was investigated (Study I). The following three studies probed possible correlates of WM performance in the realm of personality (Study II), psychological well-being (Study III) and linguistic experience (Study IV). In the final, experimental study, possible verbal WM enhancement by short-term tDCS was examined (Study V). Studies II-IV were conducted on the same large sample of adult English-speaking U.S. residents recruited on the Internet, while Studies I and V employed laboratory-based data.

Study I investigated the test-retest reliability of five commonly used executive tasks, including WM tasks. The data were passive controls' pre-post test results from three WM training studies conducted by the BrainTrain research group. This exploratory study was motivated by the scarcity of data on reliability and temporal stability of tasks tapping EF.

Study II explored the relationships between the Big Five personality traits and WM performance. Previous research suggests some links between personality traits and cognitive performance measures including WM, but the results have been inconclusive (Curtis et al., 2015). This study consisted of a systematic review and a correlational study.

Study III investigated whether everyday stress and state anxiety are related to WM performance. While both symptoms have been related to WM impairments (Eysenck et al., 2007; Schoofs, Preuß, & Wolf, 2008; Vytal et al., 2013), these relationships have been studied primarily in clinical groups or laboratory conditions. Here the focus was on everyday symptoms in a non-clinical sample.

Study IV explored the associations between bilingual experience and WM performance. Two analysis approaches were applied: comparisons between monolingual and early vs. late bilingual participants on WM tasks, as well as associations between central features of bilingual experience (AoA, L2 proficiency and frequency of language switching) and WM performance within the bilingual group. If bilingualism enhances WM in bilinguals, one would expect positive results from both analysis approaches.

Study V examined whether verbal WM performance in healthy adults could be enhanced by short-term tDCS on left inferior frontal regions that are related to verbal WM. Both the direct "online" effects of stimulation, as well as the "offline" effects after stimulation were examined. In a double-blind counterbalanced within-subjects design, we examined tDCS effects on a verbal WM updating task by using anodal, cathodal and sham stimulation.

3. MATERIALS AND METHODS

3.1 Participants

For all studies, we obtained a written informed consent from the participants. The original studies included in Study I were approved by the Institutional Review Board of the Department of Psychology and Logopedics of the Åbo Akademi University. Upon arriving in the laboratory, the participants were briefed and informed they could withdraw at any point. For Studies II-IV in the present thesis, we have obtained approval from the Joint Ethics Committee at the Departments of Psychology and Logopedics at the Åbo Akademi University and the Human Research Review Board at the University of California, Riverside. The anonymous participants in this online data collection were informed that they could withdraw from the study at any time. In Study V, all participants had undergone medical screening to ensure there were no health risks related to them taking part in the study. The study was approved by the Ethics Committee for the Hospital District of Southwest Finland. The participants were briefed upon arriving in the laboratory and informed of their right to withdraw at any point.

3.1.1 Study I: Test-retest reliability of executive measures

Study I used data from passive control groups in three separate cognitive training studies (Bäckman et al., 2011; Soveri, Waris, & Laine, 2013; Bäckman et al., 2017), consisting of a total of 37 right-handed Finnish university students (*M* age = 23, *SD* = 2.4). The participants reported having no neurological or psychiatric disorders, did not have learning difficulties, and had no history of drug and alcohol abuse. In the original cognitive training studies, they took part only in the pre- and posttest sessions. Two of the original studies had used positron emission tomography in the pre- and posttest sessions (Bäckman et al., 2011; 2017).

3.1.2 Studies II-IV: Determinants of WM performance

Data for Studies II-IV stemmed from a single cross-sectional Internet-based experiment using the in-house SOILE testing platform developed in the BrainTrain project. The participants were adult U.S. residents, recruited through the microtask website Amazon Mechanical Turk (MTurk). They were selected on the basis of their high work approval ratings (95% or higher) and the number of completed tasks (more than 100, but less than 1000). They were also asked about past participation in similar psychological studies, and 83.9% reported never having done so. Overall, 711 participants completed an extensive battery of WM tasks and a questionnaire probing their background, health, lifestyle and hobbies. The participants received $10 as reimbursement.

For the Studies II-IV, we excluded participants due to being multivariate outliers on WM task performance, having missing values on the tasks, admitting the use of external aids on WM tasks, and/or having spent over 24 hours on completing the study. We also excluded participants who currently were suffering from moderate, severe or very severe depressive symptoms according to the QIDS-SR16 scale (Rush et al., 2003). Thus, the final sample included 503 participants (*M* age = 34.15, *SD* = 10.57).

For the purposes of Study IV, which investigated the relationships between multilingualism and WM, we excluded participants who had missing data on language-related questions. We also excluded those who declared that English was their L2, since we wanted the participants to complete tasks in their first language (see Grundy & Timmer, 2017; Lehtonen et al., 2018). Language proficiency was not an exclusion criterion. This resulted in a final sample of 485 participants, out of which 220 were monolingual and 265 spoke a second language. We categorized the L2 speakers as either early bilinguals (having learned L2 during the first 12 years of their life; 115 participants) or late bilinguals (150 participants).

3.1.3 Study V: Enhancing WM with tDCS

For Study V, we recruited 34 healthy, right-handed young adults (university students or graduates) (*M* age = 22.6, *SD* = 2.42). The final decision about enrollment was made by a licensed physician. This led to the exclusion of candidates who had reported suffering from neurological or psychiatric disorders, taking medication affecting the central nervous system, having family history of epilepsy, having suffered a traumatic brain injury, being diagnosed with learning difficulties, being pregnant during the study and drug use in the last four months. One person resigned due to headache experienced while receiving tDCS, which resulted in the final sample of 33 participants. Prior to the study, each participant filled in a background information questionnaire, which included questions about age, gender, and language use. After each session of the study, the participants completed two questionnaires on side effects and possible long-term effects. They also assessed the likelihood of having received active stimulation or sham. The participants received 45€ as reimbursement.

3.2 Methods

3.2.1 Study I

Procedure
Our sample consisted of participants enrolled in the passive control groups in cognitive training studies, that is, those who were not engaged in cognitive tasks between the pre- and posttest. We pooled the passive control samples from three

independent studies (Bäckman et al., 2011; Soveri, Waris, & Laine, 2013; Bäckman et al., 2017). In each of the original cognitive training studies included in the present Study I, the participants filled in a background questionnaire at pretest. After that they completed a battery of cognitive tasks. Since the task batteries were not identical in the original studies, we used only tasks for which we had data from at least 20 participants. The testing sessions in all studies were of comparable length (2-3 hours). The posttest sessions took place after 3 (Soveri et al., 2013) to 6 weeks (Bäckman et al., 2011; 2017).

Tasks

The following sub-chapter briefly describes the tasks used in Study I. More detailed information on the tasks can be found in the articles included in the thesis.

Simon task

The Simon task provides a measure of executive control (Lu & Proctor, 1995). In this task, the participant is asked to press a specific key (left or right) depending on the color of the stimulus presented on a screen (for example, red and blue; Simon, 1990). While the location of the item is irrelevant to the task, studies show that participants are faster to respond in so-called congruent trials where the stimulus location matches the assigned response key (for example, seeing an item on the left side of the computer screen when the correct response is the left key press). In incongruent trials, the item location and correct response do not match spatially — for example, a stimulus presented on the right requires pressing the left key. The Simon effect – the difference in RT between incongruent and incongruent trials – is a robust phenomenon (Simon, 1990; Lu & Proctor, 1995). We used three RT measures in the test-retest reliability analyses: RTs for congruent trials, RTs for incongruent trials, and the Simon effect in RTs (the difference between RTs on incongruent and congruent trials).

N-back task

The n-back task requires constant monitoring and updating of WM contents in order to compare the currently seen stimulus to one seen *n* trials back (Jaeggi et al., 2010). We used a numerical-verbal n-back task (including interleaving blocks of 1-back and 3-back) and a visuospatial n-back task (with 1-back and 2-back blocks). For the verbal n-back, we used both accuracy and RT measures: both mean accuracy and RT for 1-back and 3-back blocks as well as the n-back effect scores in accuracy and RT (1-back score subtracted from the 3-back score). We used the same measure types for the visuospatial n-back, but only RT measures were included.

Letter-memory task

This task was a variant of the running memory paradigm, where letter sequences of unknown length are presented to the participants, and the task is to remember the four last letters in each sequence. In the control condition, the sequences consisted of repetitions of a single letter (e.g., AAAAAAA). The updating condition comprised 38 sequences, the order of which was randomized separately for each participant. We used the total number of correctly recalled items and the total number of correctly recalled sequences in the reliability analysis. Whereas the running memory task was intended to involve active updating of the items held in WM, there is some research showing that participants tend to use passive recall after the sequence presentation has ended. However, regardless of the strategy used by participants, running memory task correlates well with other WM tasks and is thus a valid measure of WM performance (Broadway & Engle, 2010).

Number-letter task
This paradigm measures the ability to shift between tasks (Rogers & Monsell, 1995). Whereas performing one action repeatedly is relatively easy, switching between two tasks is effortful and involves attentional control to disengage with the previous task and meet the demands of the new one (Monsell, 2003). In our study, we used number-letter pairs, presented as in Figure 1 below. Depending on whether the pair was shown inside the top or bottom square, the participants had to decide whether the number was even or odd, or if the letter was a vowel or a consonant. Thus, the location of the pair was a cue for which task to perform. The task consisted of two single task blocks and one mixed task block where both types of tasks appeared. For the analyses of the mixed block data, we separated switch trials where the task changed, and repetition trials where the task was the same as on the preceding trial. In the reliability analysis, we used RTs on single-task trials, RTs on repetition trials, RTs on switch trials, as well as the switching cost (the difference between switching and repetition RTs) and the mixing cost (the difference between repetition RTs and single-task RTs).

Statistical analyses
To investigate the test-retest reliability, we used Pearson's product moment correlation coefficient (Pearson's *r*), the intraclass correlation coefficient (ICC), and the smallest real difference (SRD). For item-level analysis that investigated whether systematic changes in performance occurred within- and between testing sessions, we employed linear mixed effects modeling. We included session (pretest/posttest), condition (e.g., congruent/incongruent) and item number as fixed factors, and entered participants as a random factor. We conducted the analyses for RTs and accuracy measures separately.

3.2.2 Studies II-IV

Procedure

The participants received a link to the testing platform and completed the study on their personal computers. They first filled in the background questionnaire, followed by ten WM tasks. We asked the participants not to use any external aid while performing the tasks, and probed for this also after completion of the battery, having ensured the participants that their answer will not affect their reimbursement. The task order was randomized for each participant, with the exception of forward simple span task being administered immediately before the backward simple span task. The average completion time was 1h 34min.

For Study II, besides the empirical part we also performed a systematic review of studies on associations between cognitive functions and the Big Five personality traits. Our systematic search, conducted using Google Scholar, PubMed and PsycINFO databases, included all studies on healthy adult populations that reported at least one of the Big Five traits and WM measures. Studies reported in a narrative review by Curtis and colleagues (2015) were also included. Overall, the search yielded 39 studies.

Tasks

Our task battery comprised ten tasks. Each of the paradigms had a visuospatial and a numerical-verbal variant.

Simple span forward and backward

In simple span tasks, the participants were presented with item sequences (either digits or locations in the grid) of unpredictable length and asked to recall them in the order of presentation (simple span forward), or in reverse order (simple span backward). The sequence length ranged from three to nine and the presentation order of the strings was randomized. We used the total number of correctly recalled items (regardless of list length) as the dependent measure. Simple span forward is considered to measure primarily WM capacity, while recalling the sequence backward requires also the manipulation of information stored in WM. It has been observed that in the verbal domain, backward span poses more difficulty than forward span, but the difference is less clear in spatial versions of the tasks (Vandierendonck et al., 2004; Monaco et al., 2013).

Complex span

Complex span task introduces an additional interfering element to the simple span task in order to increase processing demands: the participant not only has to retain the target items in WM, but also complete interleaved unrelated tasks at the same time. In our study, participants were asked to make a true/false judgment about distractor items (arithmetic equations in the numerical-verbal variant and combining of partially filled grids in the visuospatial variant) after the presentation of each target item, and recall the target items in the correct order when the sequence had finished. The dependent measure was the total number of correctly recalled items, irrespective of list length. Despite the difference in demands, the complex span and simple span tasks tend to correlate with higher-order cognitive abilities in a similar manner, and both tasks have been considered to measure the same construct (Colom et al., 2006; Unsworth & Engle, 2007).

Running memory task
This task requires updating of the WM contents to recall the last items in a sequence of unknown length. In this study, the task was similar to the Letter Memory task (see Study I), but we used numbers and spatial locations as stimuli. The participants were presented with item sequences of unpredictable length and, at the end of each list, asked to recall the last four items in the exact order of their presentation. The 4-11 item-long sequences were shown in a random order. Overall, the task consisted of eight sequences (one trial per sequence length). We used the total number of correctly recalled items as dependent measure. Responses to lists of 4 items were excluded since they did not require updating.

N-back task
In Studies II-IV, we used the 1- and 2-back versions of the task, randomizing the stimulus sequence for every participant. The stimuli were digits (for the numerical-verbal variant) and locations within a 3x3 grid (for the visuospatial variant). We calculated the dependent measure by subtracting the proportion of false alarms from the proportion of hits (correct targets) on the 2-back task.

Statistical analyses
Prior to analyses, the WM measures were Box-Cox transformed in order to approximate normal distribution (Osborne, 2010). For the purpose of maintaining content specificity, we created three WM composite scores based on an explorative factor analysis by Waris and colleagues (2017) of the present Internet-based sample. The task measures were z-transformed and summed to create the composite scores for verbal WM, visuospatial WM, and n-back.

Study II employed hierarchical multiple linear regression. In Step 1, we entered the background variables (age, education, childhood socioeconomic status, and state anxiety). Step 2 included the Big Five personality traits. We ran a separate multiple regression analysis for each of the three WM composite scores as dependent variables.

In **Study III**, following the factor analysis conducted on WM tasks, we examined the internal structure of PPS-4 and STAI-6 using exploratory factor analyses. We used linear mixed model analysis to investigate the effects of background factors (age, education, and childhood socioeconomic status) and mental health-related measures (stress and state anxiety) on WM performance as well as the interactions of those variables with WM domain (verbal, visuospatial, n-back). We compared Model 1 (including background factors) and Model 2 (background factors and mental health) to the null model, which included only participant random effects.

For **Study IV**, we first applied a genetic matching procedure (Diamond & Sekhon, 2006) to ensure equal distribution of background variables in the monolingual and bilingual groups. Then we conducted ANCOVAs with background factors (age, education, childhood socioeconomic status) as covariates, and language group (monolingual, early bilingual, late bilingual) as the independent variable. For each of the WM composite scores, we ran a separate analysis. Prompted by the results, we also ran ANCOVAs with the same setup for visuospatial and verbal n-back tasks separately. At the next step, we conducted a hierarchical regression analysis of predictors of n-back performance in the bilingual group. For this purpose, we used a within-subjects approach and considered all bilingual participants as one group. In Step 1, we entered the background variables. Step 2 included the bilingualism-related factors: L2 AoA, L2 proficiency and frequency of language switching. Due to the multiplicity of analyses in this study, we also conducted Bayesian analyses to minimize the risk for false positives.

3.2.3 Study V

Procedure
Each participant completed three separate sessions of the study, receiving sham, anodal and cathodal stimulation. Each session consisted of three ten-minute n-back task blocks: (1) without stimulation, (2) with stimulation or sham, and (3) without stimulation. The blocks were separated by self-timed pauses. There was a washout period of 48h or more between the sessions. The order of stimulation was counterbalanced.

Brain stimulation
We used the EMS BrainStim device with rubber electrodes placed in sponge pockets soaked in saline solution. The anode covered the entire left inferior frontal gyrus (IFG, or F7 in the 10-20 system). The reference electrode was placed over the right supraorbital cortex. Both anodal and cathodal stimulation was administered with a constant current of 1.5 mA. The stimulation lasted 10 minutes. In the sham condition, the participants experienced a 40-sec electric current, ramped up and down, at the beginning and the end of the block. The stimulation was double-blind.

Tasks

N-back task

In this study, we used a numerical-verbal 3-back task. The proper task was preceded by practice runs including 1-, 2-, and 3-back. The dependent variables were the RTs for correct responses and the accuracy rates measured as d-prime values.

Statistical analyses

We investigated whether brain stimulation had an effect on participants' RTs (average RTs from correct responses for both target and non-target stimuli) and accuracy (as measured by sensitivity index, d-prime). We conducted repeated-measures ANOVAs for RTs and d-primes. To analyze individual gains, we calculated gain scores for RTs and d-primes by subtracting the performance score in the first block from the scores in the second or third block. Then we conducted mixed-model ANOVAs with stimulation type (3 levels) as a within-subject factor, separately for each gain score type. Furthermore, we divided participants into high-performers and low-performers based on median split, and conducted mixed-model ANOVAs with stimulation type and block as within-subjects factors and performance group as a between-subjects factor. The ANOVAs were conducted separately for RTs and d-primes.

As regards electric current distribution in our montage, we used COMETS2 (Lee et al., 2017) for post-hoc modeling of current flow.

4. RESULTS

4.1 Study I

Study I revealed that, in most tasks, the easiest conditions (i.e., 1-back or single task trials) had the lowest reliability, followed by subtraction scores (i.e., n-back effect or switching cost). Difficult conditions had the highest reliability. Of all tasks, Simon task had the lowest reliability, ranging from r = .37 to r = .67, which indicated marginal reliability at best. On the other hand, the Letter-memory task had the highest reliability (r = .85 in the most difficult condition). For both visuospatial and verbal n-back tasks, RT measures for difficult conditions and the n-back effect had adequate to high reliability. For the five measures used in the Number-letter task, only RTs in repetition and in switch trials reached marginal reliability (r = .77 and r = .73, respectively). The additional LME analyses showed a decrease in RTs between sessions, and for some tasks we also observed within-session learning effects. Furthermore, in the visuospatial n-back task and the Number-letter task, we found a steeper decrease in RT in the easy conditions than in the difficult ones. Finally, in the Letter-memory task, the improvement in accuracy was observed within a session, but it did not carry over from pretest to posttest.

4.2 Study II

In Study II, the systematic review gave inconclusive evidence for the associations between Big Five personality traits and WM performance. Less than 15% of the included samples provided support for the association between WM performance and Extraversion, Conscientiousness, and Agreeableness: higher Extraversion and Agreeableness were related to better WM performance, while for Conscientiousness, we found evidence for both positive and negative correlations with WM. In turn, 22.2% of the samples indicated an association between higher Neuroticism and lower WM scores. For Openness, 26.1% of the samples revealed a positive association with WM performance, while one sample (4.3%) showed an opposite direction. The hierarchical multiple regression analysis of our own data showed that a model including the personality traits explained more variance than the null model only for the n-back composite score ($\Delta F(5, 493)=3.156$, $p=.008$, $\Delta R^2=0.03$). This analysis yielded statistically significant negative associations between WM updating performance and Conscientiousness and Openness.

4.3 Study III

Study III employed linear mixed modeling to investigate the effects of stress and anxiety on WM performance. We also included background variables in the models. A likelihood ratio test showed that Model 2, which included both variables of interest and background measures, provided the best fit for our data. With the marginal R^2_{GLMM} = .045, it explained 4.5% of the variance. We found statistically significant main effects of age and education, indicating that participants who were younger or reported a higher level of education also tended to score higher on the tasks. Wefound a trend towards a small negative main effect of anxiety: higher levels of state anxiety were associated with worse performance on the WM tasks. Contrary to our predictions, WM domain did not interact with the variables of interest, but we observed an interaction between age and domain, as higher age had more negative influence on visuospatial WM and n-back tasks, but did not affect verbal WM performance in the same way.

4.4 Study IV

In Study IV, we observed that early and late bilinguals outperformed monolinguals on visuospatial WM tasks. Moreover, age was a significant covariate, with lower age being associated with better visuospatial WM performance. However, Bayesian analysis showed that our data did not provide sufficient support for the bilingual advantage hypothesis (BF_{10} = 2.46). In the verbal WM domain, we observed no significant differences between the three groups (BF_{10} = .04). As regards WM updating, late bilinguals had a significant advantage over early bilinguals and monolinguals. Bayesian analysis yielded a Bayes Factor BF_{10} = 10.98, which indicated substantial support for our result. Prompted by these findings, we investigated the associations between background variables, bilingual experience and WM updating in hierarchical regression analyses. None of the bilingualism-related measures predicted n-back task performance. Furthermore, none of our models was supported by Bayesian analysis.

4.5 Study V

Study V, we found no effects of either anodal or cathodal stimulation on RTs or accuracy rates. However, we found main effects of block, which indicated a learning effect: participants became faster ($F_{(2, 60)}$ = 91.83, p < 0.001, ηp^2 = 0.754) and more accurate ($F_{(2, 60)}$ = 11.065, p < 0.001, ηp^2 = 0.269) as they grew more familiar with the task. An additional analysis of gain scores showed no influence of stimulation on participants' individual gains. When comparing high- and low-performers, we observed a trend towards significance in the interaction between group and simulation ($F_{(2, 54)}$ = 2.85, p = 0.066, ηp^2 = 0.096) for accuracy rates: while in the low-performing group the stimulation had no effect on performance, high-performers' scores were deteriorating under tDCS and recovering to original accuracy rates in the post-stimulation block.

5. DISCUSSION

The present thesis set out to investigate three important issues related to WM, all with scarce or inconclusive previous evidence. **Study I** assessed test-retest reliability and temporal stability of five commonly used EF tasks, including WM tasks. **Studies II-IV** examined the associations between WM task scores and several background factors in a large U.S.-based sample. The background factors were related to personality, psychological well-being (stress, state anxiety and subclinical depressive symptoms), and linguistic experience. Finally, **Study V** tested WM enhancement by brain stimulation using tDCS. The main findings can be summarized as follows:

1) *Reliability of EF tasks.* Of the five commonly used EF tasks, WM tasks had the highest test-retest reliability. All in all, RT measures appeared to be more reliable indicators than accuracy measures. In some tasks, within-session learning effects were observed.

2) *Background factors and WM.* The study on personality features found a small but statistically significant negative association between WM updating performance and Conscientiousness and Openness, suggesting that high Conscientiousness and Openness are associated with a lower WM updating performance. For psychological well-being, only a trend towards negative, across-the-board association between anxiety and WM performance was found. Concerning linguistic experience, late bilinguals outperformed early bilinguals and monolinguals on WM updating. However, this effect may be related to something else than bilingualism, as updating performance did not correlate with age of acquisition, level of language proficiency, or frequency of language switching in the bilinguals.

3) *Effects of brain stimulation on verbal WM.* The stimulation experiment failed to find any tDCS-related improvement in WM updating.

5.1 Psychometric properties of WM tasks and other EF measures (Study I)

This study investigated the test-retest reliability and temporal stability of commonly used executive tasks (Simon task, verbal and visuospatial n-back tasks, Letter memory task and Number-letter task). Despite their popularity, the psychometric properties of these tasks have not been thoroughly studied. The present analyses showed that most of the tasks had moderate test-retest reliabilities. It is worth noting that the highest test-retest reliabilities were observed for WM tasks (Letter memory, visuospatial and verbal n-back), which supports their continuous use. The reliabilities for the n-back task were comparable with those reported by Hockey & Geffen (2004). The Letter memory task had higher reliability coefficients than any of the n-back tasks. The two EF tasks (Simon task and Number-letter task) had overall low reliabilities, which could be partly due to the ceiling effect we observed in those measures (especially in the Simon task). The test-retest reliabilities for RT in congruent and incongruent conditions in the Simon task were comparable to those reported earlier (Wöstmann et al., 2013; Paap & Sawi, 2016). Interestingly, Wöstmann and colleagues (2013) had reported much higher test-retest reliability for the Simon RT effect than the other two studies (Paap & Sawi, 2016; Soveri et al., 2016). This could be explained by the different, possibly more difficult version of the Simon task used by Wöstmann and colleagues (2013). In the case of the traditional Simon task, the test-retest reliability for the Simon RT effect was low, which limits the adequacy of this measure (Paap & Sawi, 2016).

The present study showed significant variation in the stability of EF task scores over time. Low test-retest reliability of a task introduces noise to the data, which in turn entails that for between-group comparisons, the likelihood of detecting an effect will become lower (Kanyongo, Brook, Kyei-Blankson, & Gocmen, 2007). However, the small sample size in our study is a considerable limitation, as undersized studies have reduced chances of finding true effects; moreover, the likelihood that a significant effect in a small study is a true effect is also reduced (Button et al., 2013). Furthermore, test-retest reliability tends to become weaker with increasing length of time between pre- and posttest, as a longer interval can yield a greater difference between the two testing situations (Heise, 1969). Our participants had an interval of 3 or 6 weeks. All in all, this study would be best considered as a preliminary estimate of task reliabilities. There should be further research on psychometric properties of EF tasks that will contribute to our understanding of the processes affecting task scores and thereby their clinical usefulness.

5.2 Determinants of WM performance (Studies II-IV)

Study II found significant but weak associations between WM updating performance and two personality traits, namely Conscientiousness and Openness. Higher values of these traits were related to worse WM updating. Previous studies of personality and cognition have given inconclusive evidence, but often reported positive associations between Openness and cognition. On the other hand, the negative relationship between WM and Conscientiousness found in our study is in line with some previous studies (Soubelet, 2011; Schell & Reilley, 2004). Moutafi and colleagues (2004) suggested that people with lower cognitive abilities might develop higher Conscientiousness in order to compensate for those low abilities by developing persistent and disciplined work habits. This compensation hypothesis could explain our finding. As for the link between higher Openness and poorer WM updating performance, no existing theories give a plausible explanation. To the contrary, Openness, and its facet Intellect in particular, is usually associated with higher cognitive abilities (DeYoung et al., 2009). The negative relationship in our study is thus likely a chance finding. All in all, we observed only minuscule associations between Conscientiousness, Openness and WM performance.

As the existing body of literature provides inconclusive evidence for direct relationships between personality traits and cognition, a more fruitful approach might be to focus on attitudes and behaviors mediating personality-cognition associations. Chamorro-Premuzic and Furnham (2004) proposed a model in which Conscientiousness, Openness, Neuroticism and Extraversion affect the self-belief structure of subjectively assessed intelligence — the subjective perception of one's intellectual capacity. In turn, subjectively assessed intelligence affects test performance (for example, by contributing to test anxiety if one believes to be incapable) along with general and crystallized intelligence. In their model, the two types of intelligence are also related to Conscientiousness and Openness. This model has not yet been validated in the field of cognitive psychology. Another promising approach would be to study the specific facets of personality traits, as they may show associations with cognition that are lost in more general trait variables. This is because some of the facets are more directly associated with cognition than others; there is evidence that, for example, Ideas and Values facets of Openness correlate with WM updating (DeYoung et al., 2009).

In summary, the research on personality traits and cognition does not show robust associations between the two constructs. It is most likely that personality traits influence cognition indirectly, mediated by the individuals' behaviors and attitudes, while the direct links addressed in our study are weak, as in previous literature.

In Study III, only a trend towards negative relationship between anxiety and WM performance was found. Anxiety was linked to WM in the same manner in all three domains (verbal WM, visuospatial WM, updating), which suggests a general rather than domain-specific relationship.

Our results, albeit non-significant, are in line with the attentional control theory proposed by Eysenck and colleagues (2007), and the recent meta-analysis on WM and anxiety by Moran (2016). We found that more anxious individuals tended to obtain lower scores in all three domains of WM performance. Similar results have been reported by Petrac and colleagues (2009), who showed that cognitive performance fluctuates in response to stress and anxiety linked to recent life events.

One significant limitation of our study is the fact that the questionnaires measuring stress and anxiety used varying time scales — State/Trait Anxiety Inventory (short form; STAI-6) measures the feelings of anxiety in the given moment, while Perceived Stress Scale (PSS-4) taps stress experienced within last month. This makes the comparisons between the two variables of interest difficult. On the other hand, the measures we chose are commonly used and have good reliability and validity. Studying the more nuanced effects of anxiety and stress is, however, an interesting and relevant avenue for further research.

Our correlational setup makes it impossible to infer about causal relationships, but nevertheless indicates that currently experienced anxiety plays a role in healthy individuals' WM performance. Our findings suggest that when probing WM performance, one could consider taking transient anxiety into account as well.

Study IV addressed the relationships of bilingual experience and WM performance. We found evidence for a bilingual advantage in visuospatial WM (for early and late bilinguals) and WM updating (for late bilinguals) performance, although only the latter was supported by Bayesian analysis. Furthermore, we investigated whether task performance was related to language proficiency, frequency of language switching, and exposure to L2 in the bilingual group, but found no significant associations. If late bilinguals' better scores on WM updating tasks were indeed related to their linguistic experience, the effect should have been observed in both between-group (ANCOVAs) and within-group (linear regression) analyses. Thus, our study failed to find consistent support for the bilingual advantage hypothesis.

The importance of this study lies in comparing not only bilinguals to monolinguals, but also two groups of bilinguals: early and late, defined by their L2 AoA. L2 AoA is considered to determine not only the length of exposure to two languages, but also the mechanism of L2 learning (Birdsong, 2006), which entails partly different cognitive systems underlying language processing in early vs. late bilinguals. The other bilingual features we explored have also been considered to be significant factors in bilingual experience. There is evidence that L2 proficiency and language switching are related to cognitive performance (Luk et al., 2011; Soveri et al., 2011; Tao et al., 2011, Abutalebi et al., 2013; Verreyt et al., 2016). However, we found no association between bilingual experience and WM updating performance. In the context of the ongoing debate about the putative effects of bilingualism on cognition, our results point to the weakness of the bilingual advantage hypothesis, namely the unspecified conditions under which the advantage should emerge. For example, the ability to switch between two languages has been proposed as naturalistic training of EF, but it has been noted that sociolinguistic factors and the interactional context of code-switching can in certain situations decrease the cognitive load of switching (Costa et al., 2009;

Green & Abutalebi, 2013; Hartanto & Yang, 2016). Thus, when studying language switching, one should take these sociolinguistic and interactional aspects into account. Our study had some limitations that could have affected the results. As regards language switching, the switching question was not precise enough to capture the nuances of language use. It thus fails to give us the kind of detailed contextual information discussed by Green and Abutalebi (2013) and Costa and colleagues (2009). It should also be noted that we observed only small variation in our measure of language switching habits. The recent study by Hartanto and Yang (2016) suggests that the context of language use (for example, whether a bilingual person is used to frequent switching while talking to other bilinguals) modulates the cognitive demands, and thus not all language switching situations will be cognitively effortful. In fact, some studies have found that participants can switch within one sentence in order to decrease the cognitive load (Gollan, Kleinman, & Wierenga, 2014; Green & Abutalebi, 2013). Our data on the participants' switching was not detailed enough to discern between the different modes of switching and approximate the cognitive demand they posed.

We cannot exclude the possibility that late bilinguals' superior performance on WM updating tasks is related to some other background factor that we did not control for, for example the circumstances of language learning (such as classroom environment or learning the language at home). It is also possible that the higher scores in the late bilinguals are related to the cognitive "boost" experienced by late language learners due to cognitive demands posed by L2 use which has not yet been automatized (Paap, 2019). On the other hand, since we cannot establish causation in our correlative data, it could also be that participants with greater WM abilities were also more motivated to learn a second language (Linck et al., 2014). Furthermore, we did not measure our participants' nonverbal intelligence, which is related to WM updating (De Simoni & von Bastian, 2018). A promising direction for future studies would be to conduct longitudinal research to examine the temporal stability of bilingualism and L2 learning effects on cognition.

5.3 Brain stimulation and its effects on WM performance (Study V)

This study tested whether it would be possible to enhance verbal WM performance through tDCS on a brain region that is known to be important for this function. In an experimentally stringent double-blind setting, anodal, cathodal and sham stimulation effects on verbal WM were compared both online and offline, but no effects of stimulation on WM performance were found. It is however important to mention that our sample consisted of healthy university students who represent highly educated people at the peak of their cognitive performance. It might be that their WM performance could not be improved much further, which perhaps would not have been the case in, for example, clinical populations. Our null results could also have been due to some methodological aspects of our study, including the montage, stimulation target site, or the fact that we only tested participants under a steady cognitive load of 3-back in a cross-over design. We did observe substantial learning effects within and between sessions in our within-subject design, and it is possible that the task became too easy for our participants. Moreover, while after-effects of stimulation are commonly reported in the literature, there is no consensus regarding the timeframe (Ohn et al., 2008), and by measuring "offline" effects immediately post-stimulation we might have missed changes in performance that arose later.

Another possible reason for the null results is that group-level analysis would have masked individual stimulation effects. Previous studies have shown that tDCS effects are moderated by multiple factors, including individual differences (Meinzer et al., 2013; Kim et al., 2014; Fertonani & Miniussi, 2017), which prompted us to investigate the inter-individual variability in responsiveness to tDCS in our sample. The additional analyses of individual gain scores in high- versus low-performers yielded no statistically significant stimulation effects. However, in the subsample of high-performers, we observed a trend towards reduced accuracy during both cathodal and anodal stimulation. Nevertheless, there is previous research indicating that some subgroups may benefit more from tDCS (for example, older participants with lower baseline performance and hyperactivity in right prefrontal cortex, as in the study by Meinzer and colleagues, 2013). Nonetheless, in spite of the growing body of evidence, we are still unable to accurately predict who would respond favorably to stimulation by tDCS.

Overall, tDCS remains a promising, albeit volatile method that requires more research to establish setups that would provide reliable beneficial effects. Possible avenues of investigation include administering tasks with varying cognitive load to avoid possible ceiling effects, probing participants' background in more detail to better understand the inter-individual variability in responsiveness to tDCS, or comparing tDCS "offline" effects after a longer duration of time.

5.4 Conclusions

The studies included in this thesis addressed three themes in WM research where previous studies have been either scarce or yielded mixed results. In light of this, the inconclusiveness of some of the present findings is perhaps not so surprising. Nevertheless, there are relevant conclusions to be drawn from the present five studies.

Concerning methodology, Studies II-IV were conducted on an Internet sample. Having obtained reliable, high-quality data outside the traditional laboratory setting is a promising development in psychology that has been notoriously criticized for studying too small and homogenous samples. Our research supports the notion that Internet is a valuable tool to reach out to considerably larger and more diverse groups of participants.

Our findings on the associations between background factors and WM converge to some extent with the test-retest reliability study. There is evidence that cognitive performance can fluctuate with changes in experienced state anxiety (Petrac et al., 2009). Given the present finding of a negative (albeit weak) relationship between anxiety and WM performance, it is quite possible that anxiety-induced performance fluctuation is one factor affecting the temporal stability of WM and other tasks. While the WM tasks had highest test-retest reliabilities in our analysis, it would still be beneficial to control for current emotional state such as levels of anxiety and stress when studying WM and other EF performance in healthy participants. We have only begun to understand the ways psychosocial and environmental factors shape cognition, and the list of background factors potentially affecting current cognitive performance in a negative or positive fashion is probably much longer.

The small effects of background factors on WM in our studies could in part be explained by the compensatory effects of motivation and the fact that everyday experiences and behaviors as measured by questionnaires are not the same as performance features tapped by cognitive tests. It has been suggested that cognitive tasks used in psychological testing, applied for a brief time period in a controlled setting, are likely to measure maximal performance in the state of high motivation, which may not reflect typical performance (Toplak, West, & Stanovich, 2013). In other words, a person who struggles with impaired WM in daily life might still achieve good scores on a WM task when participating in a study. The interrelationships of motivation and EF have been extensively studied in the context of ADHD (Toplak, West, & Stanovich, 2013), and perhaps they should be brought to focus in studies of cognitively healthy populations as well.

In Studies II, III and IV, verbal WM performance was not related to background factors, and neither was it affected by tDCS in Study V. This may in part highlight the special role of verbal WM. There is a reciprocal relationship between verbal WM and education where high initial WM performance enables greater academic success (Alloway & Alloway, 2010), but also the practice of studying and learning allows one to use their WM resources more effectively. We are probably more likely to compensate with strategy use for WM lapses in the verbal domain, and the familiarity with task materials in verbal WM tasks decreases task novelty and thereby executive demands. Novel stimuli, as opposed to familiar ones, require different modes of processing as they might be not be compatible with pre-existing cognitive structures (Förster, Liberman, & Shapira, 2009).

Returning to the task impurity problem, especially with verbal WM one can question whether we are testing some well-delineated modular WM subcomponent, or rather a multi-faceted process that is entrenched in previously learned strategies and compensation mechanisms.

In turn, WM updating as measured by n-back tasks was most affected by background variables. Furthermore, the n-back tasks showed lower test-retest reliability than the verbal WM task (Letter Memory) in our battery in Study I, which could also be related to a higher sensitivity to various background factors. It has been suggested that n-back engages processes related to general intelligence (Jaeggi et al., 2010), and that it reflects the executive component of WM as opposed to storage (Colom et al., 2006). Overall, WM updating appears to stand out as a higher-order WM component, distinct from the "what" (verbal) and "where" (visuospatial) divisions of the WM system. It could also be that the distinct features of the n-back updating paradigm that separate it from the other WM paradigms used here, namely continuous encoding and response cycles, add to its sensitivity to various factors.

Despite less-than-perfect test-retest reliability and some influences from background factors discussed above, the overall picture painted by the present studies is that healthy adult WM performance, after all, is quite robust with regard to time, some important background factors, and brain stimulation (tDCS). This conclusion concurs with the notion of moderate-to-high heritability of WM (Conway, Kane, & Engle, 2003, Alloway & Alloway, 2010). One could speculate whether WM capacity is too crucial for individual success to be driven by environmental factors. If this is the case, one could question the rationale of existing WM training methods (Soveri et al., 2018). However, even if individual WM capacity would be more or less genetically determined, it should still be possible to enhance its use through compensatory strategies (e.g., Laine et al., 2018), and thereby try to overcome capacity limitations.

REFERENCES

Abutalebi, J., Della Rosa, P. A., Ding, G., Weekes, B., Costa, A., & Green, D. W. (2013). Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex*, 49(3), 905–911. https://doi.org/10.1016/j.cortex.2012.08.018

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, 80(2), 207–245. https://doi.org/10.3102/0034654310368803

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, 106(1), 20–29. https://doi.org/10.1016/j.jecp.2009.11.003

Alloway, T. P., Gathercole, S. E., Adams, A. M., Willis, C., Eaglen, R., & Lamont, E. (2005). Working memory and phonological awareness as predictors of progress towards early learning goals at school entry. *British Journal of Developmental Psychology*, 23(3), 417–426. https://doi.org/10.1348/026151005X26804

Anastasi, A. (1982). Psychological Testing (5th ed.), *MacMillan Publishing Co*, New York

Ando, J., Ono, Y., & Wright, M. J. (2001). Genetic structure of spatial and verbal working memory. *Behavior Genetics*, 31(6), 615–624. https://doi.org/10.1023/A:1013353613591

Archer, J. A., Lee, A., Qiu, A., & Chen, S. H. A. (2018). Working memory, age and education: A lifespan fMRI study. *PLoS ONE*, 13(3), 1–19. https://doi.org/10.1371/journal.pone.0194878

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*, 28(9), 1419–1432. https://doi.org/10.1162/jocn

Bäckman, L., Nyberg, L., Soveri, A., Johansson, J., Andersson, M., Dahlin, E., Neely, A.S., Virta, J., Laine, M., Rinne, J. O. (2011). Effects of working-memory training on striatal dopamine release. *Science*, 333(6043), 718.

https://doi.org/10.1126/science.1204978

Bäckman, L., Waris, O., Johansson, J., Andersson, M., Rinne, J. O., Alakurtti, K., Soveri, A., Laine, M., Nyberg, L. (2017). Increased dopamine release after working-memory updating training: Neurochemical correlates of transfer. *Scientific Reports*, 7(1), 1–10. https://doi.org/10.1038/s41598-017-07577-y

Baddeley, A. D., & Hitch, G. (1974). Working memory. In Bower, G.H. (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47-89). Academic Press.

Baddeley, A. (1998). Recent developments in working memory. *Current Opinion in Neurobiology*, 8(2), 234–238. https://doi.org/10.1016/S0959-4388(98)80145-1

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

Bertolino, A., Blasi, G., Latorre, V., Rubino, V., Rampino, A., Sinibaldi, L., ... & Nardini, M. (2006). Additive effects of genetic variation in dopamine regulating genes on working memory cortical activity in human brain. *Journal of Neuroscience*, 26(15), 3918-3922. https://doi.org/10.1523/JNEUROSCI.4975-05.2006

Bialystok, E., & Craik, F. I. M. (2010). Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science*, 19(1), 19–23. https://doi.org/10.1177/0963721409358571

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, 56, 9-49. https://doi.org/10.1111/j.1467-9922.2006.00353.x

Bordens, K. S. & Abbott, B. B. (2005). *Research Design and Methods - A Process Approach, 6th ed.* New York, McGraw-Hill Companies.

Botto, M., Basso, D., Ferrari, M., & Palladino, P. (2014). When working memory updating requires updating: Analysis of serial position in a running memory task. *Acta Psychologica*, 148, 123–129. https://doi.org/10.1016/j.actpsy.2014.01.012

Broadway, J. M., & Engle, R. W. (2010). Validating running memory span: Measurement of working memory capacity and links with fluid intelligence. *Behavior Research Methods*, 42(2), 563–570. https://doi.org/10.3758/BRM.42.2.563

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), 1–9. https://doi.org/10.1016/j.bandc.2014.01.008

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365-376. https://doi.org/10.1038/nrn3475

Chamorro-Premuzic, T., & Furnham, A. (2004). A possible model for understanding the personality-intelligence interface. *British Journal of Psychology*, 95(2), 249–264. https://doi.org/10.1348/000712604773952458

Chee, M. W., & Choo, W. C. (2004). Functional imaging of working memory after 24 hr of total sleep deprivation. *Journal of Neuroscience*, 24(19), 4560-4567. https://doi.org/10.1523/JNEUROSCI.0007-04.2004

Christopher, G., & MacDonald, J. (2005). The impact of clinical depression on working memory. *Cognitive Neuropsychiatry*, 10(5), 379–399. https://doi.org/10.1080/13546800444000128

Colom, R., Rebollo, I., Abad, F. J., & Shih, P. C. (2006). Complex span tasks, simple span tasks, and cognitive abilities: A reanalysis of key studies. *Memory & Cognition*, 34(1), 158-171. https://doi.org/10.3758/BF03193395

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., Heij, W. La, & Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *Journal of Experimental Psychology: Learning Memory and Cognition*, 34(2), 302–312. https://doi.org/10.1037/0278-7393.34.2.302

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, 7(12), 547–552.

https://doi.org/10.1016/j.tics.2003.10.005

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769-786. https://doi.org/10.3758/BF03196772

Conway, A. R., Macnamara, B. N., & de Abreu, P. M. E. (2013). Working memory and intelligence: An overview. In Alloway, T.P., & Alloway, R. G. (Eds.), *Working Memory* (pp. 27-50). New York: Psychology Press.

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, 113(2), 135-149. https://doi.org/10.1016/j.cognition.2009.08.001

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, 50(4), 491-511. https://doi.org/10.1016/j.jml.2004.02.002

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, 19(1), 51–57. https://doi.org/10.1177/0963721409359277

Curtis, R. G., Windsor, T. D., & Soubelet, A. (2015). The relationship between Big-5 personality traits and cognitive ability in older adults — A review. *Aging, Neuropsychology, and Cognition*, 22(1), 42–71. https://doi.org/10.1080/13825585.2014.888392

De Simoni, C., & von Bastian, C. C. (2018). Working memory updating and binding training: Bayesian evidence supporting the absence of transfer. *Journal of Experimental Psychology: General*, 147(6), 829–858. https://doi.org/10.1037/xge0000453

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

DeYoung, C. G., & Gray, J. R. (2009). Personality neuroscience: Explaining individual differences in affect, behaviour and cognition. In Corr, P.J., Matthews, G. (Eds.), *The Cambridge Handbook of Personality Psychology*, 323-346. New York: Cambridge University Press.

Diamond, A., & Sekhon, J. S. (2013). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. *Review of Economics and Statistics*, 95(3), 932-945. https://doi.org/10.1162/REST_a_00318

Dick, B., Eccleston, C., & Crombez, G. (2002). Attentional functioning in fibromyalgia, rheumatoid arthritis, and musculoskeletal pain patients. *Arthritis & Rheumatism*, 47(6), 639–644. https://doi.org/10.1002/art.10800

Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: an experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning Memory and Cognition*, 36(1), 170–189. https://doi.org/10.1037/a0017891

Evans, G. W., & Schamberg, M. A. (2009). Childhood poverty, chronic stress, and adult working memory. *Proceedings of the National Academy of Sciences (USA)*, 106(16), 6545–6549. https://doi.org/10.1073/pnas.0811910106

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, 7(2), 336–353. https://doi.org/10.1037/1528-3542.7.2.336

Eysenck, M. W., Payne, S., & Derakshan, N. (2005). Trait anxiety, visuospatial processing, and working memory. *Cognition and Emotion*, 19(8), 1214–1228. https://doi.org/10.1080/02699930500260245

Fertonani, A., & Miniussi, C. (2017). Transcranial electrical stimulation: What we know and do not know about mechanisms. *Neuroscientist*, 23(2), 109–123. https://doi.org/10.1177/1073858416631966

Förster, J., Liberman, N., & Shapira, O. (2009). Preparing for novel versus familiar events: shifts in global and local processing. *Journal of Experimental Psychology: General*, 138(3), 383–399. https://doi.org/10.1037/a0015748

Fritsch, B., Reis, J., Martinowich, K., Schambra, H. M., Ji, Y., Cohen, L. G., & Lu, B. (2010). Direct current stimulation promotes BDNF-dependent synaptic plasticity: Potential implications for motor learning. *Neuron*, 66(2), 198–204. https://doi.org/10.1016/j.neuron.2010.03.035

Gevins, A., & Smith, M. E. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex*, 10(9), 829-839. https://doi.org/10.1093/cercor/10.9.829

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(1), 26-34. http://dx.doi.org/10.1037/0003-066X.48.1.26

Gollan, T. H., Kleinman, D., & Wierenga, C. E. (2014). What's easier: Doing what you want, or being told what to do? Cued versus voluntary language and task switching. *Journal of Experimental Psychology: General*, 143(6), 2167-2195. http://doi.org/10.1037/a0038006

Gonçalves, M. M., Pinho, M. S., & Simões, M. R. (2016). Test–retest reliability analysis of the Cambridge neuropsychological automated tests for the assessment of dementia in older people living in retirement homes. *Applied Neuropsychology: Adult*, 23(4), 251–263. https://doi.org/10.1080/23279095.2015.1053889

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1(2), 67-81. https://doi.org/10.1017/S1366728998000133

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, 25(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Gropper, R. J., & Tannock, R. (2009). A pilot study of working memory and academic achievement in college students with ADHD. *Journal of Attention Disorders*, 12(6), 574–581. https://doi.org/10.1177/1087054708320390

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, 33(3), 325–340. https://doi.org/10.1177/0267658316678286

Hartanto, A., & Yang, H. (2016). Disparate bilingual experiences modulate task-switching advantages: A diffusion-model analysis of the effects of interactional context on switch costs. *Cognition*, 150, 10–19. https://doi.org/10.1016/j.cognition.2016.01.016

Heise, D. R. (1969). Separating reliability and stability in test-retest correlation. *American Sociological Review*, (34-1), 93-101.

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

Hockey, A., & Geffen, G. (2004). The concurrent validity and test-retest reliability of a visuospatial working memory task. *Intelligence*, 32(6), 591–605. https://doi.org/10.1016/j.intell.2004.07.009

Horvath, J. C., Forte, J. D., & Carter, O. (2015). Quantitative review finds no evidence of cognitive effects in healthy populations from single-session transcranial direct current stimulation (tDCS). *Brain Stimulation*, 8(3), 535–550. https://doi.org/10.1016/j.brs.2015.01.400

Horvath, J. C., Forte, J. D., & Carter, O. (2015). 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. *Neuropsychologia*, 66, 213–236. https://doi.org/10.1016/j.neuropsychologia.2014.11.021

Jacobson, L., Koslowsky, M., & Lavidor, M. (2012). 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., 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

Jaeggi, S. M., Buschkuehl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory and Cognition*, 42(3), 464–480. https://doi.org/10.3758/s13421-013-0364-z

Juan, C. H., Tseng, P., & Hsu, T. Y. (2017). Elucidating and modulating the neural correlates of visuospatial working memory via noninvasive brain stimulation. *Current Directions in Psychological Science*, 26(2), 165–173. https://doi.org/10.1177/0963721416677095

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

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the n-back task: a question of construct validity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 33(3), 615–622. https://doi.org/10.1037/0278-7393.33.3.615

Kanyongo, G. Y., Brook, G. P., Kyei-Blankson, L., & Gocmen, G. (2007). Reliability and statistical power: how measurement fallibility affects power and required sample sizes for several parametric and nonparametric statistics. *Journal of Modern Applied Statistical Methods*, 6(1), 81–90. https://doi.org/10.22237/jmasm/1177992480

Katz, B., Au, J., Buschkuehl, M., Abagis, T., Zabel, C., Jaeggi, S. M., & Jonides, J. (2017). Individual differences and long-term consequences of tDCS-augmented cognitive training. *Journal of Cognitive Neuroscience*, 29(9), 1498-1508. https://doi.org/10.1162/jocn_a_01115

Kessler, Y., & Meiran, N. (2008). Two dissociable updating processes in working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 34(6), 1339–1348. https://doi.org/10.1037/a0013078

Kim, S., Stephenson, M. C., Morris, P. G., & Jackson, S. R. (2014). TDCS-induced alterations in GABA concentration within primary motor cortex predict motor learning and motor memory: A 7T magnetic resonance spectroscopy study. *NeuroImage*, 99, 237–243. https://doi.org/10.1016/j.neuroimage.2014.05.070

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, 55(4), 352.

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

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, 302(5648), 1181-1185. https://doi.org/10.1126/science.1088545

Lee, C., Jung, Y. J., Lee, S. J., & Im, C. H. (2017). COMETS2: An advanced MATLAB toolbox for the numerical analysis of electric fields generated by transcranial direct current stimulation. *Journal of Neuroscience Methods*, 277, 56–62. https://doi.org/10.1016/j.jneumeth.2016.12.008

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, 144(4), 394-425. http://dx.doi.org/10.1037/bul0000142

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current Directions in Psychological Science*, 20(4), 240–245. https://doi.org/10.1177/0963721411415340

Lu, C. H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review*, 2(2), 174-207. https://doi.org/10.3758/BF03210959

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, 17(8), 391–400. https://doi.org/10.1016/j.tics.2013.06.006

Luk, G., De Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism: Language and Cognition*, 14(4), 588–595.

https://doi.og/10.1017/S1366728911000010

Luo, L., Craik, F. I. M., & Moreno, S. (2013). Bilingualism interacts with domain in a working memory task: Evidence from aging. *Psychology and Aging*, 28(1), 28–34. https://doi.org/10.1037/a0030875

Macizo, P., Bajo, T., & Cruz Martín, M. (2010). Inhibitory processes in bilingual language comprehension: Evidence from Spanish-English interlexical homographs. *Journal of Memory and Language*, 63(2), 232–244. https://doi.org/10.1016/j.jml.2010.04.002

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*, 28(8), 1063–1089. https://doi.org/10.1162/jocn_a_00956

McEwen, B. S., & Morrison, J. H. (2013). The brain on stress: vulnerability and plasticity of the prefrontal cortex over the life course. *Neuron*, 79(1), 16–29. https://doi.org/10.1016/j.neuron.2013.06.028

Meinzer, M., Lindenberg, R., Antonenko, D., Flaisch, T., & Flo, A. (2013). Anodal transcranial direct current stimulation temporarily reverses age-associated cognitive decline and functional brain activity changes. *Journal of Neuroscience*, 33(30), 12470–12478. https://doi.org/10.1523/JNEUROSCI.5743-12.2013

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, 130(4), 621–640. https://doi.org/10.1037/0096-3445.130.4.621

Monaco, M., Costa, A., Caltagirone, C., & Carlesimo, G. A. (2013). Forward and backward span for verbal and visuo-spatial data: Standardization and normative data from an Italian adult population. *Neurological Sciences*, 34(5), 749–754. https://doi.org/10.1007/s10072-012-1130-x

Monsell, S. (1996). Control of mental processes. In: Bruce, V. (Ed.), *Unsolved Mysteries of the Mind: Tutorial Essays in Cognition* (pp. 93–148). Hove: Taylor & Francis

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, 114(2), 187–202. https://doi.org/10.1016/j.jecp.2012.09.002

Moran, T. P. (2016). Anxiety and working memory capacity: A meta-analysis and narrative review. *Psychological Bulletin*, 142(8), 831–864. https://doi.org/10.1037/bul0000051

Moutafi, J., Furnham, A., & Paltiel, L. (2004). Why is Conscientiousness negatively correlated with intelligence? *Personality and Individual Differences*, 37(5), 1013–1022. https://doi.org/10.1016/j.paid.2003.11.010

Nguyen, H. T., Quandt, S. A., Grzywacz, J. G., Chen, H., Galván, L., Kitner-Triolo, M. H., & Arcury, T. A. (2012). Stress and cognitive function in Latino farmworkers. *American Journal of Industrial Medicine*, 55(8), 707–713. https://doi.org/10.1002/ajim.22035

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., Liebetanz, D., Antal, A., Lang, N., Tergau, F., & Paulus, W. (2003). Modulation of cortical excitability by weak direct current stimulation–technical, safety and functional aspects. In Paulus, W., Tergau, F., Nitsche, M.A., Rothwell, J.G., Ziemann, U., Hallett, M. (Eds.), *Supplements to Clinical Neurophysiology* (Vol. 56, pp. 255-276). https://doi.org/10.1016/S1567-424X(09)70230-2

Nunnally, J. (1978). *Psychometric Methods*. New York: McGraw-Hill.

Oberauer, K., Süß, H. M., Wilhelm, O., & Wittman, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, 31(2), 167–193. https://doi.org/10.1016/S0160-2896(02)00115-0

Oberauer, K., Wilhelm, O., Schulze, R., & Süß, H. M. (2005). Working memory and intelligence -Their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, 131(1), 61–65. https://doi.org/10.1037/0033-2909.131.1.61

Ohn, S. H., Park, C. Il, 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

Osborne, J. W. (2010). Improving your data transformations: Applying the Box-Cox transformation. *Practical Assessment, Research & Evaluation*, 15(12). Available online: http://pareonline.net/getvn.asp?v=15&n=12.

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

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66(2), 232–258. https://doi.org/10.1016/j.cogpsych.2012.12.002

Paap, K. R., & Oliver, S. (2016). The role of test-retest reliability in measuring individual and group differences in executive functioning. *Journal of Neuroscience Methods*, 274, 81–93. https://doi.org/10.1016/j.jneumeth.2016.10.002

Paap K.R., (2019). The bilingual advantage debate: Quantity and quality of the evidence. In: Schwieter JW (Ed.), *The Handbook of the Neuroscience of Multilingualism* (pp. 701-735). Hoboken, NJ: Wiley-Blackwell

Park, S., & Gooding, D. C. (2014). Working memory impairment as an endophenotypic marker of a schizophrenia diathesis. *Schizophrenia Research: Cognition*, 1(3), 127–136.

https://doi.org/10.1016/j.scog.2014.09.005

Paunonen, S. V., & Ashton, M. C. (2001). Big Five predictors of academic achievement. *Journal of Research in Personality*, 35(1), 78–90. https://doi.org/10.1006/jrpe.2000.2309

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., & Fazio, F. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping*, 19(3), 170–182. https://doi.org/10.1002/hbm.10110

Petrac, D. C., Bedwell, J. S., Renk, K., Orem, D. M., & Sims, V. (2009). Differential relationship of recent self-reported stress and acute anxiety with divided attention performance. *Stress*, 12(4), 313–319. https://doi.org/10.1080/10253890802380714

Prior, A., & Gollan, T. H. (2011). Good language-switchers are good task-switchers: Evidence from Spanish–English and Mandarin–English bilinguals. *Journal of the International Neuropsychological Society*, 17(4), 682-691. https://doi.org/10.1017/S1355617711000580

Prior, A., & Macwhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, 13(2), 253–262. https://doi.org/10.1017/S1366728909990526

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality. *Perspectives on Psychological Science*, 2(4), 313–345. https://doi.org/10.1111/j.1745-6916.2007.00047.x

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124(2), 207-231. https://doi.org/10.1037/0096-3445.124.2.207

Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., ... & Thase, M. E. (2003). The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*, 54(5), 573-583. https://doi.org/10.1016/S0006-3223(02)01866-8

Salazar-Villanea, M., Liebmann, E., Garnier-Villarreal, M., Montenegro-Montenegro, E., & Johnson, D. K. (2015). Depressive symptoms affect working memory in healthy older adult Hispanics. *Journal of Depression & Anxiety*, 4(4). https://doi.org/10.4172/2167-1044.1000204

Schell, K. L., & Reilley, S. P. (2004). Quality control pharmacy tasks: Big five personality model and accuracy of error detection. *Psychological Reports*, 94(3), 1301-1311. https://doi.org/10.2466/pr0.94.3c.1301-1311

Schmiedek, F., Hildebrandt, A., Lövdén, M., Wilhelm, O., & Lindenberger, U. (2009). Complex span versus updating tasks of working memory: the gap is not that deep. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 35(4), 1089–1096. https://doi.org/10.1037/a0015730

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2014). A task is a task is a task: putting complex span, n-back, and other working memory indicators in psychometric context. *Frontiers in Psychology*, 5(December). https://doi.org/10.3389/fpsyg.2014.01475

Schoofs, D., Preuß, D., & Wolf, O. T. (2008). Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology*, 33(5), 643–653. https://doi.org/10.1016/j.psyneuen.2008.02.004

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, 125(1), 4-27. https://doi.org/10.1037/0096-3445.125.1.4

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In Proctor R.W., Reeve, G.T. (Eds.), *Advances in Psychology* (Vol. 65, pp. 31-86). Elsevier B.V.

Sliwinski, M. J., Smyth, J. M., Hofer, S. M., & Stawski, R. S. (2006). Intraindividual coupling of daily stress and cognition. *Psychology and Aging*, 21(3), 545-557. https://doi.org/10.1037/0882-7974.21.3.545

Sommerfield, A. J., Deary, I. J., McAulay, V., & Frier, B. M. (2003). Moderate hypoglycemia impairs multiple memory functions in healthy adults. *Neuropsychology*, 17(1), 125–132. https://doi.org/10.1037/0894-4105.17.1.125

Soubelet, A. (2011). Age-cognition relations and the personality trait of Conscientiousness. *Journal of Research in Personality*, 45(6), 529–534. https://doi.org/10.1016/j.jrp.2011.06.007

Soveri, A., Lehtonen, M., Karlsson, L. C., Lukasik, K., Antfolk, J., & Laine, M. (2018). Test–retest reliability of five frequently used executive tasks in healthy adults. *Applied Neuropsychology: Adult*, 25(2). https://doi.org/10.1080/23279095.2016.1263795

Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. *Psychonomic Bulletin and Review*, 24(4), 1077–1096. https://doi.org/10.3758/s13423-016-1217-0

Soveri, A., Rodriguez-Fornells, A., & Laine, M. (2011). Is there a relationship between language switching and executive functions in bilingualism? Introducing a within-group analysis approach. *Frontiers in Psychology*, 2(August), 1–8. https://doi.org/10.3389/fpsyg.2011.00183

Soveri, A., Waris, O., & Laine, M. (2013). Set shifting training with categorization tasks. *PLoS ONE*, 8(12). https://doi.org/10.1371/journal.pone.0081693

Stagg, C. J., Antal, A., & Nitsche, M. A. (2018). Physiology of Transcranial Direct Current Stimulation. *Journal of ECT*, 34(3), 144–152. https://doi.org/10.1097/YCT.0000000000000510

Stocco, A., & Prat, C. S. (2014). Bilingualism trains specific brain circuits involved in flexible rule selection and application. *Brain and Language*, 137, 50–61. https://doi.org/10.1016/j.bandl.2014.07.005

Szmalec, A., Verbruggen, F., Vandierendonck, A., & Kemps, E. (2011). Control of interference during working memory updating. *Journal of Experimental Psychology: Human Perception and Performance*, 37(1), 137-151. http://dx.doi.org/10.1037/a0020365

Tao, L., Marzecová, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: The role of age of acquisition. *Frontiers in Psychology*, 2(June), 1–19.

https://doi.org/10.3389/fpsyg.2011.00123

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, 54(2), 131-143. https://doi.org/10.1111/jcpp.12001

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, 114(1), 104–132. https://doi.org/10.1037/0033-295X.114.1.104

Vandierendonck, A., Kemps, E., Fastame, M. C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology*, 95(1), 57–79. https://doi.org/10.1348/000712604322779460

Verreyt, N., Woumans, E., Vandelanotte, D., Szmalec, A., & Duyck, W. (2016). The influence of language-switching experience on the bilingual executive control advantage. *Bilingualism: Language and Cognition*, 19(1), 181–190. https://doi.org/10.1017/S1366728914000352

von Bastian, C. C., & Oberauer, K. (2013). Effects and mechanisms of working memory training: a review. *Psychological Research*, 78(6), 803–820. https://doi.org/10.1007/s00426-013-0524-6

von Stumm, S., Chamorro-Premuzic, T., & Ackerman, P. L. (2011). Re-visiting intelligence–personality associations: vindicating intellectual investment. In von Stumm, S., Chamorro-Premuzic, T., & Furnham, A. (Eds.), *The Wiley-Blackwell Handbook of Individual Differences* (pp. 217-241). Malden, MA: Wiley-Blackwell.

Vytal, K. E., Cornwell, B. R., Letkiewicz, A. M., Arkin, N. E., & Grillon, C. (2013). The complex interaction between anxiety and cognition: insight from spatial and verbal working memory. *Frontiers in Human Neuroscience*, 7(March), 1–11. https://doi.org/10.3389/fnhum.2013.00093

Vytal, K., Cornwell, B., Arkin, N., & Grillon, C. (2012). Describing the interplay between anxiety and cognition: From impaired performance under low cognitive load to reduced anxiety under high load. *Psychophysiology*, 49(6), 842–852. https://doi.org/10.1111/j.1469-8986.2012.01358.x

Waris, O., Soveri, A., Ahti, M., Hoffing, R. C., Ventus, D., Jaeggi, S. M., … Laine, M. (2017). A latent factor analysis of working memory measures using large-scale data. *Frontiers in Psychology*, 8(June), 1062. https://doi.org/10.3389/fpsyg.2017.01062

Waters, G. S., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, and Computers*, 35(4), 550–564. https://doi.org/10.3758/BF03195534

Wöstmann, N. M., Aichert, D. S., Costa, A., Rubia, K., Möller, H. J., & Ettinger, U. (2013). Reliability and plasticity of response inhibition and interference control. *Brain and Cognition*, 81(1), 82–94. https://doi.org/10.1016/j.bandc.2012.09.010

Wu, Y. J., Tseng, P., Chang, C. F., Pai, M. C., Hsu, K. Sen, Lin, C. C., & Juan, C. H. (2014). Modulating the interference effect on spatial working memory by applying transcranial direct current stimulation over the right dorsolateral prefrontal cortex. *Brain and Cognition*, 91, 87–94. https://doi.org/10.1016/j.bandc.2014.09.002

ISBN 978-952-12-3802-4