EEG Frequency Band Power and Pupil Dilation as Measures for Executive Working Memory Load in Task Settings of Basic and Applied Research

Dissertation
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Christian Scharinger, M.A.
aus Villingen-Schwenningen

Tübingen
2015

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Prüfung:	30.10.2015
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Peter Gerjets
2. Berichterstatter:	Prof. Dr. Torsten Schubert

for

Paula Schreitmüller

just three words:

curiosity – discipline – realism

Acknowledgments

Doing a PhD seldom is an easy job. Asking interdisciplinary questions and addressing topics from basic to applied research does not make things any easier. If you want to succeed in such a project, you need the support of the right people at the right time. Luckily, I had these people. Named or unnamed below, I would like to express those people my deep thankfulness!

First of all, I would like to thank my supervisors Prof. Dr. Peter Gerjets and Prof. Dr. Torsten Schubert. Peter Gerjets served as a valuable, rich source of ideas and suggestions at each stage of the PhD project. Given his unbound interdisciplinary thinking, I really learned what freedom of science can mean. Peter, thank you for stimulating discussions and for giving me the opportunity to pursuit my research interests! Torsten Schubert stepped in the PhD project just at the right time when I really began to experience the difficulties of doing interdisciplinary research. He helped me to find suitable journals for the basic research topics of my PhD project and to develop an adequate writing style. Torsten, many thanks for your willingness to join and support the PhD project and especially for your always appreciating and motivating feedback!

My deep appreciations also go to Dr. Alexander Soutschek and Dr. Yvonne Kammerer. They introduced me in the tools of doing science. Many thanks for always being available and for your help in analyzing and interpreting data, writing manuscripts, and dealing with reviewers, to name just some examples!

I would also like to thank Carmen, Hannah, Gabi, Romy, Claudi, and all other colleagues (especially those of AG1) for helpful discussions, support, and, most important, for sharing lunch times. If not already finished, I wish you all the best for your PhDs.

Yvonne, Peter, and Sven were willing to read earlier versions of my PhD manuscript. Thank you so much for giving me very constructive feedback that really helped me to improve the writing and argumentation. I would also like to thank Prof. Dr. Markus Huff for reading my annual PhD reports and providing helpful feedback.

Very helpful for pursuing a PhD project is also an adequate funding. Luckily, our research proposal for a project within the second funding period of the Leibniz ScienceCampus Tübingen got accepted, thus extending my funding for the entire PhD period. I would like to thank my co-authors of this research proposal, Carina Walter, Gabi Cierniak, Prof. Dr. Peter Gerjets, Prof. Dr. Martin Bogdan, and Prof. Dr. Wolfgang Rosenstiel.

Last not least I would like to thank my parents, brother, and friends for any support and, most of all, for sometimes not asking me too much about the ongoing PhD project.

One final thank goes to Prof. Dr. Stefan Schweinberger and Dr. Romi Zäske. Your enthusiasm of doing science finally convinced me to actually start a PhD project!

Summary

Performing a task, for example the selection of links during hypertext reading, raises cognitive demands, that is, induces cognitive load. Cognitive load may be attributed to working memory (WM) functioning. Core executive functions (EFs), like updating, shifting, and inhibition, may define the executive (i.e., working) part of WM and thus may be essential for WM functioning. Consequently, cognitive load may result out of the demands on core EFs during task performance. Brain oscillatory activity as captured in the electroencephalogram (EEG) by frequency band power in time-frequency representations (TFRs) might serve as a measure for demands on core EFs. Especially, EEG alpha frequency band power could serve as a global measure of the cognitive load-situation which grounds in demands on core EFs. Pupil dilation, which can be more efficiently acquired as compared to the EEG, might also be used as a measure of the overall load-situation.

Yet, the role of core EFs for WM functioning has been rarely studied to date, and the interplay of core EFs and their relation to WM is still matter of debate. Most important, the use of the physiological measures EEG alpha frequency band power and pupil dilation for assessing demands on core EFs and the global cognitive load-situation in complex, real-world task settings of hypermedia environments have to be studied further. Especially, possible non-cognitive factors have to be ruled out which might confound the physiological measures. In the current doctoral thesis, we therefore addressed in three studies ranging from basic to applied research 1) the relationship of core EFs and WM, 2) the interplay of core EFs, and 3) the use of EEG alpha frequency band power and pupil dilation as measures for increased demands on core EFs in a task setting of applied research, namely for hypertext reading and link selection.

In Study 1, we addressed the relationship of core EFs and WM by comparing two commonly used WM tasks, an n-back and an operation span task, and a simple digit span task that is generally considered as a short-term memory task. Conceptually, the two WM tasks should comprise demands on all three core EFs, whereas the simple digit span task might mainly demand the EF updating. Overall, the outcomes of Study 1 revealed that EEG correlates were more similar between the n-back and the operation span task as compared to

the simple digit span task, thus confirming the conceptual similarities between the two WM tasks.

In Study 2, we addressed the interplay of core EFs like updating and inhibition which might be due to a common underlying factor of controlled attention. We manipulated demands on two core EFs, updating and inhibition, within one single WM task. This was done by using congruent and incongruent flanker items (inhibitory demands) as stimuli in an n-back task paradigm (WM updating demands). The outcomes of Study 2 revealed a decreased flanker interference effect under severe load on updating for most of the load-related measures (P300, alpha frequency band power, pupil dilation), indicating the activity of an underlying common network structure which might serve processes of controlled attention and thus might enhance inhibitory control under severe load on WM updating.

In Study 3 a)-c), we addressed the research question whether alpha frequency band power and pupil dilation could serve as comparable measures for demands on core EFs in a complex, real-world task setting of hypertext reading and link selection. Importantly, we carefully ruled out possible perceptual-motor confounds which often hamper the interpretability of hypertext studies using physiological measures. Overall, our results showed that both physiological measures can be used to assess changes in the load-situation during link-selection processes. Surprisingly however, albeit showing a comparable result pattern, the two measures did not correlate with one another. At this point we can only speculate about this rather unexpected outcome. Clearly, more research is necessary on this.

In sum, we were interested in alpha frequency band power as an overall measure of cognitive load which may be grounded in demands on core EFs. Thus, alpha frequency band power served as dependent measure in all three studies and was supplemented by beta frequency band power, theta frequency band power, P300, and pupil dilation in some studies. Each study might serve as an initial step for conducting further research on core EFs using physiological measures in each specific task setting we used. Overall, the outcomes underline the use of physiological measures like alpha frequency band power and pupil dilation for studying core EFs in task settings of basic and applied research. EEG alpha frequency band power can be used as a measure of cognitive load which grounds in demands on core EFs.

Zusammenfassung

Das Ausführen einer komplexen Aufgabe, wie beispielsweise das Auswählen von Links beim Lesen einer Hypertextseite, stellt kognitiven Anforderungen und führt damit zu erhöhter kognitiver Belastung. Kognitive Belastung kann der Funktion des Arbeitsgedächtnisses zugeschrieben werden. Grundlegenden exekutiven Funktionen (EF) wie 'updating', 'shifting' und 'inhibition' definieren den exekutiven Teil des Arbeitsgedächtnisses. Sie sind damit zentral für die Funktion des Arbeitsgedächtnisses. Folglich sollte sich kognitive Belastung auf Anforderungen an EF während der Durchführung einer komplexen Aufgabe zurückführen lassen. Oszillatorische Hirnaktivität, die sich beispielsweise mittels des Elektroenzephalogramms (EEG) als Frequenzbandpower über die Zeit in Zeit-Frequenz-Darstellungen erfassen lässt, könnte zur Messung der Anforderungen an EF dienen. Insbesondere die EEG alpha-Frequenzbandpower könnte als ein Maß der globalen kognitiven Belastungssituation dienen, die in Anforderungen an EF begründet ist. Die Pupillendilatation, die im Vergleich zu EEG-Daten effizienter zu erheben ist, könnte gleichfalls als ein Maß der globalen Belastungssituation genutzt werden.

Allerdings wurde die Rolle, die EF für das Funktionieren des Arbeitsgedächtnisses spielen, bisher selten experimentell untersucht, und das Zusammenspiel der EF und ihr Zusammenhang mit dem Arbeitsgedächtnis wird nach wie vor kontrovers diskutiert. Vor allem scheinen weitere Untersuchungen nötig, die den Einsatz der physiologischen Maße EEG alpha-Frequenzbandpower und Pupillendilatation als Maße der Belastung von EF in komplexen, möglichst realistischen Aufgaben in hypermedialen Umgebungen zum Gegenstand haben. Insbesondere sind in diesem Kontext Faktoren auszuschließen, die die physiologischen Maße beeinflussen können, jedoch nicht kognitiver Natur sind. In der vorliegenden Dissertation adressierten wir deswegen in drei Studien folgende Forschungsfragen, die von der Grundlagenforschung zur Anwendungsforschung reichen: 1) der Zusammenhang von EF und dem Arbeitsgedächtnis, 2) das Wechselspiel zwischen EF und 3) den Einsatz von EEG alpha-Frequenzbandpower und Pupillendilatation als Maße für erhöhte Anforderungen an EF in einem Aufgabensetting der Anwendungsforschung, nämlich für die Auswahlprozesse von (Hyper-) Links während des (Hypertext-) Lesens.

In Studie 1 adressierten wir die Beziehung von EF und dem Arbeitsgedächtnis, indem wir zwei häufig genutzte Arbeitsgedächtnisaufgaben, eine sogenannte 'n-back'-Aufgabe und eine 'operation span'-Aufgabe, und eine Kurzzeitgedächtnisaufgabe, eine 'digit span'-Aufgabe, miteinander verglichen. Die beiden Arbeitsgedächtnisaufgaben sollten konzeptuell betrachtet Anforderungen an alle drei EF stellen, wohingegen die 'digit span'-Aufgabe nur die EF 'updating' erfordern sollte. Insgesamt zeigten die Ergebnisse von Studie 1, dass die untersuchten EEG-Korrelate ähnlicher waren zwischen der 'n-back'- und der 'operation span'-Aufgabe, was die konzeptuelle Ähnlichkeit dieser Arbeitsgedächtnisaufgaben bestätigte.

In Studie 2 adressierten wir den Zusammenhang zwischen den EF 'updating' und 'inhibition', der auf einen möglicherweise zugrundeliegenden gemeinsamen Faktor von kontrollierter Aufmerksamkeit zurückzuführen ist. Wir manipulierten hierzu die Anforderungen an zwei EF, 'updating' und 'inhibition' innerhalb einer Arbeitsgedächtnisaufgabe. Dies wurde erreicht durch die Nutzung kongruenter und inkongruenter Flanker-Stimuli (Anforderungen an die inhibitorische Kontrolle) als Stimuli in einem 'n-back'-Aufgabenparadigma (Anforderungen an 'updating'). Die Ergebnisse von Studie 2 zeigten einen verringerten Flanker-Interferenzeffekt unter hoher 'updating'-Belastung. Dies zeigte sich auf den meisten der untersuchten Maße (P300, alpha-Frequenzbandpower, Pupillendilatation). Die Ergebnisse weisen auf die Aktivität eines zugrundeliegenden gemeinsamen neuronalen Netzwerks hin, das möglicherweise der Aufmerksamkeitskontrolle dient und das, wenn einmal aktiviert (z.B. durch die 'updating'-Belastung), die inhibitorische Kontrolle verstärken kann.

In Studie 3 a)-c) adressierten wir die Forschungsfrage ob alpha-Frequenzbandpower und Pupillendilatation vergleichbar als Maße für Anforderungen an EF nutzbar sind und zwar in einem komplexen, realistischen Aufgabensetting bestehend aus Hypertextlesen und der Auswahl von Links. Insbesondere wurden in dieser Studie mögliche konfundierende Faktoren perzeptueller oder motorischer Art, die oft die Interpretierbarkeit physiologischer Maße in Hypertextstudien schwächen, ausgeschlossen. Insgesamt zeigten die Ergebnisse, dass beide physiologische Maße genutzt werden können, um Veränderungen in der Belastungssituation während der Linkauswahl zu detektieren. Überraschender Weise jedoch korrelierten die beiden Maße nicht miteinander, obwohl beide ein ähnliches Ergebnismuster zeigten. An dieser Stelle können über dieses eher unerwartete Ergebnis nur Spekulationen angestellt und auf zukünftige Forschung verwiesen werden.

Zusammengefasst: In der vorliegenden Dissertation untersuchten wir die EEG alpha-Frequenzbandpower als ein globales Maß kognitiver Belastung, die sich möglicherweise auf die Belastung grundlegender EF zurückführen lässt. Alpha-Frequenzbandpower fungierte dementsprechend als abhängige Variable in allen drei Studien und wurde ergänzt um die Untersuchung der beta- und theta-Frequenzbandpower, der P300 und der Pupillendilatation in einigen der Studien. Jede Studie kann als Ausgangspunkt für weitere Forschung dienen, in der die beschriebenen physiologischen Maße und Aufgabensettings genutzt werden, um EF weiter zu untersuchen. Insgesamt unterstreichen die Ergebnisse der Dissertation den Nutzen physiologischer Maße wie die alpha-Frequenzbandpower und die Pupillendilatation um EF in Aufgabensettings der Grundlagen- und Anwendungsforschung zu untersuchen. Die EEG alpha-Frequenzbandpower scheint als ein Maß der kognitiven Belastung genutzt werden zu können, die auf unterschiedliche Anforderungen an grundlegende EF zurückzuführen ist.

Content

Abbreviations

EF	executive function
WM	working memory
STM	short-term memory
LTM	long-term memory
EEG	electroencephalogram
ERP	event-related potential
ERD	event-related desynchronization
ERS	event-related synchronization
TFR	time-frequency representation
fMRI	functional magnetic resonance imaging
PET	positron emission tomography

1. Introduction and Theoretical Framework

In frameworks of applied psychological research such as, for example, the 'Cognitive Load Theory' (CLT; Plass, Moreno, & Brünken, 2010; Sweller, 1994; Sweller, van Merrienboer, & Paas, 1998) or the 'Multiple Resource Theory' (Wickens, 2002, 2008), the hypothesized limited capacity of working memory (WM) respective limited attentional resources are central for constituting the individual experience of *cognitive load* and thus for constraining the individual performance in real-world task settings. WM defines the mental work space at the intersection of perception and memory where currently perceived information and information recalled from long-term memory is temporarily kept activated, processed, and integrated under the focus of attention (Baddeley, 1992; Cowan, 1999; Engle, Kane, & Tuholski, 1999; Oberauer, Süß, Wilhelm, & Wittman, 2003). Structurally, WM may be divided into codality specific short-term memory (STM) storage components (e.g., a component for the verbal domain and a component for the visuospatial domain) which are under control of an attention-related, central-executive component which is responsible for the processing (i.e., working) aspect of WM (Baddeley, 2007, 2012; Baddeley & Hitch, 1974; Baddeley & Logie, 1999). Functionally, WM can be defined as closely intertwined cognitive processes of attention and memory (Cowan, 2010; Cowan et al., 2005; Engle, 2002; Fougnie, 2008; Oberauer, 2009). To conclude, the term cognitive load may be defined as the load imposed on WM during the execution of a task. Comparably to physiological load, cognitive load increases for increased task complexity as well as for increased time-pressure while completing a task (Galy, Cariou, & Mélan, 2012).

The CLT is a framework which is primarily used to guide instructional design in educational research (e.g., Sweller et al., 1998), but it may also be used to address concepts of human-computer interaction with respect to the evaluation and optimization of the cognitive load-situation therein (Hollender, Hofmann, Deneke, & Schmitz, 2010). With the advent and rise of the 'World Wide Web', that is, the constantly growing information space in the internet consisting of text documents (i.e., hypertext webpages) which are interconnected via hyperlinks, hypermedia environments and specific tasks therein like hypertext reading or web searching have become important tasks of today's daily life. Consequently, hypertext reading (for a review see DeStefano & LeFevre, 2007) or web-searching (e.g., Gwizdka, 2010) increasingly gather research interest, especially with a focus on cognitive load. Notably, hypertext reading may impose additional demands on WM as compared to normal, linear text reading (DeStefano & LeFevre, 2007). The CLT generally identifies WM capacity (i.e., memory storage limitations) for limiting performance in complex, real-world task settings like hypermedia environments (Hollender et al., 2010; Plass et al., 2010; Scheiter & Gerjets, 2007; Schüler, Scheiter, & Genuchten, 2011; Sweller et al., 1998). However, studies in the area of hypermedia research to date mainly consider WM either in a generic sense (i.e., as central bottleneck) or focus rather on the codality specific, memory aspects of WM (i.e., the differentiation between verbal WM and visuospatial WM; see e.g., Schüler et al., 2011), without specifically addressing the executive component of WM. In contrast, the 'Multiple Resource Theory' which primarily focusses on load-situations during multi-tasking identifies attentional resources as performance limiting factor (Wickens, 2002, 2008).

In the present doctoral thesis, I will foster a view on WM which highlights the role of core executive functions (EFs) for WM functioning (see Chapters 1.1 – 1.3). This view on WM may in a way link the two assumptions cited above concerning the role of WM and attention for constraining performance in complex task settings and thus for defining cognitive load. Core EFs (Diamond, 2013) like *updating*, *shifting*, or *inhibition* (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000); for a detailed definition see Chapter 1.1) describe attention-related cognitive processes which are constituent for other higher-order, rather elusive EFs like goal-pursuit or planning (Jurado & Rosselli, 2007) and which are, consequently, essential for any goal-directed, conscious behavior of the daily life (Banich, 2009; Diamond, 2013; Fuster, 2000; Goldstein & Naglieri, 2014; Jurado & Rosselli, 2007; Royall et al., 2002). Notably, core EFs are also essential for WM functioning as, conceptually, they may constitute the "working" part of WM (e.g., Baddeley, 1996; Bledowski, Kaiser, & Rahm, 2010; Bledowski, Rahm, & Rowe, 2009; Engle, 2002; Engle & Kane, 2004; see Chapter 1.2). For example, the 'central-executive' component of Baddeley's WM model mentioned above can be fractionated into different core EFs that are responsible for the processing (i.e., working) aspect of WM and the control of subordinated STM storage components (Baddeley, 1996, 2007; see also 1.2.1). Thus, an important part of cognitive load may originate in demands on different core EFs. Yet, to date, cognitive load has not been studied under a perspective on core EFs. Therefore, in the current doctoral thesis I was interested in the contribution of demands on different core EFs on the global load-situation in basic and applied task settings (see Chapter 1.3) and the measurement thereof (see Chapter 1.4). The overall, conceptual assumption of the current thesis is that cognitive load may originate in WM load which in turn may be mainly grounded in demands on core EFs.

A central topic of any framework in applied psychological research that deals with cognitive load is how to measure best the cognitive load-situation (Brünken, Steinbacher, Plass, & Leutner, 2002; Jong, 2009; Paas, Tuovinen, Tabbers, & Van Gerven, 2003; Parasuraman, 2011). Principally, there exist three different methodologies or techniques for assessing the cognitive load-situation in complex, real-world task settings (see, e.g., Cain, 2007, and Gawron, 2008, for recent comprehensive reviews of different load-assessment techniques). The load-situation, for example, can be assessed rather indirectly, by using subjective rating scales afterwards (e.g., the NASA-TLX; Hart & Staveland, 1988), or more directly, by using a dual-task methodology and assessing performance of the secondary task to infer the load-situation of the primary task (e.g., Brünken, Plass, & Leutner, 2003). Yet, most promising for a direct assessment of the load-situation in real-world task settings might be electrophysiological measures such as the electroencephalogram (EEG) and especially EEG frequency band power (Antonenko, Paas, Grabner, & Gog, 2010; Parasuraman, 2003; Parasuraman & Wilson, 2008) or pupil dilation (Beatty & Lucero-Wagoner, 2000). The advantages of these latter physiological measures[1] are that they allow a rather unobtrusive, objective, and direct (i.e., online) assessment of the load-situation, in contrast to the other load-assessment techniques mentioned above that are rather disruptive, potentially interfering with the performance of the primary task, and, especially reported for the subjective ratings, often may not reflect the actual load-situation (e.g., Kretzschmar et al., 2013).

---

[1] For reasons of brevity and readability, I will only subsume and refer to EEG measures and pupil dilation as 'physiological measures' henceforth. In contrast, measures that rely on brain imaging techniques like functional magnetic imaging (fMRI) or positron-emission tomography (PET) will explicitly not be subsumed under the term 'physiological measures' in the sense I will use this term throughout the current thesis. These latter measures will be referred to as 'brain imaging'. Thus, in the current thesis the term 'physiological measures' will be used for contrasting measures that give more information on the time domain (EEG and pupil dilation) from measures that give more information on the spatial domain (fMRI, PET). Clearly, this use of the term 'physiological measures' may sound rather artificial as processes that are captured via brain imaging technologies are clearly physiological in nature as well. However, this labeling might be justified for reasons of brevity as only EEG and pupil dilation served as dependent measures in the current thesis.

Despite these advantages, physiological measures like EEG frequency band power or pupil dilation so far have rather scarcely been used in hypermedia research (e.g., Antonenko & Niederhauser, 2010; Di Stasi, Antolí, Gea, & Cañas, 2011; Fitzsimmons, Drieghe, Weal, & Drieghe, 2013; Gerlic & Jaušovec, 1999). Furthermore, the few studies that used EEG frequency band power often compare task conditions that are not free of perceptual-motor confounds (e.g., Antonenko & Niederhauser, 2010; see Gerjets, Walter, Rosenstiel, Bogdan, & Zander, 2014 for discussing the problematics of perceptual-motor confounds in applied task settings). Thus, these studies fail to establish any clear relation of observed outcomes in physiological measures and underlying, genuinely cognitive processes such as, for example, core EFs. Furthermore, in basic WM research and theoretical models of WM the role of core EFs has rather recently begun to be addressed (e.g., Baddeley, 1996, 2007, 2012; see also Chapter 1.2), and, moreover, research on EFs has just started to be extended beyond a purely neuropsychological, clinical context that mainly focused on disorders of the frontal lobes (see e.g., Goldstein & Naglieri, 2014). Clearly, more research on these topics will be necessary.

In order to advance the understanding of core EFs and WM functioning and their contribution to cognitive load-situations in task settings of basic and applied research like WM tasks and hypertext reading, in the current doctoral thesis we studied the executive component of WM and the core EFs therein and compared the sensitivity of certain physiological measures thereof. The measures we applied have been previously reported as potential measures for assessing the cognitive load-situation in complex task settings (Antonenko et al., 2010; Just, Carpenter, & Miyake, 2003; Parasuraman, 2003). Using (electro-) physiological measures like EEG frequency band power, the P300 event-related potential (ERP), or pupil dilation, we examined typical WM tasks with respect to core EFs therein (Study 1), the interplay of two core EFs, namely updating and inhibition, within one single WM task (Study 2), and the use of alpha frequency band power and pupil dilation as potential measures for demands on core EFs in complex task settings like hypertext reading (Study 3 a-c). The main goal of the current doctoral thesis was to examine EEG alpha frequency band power (for a detailed description see 1.4.1.1) as a global measure of cognitive load that may be used in task settings of basic (e.g., WM tasks) and applied research (e.g., hypertext reading). As we assumed that cognitive load may be grounded in WM load and more specifically in demands on core EFs, we studied the sensitivity of EEG alpha frequency band power for demands on core EFs (Study 1, Study 2) and the use of this measure in task settings of basic (i.e., WM tasks; Study 1, Study 2) and applied research (i.e., a task of link selection in hypertext reading; Study 3). Therefore, EEG alpha frequency band power served as main measure in all three studies that were conducted as part of this thesis. Especially, Study 3 addressed the use of EEG alpha frequency band power as a measure of the load-situation in a complex, more real-world task setting of hypertext reading when potential perceptual-motor confounds have been carefully ruled out (see follow-up experiments described in Study 3). This way, we thought of tracing back EEG alpha frequency band power to cognitive processes, that is, core EFs, and to ensure that this measure, if applied in complex task settings, might not only reflect non-cognitive factors like perceptual-motor confounds. These potential confounding factors cannot be ruled out as an alternative explanation in studies of hypermedia research that have been conducted to date (e.g., Antonenko & Niederhauser, 2010; Antonenko et al., 2010; Gerlic & Jaušovec, 1999).

Apart from this main research goal of studying alpha frequency band power as a measure for the global load-situation which may originate in demands on core EFs, several additional research questions have been addressed as part of this thesis (mostly on the level of the single studies). These additional research questions can be categorized as broadly falling into two research directions, one interested in a more in-depth understanding of the interplay of core EFs in WM (Study 1 and Study 2) and the other interested in the comparison of different measures of cognitive load (i.e., demands on core EFs) like alpha frequency band power and pupil dilation in tasks of basic and applied research, that is, in a WM task (Study 2) or in a task of hypertext reading (Study 3). Furthermore, in Study 1 we analyzed EEG beta frequency band power as an additional measure potentially reflecting WM load. The different physiological measures are described in more detail in Chapter 1.4, and the different research questions are summarized in Chapter 1.5. The individual research questions will be (re)addressed in different parts throughout this thesis when corresponding tasks, measures, or concepts are discussed. Table 1 shows the different research questions at a glance and highlights their relevance for corresponding research directions.

Table 1. Overview of the Research Questions.

Research Question	Addressed in	Relevant for
Overall Research Question:		
• Examination of EEG alpha frequency band power as a measure of cognitive load which may ground in WM load and, more specifically, in demands on core EFs, in task settings of basic and applied research.	Study 1, Study 2, Study 3	basic and applied research; research on EFs, WM, and hypermedia
Additional Research Questions:		
A) With Focus on Measures		
• Comparison of EEG alpha frequency band power and pupil dilation as measures of cognitive load (i.e., demands on core EFs) in task settings of basic (i.e., a WM task) and applied research (i.e., hypertext reading).	Study 2, Study 3	basic and applied research; especially hypermedia research
• EEG beta frequency band power as a measure for WM load.	Study 1	basic research; especially WM research
• Use of alpha frequency band power in an eye fixation-related methodology, comparably to eye fixation-related potentials (EFRPs).	Study 3 a)	basic and applied research
B) With Focus on Core EFs		
• Comparison of classical WM tasks: n-back, operation span, and digit span tasks with respect to EEG correlates (P300, alpha, beta frequency band power).	Study 1	basic research: research on WM and core EFs
• Examination of the interplay of core EFs in a single WM task (i.e., without perceptual-motor confounds) using EEG measures (P300, theta, alpha frequency band power) and pupil dilation.	Study 2	basic research: research on WM and core EFs
• Examination of increased cognitive load during link selection in hypertext reading using EEG alpha frequency band power and pupil dilation.	Study 3	applied research: hypermedia research; hypertext reading

The further organization of this thesis is as follows: At first, I will give some more detailed information on EFs with a special focus on the core EFs updating, shifting, and inhibition (Chapter 1.1). In a second step, the role of core EFs in recent models of WM will be discussed (Chapter 1.2), and two classical WM tasks and a hypertext-reading situation will be conceptually analyzed with respect to core EFs therein (Chapter 1.3). Finally, the measures used to assess demands on core EFs will be presented (Chapter 1.4), and a brief summary of the three studies that were part of this thesis and the research questions thereof will be given (Chapter 1.5). In Chapter 2 the complete manuscripts of the tree studies are given as submitted to the journals. The following general discussion section comprises the discussion of the main outcomes (Chapter 3.1), some general limitations of the current research (Chapter 3.2), as well as implications for future research (Chapter 3.3), and ends with some concluding remarks (Chapter 3.4).

1.1 Core Executive Functions (EFs)

Although EFs have been extensively studied in neuroimaging or clinical research, to date any unique, overarching framework of EFs with respect to amount, manner, and labeling of different EFs is still missing. Instead, a variety of individually different definitions of EFs are used as reviewed recently by Goldstein and colleagues (Goldstein, Naglieri, Princiotta, & Otero, 2014). Therefore, these authors concluded that the term EFs might be regarded as a kind of "umbrella term" under which diverse cognitive functions are subsumed which are required for complex, conscious task performance in situations that afford adaptive behavior. The complexity of cognitive functions thereunder ranges from rather basic, core EFs like updating, shifting, or inhibition (Diamond, 2013; Miyake et al., 2000) to higher-order, rather elusive EFs like goal formation, or planning (Jurado & Rosselli, 2007). In the current doctoral thesis I will focus on the core EFs.

The EF *updating* most closely resembles processes of WM functioning as it refers to processes of retrieval, transformation, and substitution of WM content (Ecker, Lewandowsky, Oberauer, & Chee, 2010). Noteworthy, some conceptualizations of EFs use the terms WM and updating rather interchangeable (e.g., Diamond, 2013). This raises the question concerning the definition of WM and its status as being one EF amongst others or as incorporating different core EFs within a central-executive component (e.g., Baddeley, 1996; see Chapter 1.2 for a discussion of current models of WM and their relation to core EFs). The EF named *shifting* labels processes of directing the (attentional) focus towards newly relevant information for processing and task performance (e.g., task *shifting* in dual task paradigms; Monsell, 2003). The EF *inhibition* refers to processes of suppressing information that is not (or no longer) relevant for the current processing step in WM. Inhibition is a rather multi-facet cognitive construct (Nigg, 2000). Depending on the concrete stage of the information processing chain, inhibitory control (i.e., inhibition) might describe different processes. For example, demands on inhibitory control can arise on the stage of perception due to interfering perceptual information, on the stage of cognitive processing due to interfering memory information, or on the response stage due to dominant (overlearned) response tendencies (e.g., the tendency to respond to the color word and not to the color in a Stroop color task; Caldas, Machado-Pinheiro, Souza, Motta-Ribeiro, & David, 2012; MacLeod, 1991; Stroop, 1935). In sum, up to eight specific forms of inhibition can be differentiated (Nigg, 2000). Nevertheless, as for the other core EFs, we will use the term inhibition in a rather generic sense throughout this thesis without any explicit differentiation of possible sub-processes thereunder (e.g., Friedman & Miyake, 2004; Nigg, 2000; see Macleod, 2007 for an extensive discussion of inhibition). This generic view on inhibition and the other core EFs is in line with literature (e.g., Diamond, 2013).

Initially, the three core EFs updating, shifting, and inhibition have been identified by Miyake and colleagues (2000) using a correlational research methodology. In a latent-variable analysis using behavioral performance measures of a variety of simple tasks that were supposed to specifically load on single EFs and complex tasks that were supposed to incorporate different EFs, these authors found three factors, namely updating, shifting, and inhibition, that each contributed differently to the performance in the complex tasks, yet all showed some common underlying factor. This common underlying factor of the three EFs has been hypothesized to be attributable to processes of controlled attention.

The findings by Miyake and colleagues (2000) have partly been corroborated further by neuroimaging research showing that several frontal and parietal brain areas are commonly activated by all core EFs but some certain brain areas seem to be rather specifically activated by single core EFs (e.g., the intraparietal sulcus for inhibition, the left mid-dorsolateral prefrontal cortex for shifting, or the posterior ventral frontal regions for updating) as summarized by a meta-analysis conducted by Nee and colleagues (Nee, Brown, & Askren, 2013). Noteworthy however, differences in brain activation with respect to different core EFs seem to be rather subtle, affecting very specific and small brain areas, and are still matter of debate (see, e.g., Collette, Hogge, Salmon, & Van der Linden, 2006; Collette & Van Der Linden, 2002; Collette et al., 2005; Nee et al., 2013; Owen, McMillan, Laird, & Bullmore, 2005; Smith & Jonides, 1999). In sum, brain imaging studies on core EFs revealed the activity of a variety of prefrontal and parietal brain areas during tasks demanding EFs (e.g., Collette, Hogge, Salmon, & Van der Linden, 2006; Collette & Van Der Linden, 2002; Owen, McMillan, Laird, & Bullmore, 2005). These observations led to the view that the physiological underpinnings of EFs may be the complex interaction and integration of various different neuronal networks throughout the brain (Otero & Barker, 2014). Thus, the activity of prefrontal-parietal networks might be constituent for EFs (e.g., Chung, Weyandt, & Swentosky, 2014; Collette & Van Der Linden, 2002; Fuster, 2000; Otero & Barker, 2014).

As a short side-note, because of the obviously distributed neuronal networks which underlie EFs, rather than focusing on specific brain regions future research on EFs might focus more on this network character, that is, the oscillatory activity associated with executive functioning. For this line of research, the EEG can proof to be a valuable methodological alternative compared to the rather static brain imaging methods like functional magnetic resonance imaging (fMRI) or positron-emission tomography (PET) which have been used in the studies cited above. EEG allows to capture the oscillatory activities of neuronal networks of the brain with high time-resolution and thus might reflect more directly the neuronal coupling or decoupling of different brain areas (e.g., Buzsáki & Draguhn, 2004, see also Chapter 1.4).

Developmental studies also corroborate the possibility of separating the three core EFs by showing different developmental trajectories for different core EFs. Thus, for example, the capability of inhibition occurs early in life and develops most strongly during preschool age, whereas updating and shifting improve in a linear way but most strongly above about five years of age, that is, during the first school years (Best & Miller, 2010; Best, Miller, & Jones, 2010; Garon, Bryson, & Smith, 2008). Interestingly, controlled attention which has been hypothesized as the common underlying factor of the core EFs (Miyake et al., 2000) and which may reside in an anterior attention network (Petersen & Posner, 2012) shows a comparably developmental trajectory consisting of a strong improvement during late preschool and early school years (Garon et al., 2008; Posner & Rothbart, 2007). In addition, when comparing the developmental trajectories of EFs and the frontal lobes, a close connection between EFs and the frontal lobes can also be observed. Both maturate until the early days of adulthood (Best et al., 2010; Garon et al., 2008; Stuss, 1992) and both are prune to age-related decline (Burke & Barnes, 2006; Fisk & Sharp, 2004). In line with the hypothesized possibility of separating these core EFs, the age-related decline has been reported to selectively affect different core EFs (see meta-analyses by Verhaeghen, 2011). Comparably, damages or disorders of the frontal lobes often result in impaired performance in EFs (e.g., Alvarez & Emory, 2006; Royall et al., 2002). In sum, developmental studies on EFs are in line with the differentiable yet unitary account on EFs as proposed by Miyake and colleagues (2000). Controlled attention might constitute the common underlying factor of the three core EFs with the biological underpinnings of this factor possibly lying in an anterior attention network.

Noteworthy, there is some debate concerning the conceptual differentiation of core EFs. For example, the question is raised whether a fourth EF, "multitasking" (i.e., the ability to coordinate task performance of two simultaneously presented tasks), should be added to a taxonomy of core EFs (Strobach, 2014), or whether the EF inhibition might be conceptualized as the common underlying factor of updating and shifting, thus reducing the number of core EFs to two (Miyake & Friedman, 2012). Furthermore, a close connection between updating and shifting might exist as shifting might be considered to comprise similar sub-processes as updating, yet referring to procedural WM content whereas updating refers to declarative WM content (Oberauer, 2009; see 1.2.4). However, in the current thesis we restricted our research focus on the classical three core EFs, that is, updating, shifting, and inhibition, which we addressed under a rather generic view (i.e., not differentiating between possible sub-processes thereunder). This might be justified for reasons of brevity, and more importantly, as the classical definition of the three core EFs is still prevalent.

Most research on core EFs during the last decades has been done using correlational research methodologies either to identify core EFs in complex tasks (e.g., Miyake et al., 2000), or, in individual differences research, to compare performance in core EFs and, for example, scholastic achievement (e.g., St Clair-Thompson & Gathercole, 2006). Furthermore, core EFs have been addressed in developmental studies as those cited above (e.g., Best et al., 2010) as well as in training studies that just recently have become increasingly popular. Although often yielding mixed results, these training studies revealed that the training of core EFs might at least partly improve the performance in complex tasks (e.g., Dahlin, Bäckman, Neely, & Nyberg, 2009; Karbach & Kray, 2009; Klingberg, 2010; Morrison & Chein, 2011; Rabipour & Raz, 2012; Salminen, Strobach, & Schubert, 2012; Titz & Karbach, 2014; but see Shipstead, Redick, & Engle, 2010), thus underlining the importance of core EFs for complex cognition. Finally, fMRI and PET studies have been conducted to study core EFs (e.g., Bledowski et al., 2009, 2010; Collette et al., 2006; Collette

& Van Der Linden, 2002; Owen et al., 2005). So far, other physiological measures like EEG frequency band power or pupil dilation have been rather seldom used to study core EFs (see Chapter 1.4). Thus, one objective of the current doctoral thesis was to strengthen this research methodology (i.e., EEG and pupil dilation) for studying core EFs, as these measures might be well suited to capture the time course of cognitive load in task settings of basic and applied research which might be grounded in demands on core EFs[2].

---

[2] For an interesting, recent discussion of neuroimaging and electrophysiological data for studying cognitive processes see Axmacher and colleagues (Axmacher, Elger, & Fell, 2009).

1.2 The Role of Core EFs in Current Theories of Working Memory (WM)

To date, the connection of EFs and WM is somehow ambiguously defined, apparently depending on the authors' primary research focus. Thus, authors primarily stemming from EF research (e.g., Diamond, 2013) often define WM as one EF among others, whereas authors primarily stemming from WM research define EFs as a central aspect of WM functioning and thus as being a part or component of WM (e.g., Baddeley, 1996). In the current thesis we will favor the latter view. In the following I will briefly review some of the currently most prevalent models of WM in cognitive science under a special focus on the role of core EFs therein (for a more general and extensive discussion of recent models of WM see Miyake & Shah, 1999; Osaka, Logie, & D'Esposito, 2007).

1.2.1 The multi-component model of WM

The multi-component model of WM by Baddeley and colleagues (Baddeley & Hitch, 1974; Baddeley & Logie, 1999; Baddeley, 1992, 2003, 2012) still can be regarded as one of the most prevalent and most influential theories of WM. The multi-component model was one of the first conceptualizations of a WM construct. The initial formulation of the model in 1974 (Baddeley & Hitch, 1974) might be seen to mark a turning point in memory research by shifting the prevalent research focus from a rather passive STM storage system of limited capacity (Atkinson & Shiffrin, 1968) to an active processing system consisting of different components, namely the WM. According to Baddeley and Hitch (1974) WM consists of three components, two of which are memory-related, the so-called phonological loop and the visuospatial sketchpad, and one is attention-related, the so-called central-executive. The memory components were conceptualized as codality-specific, passive storage structures of limited capacity for verbal or visuospatial material, thus closely resembling the former STM. In contrast, the central-executive was conceptualized to supervise and coordinate the storage components, thus resembling the supervisory attentional system (SAS) proposed by Norman and Shallice (Norman & Shallice, 1986). In a later revision of the multi-component model, a forth component, the so-called episodic buffer was added to provide a unified storage component for processing task material irrespective of codality (Baddeley, 2000).

Over the years, the central-executive component has also been developed further (Baddeley, 1996, 2007). Starting from a conceptualization as a rather unitary control component, the central-executive later was fractionated into several attention-related EFs which were hypothesized to be essential for WM functioning (Baddeley, 1996, 2002, 2007). According to Baddeley (1996) apart from supervising and coordinating the storage systems, the EFs of the central-executive are (a) the ability to focus attention against potentially distracting, irrelevant information, (b) the ability to switch attention between two or more stimulus sources or actions, and (c) the ability to divide attention in order to perform two tasks simultaneously. Noteworthy, the EFs of the central-executive as formulated by Baddeley (1996, 2002) seem to closely resemble the core EFs updating, shifting, and inhibition (Miyake et al., 2000; Diamond, 2013). The coordination of the storage components comprises processes which are subsumed under the EF updating, and the other processes described by Baddeley closely resemble the definitions of the EFs inhibition and shifting (compare Chapter 1.1).

The component view on WM by Baddeley and colleagues stimulated a wealth of neuroimaging studies that searched for brain areas associated with specific WM components (see e.g., Baddeley, 2007; Collette & Van Der Linden, 2002; D'Esposito, 2001; D'Esposito et al., 1995; Duncan & Owen, 2000; Nee et al., 2012; Smith & Jonides, 1997; Wager & Smith, 2003). Although initially neuroimaging studies seemed to corroborate the multi-component model of WM (e.g., Smith & Jonides, 1997), over the years it turned out that in sum the component view of WM is only weakly supported by neuroimaging results. For example, WM tasks which load on the phonological loop show brain activity primarily in left frontal and temporoparietal brain regions (e.g., Broca's area), but only for low to moderate WM load, whereas tasks which load on the visuospatial sketchpad show rather right-lateralized activity (Baddeley, 2003; Smith & Jonides, 1997, 1999; Wager & Smith, 2003). Thus, the spatial differentiation within the brain of codality-specific WM components (i.e., the phonological loop, the visuospatial sketchpad) holds only true under certain, specific circumstances.

The frontal lobes generally show more activity for WM tasks which require executive processing than for simple STM storage tasks (Wager & Smith, 2003). These observations indicate the location of a central-executive component in the frontal lobes. More specifically, Duncan and Owen (2000) identified the mid-dorsolateral, mid-ventrolateral and dorsal anterior cingulate cortex as commonly showing activity for a variety of different tasks which required executive processing, which might attribute especially these structures to a central-executive component (Baddeley, 2007). Yet, other authors reported the activity of rather diverse frontal and parietal brain areas for tasks requiring EFs (Collette & Van der Linden, 2002). Therefore, over the years of neuroimaging research, a new view emerged on the conception of WM. Instead of focusing on structural aspects of WM and thinking about WM in terms of differentiable, specific brain areas (i.e., different components that have biological underpinnings in specific, differentiable brain areas), a more functional view on WM emerged that attributed WM to the activity of fronto-parietal neuronal network structures without identifying a specific brain region as "the WM" (Collette & Van der Linden, 2002; Postle, 2006; Zimmer, 2008). Thus, the current neurobiological view on WM is one that hypothesizes WM as rather flexible, distributed fronto-parietal networks which rely on existent brain structures associated with perception, memory, and attention, instead of hypothesizing WM as a set of fixed brain regions which are specifically connected to single WM components (D’Esposito, 2007; D’Esposito & Postle, 2015; Postle, 2006). This functional view on neurobiological underpinnings of WM is also reflected in the newer theoretical models of WM that I will review below.

1.2.2 The embedded process model of WM

The embedded process model of WM proposed by Cowan (Cowan, 1988, 1995, 1999) can be seen as one representative of a functional view on WM. It emphasizes the link between long-term memory (LTM) and attention for WM functioning. According to Cowan, WM can be defined as an activated subset of LTM in the focus of attention, that is, without assuming any specific STM components as in the conceptualization of Baddeley and colleagues (e.g., Baddeley & Hitch, 1974; Baddeley & Logie, 1999). Thus, Cowan describes a unitary system of WM with no codality-specific temporary memory stores. In doing so, his model overcomes a limitation of Baddeley's multi-component model that currently defines only two memory components for verbal and visuospatial task materials but would have to define additional components for dealing with other sensory information (e.g., tactile information or odors). According to Cowan, the amount of information in the activated subset of LTM is limited by time (if no longer under the focus of attention). This is a comparable operationalization as for the capacity limitations of the STM storage components in the model by Baddeley and colleagues, where capacity limitations are also defined by the temporary decay when no processes of rehearsal take place.

Additionally, in the model by Cowan the focus of attention is hypothesized to be capacity-limited to about four chunks of information (Cowan, 2001). Cowan assumes the focus of attention to be either under voluntary control of a central-executive system or under involuntary control of the attentional orienting system. However, despite this connection of WM to a central-executive system which reminds of Baddeley’s central-executive component, the embedded process model of WM does not provide any further information on the central-executive system, its potential fractionation, or the definition of specific processes therein. For example, Cowan (1999) speaks rather vaguely about the executive system as "a set of processes influenced by instructions or incentives" (Cowan, 1999, p. 67).

Thus, as main benefits of the model by Cowan may be regarded, that the model addresses topics which are problematic or which have been underrepresented in the multi-component model of WM, like, at least in earlier versions of the model, the role of LTM (see Baddeley, 2012). However, apart from emphasizing the role of attention for WM functioning, the embedded process model by Cowan seems to be of rather limited benefit with respect to defining the role of separate core EFs in WM.

1.2.3 The executive attention framework

Engle and colleagues proposed a framework of WM that was developed out of individual differences research (Engle, 2002; Engle & Kane, 2004; Engle, Kane, & Tuholski, 1999; Kane & Engle, 2002) and thus mainly focus on limiting factors of WM which may induce individual differences in cognition. They identified controlled attention (later renamed in executive attention; see Engle, 2002) as the central, limiting factor in WM, that is, as the factor which defines the overall capacity limits of WM. Comparable to the focus of attention described by Cowan (1995), controlled attention is hypothesized to be domain-independent, of limited capacity, and operating on LTM traces (Engle et al., 1999). Thus, the memory aspect of WM in this framework is conceptualized as activated parts of LTM. Furthermore, the differentiation of WM with respect to verbal or visuospatial task materials is met by assuming codality-specific codes (Engle et al., 1999) rather than codality-specific structures like specific WM storage components as hypothesized in the multi-component model of Baddeley and colleagues (see 1.2.1).

For the working aspect of WM, processes of controlled attention are hypothesized to activate memory representations, bring them into the focus of attention and, most important, maintain them in the focus of attention in the face of interference and distraction. Thus, Engle and colleagues describe a dual process model of executive attention. Executive attention is hypothesized to be necessary for maintaining the information in active memory as well as for the resolution of conflict (e.g., in case of two competing memory traces). This latter aspect of executive attention emphasizes particularly the role of the EF inhibition for WM functioning. Consequently, inhibitory capabilities play a central role in defining WM capacity limitations and thus in establishing individual differences in cognition (Kane & Engle, 2002). Anatomically, Engle and colleagues link individual differences in WM capacity (i.e., the ability to control attention) to the functioning of the prefrontal cortex. This resembles the link between the frontal cortex and EFs as described in Chapter 1.1.

In sum, the executive attention framework of WM by Engle and colleagues may be regarded as a functional WM model which highlights the role of executive attention for WM functioning. Thus, it might be seen to bridge research on EFs (focusing on inhibitory control) as well as research on attention (i.e., executive attention; Gazzaley & Nobre, 2012; Vandierendonck, 2014) under a framework of WM.

1.2.4 The functional model of WM

The functional model of WM proposed by Oberauer and colleagues (Oberauer, 2009; Oberauer, Süß, Wilhelm, & Wittman, 2003) comprises aspects of all three models of WM reviewed above. It may be seen as the most global and at the same time most specific account of WM to date (Baddeley, 2012). Central to the functional model of WM is the division of WM into two facets, namely the content domain and cognitive functions (Oberauer et al., 2003). This division resembles the differentiation of WM in memory components and a central-executive component by Baddeley and colleagues (see 1.2.1). In the later version of the functional model (Oberauer, 2009), this dichotomy of functional and content aspects of WM became specified as the procedural part of WM and the declarative part of WM. Oberauer (2009) hypothesized each of the two parts of WM to comprise three embedded components. These three components are (1) an activated part of long-term memory, (2) a component where new structural representations are created by dynamic binding, called "region of direct access" for declarative WM or "bridge" for procedural WM, and (3) a mechanism for selecting a single element, namely the "focus of attention" for declarative WM and the "response focus" for procedural WM. The "bridge" in this conceptualization of WM may be regarded as closely resembling the central-executive component of the multi-component model of WM (see, Baddeley, 2012).

A detailed discussion of the functional model of WM is beyond the scope of the current thesis. Instead, some aspects of the model that relate to the discussion of core EFs in WM will be highlighted. According to Oberauer (2009), executive processes (i.e., EFs) come into play in WM for the flexible reconfiguration of WM which is regarded to be a general purpose mechanism that has to be flexibly adapted to fulfill the specific requirements of a current task. With respect to the core EFs, Oberauer (2009) defines updating to work on the declarative part of WM (i.e., replacing old declarative WM content by new one), whereas inhibition and shifting (called 'switching' by Oberauer, 2009) are hypothesized to work on the procedural part of WM (i.e., on WM functions). In this definition, shifting can be regarded as a form of updating, yet operating on the procedural part of WM, i.e., shifting describes processes of replacing the current task set in the "bridge" by a new one. Finally, inhibition refers to processes of establishing and consolidating a new task set in the "bridge" against distracting procedural information of old tasks sets that, for example, stem from habitual response tendencies in LTM.

In sum, the functional model of WM proposed by Oberauer (2009) and the differentiation in a declarative part of WM and a procedural part of WM may serve as both, a thorough definition of WM functioning as well as a definition of the core EFs thereunder. However, as Baddeley (2012) stated, an extensive experimental investigation of all of the different aspects of the functional model is still to come.

1.2.5 Conclusion: Core EFs and models of WM

To sum up, the review of four of the currently most prevalent cognitive theories of WM revealed a central position of core EFs for WM functioning. The three core EFs updating, shifting, and inhibition may be subsumed under a central-executive component of WM (e.g., Baddeley, 2007) or may be incorporated in processes of controlled or executive attention (Engle, 2002). A differentiable view on WM as consisting of a declarative and a procedural part (Oberauer, 2009) may help to capture subtle differences and commonalities in the definition of the three core EFs (see 1.2.4). However, whether the conceptual considerations concerning the relation of WM and core EFs as proposed by the models reviewed above might hold experimental examinations still has to be shown. Therefore, in the current doctoral thesis as one research question the relation of core EFs and WM has been addressed by comparing different WM tasks with respect to core EFs therein (Study 1) and by manipulating different core EFs within one single WM task (Study 2).

1.3 Core EFs in Tasks of Basic and Applied Research: Conceptual Task Analyses

To further elaborate on the relationship between core EFs and WM in the following conceptual task analyses of two classical families of WM tasks, namely n-back and WM span tasks will be given. The n-back task has been used in Study 1 and Study 2 of the current thesis. WM span tasks have been used in Study 1 for comparison with the n-back task by means of EEG measures. Finally, (see 1.3.3) demands on core EFs during hypertext reading will be conceptually analyzed. Consequently, Study 3 examined the sensitivity of EEG alpha frequency band power and pupil dilation for link selection processes which might increase demands on core EFs during hypertext reading.

1.3.1 The n-back task

The n-back task has been mainly used in neuroimaging studies of cognitive neuroscience research and is commonly regarded as a typical task loading on WM updating (Gevins, Smith, McEvoy, & Yu, 1997; Jonides et al., 1997; Krause, Pesonen, & Hämäläinen, 2010; Krause et al., 2000; Owen et al., 2005; Palomäki, Kivikangas, Alafuzoff, Hakala, & Krause, 2012; Pesonen, Hämäläinen, & Krause, 2007). In a typical n-back task a temporal sequence of stimuli is presented and participants have to decide via key-presses for each stimulus whether it was similar or different to the stimulus they saw n-steps back with respect to a certain stimulus dimension (i.e., they have to perform a binary match/no-match decision). The stimulus dimension (e.g., identity, color, or semantic content) the participants have to react to as well as the n-back level (e.g., 0-back, 1-back, 2-back) is announced at the beginning of a sequence. The 0-back task condition can be regarded as a pure matching task, as in this task condition one specific n-back target stimuli is shown at the beginning of the sequence, and during the following stimuli sequence participants have to press the match-key each time they saw this specific stimulus and otherwise the no-match-key. All other n-back conditions above the 0-back are hypothesized to require processes of WM updating, as beginning with the 1-back task condition participants have to base their match/no-match decision each time by comparing the current stimulus with the stimulus they saw at position n-back in the sequence.

Noteworthy, an EEG event-related potential (ERP) study by Watter and colleagues (Watter, Geffen, & Geffen, 2001) indicated that the n-back task might also incorporate processes of shifting, as the n-back task seems to incorporate aspects of a dual task. The authors observed a decrease of the P300 mean amplitude for increased n-back levels comparably to the decrease of the P300 amplitude of the secondary task in dual task studies when the task demands of the primary task increased (e.g., Allison & Polich, 2008; Kok, 2001; see 1.4.1.4 for a detailed discussion of the P300). Furthermore, Watter and colleagues (2000) observed a significant increase in P300 latency only between the 0-back task condition and the following n-back task conditions, but not between the higher n-back levels. This was interpreted to underline the differentiable character of the 0-back task condition as a purely matching task, whereas the higher n-back levels seem to put demands on other cognitive processes (i.e., WM updating). This leads to the necessity to shift the attentional focus between the different sub-processes of matching and WM updating for the higher n-back levels (see also Chen, Mitra, & Schlaghecken, 2008). Thus, in an n-back task for increased n-back load levels controlled attention might have to be divided between different processes (updating and shifting) which results in a decrease of the P300 amplitude as a measure of attentional demands (Kok, 2001; Polich, 2007).

Conceptually, the n-back task might also necessitate processes of inhibition. For example, the dual decision (match/no-match) that participants have to make in an n-back task always requires them to suppress (i.e., inhibit) the inadequate (i.e., incorrect) response tendency. Furthermore, in the higher n-back levels (above the 1-back), this inadequate response tendency might be fostered by so-called lure trials, that is, n-back stimuli that match previous stimuli but at an inadequate position within the sequence (e.g., 1-back lures within a 2-back task; see Szmalec, Verbruggen, Vandierendonck, & Kemps, 2011).

1.3.2 WM span tasks

WM span tasks have been mainly used in individual differences research to assess WM capacity (i.e., the amount of stimuli correctly recalled). This line of research usually observes a strong correlation between performance in complex span tasks (i.e., WM capacity) and several higher order cognitive abilities like reading, mathematics, problem solving, or goal pursuit (e.g., St Clair-Thompson, 2007; Yuan, Steedle, Shavelson, Alonzo, & Oppezzo, 2006).

Because of their dual-task nature, complex span tasks can be regarded to particularly put demands on the EF shifting. Complex span tasks like the reading span task (Daneman & Carpenter, 1980) or the operation span task (Turner & Engle, 1989) consist of a processing sub-task (reading and verifying a sentence, respectively calculating and verifying an equation) and a short-term memorization sub-task (memorize words, digits, or letters). These sub-tasks of processing and memorization are presented in alternation until after three to mostly seven or eight alternations (i.e., three to eight stimuli to be memorized) the participants have to recall the remembered stimuli in correct serial order. Thus, the alternation between the processing sub-task and the memorization sub-task clearly requires processes of shifting between different task sets.

Conceptually, it might be plausible to assume that complex span tasks also incorporate processes of WM updating due to the memorization sub-task. In principal, the WM updating demands in complex span tasks might be comparable to the WM updating demands in n-back tasks, although in an n-back task participants have to continuously update their WM content (i.e., load a new item in WM and 'forget' an old item), whereas in a span task updating seems to involve first of all processes of adding a new item for temporary storage in memory. Thus, on a first, shallow view the WM updating demands of span tasks and n-back tasks might occur to be differently, as in the span tasks the replacement (i.e., forgetting) of old memory items seem not to be necessary. However, because of the shifting between the sub-tasks and, furthermore, because of potential strategies of chunking (i.e., the combination of items for better memorization), complex span tasks might comprise comparable WM updating processes of loading new and replacing old, no longer relevant items in WM (e.g., items before chunking) as the n-back task.

Also conceptually, complex span tasks might incorporate processes of inhibition. Comparably to the n-back task, a dual choice decision has to be made for the processing sub-task (i.e., the decision whether the sentence respective the equation is correct or not). Furthermore, because of the shifting between two sub-tasks it might be plausible to assume that the sub-task currently not relevant for task performance has to be temporarily suppressed. In Study 1 of the current thesis an operation span task was used as complex span task. In order to create a complex span task that was conceptually as similar as possible to an n-back task, participants in this task had not simply to memorize an unrelated item during the memorization sub-task (as in typical complex span tasks, for example, Conway et al., 2005; Daneman & Carpenter, 1980; Turner & Engle, 1989), but to remember a given digit that served as possible result for the preceding equation, that is, that matched or mismatched the result that participants had mentally calculated during the processing sub-task before. In doing so, demands on inhibitory control were also increased as compared to a standard complex span task where unrelated stimuli have to be memorized. This is because in cases when the digit presented during the memorization sub-task was the wrong result for the previous equation (i.e., mismatched the result participants have mentally calculated), participants would have to actively suppress the formerly calculated result that interfered with the digit that had to be remembered.

In contrast to complex span tasks which may incorporate demands on all three core EFs, a simple span task like the digit span task, where participants simply have to recall a given sequence of three to eight digits in correct order, may mainly demand WM updating[3]. In order to empirically examine these conceptual considerations of the role of core EFs in n-back and WM span tasks, Study 1 was conducted. Using EEG measures an n-back task, an operation span task, and a simple digit span task have been compared in this study (see 1.5.1).

1.3.3 Core EFs in hypertext reading and link selection

The core EFs updating, inhibition, and shifting may not only define central aspects of WM functioning but they may also be highly relevant for tasks in hypermedia environments like hypertext reading and link selection. In the following a typical hypertext reading situation will be recapitulated with respect to demands on core EFs therein.

---

[3] I will refer to the digit span task rather interchangeable as STM or WM task. Traditionally, the digit span task is regarded as a purely STM task that only puts demands on the memory components of WM (i.e., STM). However, at least when the STM capacity limits are exceeded (i.e., when more than about four digits have to be memorized; Cowan, 2010) this task will also require central-executive processes, that is, the core EF updating.

In a typical hypertext webpage, for example, a Wikipedia article (http://wikipedia.org), the reader is confronted at certain positions in the text with hyperlinks, that is, often blue-colored, marked words which lead to other webpages. Each time the reader gets to a hyperlink, the reading process is interrupted and decision processes have to take place (DeStefano & LeFevre, 2007). Thus, the EF shifting is required, as the reader has to perform a task set shifting, that is, a shift from purely text reading processes to hyperlink selection processes. These hyperlink selection processes then might put demands on the EF inhibition, as the reader has to ignore the hyperlink when deciding to continue reading without clicking on it. These demands on inhibitory control might be especially high, if the hyperlink is of special interest for the reader but not relevant for the current information gathering process that had been the initial reason for starting reading the hypertext page (Salmerón, Cerdán, & Naumann, 2015). Finally, the EF updating might be continuously demanded, as during reading and text comprehension the situation model generated out of the text in combination with prior knowledge has to be continuously updated (Kintsch, 1988; Kintsch, Patel, & Ericsson, 1999; Zwaan & Radvansky, 1998).

In sum, when reading a hypertext all three core EFs may be demanded. The demands on core EFs might moreover vary, depending on the complexity of the text itself (which should mainly affect updating-demands[4]), or the amount and manner of hyperlinks presented (which should additionally affect inhibition and shifting-demands). Importantly, the demands on core EFs should already occur during reading of the initial hypertext page, that is, before accessing (via the hyperlinks) subsequent webpages which will put additional load on the reader due to the additional content to be processed and the navigation-step that has to be remembered.

To date, hypertext research prevalently examined the entire process of link selection and accessing subsequent webpages (for a review see DeStefano & LeFevre, 2007), thus confounding the load-situation of purely link selection (and the EFs therein) by the content of additional webpages and potential effects of disorientation the reader might experience (e.g.,

---

[4] Noteworthy, the relevance of updating for text comprehension is mainly drawn out of the text comprehension model proposed by Kintsch and colleagues (e.g., Kintsch, 1991). Other models of text comprehension like the 'Structure Building Framework' proposed by Gernsbacher and colleagues also highlight the role of inhibition and shifting for creating and updating a text model, that is, for text comprehension (Gernsbacher & Foertsch, 1999; Gernsbacher, 1991).

Amadieu, van Gog, Paas, Tricot, & Mariné, 2009; Cangoz & Altun, 2012; Kim & Hirtle, 1995; Scheiter & Gerjets, 2007). Such confounding factors, like the processing of additional information (because of reaching subsequent webpages), as well as perceptual-motor differences between task conditions, have especially to be considered when using physiological load measures like EEG frequency band power or pupil dilation (see e.g., Gerjets et al., 2014 for a recent and comprehensive discussion of potentially confounding factors in studies that used EEG measures in complex task settings).

For example, Antonenko and Niederhauser (2010) studied the influence of *leads* on the cognitive load-situation in hypertext reading and hyperlink selection. The leads were previews which consisted of a few sentences summarizing the content of the hyperlinked webpages. Some hyperlinks of the primary hypertext page were presented with leads, some without. The leads appeared as mouse-over popup balloons when the mouse-pointer was over the hyperlink. The authors compared EEG frequency band power for data epochs of the hyperlinks with leads and the hyperlinks without leads and observed less cognitive load (i.e., less decrease of EEG alpha frequency band power) in the lead as compared to the no-lead hyperlink situations. Although in line with the authors' hypothesis, the results might be severely confounded. First, the lead condition was perceptually different from the no-lead condition. Second, the data epochs used for analysis contained different content between the two conditions because of the additional text of the leads which also might have influenced the EEG data.

Thus, in the current thesis (Study 3 a-c) we focused on the initial load-situation during hyperlink selection that has not explicitly been studied up to now. Importantly, we avoided potential confounds due to the presentation of subsequent webpages as well as confounds due to perceptual-motor differences. This was achieved by creating a hypertext-like reading situation using 'simulated' hyperlinks without any link-functionality (i.e., without leading to subsequent pages) and by conducting two additional follow-up experiments to exclude possible perceptual-motor confounds (see 1.5.3 for a more detailed discussion of Study 3). The main research question addressed in Study 3 was the comparison of EEG alpha frequency band power and pupil dilation with respect to their sensitivity for increased load-situations during hyperlink selection in a natural, online-text reading scenario. In the following chapter, the physiological measures will be presented in more detail.

1.4 Electrophysiological Measures and Pupil Dilation for Demands on Core EFs

The advantages of using electrophysiological measures or pupil dilation for capturing cognitive load (i.e., demands on core EFs) have already been outlined in the introductory section of this thesis. In short, they may allow a rather unobtrusive, objective, direct, and potentially even online assessment of the load-situation in tasks of basic or applied research (e.g., Antonenko et al., 2010; Just et al., 2003). However, with respect to core EFs electrophysiological measures or pupil dilation have been seldom used to study the related cognitive processes under a framework of core EFs. Most research on EFs or WM so far has been done using neuroimaging techniques like fMRI or PET (e.g., Nee, Brown, Askren, 2013; Owen et al., 2005; Smith & Jonides, 1999). Noteworthy, as mentioned in Chapter 1.2, in WM research there has been a shift from focusing on structural aspects of WM (that might be well captured by neuroimaging techniques like fMRI) to focusing on functional aspects of WM. These functional aspects might be better captured using a electrophysiological technique like the EEG that allows capturing the neuronal oscillatory activity with a high time resolution. Furthermore, neuroimaging studies on EFs showed the distributed activity of larger amounts of different frontal and parietal brain areas, leading to the assumption that the biological underpinnings of EFs might be fronto-parietal network structures that are flexibly activated (Collette & Van der Linden, 2002; see also Chapter 1.1). EEG frequency band power that captures the oscillatory activation and deactivation of neuronal networks might therefore be an adequate measure of EFs. In addition, pupil dilation may serve as an overall measure of a current load-situation, including cognitive and emotional aspects thereof (Beatty & Lucero-Wagoner, 2000). Especially in task settings of applied research, pupil dilation might serve as a load-measure that can be acquired more efficiently than EEG measures (cf. 1.4.2), given the problem of its vulnerability to changes in ambient or stimulus lighting conditions can be dealt with. In the following, I will describe the different measures in more detail.

1.4.1 Electrophysiological (EEG) measures

In general, the EEG allows capturing the ongoing oscillatory activity of neuronal networks of the cortex (i.e., synchronously oscillating cell assemblies of pyramidal cells in the neocortex that have a specific horizontal orientation with respect to the scalp) with a time resolution in the range of milliseconds by recording and amplifying the thereby generated electrical current from the scalp (Kappenman & Luck, 2012). Thus, the EEG may provide an unobtrusive methodology to track instant changes in oscillatory activity due to cognitive processes. In doing so, the EEG may serve as an ideal measure to continuously track the current load-situation. In principal, there are two ways in analyzing EEG data to detect changes in oscillatory activity due to internal (e.g., thoughts) or external events (e.g., stimuli): the event-related potentials (ERPs) and the event-related synchronization (ERS) and desynchronization (ERD).

First, ERPs (e.g., Kappenman & Luck, 2012) are calculated by simply averaging several data epochs time-locked to the same event. This averaged (and baseline-corrected) signal, the ERP curve, provides a picture of the evoked (i.e., time- and phase-locked) neuronal activity associated with the event, prune of the general EEG activity that in this research account is considered as noise (Bastiaansen, Mazaheri, & Jensen, 2012). Within the ERP curve different positive or negative deflections (i.e., components) can be visually identified. These ERP components have been associated with fairly specific cognitive processes, with the earlier components (< 100 ms) being mainly associated with perceptual processing (e.g., perceptual encoding) and the later components with higher order cognitive processes (see Kappenman & Luck, 2012 for an extensive discussion of ERPs).

Second, the event-related synchronization (ERS) and desynchronization (ERD) can be analyzed, that is, the phasic changes of frequency band power induced by an event (Bastiaansen et al., 2012). An increase in frequency band power (i.e., an ERS) is due to the synchronized activity of large neuronal cell assemblies which oscillate in synchrony as a result of an event (e.g., a stimulus), whereas a decrease in frequency band power reflects a desynchronization (i.e., a decoupling of large neuronal cell assemblies into smaller ones of different oscillatory activity). Thus, frequency band power may be regarded to reflect the size and the activity of the underlying neuronal networks.

An ERD/ERS analysis of oscillatory activity can be done by transforming the EEG signal of data epochs time-locked to events from the time-domain to the frequency-domain, that is, by calculating the frequency spectrum (e.g., by using a fast-fourier transform, FFT).

By calculating the frequency spectrum for a certain time window which is moved across each time point of the source EEG signal (e.g., by using an FFT and a moving analysis window) a time-frequency representation (TFR) of the original EEG data can be created. A TFR captures the frequency band power over time and for a certain frequency range and allows analyzing the time course of changes in frequency band power as a consequence of an event. Normally, TFRs are calculated which do not represent absolute power values but relative values with respect to a certain baseline. Common baselines are a pre-stimulus time period (i.e., the time directly adjacent to the event comparably the one used for the ERP calculation), but a rest-baseline (i.e., some specific condition) might also be used (e.g., Stipacek, Grabner, Neuper, Fink, & Neubauer, 2003). The so called ERD/ERS% formula given by Pfurtscheller and Lopes da Silva (1999) is one possibility to calculate such relative frequency band power values. This formula calculates the percentage of change of frequency band power between a baseline and a test condition (Pfurtscheller & Lopes da Silva, 1999; see also Antonenko et al., 2010).

Traditionally, EEG frequency band power has been divided into several different frequency bands which might be associated with different (but rather broadly defined) categories of cognitive functioning (Krause, 2003). The theta (4 – 8 Hz), alpha (8 – 13 Hz), and recently the beta (14 – 30 Hz) frequency bands have been associated with cognitive processes related to attention and memory, and especially WM (Engel & Fries, 2010; Gevins & Smith, 2000; Klimesch, 1999; Klimesch, Schack, & Sauseng, 2005; Krause et al., 2000; Krause et al., 2010). Klimesch (1999) proposed a further subdivision of the alpha frequency band into an upper part (10 – 13 Hz) which might reflect rather semantic aspects of memory processing and lower parts (6 – 8 Hz and 8 – 10 Hz) which might reflect rather attentional aspects. Noteworthy however, the association of different frequency bands and cognitive processes is still matter of debate and a direct one-to-one relation between single cognitive processes and specific frequency bands might not be possible (Krause, 2003). Furthermore, as the frequency bands seems to vary with respect to exact position and band-width due to individual differences (Klimesch, 1999), recent studies have analyzed a rather broad frequency range (e.g., 4 – 30 Hz) in TFRs and defined the traditional frequency bands rather flexible and data-driven to categorize different oscillatory observations (e.g., Krause et al., 2010; Palomäki et al., 2012; Pesonen et al., 2007).

In sum, analyzing ERPs might have the advantage that different ERP components have been extensively studied and thus might be associated with rather specific cognitive processes (see e.g., Kappenman & Luck, 2012). On the contrary, EEG frequency band power and potentially associated cognitive processes are still rather controversially discussed and might require further research (e.g., Krause, 2003). However, the advantage of using frequency band power as a measure of EFs might be that, when averaged over trials, it captures not only the evoked (i.e., time- and phase-locked) oscillatory activity as ERPs do, but also the induced (i.e., time- but not phase-locked) oscillatory activity (see Bastiaansen et al., 2012, p. 33, Fig. 2.3 for a comprehensible graphical explanation of the difference between phase-locked, i.e., evoked, and non-phase-locked, i.e., induced, EEG activity. The simple averaging technique of calculating ERPs potentially cancels out non-phase-locked oscillatory activity, whereas the calculation of frequency band power before averaging separates the power from the phase information of the signal. Consequently, amplitudes of different phases, i.e., the power of non-phase-locked oscillations is preserved using this technique.). Thus, in short, analyzing time-frequency representations of frequency band power might contain more information about oscillatory activity in the brain than ERPs (Bastiaansen et al., 2012). Furthermore, time-frequency representations of frequency band power may allow capturing the oscillatory activity more directly than ERPs, as in principal they do not require the averaging of data epochs for analysis and thus may even be used for an online analysis of changes of oscillatory activity due to changes of the load-situation. Therefore, frequency band power might be a valuable measure especially for the analysis of complex, real-world task material which do not provide high amounts of trials (e.g., Antonenko et al., 2010).

The current thesis primarily focus on oscillatory activity in the alpha frequency band (i.e., alpha frequency band power), as oscillatory activity within this frequency band is generally most pronounced in the EEG and currently best studied with respect to cognitive functioning. Thus, alpha frequency band power was used as dependent measure in all three studies. Additionally, beta frequency band power (Study 1) and theta frequency band power (Study 2) were analyzed. Beta frequency band power was analyzed in Study 1, as the simple digit span task used in this study allowed to disentangle cognitive processes and motor confounds which both affect this frequency band (see also 1.4.1.3). In contrast, theta frequency band power was analyzed in Study 2 as this frequency band power has been especially associated with cognitive control (e.g., Sauseng, Griesmayr, Freunberger, & Klimesch, 2010). The P300 event-related potential was used in Study 1 and Study 2 as a measure of the attentional distribution during increased demands on EFs (see 1.4.1.4 for a discussion of the P300) and pupil dilation was used as an additional, overall measure of the load-situation in Study 2 and Study 3 (see 1.4.2 for a discussion of pupil dilation). In the following, I will describe each of these measures in more detail, focusing on their sensitivity to demands on core EFs and their suitability to be used in complex, real-world task settings like hypertext reading.

1.4.1.1 Alpha frequency band power

The alpha frequency band was one of the first frequency bands that have been identified in the human EEG (Berger, 1929). Alpha frequency band power is observed maximally at about 10 Hz (the so-called alpha peak) over parietal-occipital brain regions in the EEG of healthy young adults if they are in a relaxed yet attentive state (Klimesch, 1999). Alpha seems to reflect activity of thalamo-cortical and cortico-cortical brain networks (Başar, 2012; Klimesch, 1999; Krause, 2003). Interestingly, alpha oscillations of the brain share some commonalities with core EFs with respect to developmental traces, brain disorders, or higher cognitive functions (see Chapter 1.1 for the discussion of the core EFs). For example, alpha commonly is not observed in the spontaneous EEG of humans below age of three (Başar & Güntekin, 2012), increases then until after puberty, and slowly decays with age (Klimesch, 1999). Alpha power may be related to intelligence. For example, a higher alpha frequency band power of the spontaneous EEG during rest periods has been reported to correlate positively with the performance in intelligence tests (Doppelmayr, Klimesch, Stadler, Pöllhuber, & Heine, 2002). Furthermore, alpha oscillatory activity has been observed to be reduced due to cognitive impairment (e.g., schizophrenia or Alzheimer disease; Başar & Güntekin, 2012).

The first functional correlate of alpha that had been identified was the so-called "alpha blocking", first reported by Berger (1929). Berger observed an alpha ERD (i.e., a reduction in alpha band power) when participants that have previously closed their eyes, opened their eyes. Since then, the alpha ERD has been observed as a general phenomenon when participants switch from a task that require less cognitive processing (e.g., a rest period) to a task that requires more cognitive processing (e.g., mental calculation). This leads to the long-standing prevalent hypothesis that alpha oscillatory activity might be regarded as reflecting a kind of 'cortical idling state' of the brain (Pfurtscheller, Stancák, & Neuper, 1996). In this interpretation, the alpha ERD signals that the activity of a global rest-network breaks apart when specific brain networks are needed for task fulfillment. However, recently this prevalent hypothesis has been questioned by several authors that attributed alpha ERS rather than alpha ERD the active role for cognition (Başar, 2012; Klimesch, 2012; Klimesch, Sauseng, & Hanslmayr, 2007; Palva & Palva, 2007). In their interpretation alpha ERS that has been shown to occur for rather specific brain regions and task requirements (e.g., during the retention period of a STM Sternberg task; Jensen, Gelfand, Kounios, & Lisman, 2002) may reflect the active inhibition of brain networks which otherwise might cause interference with the relevant task (Händel, Haarmeier, & Jensen, 2011; Klimesch et al., 2007; Payne & Sekuler, 2014).

In sum, alpha oscillatory activity clearly shows functional correlates of cognitive processing albeit a unique functional correlate could not be established (Başar, 2012) and the manner of alpha reactivity is still matter of debate (Klimesch et al., 2007; Palva & Palva, 2007). With respect to the core EFs updating, shifting, and inhibition, several studies observed an alpha ERD for increased demands on these EFs.

For example, several studies (e.g., Gevins & Smith, 2000; Gevins et al., 1997; Krause et al., 2000; Krause et al., 2010; McEvoy, Pellouchoud, Smith, & Gevins, 2001; Palomäki et al., 2012; Pesonen et al., 2007) used the n-back task paradigm which may be regarded as predominantly loading on WM updating (but see our discussion on the n-back task and underlying core EFs in 1.3.1). These studies consistently observed an alpha ERD at parietal electrodes for increased load on WM updating, that is, for increased n-back levels. Furthermore, TFRs revealed not only a more pronounced decrease of alpha frequency band power (i.e., increased alpha ERD) but also timely longer lasting alpha ERDs for the higher n-back levels (e.g., Pesonen et al., 2007). Generally, the increased alpha ERD for increased n-back load levels has been shown consistently in literature for a wide variety of different stimuli with respect to codality and complexity, ranging from simple stimuli like digits or letters (e.g., Krause et al., 2000) over visuospatial stimuli (i.e., stimuli locations; e.g., Gevins & Smith, 2000; Gevins et al., 1997) to complex stimuli like websites (e.g., Palomäki et al., 2012). Thus, the alpha ERD seems to be rather codality-free. This is in line with the conceptualization of EFs which are also hypothesized to be codality-unspecific (e.g., Baddeley & Logie, 1999).

Furthermore, an alpha ERD has been observed for a simple span task (Stipacek et al., 2003) which is also hypothesized to load on WM updating (see 1.3.2). Interestingly, an alpha ERD may be rather specifically associated with load on WM updating, that is, for continuous updating of WM content. In contrast, for rather passive load on WM storage components (i.e., during the retention interval of a Sternberg WM task) an alpha ERS has been reported (e.g., Jensen et al., 2002). However, an alpha ERD has also been observed in studies that raised demands on the EF inhibition (e.g., in a Stroop task; Hanslmayr et al., 2008), and studies that raised demands on the EF shifting (Sauseng et al., 2006). Thus, an alpha ERD is not unique to demands on the core EF updating but may reflect basic executive demands more generally.

In sum, alpha oscillatory activity has been rather extensively studied for load on WM updating. These studies commonly observed an alpha ERD for demands on the EF updating. An alpha ERD was also observed in studies that put demands on the EFs inhibition or shifting. Thus, alpha ERD seems to be a rather general correlate of demands on EFs. However, the question remains whether alpha ERD might show a general sensitivity for different core EFs. This research question was part of Study 2 where we manipulated load on WM updating and demands on inhibitory control within one single task.

Apart from basic research, alpha ERD has also been used as a rather global measure of the overall load-situation in some complex, real-world task settings (e.g., Antonenko & Niederhauser, 2010; Antonenko et al., 2010; Gerlic & Jaušcvec, 1999; Smith, Gevins, Brown, Karnik, & Du, 2001). However, many of these studies might be criticized for not adequately controlling for possible perceptual-motor confounds of the EEG data (Gerjets et al., 2014; see also 1.3.3). Thus, in Study 3 we were interested in the sensitivity and suitability of alpha ERD as a measure of the load-situation in a complex task consisting of natural text reading and link-selection processes when possible perceptual-motor confounds were ruled out.

1.4.1.2 Theta frequency band power

Oscillatory activity in the theta frequency band generally has been described to form a kind of counterpart to alpha oscillatory activity (Klimesch, 1999; Krause, 2003). For example, over the life span theta power decreases from childhood to adulthood and then increases again in the older age (above approx. 60 years of age) or in case of neuronal disorders (e.g., dementia), thus showing the opposite pattern as alpha does (Klimesch, 1999). Furthermore, with respect to induced oscillatory activity, studies that report an alpha ERD for load on WM updating often also report a simultaneous theta ERS that is most pronounced at mid-frontal electrodes (e.g., Gevins et al., 1997; Palomäki et al., 2012).

Generally, oscillatory activity in the theta frequency band has been associated with processes of cognitive control in WM functioning (Sauseng, Griesmayr, Freunberger, & Klimesch, 2010). Theta seems to be relevant for memory encoding as the strength of theta power during encoding of stimuli seems to predict their later successful recall (Jacobs, Hwang, Curran, & Kahana, 2006). Besides its role for cognitive control, theta may rather directly define the capacity limits of STM. This has been speculated because of the observation that the number of gamma cycles (a frequency band beyond 40 Hz) that fits into a theta cycle (i.e., the theta/gamma ratio) seems to be in accordance with the individual STM capacity as assessed by a digit span task (Kaminski, Brzezicka, & Wróbel, 2011; see also Lisman & Jensen, 2013).

In sum, theta is often observed to form a kind of counter-part to alpha (i.e., in studies showing a theta ERS and an alpha ERD). Functionally, however, it might be different from alpha. Theta might be especially related to processes of controlling brain networks (e.g., in WM functioning; Roux & Uhlhaas, 2014; Sauseng et al., 2010), whereas alpha might be primarily related to processes of cortical inhibition (Klimesch et al., 2007; Palva & Palva, 2007; Roux & Uhlhaas, 2014). We analyzed theta frequency band power in Study 2, thus allowing a comparison with the alpha ERD with respect to the general sensitivity for demands on the core EFs updating and inhibition.

1.4.1.3 Beta frequency band power

Oscillatory activity in the beta frequency band traditionally has been attributed to motor activity rather than to cognitive processing (e.g., Keinrath, Wriessnegger, Müller-Putz, & Pfurtscheller, 2006; Pfurtscheller, Zalaudek, & Neuper, 1998). A beta ERS before the motor activity (e.g., a finger movement) followed by a beta ERD shortly after the motor activity (i.e., the so-called beta-rebound) is the typical oscillatory pattern observed therein (Pfurtscheller et al., 1998).

Recently however, there has been accumulating evidence that beta band activity might not be a purely motor correlate, but might be also related to cognitive processing (Engel & Fries, 2010; Weiss & Mueller, 2012). Thus, the pattern of beta activity especially in a lower beta band range (14-20 Hz) seems to be highly comparable to the oscillatory alpha frequency band pattern. For example, studies that used the n-back task not only observed an increased alpha ERD for increased WM updating load but also an increased beta ERD (Krause et al., 2010; Pesonen et al., 2007). However, in these studies the observed beta ERD might be also attributable to motor activity as the n-back task is not free of motor activity because of the key-presses.

We analyzed oscillatory activity in the beta frequency band in Study 1. In this study we compared the EEG correlates of an n-back task, an operation span task, and a simple digit span task. As the simple digit span task required no motor activity and was hypothesized to also load on the EF updating like the two other WM tasks, in addition to other research questions (see Table 1), Study 1 was also sought to extend research on oscillatory activity in the beta frequency band which might also be attributable to WM load and demands on core EFs.

1.4.1.4 The event-related potential P300

The P300 can be observed as a positive deflection in the ERP curve, peaking in a time-window between 250 ms to 500 ms after stimulus onset and being maximal over parietal electrode positions (Polich, 2007)[5]. A distributed fronto-parietal network in the brain that

---

[5] More precisely, the P300 might be differentiated into a P3a component, peaking maximally over frontal electrodes and the P3b component maximally peaking over parietal electrode positions (Polich, 2007). In the

serves attention and memory processes is hypothesized as neuronal generator of the P300 (Linden, 2005). In line with this assumption, the P300 has been linked to WM functioning. Its elicitation may reflect context-updating processes which occur in WM (Polich, 2007). For example, a P300 is elicited in a so-called oddball paradigm for deviant stimuli (i.e., stimuli that are different from the standard context) within a sequence of standard stimuli (i.e., stimuli generating the standard context). P300 latency may index stimulus complexity (Kutas, Mccarthy, & Donchin, 2007), whereas P300 amplitude may reflect attentional factors (Polich, 2012). Thus, P300 amplitude has been observed to be the larger, the less frequent a deviant in an oddball paradigm occurred (i.e., the more attention is attracted by the new stimulus).

In dual task situations, where a P300 is elicited in a secondary task (e.g., an oddball task) and task demands of the primary task (e.g., text reading) increase, the P300 amplitude has been observed to decrease as a matter of task demands of the primary task. This decrease in amplitude is commonly interpreted as showing that attention is deducted from the secondary task by the increased demands of the primary task that binds attentional resources (Kok, 2001). Subsequently, in neuro-ergonomic research this dual task methodology has been established to estimate the cognitive load of a primary task by measuring the P300 amplitude of a secondary task (Fu & Parasuraman, 2008; Schultheis & Jameson, 2004; Wickens, Kramer, Vanasse, & Donchin, 1983).

Interestingly, in basic research a study by Watter and colleagues (Watter et al., 2001) showed a decreased P300 amplitude for increased n-back load levels (see also 1.3.1). This was interpreted by the authors such that attention might be drawn away from a purely matching sub-task (i.e., the 0-back task) to processes of WM updating in the n-back levels above 0-back. Watter and colleagues (2001) concluded that the n-back task might thus be regarded as a dual task with a purely matching sub-task and a WM sub-task. As in dual task settings especially the EF shifting is demanded, a possible conclusion may be that the observed decrease in P300 amplitude may be associated with this EF. If demands on shifting increase because of the task set of one task becoming more complex, the P300 amplitude

---

current thesis we focused on the P3b only, that resembles the classical P300. Thus, I will use the term P300 rather than P3b.

decreases. Thus, we would expect to observe a decrease in P300 amplitude for all tasks that comprise a dual task character. In this vein, we analyzed the P300 amplitude in Study 1 and Study 2 as a potential index of the EF shifting.

1.4.2 Pupil dilation

Pupil dilation might serve as an overall physiological measure for load on EFs. Given that pupil dilation is of comparable sensitivity and validity as EEG alpha frequency band power to measure load on EFs and given technical developments that pupil dilation for cognitive processes can be securely disentangled from pupil dilation due to changing light conditions, the advantage of using pupil dilation would be that it is the more efficient technique than the EEG. Pupil dilation can be easily measured by a remote eye-tracking system and might in future even be measurable using the increasingly prevalent webcams in mobile computer devices. EEG on the contrary requires (at least to date) a thorough preparation of the electrodes for measurement and a rather expensive technical equipment for amplifying the signal recorded from the scalp. Thus, especially in task settings of applied research, pupil dilation might be a more efficiently usable measure than the EEG for capturing changes in cognitive load.

From the 1960s on, pupil dilation has been studied as an overall measure of cognitive processing demands and levels of arousal (for recent reviews, see, Andreassi, 2007; Beatty & Lucero-Wagoner, 2000; Laeng, Sirois, & Gredeback, 2012). The pupil has been shown to dilate for emotional stimuli (Granholm & Steinhauer, 2004; Hess & Polt, 1960; Partala & Surakka, 2003) as well as for cognitive processing demands like mental calculation (Hess & Polt, 1964), sentence processing (Just & Carpenter, 1993; Sevilla, Maldonado, & Shalóm, 2014), encoding and recall of memory (Goldinger & Papesh, 2012; Van Gerven, Paas, Van Merriënboer, & Schmidt, 2004), visual target detection (Privitera, Renninger, Carney, Klein, & Aguilar, 2010), or decision making (Einhäuser, Koch, & Carter, 2010), to name but a few examples.

Mental effort has been proposed as the common underlying factor that causes pupil dilation in cognitive tasks (Beatty, 1982; Kahneman, 1973). Although the changes in pupil diameter due to cognitive processing demands are rather small (normally less than 1 mm irrespective of baseline pupil diameter), this measure has been proven to be very reliable (Beatty & Lucero-Wagoner, 2000; but see Van Gerven et al., 2004). The pupil dilation normally starts with a short delay of about 200 milliseconds after stimulus onset, reaching its maximal amplitude (dilation) at about 1000 milliseconds (Andreassi, 2007).

Anatomically, pupil dilation is caused by two opposing muscle fibers: Circular muscle fibers under control of the parasympathetic system produce pupil constriction when innervated, whereas radial muscle fibers under control of the sympathetic system produce pupil dilation when innervated. Recently, a direct connection between pupil dilation and the activity of the locus coeruleus in the brain that is central for the noradrenergic system has been proposed, based on the results of fMRI outcomes (Laeng et al., 2012). This indicates a close connection between pupil dilation and general states of arousal. Furthermore, some recent studies showed correlations between pupil dilation and ERPs like the P300 (Murphy, Robertson, Balsters, & O'Connell, 2011) and the N400 (Kuipers & Thierry, 2010, 2013), and (on a single subject basis) pupil dilation and pre-stimulus alpha power (Hong, Walz, & Sajda, 2014). Thus, as Beatty and Lucero-Wagoner (2000) already concluded in their comprehensive review on pupil dilation, pupil dilation might serve as an easily to acquire, representative measure of the overall brain activity (see also Hartmann & Fischer, 2014).

With respect to core EFs rather few studies have used pupil dilation as dependent measure. Nevertheless these studies consistently showed increased pupil dilation for increased load irrespective of the core EF loaded. Pupil dilation has been observed for increased load on WM updating in n-back tasks (Brouwer, Hogervorst, Holewijn, & van Erp, 2014; Ewing & Fairclough, 2010; Karatekin, Marcus, & Couperus, 2007) or WM span tasks (Cabestrero, Crespo, & Quirós, 2009; Kahneman & Beatty, 1966; Peavler, 1974). Furthermore, pupil dilation has been observed for increased demands on inhibitory control in a Stroop task (Brown et al., 1999; Laeng, Ørbo, Holmlund, & Miozzo, 2011; Siegle, Steinhauer, & Thase, 2004), and in a flanker task (van Bochove, Van der Haegen, Notebaert, & Verguts, 2013). Finally, pupil dilation has been observed for increased demands on shifting due to dual tasking situations (Karatekin, Couperus, & Marcus, 2004). In sum, these studies underline the above stated character of pupil dilation as an overall measure of cognitive processing load.

However, as all cited studies used pupil dilation separately in different task settings with different task materials, the research question remains whether pupil dilation would show a genereal sensitivity for different EFs. Furthermore, the research question whether pupil dilation and alpha frequency band power would comparably react to demands on different core EFs has also not been adressed explicitely yet. Study 2 of the current thesis aimed to shed light on these research questions.

With respect to complex, real-world task materials of digital information environments, only a few studies used pupil dilation to measure cognitive processing load and to compare different task conditions (Di Stasi, Antolí, Gea, Cañas, 2011; Iqbal, Zheng, & Bailey, 2004; Oliveira, Aula, & Russell, 2009). For example, Oliveira and colleagues (2009) studied the relevance of web search results using pupil dilation. They found that pupil dilation was increased for the selection of web search results that were considered relevant as compared to less relevant search results. Di Stasi and colleagues (2011) recorded participants pupil diameter, while participants had to perform two shopping tasks on a commercial website, either a goal-oriented search task (find and buy a specific object) or an experience-oriented search task (freely browse through the websites and possibly buy objects of own choice). As baseline in both task conditions served an initial free exploration phase of the website that lasted two minutes, without buying objects. The authors found that in both task conditions the eye pupils significantly dilated from baseline when participants actually began one of the tasks. However, subjective rating scores revealed that the goal oriented search task was experienced as being more difficult as the browsing task. This subjective difference in difficulty between tasks was not reflected in pupillary results. Thus, these findings might indicate that pupil dilation data might be of only limited validity as a measure of load on EFs in a rather complex task setting and further research will be needed on this topic.

Especially in the context of complex, real-world task materials a disadvantage of using pupil dilation as a measure of the overall processing load has to be mentioned. As outlined above, the pupil dilation for cognitive processes causes rather small (yet reliable) changes in pupil size (i.e., less than 1 mm). In contrast, the pupil dilation for adapting to the light situation of the environment is rather large, ranging from a pupil diameter of about 3 mm (sun light) to a pupil diameter of about 7 mm (darkness; see Beatty & Lucero-Wagoner, 2000). Thus, changes in the ambient light condition or stimulus brightness are problematical confounding factors in studies using pupil dilation as (the only) dependent measure, especially if these studies are not controlling for these potential confounds between task conditions. However, technical developments (e.g., using specific computational algorithms to disentangle pupil dilation for cognitive processing from pupil dilation due to the light conditions) have been proposed that might allow using pupil dilation in more unconstraint task settings despite cofounding light conditions (e.g., Jainta & Baccino, 2010).

In Study 3 of the current thesis, we were therefore interested in a direct comparison of pupil dilation and alpha frequency band power as measures for an increased load-situation due to processes of link selection during text reading (i.e., in complex, real-world task material), while controlling for possible perceptual-motor confounds.

1.5 Summary: Research Questions of the Current Doctoral Thesis

Core EFs may play a central role for WM functioning (see 1.2, 1.3.1, and 1.3.2) as well as for complex cognitive tasks like hypertext reading (see 1.3.3). However, their relationship with WM and their hypothesized intertwined yet separable features (Miyake et al., 2000) have to be studied further. Thus, in Study 1 (see 1.5.1) we were interested in whether the assumed conceptual commonalities with respect to core EFs between different classical WM tasks (i.e., an n-back, an operation span, and a digit span task) would be identifiably using EEG measures like alpha frequency band power, beta frequency band power, and the P300. In Study 2 (see 1.5.2) we were interested in the interplay of the core EFs WM updating and inhibition when being manipulated within a single WM task. We addressed this research question using a variety of physiological measures like EEG alpha and theta frequency band power, the P300, and pupil dilation. These measures have been previously proposed as possible measures of cognitive load (e.g., Antonenko et al., 2010; Just et al., 2003; Parasuraman, 2003). Finally, in Study 3 (see 1.5.3) we were interested in whether pupil dilation and EEG alpha frequency band power were comparably measures for increased demands on EFs in complex, rather real-world task materials, that is, for link selection during hypertext reading.

The overarching research question of the current thesis was the use of EEG alpha frequency band power as a general measure for demands on core EFs in basic and applied task settings, that is, to examine the use of EEG alpha frequency band power as a general measure of the global cognitive load-situation which may be grounded in demands on different core EFs. Furthermore, and more generally, the usage of the EEG as research methodology for studying core EFs was sought to be established. So far core EFs have been mainly studied using behavioral methods (e.g., correlational studies of performance scores) or using brain imaging techniques (e.g., fMRI). This may be due to the fact that the concept of EFs originally stems from clinical and neuropsychological research (Suchy, 2009). Although the EEG and alpha frequency band power have been used in basic research to assess the load-situation for example in WM tasks or with some success in applied research (see 1.4.1), to date a framework of core EFs has not been studied systematically using EEG measures. Thus, the present thesis may serve as an initial step in strengthening the use of

EEG (particularly with a focus on frequency band power) for the study of core EFs. Especially, different WM tasks like span and n-back tasks have not been compared using EEG measures yet (Study 1) and the effects of different core EFs on EEG measures have also not been explicitly compared yet (Study 2). Furthermore, we were interested in a direct comparison of the sensitivity of EEG alpha frequency band power and pupil dilation for load on different EFs within one single task (Study 2) as well as in a comparison of the sensitivity of these measures for the load-situation in a complex, real-world task of link selection during hypertext reading (Study 3).

In sum, all three studies we ran focused on core EFs and physiological measures thereof. However, the granularity of the focus on EFs varied between the three studies. Study 1 and Study 2 can be considered as basic research using classical WM tasks with a very specific focus on core EFs and their intertwined nature. In Study 1 two widely used WM task families, namely n-back and span WM tasks, were analyzed with respect to core EFs. In Study 2 the interplay of the core EFs inhibition and updating has been studied within one single WM task, namely an n-back task (that mainly loads on updating[6]) with flanker stimuli (inhibition). Study 3 can be considered as applied research using a task set of natural, online text reading with link selection processes which were expected to put demands on core EFs. In the following, I will briefly review the three studies we ran, focusing on the main research questions (see Table 1 for an overview of the different research questions). The entire manuscripts of the studies are given in Chapter 2. The main results with respect to core EFs will be presented and discussed in Chapter 3.1. See Table 2 for an overview of the three studies.

1.5.1 Study 1: The comparison of n-back and WM span tasks

In Study 1 we were interested in the relationship of the core EFs and WM. We addressed this research question by analyzing the EEG correlates of two widely used WM tasks, an n-back task and a complex operation span task, and a simple digit span task which is traditionally considered as a STM task. Both, the n-back task and the operation span task were

---

[6] In contrast to Study 1, a rather classical view on the n-back task was prevalent in Study 2 where this task was considered to mainly load on WM updating. This different view will be addressed in more detail in the general discussion section.

hypothesized to be conceptually similar with respect to core EFs, that is, both tasks were expected to comparably demand the core EFs updating, shifting and inhibition. In contrast, the simple digit span task should mainly demand the core EF updating (see 1.3.1 and 1.3.2 for a detailed description and analysis of the three tasks with respect to core EFs therein). As dependent measures we used behavioral data, EEG alpha and beta frequency band power, and the P300. Important to note, for the EEG analyses we compared the memorization sub-task of the operation span task and the digit span task with the n-back task as conceptually these task-phases were most similar with respect to timing (i.e., trial duration) and the perceptual, motor, and cognitive demands involved.

Given the sensitivity of the P300 amplitude as a measure of the attentional distribution between sub-tasks in dual-task situations that results in decreased P300 amplitudes for increased demands on shifting (see 1.4.1.4), we expected to observe decreased P300 amplitudes for increased load-levels in the n-back as well as in the operation span task but not in the digit span task. This was expected because of the dual task character of the n-back task (Watter et al., 2001) and the operation span task. For increased load levels, the demands on shifting between the operation and memorization sub-tasks and sub-processes should increase in these tasks but not in the simple digit span task that contained no sub-tasks during the memorization phase and exclusively demanded WM updating.

In line, the TFRs showing the time-courses of alpha and beta frequency band power changes were expected to be more similar between the n-back and the operation span task as compared to the digit span task, thus reflecting the core EFs within these tasks. However, as Study 1, to the best of our knowledge, was the first study analyzing EEG frequency band power in complex span tasks, and moreover, as oscillatory activity captured by EEG frequency band power in TFRs might not be easily linked to single EFs (see 1.4.1), we had no clear a priori expectations concerning the concrete differences or commonalities in the TFR patterns. Generally, we expected to observe a decreased alpha and beta frequency band power (i.e., an increased alpha and beta ERD) for increased load-levels in all tasks. However, we might expect the TFRs of the simple digit span task showing less pronounced alpha and beta ERDs as compared to the operation span task and the n-back task. This is because of, conceptually, less EFs are required for task performance in the simple digit span task as compared to the two other tasks.

A direct comparison of WM span tasks and the n-back task using EEG measures has, to the best of our knowledge, not been conducted yet. Thus, Study 1 was intended to enrich research on both tasks by adding a new, electrophysiological research methodology. So far, n-back and WM span tasks have been compared using behavioral performance data (e.g., Redick & Lindsey, 2013). These correlational studies generally found rather weak correlations between the two WM task families of n-back and complex span tasks. Given the conceptual similarities between these tasks, the observed weak correlations for behavioral performance measures are rather puzzling. To date, a convincing explanation for this observation is missing. As stated above, EEG correlates may allow investigating further whether the hypothesized commonalities between n-back and complex span WM tasks exist and thus may help to advance the longstanding debate concerning the relationship of n-back and complex span tasks as measures of WM functioning (e.g., Jaeggi, Buschkuehl, Perrig, & Meier, 2010).

As a minor research question, we also analyzed the beta frequency band power in Study 1, as our task setting allowed to disentangle whether a beta ERD is mainly attributable to motor activation (Pfurtscheller et al., 1998) or whether it might also reflect cognitive processing (Engel & Fries, 2010; also see 1.4.1.3 for an extensive discussion of beta frequency band power). The differentiation of the beta ERD as reflecting motor or cognitive demands was possible in Study 1 as the simple digit span task which we used required no motor response. Thus, if the simple digit span task would show a beta ERD, this beta ERD should be genuinely attributable to cognitive processing. Furthermore, if the beta ERD would occur comparably in the different tasks, we could, at least indirectly, conclude that the beta ERD reflects cognitive processing (and not motor demands) in all tasks that might be specifically related with WM demands. Thus, Study 1 might valuably contribute to the current debate on beta frequency band power (see also 1.4.1.3).

1.5.2 Study 2: The interplay of core EFs

In Study 2 we were interested in the interplay of core EFs. We studied the relationship of two core EFs, namely updating and inhibition, by specifically manipulating demands on these EFs within one single WM task. This was done by combining an n-back task (see 1.3.1) and a flanker task (Eriksen, 1995; Sanders & Lamers, 2002), that is, by using an n-back task which mainly loads on WM updating with congruent and incongruent flanker stimuli for additionally demanding inhibitory control. Participants performed different n-back load levels (0-back to 2-back) and base their n-back comparison on a central letter while ignoring the congruent or incongruent flankers surrounding the central letter.

As outlined in Chapter 1.1 according to Miyake and colleagues (2000), the core EFs updating, shifting, and inhibition may be partly separable but may share a common underlying factor which has been associated with controlled attention. Traditionally, EFs have been studied using correlational approaches that focus on behavioral performance measures (e.g., latent-variable analyses, Miyake et al., 2000), or using neuroimaging approaches (e.g., fMRI, Nee et al., 2013) that focus on specific activations of brain areas. Both approaches have the disadvantage that the EFs of interest are confounded by the specific tasks used (e.g., Stroop task for inhibition, or an n-back task for updating). Therefore, in Study 2 we proposed a new experimental approach, namely the manipulation of different EFs within one single task, thus avoiding perceptual-motor confounds (see, e.g., Gerjets et al., 2014) that otherwise might impede a direct comparison of core EFs.

As a further innovation, in contrast to most studies on EFs, we focused on EEG correlates of EFs. This was done because the EEG provides a high time-resolution that allows a detailed examination of the oscillatory activity during demands on EFs (see 1.4.1). We analyzed the theta and the alpha frequency band power in TFRs to capture the dynamics of oscillatory activity when the EFs inhibition and updating are demanded within one single task. Furthermore, we analyzed the P300 amplitude as a potential measure of the distribution of attentional demands (see 1.4.1.4). Additionally, we analyzed pupil dilation as a measure of the overall processing load (see 1.4.2). With respect to pupil dilation, we addressed a further research question, namely, whether pupil dilation would be of comparable sensitivity for load on different EFs as compared to EEG alpha frequency band power.

Generally, we expected to observe for increased demands on WM updating and inhibitory control increased pupil dilation, increased theta frequency band power, and increased reaction times, and decreased accuracy and decreased alpha frequency band power. Based on the assumption that EFs might be separable but share some common factor (e.g., Miyake et al., 2000) and based on results by Lavie and colleagues (Lavie, 2005, 2010; Lavie, Hirst, de Fockert, & Viding, 2004), we expected to observe an interaction between increased load on WM updating and increased demands on inhibitory control. In studies by Lavie and colleagues load on WM storage components and demands on inhibitory control were manipulated in dual-task paradigms. A flanker task was used for demanding inhibitory control. This task was integrated in the memorization phase of a Sternberg task that was used for manipulating load on STM (i.e., remember one or six items). The authors reported an increased flanker interference effect for the high as compared to the low STM load condition. Thus, in our task paradigm we also expected to observe an interaction between demands on inhibitory control and WM updating. However, as we specifically manipulated demands on core EFs which were hypothesized to reside within a single WM component, namely the central-executive (see Chapter 1.2), and to share some common underlying neuronal network structures of controlled attention, we had two contrasting hypotheses concerning the direction of the expected interaction. The interaction might be either over-additive, that is, we might observe an increased flanker interference effect for increased demands on WM updating, comparably to the results by Lavie and colleagues (2004, 2005). Such an outcome might occur because of a general depletion of attentional resources under high WM updating load (*depletion hypothesis*). Or, the interaction might be under-additive, that is, we might observe a decreased flanker interference effect for increased demands on WM updating. Such an outcome might occur because of a commonly underlying attention network that is generally activated due to demands on WM updating and that thus might facilitate inhibitory control (*facilitation hypothesis*).

1.5.3 Study 3 a)-c): EEG alpha frequency band power and pupil dilation for link selection during hypertext reading

In Study 3 a) – c) we addressed the research question whether pupil dilation and alpha frequency band power might be equally suited to measure demands on core EFs within a complex, real-world task setting of text reading and link selection. In contrast to hypertext research conducted so far that often confounds initial processes of link selection with an additionally increased load-situation due to the presentation of subsequent webpages, we were specifically interested in the load situation during link selection and thus carefully avoided such possible confounds (see also 1.3.3). This was achieved by using a carefully designed, yet rather artificial hypertext reading situation, and by conducting two follow-up experiments to exclude possible alternative explanations due to perceptual-motor confounds which might have occurred in Study 3 a) despite our careful task design.

In Study 3 a), participants had to read one long hypertext page that was presented in the web browser on the screen. At certain positions of the text, three words were set in brackets, one out of the three was the appropriate, context-matching word, and the other two were filler words. All three words were marked as hyperlinks, but without any hyperlink functionality. That is, when clicking on these marked words, no subsequent webpages were shown. Only the color of the 'hyperlinks' changed, but participants always stayed on the same single hypertext page. Participants were instructed to click on each word marked as a hyperlink that fits the context. This way, we simulated 'real' link selection processes as participants had to interrupt reading and to perform semantic decisions that might be comparably to genuine hyperlink-related decision processes (DeStefano & LeFevre, 2007). As described in Chapter 1.3.3, we expected that processes of link selection would lead to increased demands on core EFs. Thus, we expected to observe an increased alpha frequency band power ERD and an increased pupil dilation when comparing data epochs of purely text reading (baseline condition) with data epochs containing link selection.

Because of potential perceptual, motor, or structural confounds between baseline and test condition in Study 3 a), we conducted two follow-up studies trying to exclude these possible confounds. In Study 3 a) participants only in the test condition but not in the baseline condition performed a mouse-click (i.e., motor confounds). Furthermore, because of the words colored as hyperlinks, perceptual differences between baseline and test condition exist. These perceptual-motor confounds were avoided in Study 3 b). In this study participants had to perform a mouse-click in the baseline and in the test condition, and furthermore, perceptually both conditions were matched. Study 3 b) was expected to replicate the results of Study 3 a). Finally, as baseline and test condition might simply differ because of being positioned at different parts of the text (i.e., as they may contain different levels of text difficulty), Study 3 c) was conducted. In this study, baseline and test condition were kept identical with respect to link selection processes. Therefore, we expected to observe no differences between baseline and test condition in Study 3 c). This outcome would indicate that the expected load-related effects in Study 3 a) and 3 b) were genuinely related to link selection.

In the next chapter the complete manuscripts of the three studies are presented as submitted for publication. Then, in Chapter 3.1, the main outcomes of the studies will be briefly summarized and discussed.

Table 2. Overview of the Three Studies Conducted as Part of the Current Thesis.

Study	N	Research Question	Task Materials	Physiol. Measures	Results
Study 1	16	Do EEG correlates reflect the conceptual similarities and dissimilarities between an n-back WM task, a complex span WM task, and a digit span STM task with respect to demands on core EFs?	**n-back task (digits)** load-levels: 1-back - 4-back EFs associated: *updating, shifting, inhibition* **operation span task** load levels: 3 - 7 digits EFs associated: *updating, shifting, inhibition* **digit span task** load levels: 3 - 8 digits EFs associated: *updating*	P300 alpha TFR beta TFR	P300 amplitude decreased for increased demands in n-back and operation span task. This might reflect load on the EF *shifting.* Oscillatory activity as revealed by TFRs in sum more similar between n-back and operation span task as compared to the simple digit span task. Beta frequency band showed comparable oscillatory pattern as alpha frequency band, even if motor activity was absent. This indicates that beta reflects cognitive processes and not only motor processes.
Study 2	22	Do demands on core EFs interact?	**n-back task with flanker stimuli** load levels: 0-back - 2-back; congruent/incongruent flanker EFs manipulated: *updating, inhibition*	P300 theta TFR alpha TFR pupil dilation	All measures were sensitive for increased demands on *updating* and *inhibition.* Interaction between *updating* and *inhibition* as revealed by most measures: under high *updating* load the flanker interference effect was decreased or absent. Alpha frequency band power and pupil dilation showed a comparably pattern for increased demands on the EFs *updating* and *inhibition.*
Study 3 a-c	23 20 24	Are pupil dilation and alpha frequency band power comparably suitable measures for increased load on core EFs for (hyper-) link selection during natural text reading?	**text with links** presented on the screen to simulate a hypertext reading situation with link selection processes EFs associated: *updating, shifting, inhibition*	alpha frequency band power pupil dilation	Alpha frequency band power and pupil dilation are both suitable measures for increased load due to link-selection processes in a natural reading situation. Both measures were comparably sensitive, yet do not show a significant correlation. Possible perceptual-motor confounds were ruled out by two follow-up experiments.

2. Manuscripts of the Three Studies Conducted as Part of the Current Thesis[7]

2.1 Study 1: Same, same, but different: Commonalities and differences of n-back and working memory span tasks as revealed by EEG P300 and frequency band power data.

Christian Scharinger[1], Alexander Soutschek[2,3], Torsten Schubert[3], & Peter Gerjets[1]

Submitted to *Brain and Cognition*, pending major revisions

2.2 Study 2: When flanker meets the n-back: What EEG and pupil dilation data reveal about the interplay between the two central-executive working memory functions inhibition and updating.

Christian Scharinger[1], Alexander Soutschek[2,3], Torsten Schubert[3], & Peter Gerjets[1]

Psychophysiology, (2015), 1293-1304. doi: 10.1111/psyp.12500

2.3 Study 3: Pupil dilation and EEG alpha frequency band power reveal load on executive functions for link-selection processes during text reading.

Christian Scharinger[1], Yvonne Kammerer[1], & Peter Gerjets[1]

PLoS ONE, (2015), 10(6): e0130608. doi:10.1371/journal.pone.0130608

[1]Knowledge Media Research Center, Tuebingen, Germany

[2]Laboratory for Social and Neural Systems Research, University of Zurich, Zurich, Switzerland

[3]Department of Psychology, Humboldt-University of Berlin, Berlin, Germany

---

[7] Note. The manuscripts presented in the following chapters are not the final, proof-read and copy-edited versions of the articles.

2.1 Study 1

Same, same, but different: Commonalities and differences of n-back and working memory span tasks as revealed by EEG P300 and frequency band power data

Christian Scharinger[1*], Alexander Soutschek[2,3], Torsten Schubert[3], & Peter Gerjets[1]

[1] Knowledge Media Research Center, Tuebingen, Germany
[2] Laboratory for Social and Neural Systems Research, University of Zurich, Zurich, Switzerland
[3] Department of Psychology, Humboldt-University of Berlin, Berlin, Germany

*Correspondence: Christian Scharinger, Knowledge Media Research Center, Schleichstr. 6, 72076 Tuebingen, Germany. Phone: +49 (0) 7071 979-360, Email: c.scharinger@iwm-kmrc.de

Abstract

Although both, n-back tasks and complex span tasks are conceptualized as working memory (WM) tasks, behavioral performance measures normally show rather weak correlations between the two. In addition, correlations between n-back tasks and simple short-term memory span tasks are often reported to be comparably higher, thus questioning the nature of the n-back task as genuine WM task. In the current study we used the electroencephalogram (EEG) to compare n-back and span tasks. We focused on the P300 event-related potential and EEG alpha and beta frequency band power that have been shown previously to be sensitive measures of WM load in the n-back task. Our results revealed a comparable decrease of the P300 amplitude in the n-back task and in the memorization subtask of a complex operation span task for increased WM load that was absent in a simple digit span task. This might indicate the dual-task nature of the both, n-back and complex span WM tasks. The overall oscillatory pattern showing a decreased alpha and beta band power for increased WM load was comparable between the three tasks, although differences in magnitude and timing occur. We interpret these findings as indication that all three tasks might load on WM-updating processes, albeit to a different degree.

Keywords: EEG time-frequency analysis; ERD/ERS; P300; working memory; n-back task; span tasks

Highlights: > Comparison of EEG correlates of n-back, operation span, and digit span task. > Conceptually, n-back and operation span are comparable working memory tasks. > P300 amplitude decreased for increased load in the two, but not in digit span task. > Overall, EEG alpha and beta frequency band power also more similar between the two. > In sum, EEG correlates confirmed conceptual task commonalities and differences.

1. Introduction

Working memory (WM) can be defined as a central cognitive construct essential for other higher order cognitive abilities like reading, mathematics, problem solving or goal pursuit (Shipstead, Redick, & Engle, 2012; St Clair-Thompson & Gathercole, 2006; Yuan et al., 2006). For example, a typical situation requiring WM processes might be when you have been out for dinner and before paying you want to check the bill by doing quick mental calculations: You probably add the values of single positions, temporarily keeping intermediate results in mind until having calculated the final sum. In the meanwhile you might be trying to not getting distracted by ongoing conversations or even shifting back and forth between a conversation and the mental calculation.

Structurally, WM is hypothesized to consist of a processing component associated with controlled attention (the so-called central-executive component, incorporating executive functions), and short-term memory (STM) components for temporarily storage of to-be-processed information (Baddeley, 1992, 2003, 2012). Thus, in the above example the processing component might control the calculations, whereas intermediate results would be temporarily stored in the memory components. Functionally, WM is characterized by a temporary retention and manipulation of information under the focus of attention in the face of interference (Cowan, 2000; Engle, 2002; Kane & Engle, 2002; Oberauer, 2009). Amongst others, (and although sometimes slightly differently defined or named) often cited basic executive functions (EFs) in WM are updating, shifting, and inhibition (Bledowski, Kaiser, & Rahm, 2010; Miyake et al., 2000): The updating of information temporarily memorized for processing, the shifting of the attentional focus between different task demands, and the inhibition (or interference control) of information not (or no longer) relevant for the current processing step.

Often used tasks in WM research are complex span tasks (e.g., Unsworth & Engle, 2007) and n-back tasks (e.g., Gevins & Smith, 2000). Albeit originating from different research areas (individual differences research versus neuropsychological research), both, complex span tasks and n-back tasks, are conceptualized as genuine WM tasks that load on EFs. However, direct comparisons of behavioral performance measures of exemplars of the two task families revealed rather weak correlations between the two (e.g., Jaeggi, et al., 2010;

Kane et al., 2007; Miller et al., 2009; for a recent review and meta-analysis see Redick & Lindsey, 2013). Thus, in the current study we propose to apply another, namely an electrophysiological research methodology to compare n-back and span tasks. Electrophysiological measures based on the electroencephalogram (EEG) might allow us to examine whether dissociable processes are involved in n-back and span tasks and, by this, to potentially unravel reasons for the weak correlations between these tasks.

1.1. Conceptual similarities of N-back and WM Span Tasks With Respect to EFs

Complex span tasks like the reading span task (Daneman & Carpenter, 1980) or the operation span task (Turner & Engle, 1989) are dual-tasks consisting of a processing subtask and a (short-term) memory subtask. For example, in a classical operation span task, participants have to validate the results of simple equations (e.g., 3*5-10=4?) intermixed with the memorization of certain items (e.g., words, letters or digits). After three to seven trials, each trial consisting of an equation-validation paired with a memory-item, participants are asked for recall of the memorized items in correct serial order. In contrast to simple span tasks, like a digit span task where participants 'simply' have to update short-term memory content (e.g., to memorize several digits for later recall in correct serial order), in complex span tasks WM processes like shifting and inhibition are additionally required: Participants have to shift between the two subtasks and to inhibit currently irrelevant information. The amount of correctly recalled items defines the individuals' memory span and is used as measure of WM capacity. Complex span tasks are mostly used in correlational research addressing the connection between WM capacity and performance in higher order cognitive processes like reading or mathematics mostly under an individual differences research perspective (e.g., Bayliss, et al., 2003; Unsworth & Engle, 2007).

In n-back tasks (Jonides et al., 1997) stimuli are presented in a temporal sequence and participants have to decide via key-press whether a current stimulus is or is not identical to a stimulus they saw n-steps back in the sequence (e.g., in a 2-back task this would be each stimulus 2-steps back in the sequence). Thus *n* defines the WM load in this task. Depending on *n*, the n-back task differently requires WM processes like updating, shifting, and inhibition. Whereas the 0-back task condition is a simple matching task with no specific demands on EFs (as participants simply have to compare the current stimulus within the sequence with one specific stimulus that had been shown at the beginning of the sequence), in the 1-back or 2-back task condition participants have to continuously update the temporarily memorized stimuli, shift the attentional focus between the stimuli for comparison, and inhibit stimuli no longer necessary for comparison, or inhibit incorrect response tendencies for stimuli at the wrong position in the sequence (in the 2-back task condition; see for an in-depth task analysis, e.g., Chen, Mitra, & Schlaghecken, 2008; Jonides et al., 1997). There is some evidence that the n-back task might have a dual-task character (Watter, Geffen, & Geffen, 2001). Thus from a conceptual point of view the n-back task should share several processes and mechanisms with complex span tasks. The n-back task is mainly used in neuropsychological WM research (e.g., Gevins & Smith, 2000; Owen et al., 2005; Pesonen et al., 2007).

To sum up, both, complex span and n-back tasks are conceptually designed as genuine WM tasks that might comparably incorporate the EFs updating, shifting, and inhibition. Nevertheless, as mentioned above, behavioral measures generally show only weak correlations between the two tasks (mean $r \sim .20$; see Redick & Lindsey, 2013), and even higher correlations are reported between the n-back and simple span tasks (e.g., Gevins & Smith, 2000; Roberts & Gibson, 2002). This observation is challenging to interpret, as simple span tasks like the digit span task are conceptualized as STM tasks and may be regarded to mainly put demands on the EF updating.

Thus, in the current study we examined the relationship between an n-back WM task, an operation span (complex span) WM task, and a simple digit span (STM) task. These tasks were all closely matched with respect to time constraints (i.e., timing of trials), number of trials, and the overall duration. We analyzed EEG event-related potential (ERP) data (the P300) and oscillatory data (alpha and beta frequency band power) that might serve as measures for the EFs activated in the different tasks. We focused on the memorization subtask of the span tasks (see Figure 1) for which we hypothesized that in the operation span task comparable WM processes of updating, shifting and inhibition would be required as in the n-back task, whereas in the digit span task only updating processes might be required. We hypothesized that these different requirements for EFs might specifically be manifested in the

EEG, that is, in the P300 and in EEG frequency band power as captured by time-frequency representations (TFRs).

1.2. EEG Correlates of N-back and Span WM Tasks With Respect to EFs

Both, P300 and EEG frequency band power have been shown to be sensitive load measures mainly in the context of n-back tasks. However, to the best of our knowledge a direct comparison of span and n-back tasks within one single study by means of these EEG measures has not been conducted yet. In the following, we will briefly review studies that used these measures in n-back or span tasks and we will formulate hypotheses about associated EFs and potential outcomes of the measures.

1.2.1. The P300. The P300 amplitude might be specifically sensitive for dual task situations, that is, when the EF shifting is needed. For increased n-back load levels the P300 amplitude decreased (Chen & Mitra, 2009; McEvoy et al., 2001; Watter et al., 2001). Thus, the P300 amplitude in an n-back task showed a comparable outcome as in dual-task studies (e.g., Kok, 2001), leading Watter and colleagues (2001) to conclude the n-back being of a dual-task nature. In general, according to Watter and colleagues (2001), the P300 amplitude serves as an index for the internal distribution of controlled attention when different executive WM functions are required. In contrast, P300 latency seems to be unaffected by any increase in load above the 1-back load level (Watter et al., 2001).

Only few studies have addressed the P300 for measuring load in simple span tasks, yet yielding mixed results. Grune et al. (1996) studied the P300 amplitude in a digit span task. Participants had to remember sequences of seven digits, each digit presented one after another on the screen. The authors compared the averaged ERP curves for each digit position and found a decrease of the P300 amplitude for increased digit position (i.e., increased load on updating). However a closer examination of the results revealed a significant decrease of the P300 amplitude only up to the fourth digit position. Contrary to the results by Grune and colleagues, a comparable study by Gross, Metz, and Ullsperger (1992) found an increased P300 amplitude for increased load in a digit span task.

In sum, the results concerning the P300 in span tasks are sparse and conflicting for simple span tasks like digit span tasks, and, to the best of our knowledge, are non-existent for operation span tasks. However, as both, the n-back and the operation span task are (a) conceptualized as WM tasks, (b) share a dual-task nature (Watter et al., 2001) and (c) possibly require common WM processes like updating, shifting, and inhibition, we would expect to observe a comparable decrease in P300 amplitude in both tasks for increased load levels. On the contrary, the P300 latency might be unaffected by increased WM load in both tasks. For the digit span task we expected to observe either an increase or a decrease in P300 amplitude (Gross et al., 1992; Grune et al, 1996), however in case of a decrease a less pronounced one as compared to the operation span or the n-back task; this is because the digit span task requires less EFs compared with the complex span and the n-back task.

1.2.2. Time-frequency representations (TFRs). TFRs visualize the oscillatory activity measured as frequency band power values for a certain time period and frequency range (see Figure 5). Several studies reliably found a characteristic oscillatory pattern in TFRs for the n-back task when analyzing oscillatory activity at parietal-midline electrodes in a frequency range of about 4 Hz to 30 Hz (Krause, Pesonen, & Hämäläinen, 2010; Palomäki et al., 2012; Pesonen, Hämäläinen, & Krause, 2007). For example, Pesonen et al. (2007) used a verbal n-back task (visually presented letters) with the load-levels 0-back to 3-back and analyzed the oscillatory event-related desynchronization (ERD) and synchronization (ERS) pattern for a frequency range between 4 Hz to 30 Hz in the time-window 0 – 1800 ms after stimulus onset. At parietal electrode sites, they observed in the alpha frequency band (8 – 13 Hz) an ERD (i.e., a decrease in alpha frequency bad power) that became more pronounced and of longer durations in the higher n-back load levels (2-back, 3-back) as compared to the lower load levels. In the beta frequency band (14 – 24 Hz) the authors observed a comparable pattern of load-related oscillatory activity than in the alpha band, with a beta ERD that also became more pronounced and of longer durations in the higher n-back load levels. They attributed these effects as reflecting increased load on WM processes. Additionally, the authors observed an ERS in the beta frequency band close to the alpha frequency band starting about 500 ms post-stimulus onset and visible only in the low (0-back, 1-back) load conditions. The authors identified this ERS as a possible post-movement beta-rebound (cf., Pfurtscheller, Zalaudek, Neuper, 1998).

When looking at mean alpha frequency band power averaged over time, the alpha frequency band power is commonly reported to decrease with increasing n-back levels (Gevins et al., 1997; Gevins & Smith, 2000) which normally is interpreted in terms of indicating load on WM processes (Klimesch, 1999). For a digit span task, Stipacek et al. (2003) also reported a decreased alpha frequency band power for increased numbers of to-be-remembered digits (but see in contrast, yet in other task settings, reports on increased alpha frequency band power associated with increased WM load, for example, Jensen et al., 2002; Palva & Palva, 2007).

In sum, with respect to oscillatory activity in the alpha and beta frequency band, we expected to observe TFRs at parietal electrodes showing an oscillatory pattern more similar between the two WM tasks (n-back, operation span) as compared to the digit span (STM) task. However, as a direct comparison of TFRs for n-back and span tasks has not been done before, we had no clear expectations how exactly the TFR patterns of the digit span task might differ from the two WM tasks. Generally, we expected to observe a decrease of alpha frequency band power and beta frequency band power (i.e., an increased alpha and beta ERD) for increased load in all tasks. However, we might observe task specific effects in magnitude and timing of the ERDs because of the hypothesized different requirements for EFs in the current tasks.

1.3. The Current Study

To sum up, in the current study we were interested in a direct comparison of n-back and span tasks, with focus on a numerical n-back and the memorization subtask of an operation span and a digit span task, using the EEG measures P300 and alpha and beta frequency band power TFRs. Because of their conceptualization as genuine WM tasks and their comparatively assumed dual-task nature, we hypothesized that both, the n-back and the operation span task would incorporate the EFs updating, shifting, and inhibition. In contrast, we hypothesized the simple digit span task to require less EFs, potentially only updating. Thus, we expected to observe EEG patterns of the P300 and TFRs being more similar for the n-back and the operation span task as compared to the digit span task. Especially, because of the hypothesized dual-task nature of the operation span and the n-back task we expected to observe a comparable decrease in P300 amplitude for these two tasks.

2. Method

2.1. Participants

Sixteen university students (age: *M* = 25.06, *SD* = 3.32; 9 females) participated in the study and received a payment of 8,- €/h. They were all native speakers of German, right-handed according to the Edingburgh Handedness Inventory (Oldfield, 1971) and reported no neurological disorders. All participants had normal or corrected-to-normal visual acuity. The study was approved by the local ethic committee. Participants gave their written consent at the beginning of the study. Four participants had to be excluded from final data analysis because of falling below our minimal criterion of achieving at least fifteen artifact-free, correctly recalled trials for the higher load conditions in the operation span task (see section 2.5.).

2.2. Materials

2.2.1. n-back task. Eight different single digits (1-9, except 7) served as stimuli in the n-back tasks (the 7 was excluded because in contrast to all other used digits it consists of two syllables; thus potentially it may load verbal WM differently). The digits were either printed in blue (RGB-values: 51,75,177) or in yellow (51,75,177) on black background and printed in one of four different fonts (Arial; Curlz MT; Viner Hand ITC; Castellar) in 25 points font size each. This combination of two colors and four fonts resulted in eight different stimuli forms. The stimuli were presented at one of the eight outer positions of a 3x3 grid. The grid was centrally located on the screen and was marked through grey colored thin horizontal and vertical lines. The height and width of the grid was about 5.5 cm each. The assignment of digit value, position, and form (color and font) was done randomly for the current stimuli in the n-back sequences.

The stimuli were presented sequentially at a random position within the grid for 500 ms followed by 1500 ms of black screen with only the grey grid lines visible (cf., Figure 1).

Thus one trial lasted 2000 ms. A sequence of 34 trials constituted an n-back block. The first four trials of a block were excluded from data analysis. The total duration of a block (including task instruction) summed up to about 90 seconds.

During an n-back block, participants had to indicate via key-press whether, or not, the stimulus of the current trial matched the stimulus they saw n-steps back with respect to a certain stimulus dimension (i.e., digit value, location, or form). One third of the trials of a block were 'matches', that is, required participants to press the 'yes'-key ('m') as correct response, two third of the trials were 'no-matches', that is, required participants to press the 'no'-key ('x') as correct response. The sequences of 'matches' and 'no-matches' were pseudo-randomly generated with the constraint that after a maximum of three 'matches' at least one 'no-match' followed.

At the beginning of each n-back block the actual task condition and n-back level for the current block was announced. We used four n-back levels (1-back, 2-back, 3-back, 4-back) and three task conditions (digit value, position, form), albeit data analyses for comparison of the n-back and the span tasks was restricted to the numerical (digit value) n-back condition. During blocks of the task condition 'digit value', participants were instructed to focus only on the digit values and to perform the n-back comparison process only on this stimulus dimension and ignore the other two dimensions (position and form). During blocks of the task condition 'position', participants had to focus only on the position of the stimuli while ignoring digit values or forms, during blocks of the task condition 'form' they had to focus only on the form (color and font) of the stimuli while ignoring the other two dimensions. In the 1-back task condition participants had to compare the announced stimulus dimension (e.g., value, position, or form) of the current stimulus with the same stimulus dimension of the directly preceding stimulus. In the 2-back (3-back, 4-back) load condition the stimulus for comparison was the stimulus presented two (three, four) steps back in the sequence.

In total we used twelve different task conditions (digit value, position, and form, each 1-back to 4-back). Each task condition was assigned to three blocks. This resulted in 36 blocks in total. The sequence of the 36 blocks (four n-back levels and three n-back conditions) was randomly created for each participant. Importantly, all blocks were perceptually identical. Only the pre-block task instruction defined the task at hand for the participants.

Participants performed a training session of each n-back level and n-back condition (with blocks of only 20 trials) once before the actual task (36 blocks) started. The total duration of the n-back tasks, including training and breaks, summed up to about 90 minutes.

Data of the different n-back task conditions (value, position, form) were recorded to address a second research question, namely the comparison of EEG correlates for different kinds of n-back tasks. However, in the current paper we restricted our research question on the comparison of n-back and span tasks. As we only had data for verbal variants of the span tasks (operation span and digit span), we only used data of the verbal n-back task conditions ('digit value') for further analyses and comparisons.

2.2.2. digit span task. We used eight different single digits (1-9, except 7) as stimuli for the digit span task (i.e., the same digit values as in the n-back task). The digits were printed in grey color on black background in Arial font (25 points font size). Sequences of single digits were presented at the center of the screen. Each digit was shown for 500 ms, followed by a fixation-cross for 1500 ms. Thus, a digit span trial was of the same length as an n-back trial. Participants were instructed to remember the digits they saw on the screen in correct order. After three to eight trials (digits) a recall screen was shown where participants had to type in the remembered digits. Participants were allowed to take a self-paced short break after having typed in the digits. The next digit span sequence started when participants pressed the 'return'-key.

Participants did not know beforehand the length of the current digit span sequence (i.e., at which trial position the recall-screen would occur). We used 30 sequences of eight digits, and 18 sequences of seven to three digits each, respectively. At the end of these digit span sequences the recall screen was shown. The single digits of a sequence as well as the order of the sequences were randomly chosen for each participant. Importantly to note, as the higher digit span sequences incorporated trials of the lower digit span sequences (e.g., in a digit span sequence of eight digits are incorporated trials of the lower digit span sequences one to seven), the total amount of trials per load level (i.e., the current trial position in a digit span sequence) was additive by the factor 18. Thus, we had 30 trials of load level eight, 48 trials of load level seven, 66 trials of load level six and so on.

The duration of a sequence varied between six to 16 seconds, depending on length. The total duration of the digit span task (including the recall screens) summed up to about 35 minutes. Therefore, we split the digit span into two parts of about 16 minutes length that were presented in alternation with two parts of the operation span task (cf. below).

2.2.3. operation span task. The operation span task design built upon the digit span task design: Participants had to remember sequences of single digits (1-9, except 7) of different length (memorization subtask). Additionally, they had to perform simple calculations within each trial (processing subtask).

Each trial in the operation span task started with a simple equation (e.g., 3*4-8), centrally presented on the screen (Arial, 25 points font size, grey-colored, black background). The equation was shown for 2500 ms followed by an equal sign presented for 500 ms. Then a single digit (1-9, except 7) was shown for 500 ms as possible result of the equation. This digit was blue-colored to ease participants' comprehension of the task instruction, which digit they had to remember for later recall (see below). After a question mark, presented for 1500 ms, the next trial began. The total duration of an operation span trial summed up to five seconds.

The task of the participants was to indicate via key-press whether the single digit presented before the question mark was the correct result of the preceding equation or not. Assignment of response keys and length of the response window was the same as in the n-back task, as well as the probability of 1/3 matches (i.e., correct result shown) and 2/3 no-matches (i.e., wrong result shown). Additionally, like in the digit span task, participants had to remember these (colored) digits (i.e., the possible results) for later recall in correct order. After sequences of three to seven trials, a recall screen was shown, identical to the procedure of the digit span task described above.

The equation was either of the form of a multiplication combined with a subtraction (a*b-c) or a division combined with an addition (c/b+a), with 'a' and 'b' being single digits (1-

9, except 7) and 'c' being a two digits number. The third operator, 'c', was chosen to create a result of the equation that again was a one digit number (1-9, except 7).

For the operation span task, we set the maximal length of a sequence to seven trials. We had 30 sequences of length seven, and like in the digit span task, 18 sequences of length six to three, respectively. Thus, the amount of trials per load level was comparable to the digit span with 30 trials for the load level seven, 48 trials for the load six, 66 trials for the load five, and so on.

The duration of a sequence varied between 15 to 35 seconds, depending on length (three to seven trials). The total duration of the operation span task (including the recall screens) summed up to about 55 minutes. We split the operation span task into two parts of about 28 minutes length that were presented in alternation with two parts of the digit span task (cf. above).

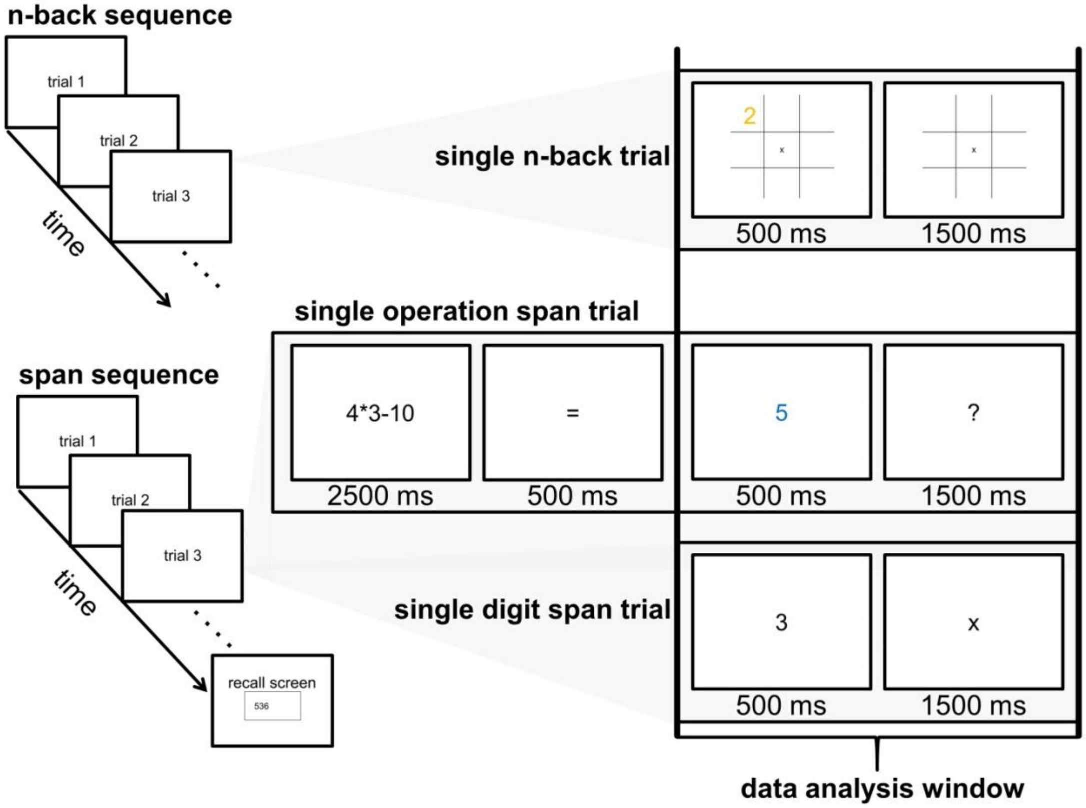


Figure 1. Schematic sequence of n-back and span task trials (left). The right side of the figure exemplarily shows single n-back, operation span and digit span trials and their timing. Marked by thick vertical lines is the 2000 ms data analysis window that we used for all tasks.

2.3. Procedure

Participants performed two experimental sessions on two different days within one week. Each session, including EEG preparation and breaks, lasted to about two and a half hours. In the first session, participants performed the n-back tasks. In the second session, the same participants performed the operation span and digit span task, each split into two parts and in alternation with one another. At the beginning of each session, participants got training trials of the sessions' tasks.

2.4. Apparatus

The study was run in a quiet room that was dimly lightened. Participants sat in a comfortable chair in front of a 17-inch monitor (iiyama ProLite E481S, 1024 x 768 pixels screen resolution, about 70 cm viewing distance) while their EEG data were recorded. EEG data were recorded from 27 electrode sites (Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, O2) positioned according to the international 10/20 system (Jasper, 1958). The right mastoid served as reference during recording. Ground electrode was positioned at FPz. Three additional electrodes were placed around the eyes for EOG recording. EEG data were recorded with the BCI 2000 toolbox (Schalk et al., 2004) at 512 Hz sampling rate (two 16 channels g.USBamp Generation 3.0 amplifiers, g.tec medical engineering, Inc.) using active electrodes (ActiCap, Brainproducts, Inc.). Impedances were kept below 5 kOhm.

2.5. Data Preprocessing and Analysis

EEG data were preprocessed and analyzed using customized analysis scripts (Matlab 2012b, MathWorks, Inc.; EEGLAB v. 11.0.5.4b, Delorme & Makeig, 2004). During preprocessing the continuous EEG data were filtered (low-pass 40 Hz, high-pass 0.5 Hz, linear finite impulse response filters). Electro-Occulogram (EOG) artifacts were corrected using independent component analysis (ICA) decompositions. Independent components (ICs) identified as EOG-ICs by visual inspection were rejected. EEG data were re-referenced to average reference.

After preprocessing, the continuous EEG data were divided into stimulus-locked epochs of 2000 ms length. The data analysis window which is covered by these epochs is marked in Figure 1 for all three tasks (n-back, operation span, digit span) in comparison. We chose a data analysis window that covered aspects of the three different tasks that were highly comparable. Within the data analysis window in all tasks participants were presented digits (500 ms) they had to remember followed by a 1500 ms retention period. In the n-back task and the operation span task participants additionally had to perform decision processes during this 2000 ms time window.

An automatic artifact removal was performed with respect to the EEG data: Epochs that exceeded ±100 µV were excluded from further analyses (Duncan et al., 2009). In doing so, epochs containing severe artifacts (e.g., muscle artifacts) were excluded. No further artifact removal or correction was performed on the EEG data.

With respect to the n-back and the span tasks, only correct trials were included in the analysis. For the span tasks, we defined correct trials as those trials, whose digits were remembered and recalled later at the correct serial position when being typed in at the recall screen. This way we wanted to ensure that participants in fact had activated memory processes during the trials. Figure 2 shows the recall accuracies for the different digit positions in the sequence (p1 to p8), that is, the recall accuracies for increasing WM load. The recall accuracy of a digit trial in the sequence was calculated as the percentage of correctly recalled digits at the specific digit position with respect to the total amount of trials of this specific digit position (i.e., we performed the first step of the partial-credit, unit scoring procedure used for the calculation of an overall span score as described by Conway et al., 2005).

We run a one way repeated-measures analysis of variance and post-hoc pairwise comparisons on these accuracy values. This analysis (see section 3.1.2. and 3.1.3.) revealed certain steps within the digit sequence where recall accuracy significantly dropped and others parts within the sequence for which recall accuracy remained quite stable. For example, the accuracy is quite comparable for remembering one to three digits and it was significantly reduced for remembering four digits. Using results of this analysis, we defined three load categories (l1 = easy, l2 = mid, l3 = difficult) and corresponding trial difficulty levels. For the operation span and the digit span task, the load category l1 was formed by the trials within the sequences of the third digit (digit position p3) to be remembered. The load category l2 was formed by trials of the fifth (operation span) or sixth (digit span) digit to be remembered (p5 or p6, respectively). The load category l3 was formed by the combined trials of the sixth and seventh digit to be remembered for the operation span (i.e., p6 + p7) and the combined trials of the seventh and eighth digit to be remembered for the digit span task (i.e., p7 + p8). We used a different load assignment of these three categories for the operation span and the digit span to match the difficulty levels between the span tasks. As in the n-back task reaction times and accuracy indicate no significant difference between 3-back and 4-back load level, we used the 1-back to 3-back load level as representative for defining load level l1 to l3 and excluded the 4-back load level from any further analysis.

In the span tasks, we combined trials of two digit positions for the high load condition (l3) to increase the total amount of correctly recalled trials which would have otherwise been too low for a concise EEG data analysis in many subjects. Nevertheless, we had to exclude four participants from further analyses. These participants showed too few correct trials in the operation span high load (l3) category (< 15 trials). All other participants had in average 43 (SD = 16) correctly remembered trials in the high load condition of the operation span, which fitted the recommendations for running P300 analysis (e.g., Duncan et al., 2009). In all other tasks and conditions the minimal trial amount was above 50 artifact-free, correctly recalled trials.

2.5.1. P300. Single trial EEG data epochs were averaged separately for each load level and participant, using a -150 ms prestimulus baseline. We calculated P300 latencies for each load level and participant separately by detecting the most positive data point within the time window 250 ms to 450 ms after stimulus onset at electrode Pz (Polich, 2007). P300 peak amplitudes were then defined for these Pz latency time points for all electrodes (Picton et al., 2000). Figure 3 shows the grand average event-related potentials at electrode Pz for the three tasks.

2.5.2. Time-frequency representations (TFRs). We calculated TFRs at electrode Pz for each task, load level and participant separately within a frequency band range from 2 Hz to 32 Hz and a time range from 0 ms to 2000 ms. The frequency band power for the TFRs was calculated using stepwise fast-fourier transforms (FFTs, 500 ms width moving windows, 10 ms steps) applied over the entire epoch length. Then the percentages of event-related desynchronization/synchronization (ERD/ERS%; Pfurtscheller & Lopes da Silva, 1999) were calculated for each data point with respect to a baseline. As baseline we used the mean frequency band power of the load level l1 (i.e., the averaged power over the entire epoch length of l1) for each task individually. Grand-average TFR plots (TFRs averaged over all participants) are given in Figure 5. Blue colors denote percent of event-related desynchronizations (ERD), red colors indicate event-related synchronizations (ERS). Pairwise comparisons of consecutive load levels within a task (rows) were performed (EEGLAB bootstrapping statistics, using false-discovery-rate corrections for multiple comparisons) and are shown in Figure 5 as small plots in between the TFR plots. Red color areas in the small plots indicate data points of statistically significant differences between two consecutive load levels ($p < .05$).

For further statistical analysis comparable to the P300 analysis, we defined two frequency bands, alpha (10 – 13 Hz) and beta (14 – 24 Hz) and two time-windows (0 – 800 ms, 800 – 1600 ms) by visual inspection of the TFRs. For each frequency band separately, the ERD/ERS% values were averaged over time and frequency range and analyzed using 3-factorial repeated-measures ANOVAs with the factors task(3), time-window(2), and load(3).

3. Results

3.1. Behavioral Data

For each task, we run separate one-factorial repeated-measures ANOVAs. Greenhouse-Geyser corrections were performed on the *p*-values where necessary. For post-hoc pairwise comparisons (t-tests, two-tailed) all *p*-values were Bonferroni corrected for multiple comparisons. Level of significance was set at $\alpha = .05$ for all analyses and partial eta-square (η_p^2) is reported as a measure of effect size. Figure 2 shows the mean recall accuracies of the span tasks (for digit positions p1 to p7 and p8, respectively) and the accuracies and reaction times of the n-back task.

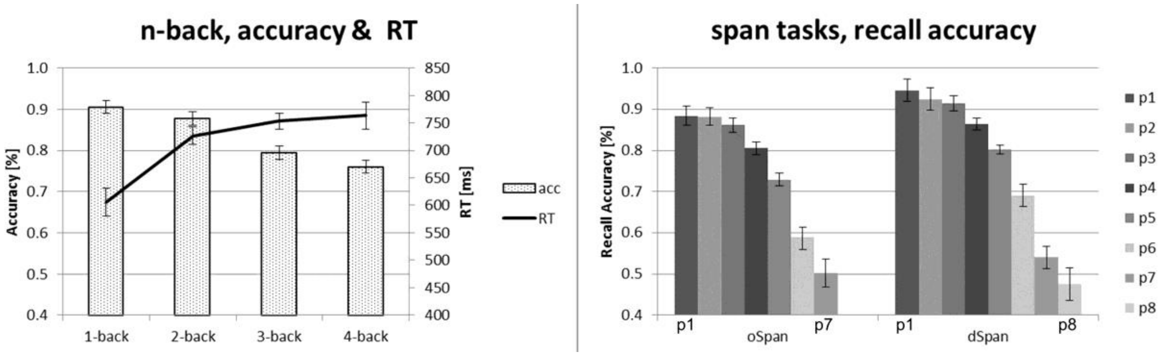


Figure 2. Behavioral data results. Mean accuracy and reaction times (RT) for the n-back task on the left-hand side. Mean recall accuracy for each digit position for the operation span (positions p1 to p7) and the digit span task (positions p1 to p8). Black error bars indicate ±1 standard error of the mean.

3.1.1. n-back task. Reaction times (RTs) were calculated only for correctly responded trials. For RTs, the one-factorial repeated-measures ANOVA revealed a main effect of load, $F(3, 33) = 9.11$; $p < .001$; $\eta_p^2 = .48$. The strongest increase in RTs could be observed for the step from the 1-back (606 ms) to the 2-back task condition (727 ms, $p = .006$). As can be seen in Figure 2, the RTs increased further on the higher load levels (3-back: 753 ms, 4-back: 763 ms), however this increase was rather marginally and statistically not significant between load levels above the 2-back level (all $p = 1.00$).

Accuracy decreased with increasing n-back load as shown by a main effect of load, $F(3, 33) = 8.93$; $p < .001$; $\eta_p^2 = .45$. Interestingly, accuracy did not differentiate between 1-back (90%) and 2-back (88%, $p = 1.00$) and between 3-back (80%) and 4-back (76%, $p = 1.00$) load level. Only between 2-back and 3-back load level there was a significant decline in accuracy ($p < .05$).

Thus it seemed reasonable to only use n-back levels 1-back to 3-back for a direct comparison with three defined load categories of the span tasks: RTs increased most between 1-back to 2-back load level, accuracies decreased significantly between 2-back and 3-back load levels. No significant difference could be observed between 3-back and 4-back load levels on neither RTs nor accuracy.

3.1.2. operation span task. In the operation span we observed a significant decrease of recall accuracy with increasing digit positions, as revealed by a main effect of load, $F(6, 66) = 37.01$; $p < .001$; $\eta_p^2 = .77$. Post-hoc pairwise comparisons showed certain steps of comparable recall accuracy followed by significant drops in accuracy. Recall accuracy for trials at digit position p1 to p3 was statistically equal (p1: 88%, p2: 88%, p3: 86%, all $p = 1$) and significantly higher as recall accuracy for trials at digit position p5 (73%, all $p < .05$). Trials at this digit position in turn showed higher accuracies than trials at digit positions p6 and p7 (p6: 59%, p7: 50%, all $p < .05$). Trials at these digit positions in turn showed quite equal accuracies ($p = .19$). Thus our classification of three load-levels and the assignment of digit position p3 to load level l1, p5 to l2 and p6 and p7 to l3 as described in the method section seemed to be justified by the results of this statistical analysis. Additionally, we checked the performance in the processing subtask of the operation span (i.e., the accuracies and reaction times for the decision, whether the given result is the correct or wrong result of the preceding equation).

The accuracy of the operation subtask (equation-result decision) numerically decreased only slowly and showed in average 77% correct responses, with a range between 81% ($SD = 12$) at trial position p1 to 68% ($SD = 20$) at trial position p7. A one-factorial repeated-measures ANOVA revealed only a trend for a main effect of load, $F(6, 66) = 2.78$; $p = .07$; $\eta_p^2 = .20$. In turn, the reaction times of the equation-result decision numerically increased for increased trial positions from p1: 744 ms ($SD = 183$) to p7: 769 ms ($SD = 142$) and an average of 720 ms ($SD = 165$). RTs of the result-decision were significantly influenced by WM load, $F(6, 66) = 2.67$; $p = .022$; $\eta_p^2 = .20$. These results nevertheless generally indicated that participants as expected were equally performing both subtasks of the operation span task, that is, the processing subtask and the memorization subtask, thus confirming the successful execution of the task.

3.1.3. digit span task. In the digit span task we also observed a significant decrease of recall accuracy with increasing digit positions, as revealed by a main effect of load, $F(7, 77) = 45.62$; $p < .001$; $\eta_p^2 = .81$. Like in the operation span, we observed certain steps of comparable recall accuracy followed by significant drops in accuracy. Recall accuracy was comparable between trials at digit positions p1, p2 and p3 (p1: 95%, p2: 92%, p3: 91%, all *p* = 1.00). Recall accuracy of these trials was significantly higher than recall accuracy for trials at position p6 (69%, all *p* < .05). Trials at this digit position in turn showed higher accuracies than trials at digit position p7 and p8 (p7: 54%, p8: 48%, all *p* < .05). Trials at digit position p7 and p8 in turn showed quite equal accuracies (*p* = .54). Thus our classification of three load-levels and the assignment of digit position p3 to load level l1, p6 to l2 and p7 and p8 to l3 as described in the method section seems to be justified by the results of this statistical analysis.

3.1.4. Correlational analysis. For correlational analysis we calculated a memory span score for the operation span and the digit span that was defined as the last memory load-level where more than two-third of the digit sequences were memorized completely in the correct order (i.e., we performed a rather restrictive all-or-nothing scoring procedure; see Conway et al., 2005). We calculated Pearson's correlation coefficients (two-tailed) for these mean recall accuracy scores of the operation span and the digit span task and the accuracies of the four n-back load levels. The results of the correlational analysis are given in Table 1.

Table 1. Pearson's correlation coefficients (two-tailed) for the mean recall accuracy scores of the operation span (oSpan) and the digit span task (dSpan) and the accuracies of the four n-back load levels. Note. + *p* < .10, * *p* < .05, ** *p* < .01

	1	**2**	**3**	**4**	**5**	**M**	**SD**	**N**
1) oSpan						5.00	1.28	12
2) dSpan	.66**					5.75	0.87	12
3) 1-back	.07	.27				0.90	0.06	12
4) 2-back	.51+	.009	.62*			0.88	0.10	12
5) 3-back	.63*	.33	.21	.20		0.80	0.08	12
6) 4-back	.51+	.50	-.47	-.19	.51+	0.76	0.09	12

With respect to a comparison of performance in the n-back task load levels and the overall performance in the span tasks, we found a positive correlation between performance in the 3-back load level and the overall performance in the operation span task (*r* = .63, *p* < .05) as well as a trend for a positive correlation between performance in the 2-back and 4-

back load levels and the overall performance in the operation span task (both *r* = .51, *p* < .07). We did not find any significant correlation between any level of the n-back task and the digit span task. However, we observed a positive correlation between overall performance in the operation span task and performance in the digit span task (*r* = .66, *p* < .01), which is consistent with other studies (e.g., Kane et al., 2004; Turner & Engle, 1989; Unsworth & Engle, 2007). Besides, the 1-back and 2-back load levels showed a significant positive correlation (*r* = .62, *p* < .05) and the 3-back and 4-back load levels showed a trend for a positive correlation (*r* = .51, *p* < .07).

3.2. P300 Data

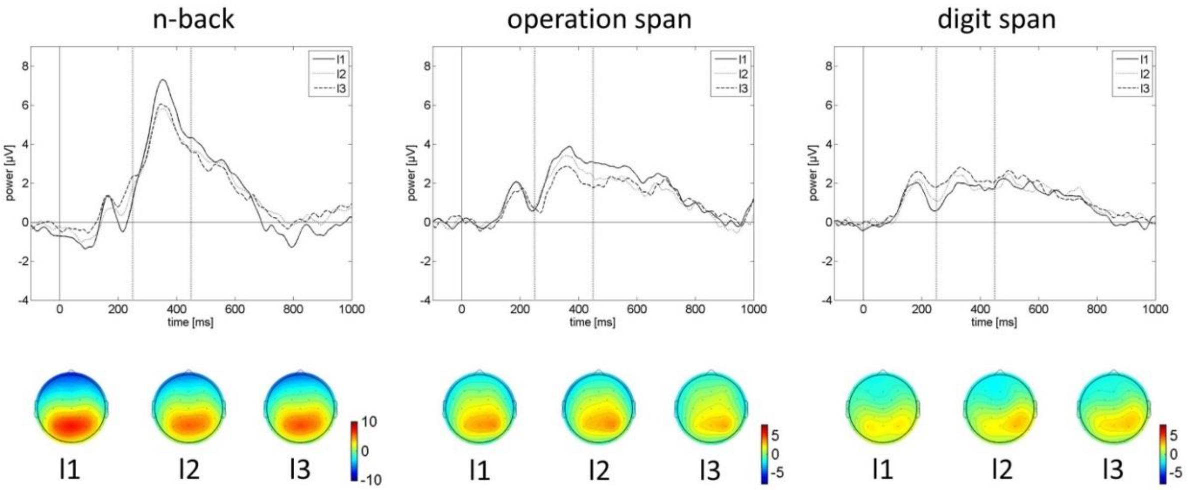


Figure 3. Event-related potential curves at electrode Pz (upper part of the figure) for the three tasks and three load levels. Lower part of the figure shows topoplots (i.e., topographically distributions of the peak amplitude values at all electrodes over the scalp).

Table 2. Mean P300 latencies [in ms] and mean P300 peak amplitudes [in µV] at electrode Pz for the three tasks and load categories.

	P300 latency [ms]					
	n-back		**oSpan**		**dSpan**	
l1	352	(15)	359	(40)	363	(65)
l2	365	(26)	355	(43)	339	(58)
l3	357	(19)	357	(28)	343	(66)

	P300 amplitude [µV]					
	n-back		**oSpan**		**dSpan**	
l1	8.09	(2.61)	3.02	(2.12)	1.70	(1.38)
l2	6.47	(2.84)	2.62	(2.00)	2.49	(1.72)
l3	6.91	(3.12)	1.95	(1.86)	2.01	(1.90)

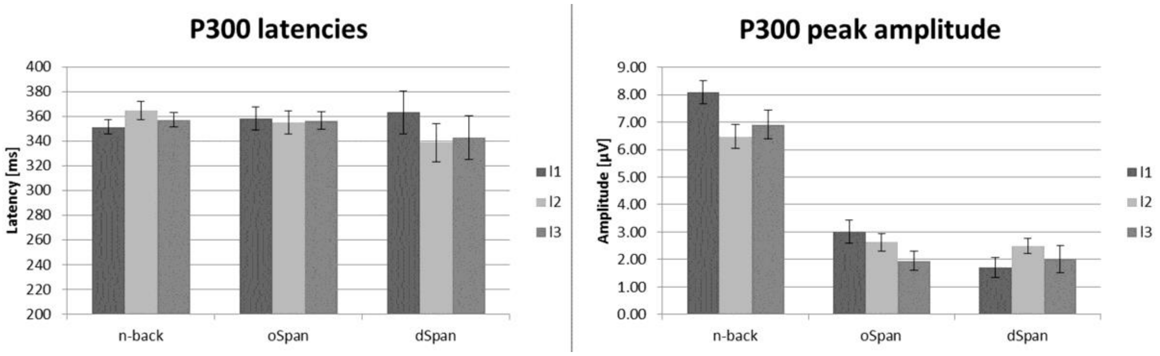


Figure 4. P300 latency and peak amplitude at electrode Pz for the three load-levels of the n-back, operation span (oSpan), and digit span (dSpan) task. Black error bars indicate ±1 standard error of the mean.

Grand-average ERPs are shown in Figure 3. Peak latency and amplitude data for the P300 at electrode Pz are given in Table 2. We ran two separate two-factorial repeated measures ANOVAs with the factors task (n-back, operation span, digit span) and load (l1, l2, l3) for the peak latencies and peak amplitude values respectively.

The peak latencies did not differ, neither between the load levels of a single task, nor between the three different tasks ($F(4, 44) = .73$; all $p > .50$). For the peak amplitudes however, we found a main effect of task, $F(2, 22) = 45.75$; $p < .001$; $\eta_p^2 = .81$ and a main effect of load, $F(2, 22) = 4.17$; $p = .029$; $\eta_p^2 = .28$. These main effects were modulated by a significant interaction, $F(4, 44) = 5.43$; $p = .001$; $\eta_p^2 = .33$. The main effect of task was not affected by this interaction: On all three load levels, the peak amplitude of the n-back task was significantly higher than the peak amplitudes of the operation span and digit span task (l1: 8.10 vs. 3.02 and 1.70 µV, $p < .001$; l2: 6.47 vs. 2.62 and 2.49 µV, $p < .001$; l3: 6.91 vs. 1.95 and 2.01 µV, $p < .001$). The peak amplitudes of the operation span and digit span did not differ. However, depending on task the P300 peak amplitude showed a differently load related decrease (n-back, operation span) or no decrease at all (digit span). In the n-back task, the P300 amplitude decreased for increasing load, as expected. In line with the results of the reaction time data, the strongest decrease in amplitude could be observed between the 1-back and 2-back load condition (-1.62 µV, $p < .001$). No significant difference occurred between the 2-back and the 3-back load level ($p > .4$). In the operation span, we also observed a significant decrease in P300 peak amplitude for increased load levels. In the operation span task the most pronounced decrease in peak amplitude was between load levels l1 and l3 (-1.07 µV, *p* = .024). In the digit span task however, we did not observe any decrease in amplitude for increased load levels. In fact, between certain load levels, the amplitude in the digit span rather seemed to increase for increased load (l1 to l2: +.79 µV, l2 to l3: -.44 µV; all n.s.).

3.3. Time-Frequency Representations (TFRs)

The TFR plots (cf. Figure 5) show the ERD/ERS% values (i.e., the oscillatory activity) over time for the three tasks and three load levels. All tasks show an ERD in the alpha frequency band and beta frequency band in a time-window of about 0 - 800 ms post-stimulus onset that increased with increased load levels. Furthermore, the alpha frequency band showed ERD effects which lasted longer in time for larger load conditions compared to smaller load conditions in the n-back and the operation span task. In the digit span task, the alpha ERD seemed to be more focused in time (i.e., maximally in the first time-window, 0 – 800 ms post-stimulus onset). The oscillatory pattern of the n-back resembled the TFR results described by Pesonen et al. (2007). Interestingly, the ERS in the low-load (l1, 1-back) condition in the beta frequency band starting about 800 ms after stimulus onset that might be attributed to a post-motoric beta rebound effect (Pesonen et al., 2007) was only visible in the n-back task.

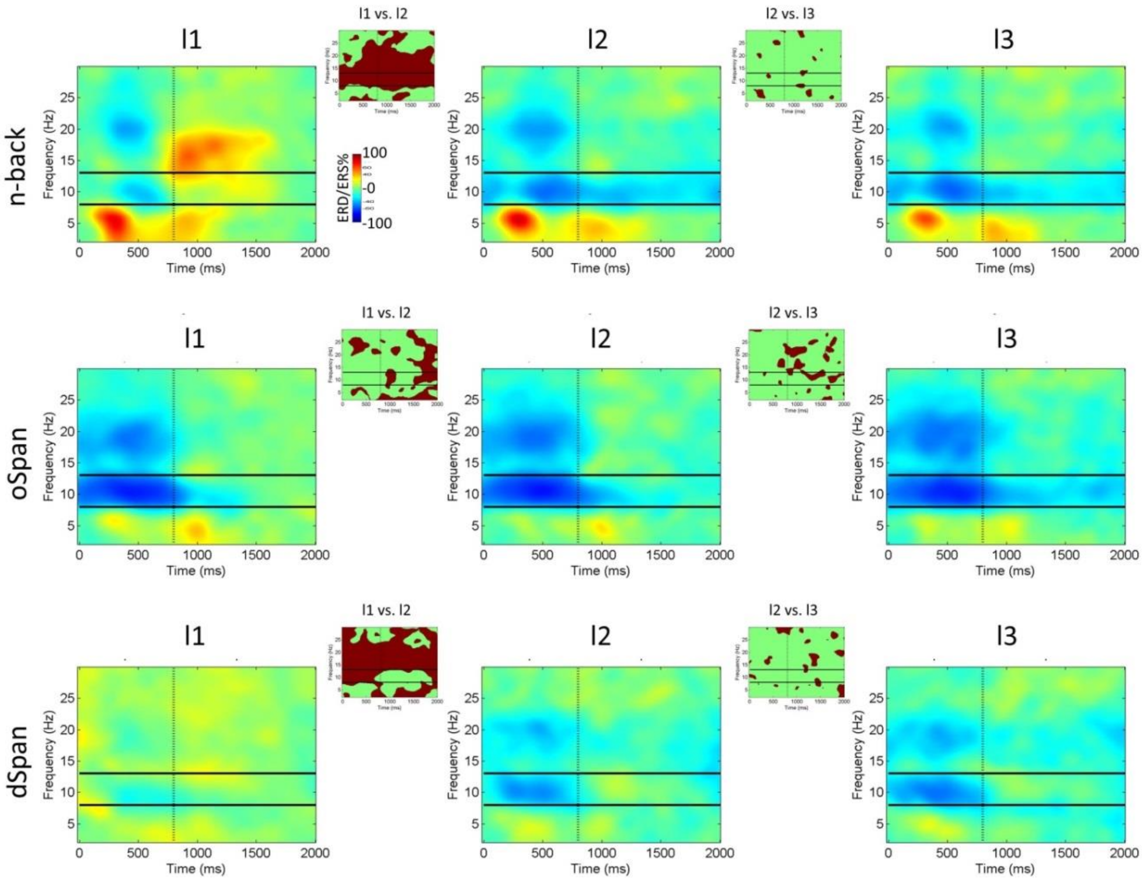


Figure 5. Time-frequency representations (TFRs) of the ERD/ERS% values of the n-back, operation span (oSpan) and digit span (dSpan) task in a frequency range from 2 Hz to 32 Hz and a time range from 0 to 2000 ms post-stimulus onset. From left to right increasing load levels (l1 to l3). The black horizontal lines denote the frequency band borders of the alpha frequency band (8 – 13 Hz). The dotted vertical line at 800 ms marks the border between the two time-windows (0 – 800 ms and 800 – 1600 ms) that were used for further analysis. Blue colors signal event-related desynchronization (ERD), red colors signal event-related synchronization (ERS) measured in percent with respect to a baseline condition. The small plots in between the larger TFRs indicate statistically significant differences in ERD/ERS% values between two adjacent load levels ($p < .05$).

For a more in-depth statistical analysis, we performed two separate three-factorial analyses of variance for the averaged ERD/ERS% data of the alpha frequency band (8 -13 Hz) and the beta frequency band (14 – 24 Hz; cf. methods section) with the factors task (n-back, oSpan, dSpan), time window (0 – 800 ms, 800 – 1600 ms) and load (l1 to l3). The averaged ERD/ERS% data are given in Figure 6.

3.3.1 alpha band mean ERD/ERS%. The three-factorial repeated-measures ANOVA revealed significant main effects for all three factors, task: $F(2, 22) = 13.57$; $p < .001$; $\eta_p^2 = .55$, time-window: $F(1, 11) = 22.84$; $p < .001$; $\eta_p^2 = .68$, and load: $F(2, 22) = 13.87$; $p = .001$; $\eta_p^2 = .56$. Overall, the operation span showed the largest alpha ERD (-40.69%), the digit span (-14.52%) and n-back (-20.42%) showed a comparably lower ERD. The alpha ERD was overall more pronounced in the time window 0 – 800 ms (-37.25% vs. -13.17%). Increased load levels led to significantly increased alpha ERD (i.e., decreased alpha frequency band power values). These effects however were further qualified by a two-way interaction task X load: $F(4, 44) = 4.08$; $p = .043$; $\eta_p^2 = .27$ and a three-way interaction task X time-window X load: $F(4, 44) = 9.36$; $p = .001$; $\eta_p^2 = .46$. To resolve this three-way interaction, we run additional two-factorial repeated-measures ANOVAs with the factors task(3) and load(3) for each time-window separately.

In the first time-window (0 – 800 ms) the two-factorial repeated-measures ANOVA revealed a main effect of task, $F(2, 22) = 17.05$; $p < .001$; $\eta_p^2 = .61$ and a main effect of load, $F(2, 22) = 20.33$; $p < .001$; $\eta_p^2 = .65$ as well as a significant interaction between these factors, $F(4, 44) = 7.30$; $p = .007$; $\eta_p^2 = .40$. In the n-back and in the digit span task the alpha ERD increased significantly from load-level l1 (n-back: -13.61%, digit span: -6.31%) to load-level l2 (-36.21%, -34.32%) and l3 (-37.79%, -38.43%) with the latter two showing no significant differences. In the operation span the alpha ERD values did not vary significantly for the different load levels (l1: -56.23%, l2: -56.35%, l3: -55.97%) and was on all load-levels more pronounced as in the n-back (-13.62%, -36.21%, -37.79%) or digit span task (-6.31%, -34.32%, -38.43%), with the latter two showing comparable (i.e., statistically not-significant different) alpha ERD values on all load-levels. This result might indicate that in the first time-window of the operation span memorization subtask additional cognitive processes might be loaded because of still ongoing processes associated with the equation calculation of the preceding processing subtask.

In the second time window (800 – 1600 ms), the ANOVA revealed a trend for a main effect of task, $F(2, 22) = 3.53$; $p = .07$; $\eta_p^2 = .24$, a main effect of load, $F(2, 22) = 8.99$; $p = .007$; $\eta_p^2 = .45$ and again an interaction between the two, $F(4, 44) = 4.50$; $p = .031$; $\eta_p^2 = .29$. Like in the first time-window, in the n-back task the alpha ERD increased significantly for increasing load from load-level l1 (18.36%) to l2 (-28.72%) and l3 (-24.54%), not differentiating between the latter two. However the digit span showed no load-effect in the second time window (l1: 2.38%, l2: -2.57%, l3: -7.88%). In contrast, the operation span which showed no load-effect in the first time-window, showed a significantly increased alpha ERD for load-level l3 (-31.96%) as compared to load level l1 (-18.71%). In the second time-window only for load-level l1 the alpha ERD of operation span was significantly larger as the alpha ERD of the n-back and the digit span task (-18.71% vs. 18.36% and 2.38%). For load-level l2 (operation span: -24.90, n-back: -28.72, digit span: -2.57) and l3 (operation span: -31.96%, n-back: -24.54%, digit span: -7.88%) the alpha ERD did not differ between tasks.

3.3.2. beta band mean ERD/ERS%. The overall picture for beta band effects was comparable to the alpha band effects. The three-way ANOVA revealed a main effect of task, $F(2, 22) = 9.81$; $p = .001$; $\eta_p^2 = .47$, a main effect of time-window, $F(1, 11) = 58.99$; $p < .001$; $\eta_p^2 = .84$, and a main effect of load, $F(2, 22) = 34.73$; $p < .001$; $\eta_p^2 = .76$. The operation span (-23.83%) showed overall a larger beta ERD as compared to the n-back (-7.96%) and the digit span (-10.58%), with the latter two showing comparable ERD values. The ERD was more pronounced in the first time window (-24.31% vs. -3.94%) and for the higher load-levels (l2: -18.66% and l3: -20.24%) as compared to the lowest load-level (l1: -3.47%) and the second time-window. However these main effects were qualified further by significant two-way and three-way interactions: interaction task X load: $F(4, 44) = 7.20$; $p < .001$; $\eta_p^2 = .40$, interaction task X time window: $F(2, 22) = 14.36$; $p < .001$; $\eta_p^2 = .57$, interaction time window X load: $F(2, 22) = 5.10$; $p = .015$; $\eta_p^2 = .32$, and a three-way interaction task X time window X load: $F(4, 44) = 6.63$; $p < .001$; $\eta_p^2 = .38$. To resolve this three-way interaction, we run two additional two-factorial repeated-measures ANOVAs, one for each time-window, as described for the alpha band.

In the first time-window (0 – 800 ms) the two-factorial repeated-measures ANOVA revealed a main effect of task, $F(2, 22) = 15.93$; $p < .001$; $\eta_p^2 = .59$, a main effect of load, $F(2, 22) = 30.20$; $p < .001$; $\eta_p^2 = .73$ and an interaction between these factors, $F(4, 44) = 7.41$; $p = .001$; $\eta_p^2 = .40$. For the n-back task, load-level l2 (-25.17%) showed a significantly increased beta ERD as compared to load-level l1 (-12.96%). For the operation span we observed a tendency for an increased beta ERD for load-level l3 (-40.23%) as compared to l1

(-35.21%, p = .076). The digit span showed the same outcome as for the alpha frequency band ERD: beta ERD was significantly increased for load-levels l2 (-23.29%) and l3 (-24.59%) as compared to load level l1 (2.30%) with no significant difference between load-levels l2 and l3. Overall, the operation span showed a more pronounced beta ERD (-37.82%) as the digit span (-15.20%) and the n-back (-19.91%), this pattern being consistently found for all load-levels.

In the second time-window (800 – 1600 ms) the two-factorial repeated-measures ANOVA revealed a main effect of task, $F(2, 22) = 5.53$; $p = .011$; $\eta_p^2 = .34$, a main effect of load, $F(2, 22) = 29.29$; $p < .001$; $\eta_p^2 = .73$, and a significant interaction between the two, $F(4, 44) = 6.77$; $p = .005$; $\eta_p^2 = .38$. In the n-back task in load-level l1 we observed a beta ERS rather than a beta ERD (as indicated by the positive ERD/ERS% value: 23.64%). This ERS turned into a beta ERD for load-levels l2 (-7.36%) and l3 (-4.30%), with the latter two being statistically not different. This pronounced beta ERS load-level l1 could be only observed for the n-back task. As a consequence, at load level l1 ERD/ERS% values differed significantly between the n-back task (23.64%) and the operation span (-6.31%) and digit span (7.73%) task, with the latter two being statistically not different. At load levels l2 (n-back: -7.36%, operation span: -7.12%, digit span: -11.00%) and l3 (-4.30%, -16.07%, -14.63%) the beta ERD values did not differ significantly between the tasks. The operation span showed significantly more pronounced beta ERD at load level l3 as compared to l1 and l2 (the latter two being of comparable magnitude). The digit span showed significantly increased beta ERD values for load levels l2 and l3 in comparison to load level l1 (with l2 and l3 being of equal magnitude).

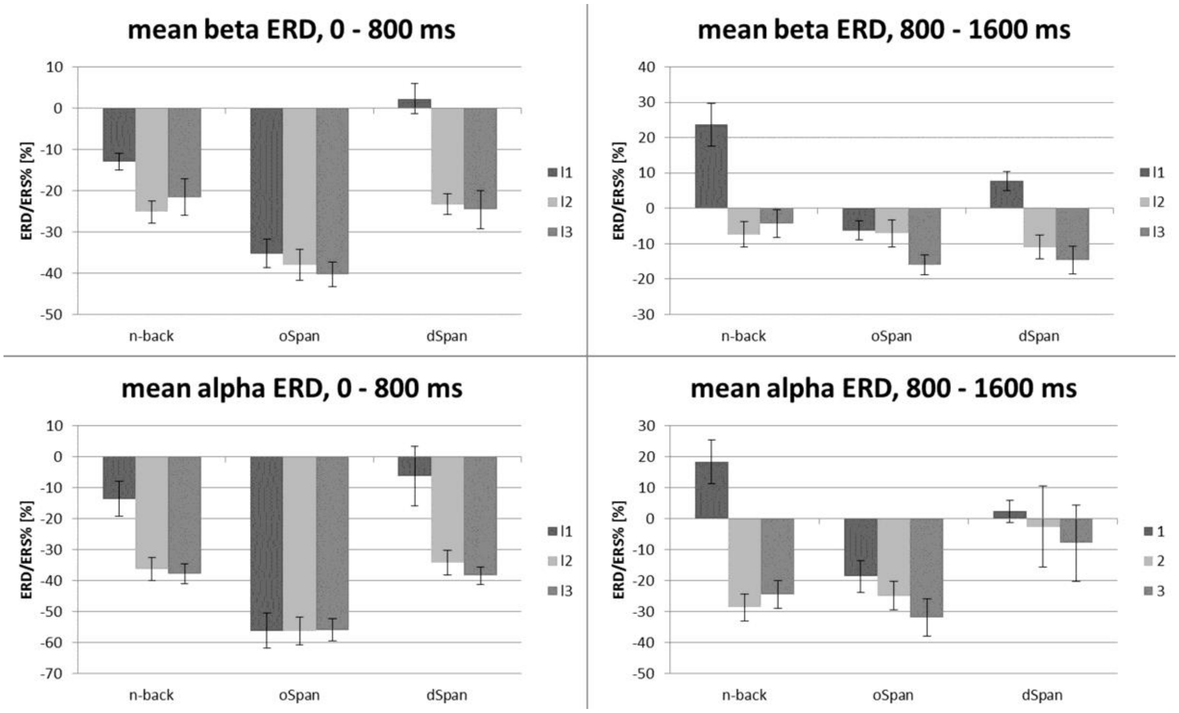


Figure 6. Mean ERD/ERS% values for the n-back, operation span (oSpan) and digit span (dSpan) task for the alpha (8 – 13 Hz) and beta (14 – 24 Hz) frequency band and the two time-windows (0 – 800 ms, 800 – 1600 ms post-stimulus onset). Black error bars indicate ±1 standard error of the mean.

4. Discussion

In the current study we were interested in a direct comparison of a simple digit span (STM) task and two classical WM tasks, a verbal n-back and an operation span task, using the P300 and EEG frequency band power changes as expressed by ERD/ERD% values in TFRs. Studies using correlational designs for comparison of behavioral performance measures of span and n-back tasks normally reported rather low correlations (with *r* ~ .20) between these two (Jaeggi et al., 2010; Kane et al., 2007; Redick & Lindsey, 2013). We used carefully designed tasks with respect to timing of the stimuli presentation as well as the trial numbers of the tasks and observed rather high correlations between the operation span task performance scores and the n-back task accuracy (*r* between .50 and .60 for the 2-back, 3-back, and 4-back load level accuracies and the overall operation span performance).

The used EEG measures added valuable information about underlying WM processes (i.e., EFs) that are involved in the different task of the current study. The main outcomes concerning the EEG measures were A) with respect to P300: A comparable decrease of P300 amplitude in the n-back and operation span task that was absent in the digit span task, suggesting that n-back and operation span expose higher demands on EFs than the digit span task; and B) with respect to TFRs: Comparatively oscillatory patterns of increased alpha and beta frequency band ERDs for all tasks, albeit differences in exact timing and magnitude could be observed that might indicate (in line with the P300 results) that more EFs are required and are more intensively demanded in the n-back and operation span task as compared to the digit span task. In the following, we will discuss these results in the light of our hypotheses for each measure separately followed by a general conclusion section.

4.1. Behavioral Data

The behavioral data of the n-back and the span tasks are in accordance with literature and confirmed the successful load manipulation. For increasing WM load due to increasing n-back levels or to-be-remembered digits, the accuracy decreased and the reaction times in the n-back increased (e.g., Gevins & Smith, 2000; Turner & Engle, 1989).

In contrast to literature, we observed noticeably higher correlations between n-back and complex span performance (see Table 1) as usually reported (cf. the review by Redick & Lindsey, 2013, reporting a mean correlation coefficient of about $r = .20$). These different outcomes might be mainly due to the careful task design in our study with respect to comparable timing and amount of trials (i.e., local and global timing) between n-back and span tasks. Such careful, comparable task designs with respect to timing are often lacking in correlational studies of purely behavioral performance measures (e.g., Conway et al., 2005; Jaeggi et al., 2010). For example, in the current study the amount of trials in the span tasks exceeds the amount of trials normally used by a factor of about ten. Especially, in order to have data epochs of comparable length and timing in all tasks for later EEG data analysis, we put severe and comparable time-constraints on all tasks. These time-constraints are normally lacking in complex span tasks (e.g., Turner & Engle, 1989). In line with theoretical considerations about the influence of time and time constraints on WM load (see the 'time-based resource sharing model' proposed by Barrouillet, Bernardin, & Camos, 2004;

Barrouillet et al., 2007), we hypothesize the observed rather high correlations between the n-back and the span tasks might be grounded in comparably severe time-constraints in all tasks.

4.2 P300 Data

As the n-back and the operation span task were conceptualized as genuine WM tasks and were hypothesized to share a dual-task character (see Watter et al., 2001), we hypothesized these tasks to comparably require EFs updating, shifting, and inhibition. In contrast to the n-back and operation span task, the digit span task seems to involve less demands on EFs, and potentially it requires only the updating function. Therefore, we expected the P300 amplitude to decrease for increased load in the n-back and the operation span task in a comparable manner. In contrast, for the digit span task we expected an outcome of the P300 amplitude being distinct. As P300 latencies have been shown to be not influenced by increased load in the n-back task above the 1-back load level (McEvoy et al., 2001; Watter et al., 2001), we expected P300 latencies not to differentiate between increased WM load in the n-back task and the operation span task.

Our hypotheses concerning the P300 were completely fulfilled. P300 latencies did not vary significantly between load levels of all three tasks. The n-back task and the operation span task evoked a P300 with the amplitude significantly decreasing for increased WM load. Furthermore, also the absolute P300 amplitude in the n-back task was significantly higher than in the span tasks, the reductions of the P300 amplitude between the low load condition and the high load condition in both, the n-back and the operation span task were of comparable magnitude. The digit span task in contrast showed only a weak P300 effect with an amplitude that rather increased than decreased for increased load levels (albeit not significantly).

In sum, the P300 data showed the expected similarity between the n-back and the operation span task. In our view, these findings point to both, the dual-task nature of the n-back task and to similarities between the n-back task and the operation span task as WM tasks that comparably load on EFs. When WM processes are increasingly involved, internal attention seems to be more wide-spread distributed over different WM processes as indicated by the decreased P300 amplitude for increased WM load. Thus, we conclude that in an n-back task and the memorization subtask of an operation span task common EFs like updating, shifting and inhibition are required, potentially to a similar degree, whereas the digit span task requires less EFs, potentially mainly updating.

4.3 Time-Frequency Representations (TFRs)

With respect to the TFRs we expected to observe an oscillatory pattern being more similar between the n-back and the operation span as compared to the digit span. We hypothesized the alpha ERD (and beta ERD) to indicate load on WM processes (Engel & Fries, 2010; Gevins et al., 1997; Gevins & Smith, 2000; Pesonen et al., 2007; Stipacek et al., 2003). Thus, we expected to observe in the n-back task and in the operation span task comparable, more pronounced alpha ERDs (and beta ERDs) as conceptually (and indicated by the P300 amplitude effects reported above) in these tasks more EFs might be demanded as in the digit span task.

Our results however yielded mixed results concerning this hypothesis. In the alpha frequency band the n-back task and operation span task showed significant load effects in a later time-window (800 – 1600 ms post-stimulus onset) whereas the digit span task showed a significant load effect only in an earlier time window (0 – 800 ms post-stimulus onset). Additionally, the n-back task and the operation span task showed timely longer lasting alpha ERDs for increased load levels. This might be related to the increased updating and shifting demands in the n-back task and the operation span task compared to the digit span task. From a task analytical point of view this interpretation is plausible: The observed P300 data indicated that both, n-back task and operation span task may be regarded as dual-tasks. Thus, in both tasks additionally to purely WM updating, processes of shifting (between the subtasks) and inhibition (e.g., of the currently not active subtask component) are required. In the digit span task during the first time window the memory updating process might take place. After about 800 ms in the simple digit span task a purely retention phase might take place which might not be associated with an increased alpha ERD (or even, depending on task, with an alpha ERS, cf. Jensen et al., 2002; Palva & Palva, 2007).

Furthermore, in the operation span task and the n-back task the retention phase might start later and might not be free of additional processes because of the dual-task nature of the both tasks: The updating process might be timely prolonged as in the first time window the stimuli-comparison processes of the dual-task take place. After these comparison processes took place, WM is updated with the new stimulus. This interpretation might be corroborated further by the beta ERS effect in the n-back task in the low load condition. This ERS in the second time window might be interpreted as a beta rebound of previous motoric activity (Pfurtscheller, Stanca, & Neuper, 1996) during the first time window (key-press), thus signaling the end of thereby associated processes. This interpretation holds also for the operation span task, although we did not observe a beta-rebound in this task. However, the average reaction times that range in both tasks within the first time window, might also indicate that certain processes of stimulus comparison are terminated in the second time window and a shift to updating and memorization processes might be necessary.

In the beta frequency band, the load-associated effects resembled those of the alpha frequency band albeit showing less differences between tasks and time-windows. Beta ERD has recently begun to be interpreted as being associated with cognitive processes like WM processes and not only with motor activity (Engel & Fries, 2010; Weiss & Mueller, 2012). Beta ERD effects for increased load in an n-back task paradigm have been consistently observed in several studies (Krause et al., 2010; Palomäki et al., 2012; Pesonen, et al., 2007). Although these authors hypothesized the beta ERD effects as being cognitively induced, they could not rule out a purely motor explanation as their n-back task required overt motor activity (key-press). Our data show comparable beta ERD effects for all three tasks. As the digit span task required no motor activity, we may hypothesize the beta ERD truly reflecting WM load associated processes. Furthermore, we may hypothesize that the observed beta ERD may reflect specifically processes of WM updating, as conceptually the EF updating is the only common EF that might be involved in the digit span task as well as in the operation span task and the n-back task.

Interestingly, the operation span task showed more pronounced alpha ERD and beta ERD in the first (and to lesser extend in the second) time window as compared to the n-back and the digit span task. This might indicate that in the operation span task WM processes are loaded most, or other cognitive processes associated with the calculation subtask are still active during (or at least at the beginning) of the memorization subtask. In line with this interpretation is the missing beta ERS effect in the operation span task as compared to the n-back task. This beta ERS vanishes in the n-back task for higher load levels and might not be present in the operation span task even in the low load condition because of a generally increased load-situation in this task.

In sum, the overall pattern of alpha band and beta band ERD was comparable across all three tasks but yielded differences in magnitude and timing between the tasks that might indicate differences in demands on EFs. In the digit span task less EFs (potentially only updating) seem to be demanded as indicated by changes in oscillatory activity mainly in an early time-window. In contrast in the n-back and operation span task more EFs seem to be demanded (additionally to purely updating, shifting and inhibition might be required) as indicated by longer lasting and more pronounced changes in oscillatory activity. However, put the other way around, if alpha and beta ERD reflect load on WM processes as hypothesized, all three tasks might be regarded as WM tasks.

5. Conclusion

To sum up, in the current study we were interested in a direct comparison of two tasks conceptualized as WM tasks, an n-back task and a complex operation span task, and a STM task, a simple digit span task. Our study revealed a comparable character of the n-back and the operation span task as dual-tasks as indicated by a comparable decrease of the P300 amplitude in both tasks that was different from the simple digit span task. Additionally more pronounced and longer lasting effects in the TFRs indicated that these two tasks exposed more demands on EFs than the simple digit span task. These results are consistent with our assumption that the operation span and then-back tasks require common EFs like updating, shifting, and inhibition to similar degree, whereas the digit span task might only require the EF updating.

Consistent with this assumption, we found rather high correlations between performance measures of the n-back and the operation span task. We hypothesized this unexpected outcome to be potentially caused by the carefully designed tasks we used with comparably time-constraints in the two paradigms. Therefore, the current behavioral and EEG results pointing to commonalities between the n-back and the operation span task (and in case of the digit span task: differences) with respect to EFs, call for future studies that control for similar time-constraints when comparing the performance of subjects and the involved processes in n-back and span tasks.

Acknowledgements

This research was funded by the Leibniz ScienceCampus Tübingen "Informational Environments".

References

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 555–559.

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

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29. doi:10.1146/annurev-psych-120710-100422

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology. General*, *133*(1), 83–100. doi:10.1037/0096-3445.133.1.83

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(3), 570–85. doi:10.1037/0278-7393.33.3.570

Bayliss, D. M., Jarrold, C., Gunn, D. M., & Baddeley, A. D. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology: General*, *132*(1), 71–92. doi:10.1037/0096-3445.132.1.71

Bledowski, C., Kaiser, J., & Rahm, B. (2010). Basic operations in working memory: Contributions from functional imaging studies. *Behavioural Brain Research*, *214*(2), 172–179. doi:10.1016/j.bbr.2010.05.041

Chen, Y.-N., & Mitra, S. (2009). The spatial-verbal difference in the n-back task: An ERP study. *Acta Neurologica Taiwanica*, *18*(3), 170–179.

Chen, Y.-N., Mitra, S., & Schlaghecken, F. (2008). Sub-processes of working memory in the n-back task: An investigation using ERPs. *Clinical Neurophysiology*, *119*(7), 1546–1559. doi:10.1016/j.clinph.2008.03.003

Conway, A. R. A., 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.

Cowan, N. (2000). Processing limits of selective attention and working memory: Potential implications for interpreting. *Interpreting*, *5*(2), 117–146. doi:10.1075/intp.5.2.05cow

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. doi:10.1016/j.jneumeth.2003.10.009

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., Polich, J., Reinvang, I., & Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, *120*(11), 1883–1908. doi:10.1016/j.clinph.2009.07.045

Engel, A. K., & Fries, P. (2010). Beta-band oscillations - signalling the status quo? *Current Opinion in Neurobiology*, *20*(2), 156–165. doi:10.1016/j.conb.2010.02.015

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

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. doi:10.1093/cercor/10.9.829

Gevins, A., Smith, M. E., McEvoy, C. L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, *7*(4), 374–385.

Gross, B., Metz, A., & Ullsperger, P. (1992). Die P300-Komponente des ereigniskorrelierten Hirnpotentials in einem Kurzzeitgedächtnisparadigma. *Zeitschrift für experimentelle und angewandte Psychologie*, *39*(1), 56–67.

Grune, K., Metz, A., Hagendorf, H., & Fischer, S. (1996). Information processing in working memory and event-related brain potentials. *International Journal of Psychophysiology*, *23*, 111–120.

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. doi:10.1080/09658211003702171

Jasper, H. H. (1958). The ten±twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371–375.

Jensen, O., Gelfand, J., Kounios, J., & Lisman, J. E. (2002). Oscillations in the alpha band (9-12 Hz) increase with memory load during retention in a short-term memory task. *Cerebral Cortex (New York, N.Y. : 1991)*, *12*(8), 877–882.

Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., & Koeppe, R. A. (1997). Verbal working memory load affects regional brain activation as measured by PET. *Journal of Cognitive Neuroscience*, *9*(4), 462–475.

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. doi:10.1037/0278-7393.33.3.615

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology. General*, *133*(2), 189–217. doi:10.1037/0096-3445.133.2.189

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, *29*(2-3), 169–195.

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, *38*(3), 557–577. doi:10.1017/S0048577201990559

Krause, C. M., Pesonen, M., & Hämäläinen, H. (2010). Brain oscillatory 4-30 Hz electroencephalogram responses in adolescents during a visual memory task. *Neuroreport*, *21*(11), 767–771. doi:10.1097/WNR.0b013e32833bfcbb

McEvoy, L. K., Pellouchoud, E., Smith, M. E., & Gevins, A. (2001). Neurophysiological signals of working memory in normal aging. *Brain Research. Cognitive Brain Research*, *11*(3), 363–376.

Miller, K. M., Price, C. C., Okun, M. S., Montijo, H., & Bowers, D. (2009). Is the n-back task a valid neuropsychological measure for assessing working memory? *Archives of Clinical Neuropsychology*, *24*(7), 711–717. doi:10.1093/arclin/acp063

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (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.

Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation*, *51*, 45–100. doi:10.1016/S0079-7421(09)51002-X

Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, *9*, 97–113.

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.

Palomäki, J., Kivikangas, M., Alafuzoff, A., Hakala, T., & Krause, C. M. (2012). Brain oscillatory 4-35 Hz EEG responses during an n-back task with complex visual stimuli. *Neuroscience Letters*, *516*(1), 141–145. doi:10.1016/j.neulet.2012.03.076

Palva, S., & Palva, J. M. (2007). New vistas for alpha-frequency band oscillations. *Trends in Neurosciences*, *30*(4), 150–158. doi:10.1016/j.tins.2007.02.001

Pesonen, M., Hämäläinen, H., & Krause, C. M. (2007). Brain oscillatory 4-30 Hz responses during a visual n-back memory task with varying memory load. *Brain Research*, *1138*, 171–177. doi:10.1016/j.brainres.2006.12.076

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology*, *110*(11), 1842–1857.

Pfurtscheller, G., Zalaudek, K., & Neuper, C. (1998). Event-related beta synchronization after wrist, finger and thumb movement. *Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control*, *109*(2), 154–160. doi:10.1016/S0924-980X(97)00070-2

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*, 127–152.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. doi:10.1016/j.clinph.2007.04.019

Redick, T. S., & Lindsey, D. R. B. (2013). Complex span and n-back measures of working memory: a meta-analysis. *Psychonomic Bulletin & Review*, *20*(6), 1102–1113. doi:10.3758/s13423-013-0453-9

Roberts, R., & Gibson, E. (2002). Individual differences in sentence memory. *Journal of Psycholinguistic Research*, *31*(6), 573–598.

Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. IEEE Transactions on Bio-Medical Engineering, 51(6), 1034–43. doi:10.1109/TBME.2004.827072

Schmiedek, F., Hildebrandt, A., Lövdén, M., Lindenberger, U., & Wilhelm, O. (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. doi:10.1037/a0015730

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

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Experimental Psychology*, *59*(4), 745–759. doi:10.1080/17470210500162854

Stipacek, A., Grabner, R. H., Neuper, C., Fink, A., & Neubauer, A. C. (2003). Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load. *Neuroscience Letters*, *353*(3), 193–196. doi:10.1016/j.neulet.2003.09.044

Turner, M., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154. doi:10.1016/0749-596X(89)90040-5

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, *133*(6), 1038–1066. doi:10.1037/0033-2909.133.6.1038

Watter, S., Geffen, G. M., & Geffen, L. B. (2001). The n-back as a dual-task: P300 morphology under divided attention. *Psychophysiology*, *38*(6), 998–1003.

Weiss, S., & Mueller, H. M. (2012). “Too Many betas do not Spoil the Broth”: The Role of Beta Brain Oscillations in Language Processing. *Frontiers in Psychology*, *3*(June), 201. doi:10.3389/fpsyg.2012.00201

Yuan, K., Steedle, J., Shavelson, R., Alonzo, A., & Oppezzo, M. (2006). Working memory, fluid intelligence, and science learning. *Educational Research Review*, *1*(2), 83–98. doi:10.1016/j.edurev.2006.08.005

2.2 Study 2

When flanker meets the n-back: What EEG and pupil dilation data reveal about the interplay between the two central-executive working memory functions inhibition and updating

Christian Scharinger[1*], Alexander Soutschek[2,3], Torsten Schubert[3], & Peter Gerjets[1]

[1]Knowledge Media Research Center, Tuebingen, Germany
[2]Laboratory for Social and Neural Systems Research, University of Zurich, Zurich, Switzerland
[3]Department of Psychology, Humboldt-University of Berlin, Berlin, Germany

*Correspondence: Christian Scharinger, Knowledge Media Research Center, Schleichstr. 6, 72076 Tuebingen, Germany. Phone: +49 (0) 7071 979-360, Email: c.scharinger@iwm-kmrc.de

Abstract

We investigated the interplay between inhibition and updating, two executive working memory (WM) functions. We applied a novel task paradigm consisting of flanker stimuli presented within an n-back task and studied the interaction between inhibitory demands and load on WM updating using behavioral measures, electroencephalograph (EEG), and pupil dilation. In contrast to studies that examine the interaction between inhibitory demands and load on WM storage components, the current task paradigm allowed testing the interaction between the executive WM components updating and inhibition. We found a reduced flanker interference effect for the highest (2-back) updating load condition compared to lower updating load conditions on most measures. We interpret these findings as indicating that inhibitory control and WM updating are closely intertwined executive functions. Increased load on updating seemed to result in an overall more activated attentional network thus enhancing inhibitory control, such that task performance is less susceptible to distracting information.

Introduction

Working memory (WM) is functionally defined by the temporary storage and manipulation of relevant information during performance of cognitive tasks, like reading, problem solving or calculating (Shah & Miyake, 1999). Defining WM as the interplay of attentional processes and memory structures seems to encompass the minimal commonalities between different frameworks of WM (Baddeley, 2003; Cowan et al., 2005; Engle, 2002; Oberauer, 2009). According to the multi-component model of WM (Baddeley & Hitch, 1974; Baddeley, 2012), which is currently a leading model in the field, WM may consist of at least three structural components: two memory-related storage components and an attention-related supervisory component, the central-executive. The central-executive component is hypothesized to be divided into executive WM functions (EFs) that serve the coordination and manipulation of information in WM (Baddeley, 1996, 2007).

Three core EFs are commonly differentiated: updating, shifting, and inhibition (Diamond, 2013; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). *Updating* refers to processes of retrieval, transformation, and substitution of WM content (Ecker, Ullrich, Lewandowsky, Oberauer, & Che, 2010). *Shifting* indicates processes of directing attentional focus toward newly relevant information (e.g., task shifting in dual task paradigms; Monsell, 2003). *Inhibition* refers to processes of suppressing information that is not (or no longer) relevant for the current processing step in WM (Diamond, 2013). The n-back task is regarded by many authors as a typical task loading on WM updating (Jonides et al., 1997; Miyake et al., 2000). In an n-back task temporal sequences of stimuli are presented and participants have to indicate via key press whether a stimulus was same as or different from the stimulus they saw n-steps back. Thus, depending on *n*, this task relies on WM content to be continuously updated. The flanker task is a typical task demanding inhibitory control (Eriksen, 1995; Sanders & Lamers, 2002). In a letter flanker task, participants have to focus on a central letter and ignore surrounding letters (i.e., the flankers) that could be identical (i.e., congruent) or different (i.e., incongruent) to the central letter. In case of incongruent flankers, inhibitory control processes are loaded to overcome interference effects. However, although these EFs can be manipulated separately using specific tasks, a latent variable analysis by Miyake et al. (2000) revealed that updating, shifting, and inhibition share a common underlying process. These findings are corroborated by neuroimaging studies observing overlapping frontal and parietal activations for these EFs (Collette et al., 2005; Nee et al., 2013). All EFs rely on attentional processes, and controlled attention might constitute the commonality that updating, shifting, and inhibition share (Miyake et al., 2000; Engle, 2002).

The interplay of attentional demands and load on WM storage components has been studied by several authors (see Fougnie, 2008 and Vandierendonck, 2014 for comprehensive reviews). For example, Lavie and colleagues (Lavie, Hirst, de Fockert, & Viding, 2004; Lavie, 2005, 2010) studied the influence of load on WM storage on the performance of inhibitory control. In dual task studies, these authors used a flanker task to induce processing load on inhibitory control and a Sternberg task (Sternberg, 1966) to induce either low or high load on WM storage (i.e., participants had to remember one or six items and to perform the flanker task during the retention interval). The authors generally observed an increased (over-additive) flanker interference effect in the high as compared to the low WM load condition. This was interpreted as depletion of common attentional control processes necessary in both tasks, resulting in less control over suppressing the flanker interference in the high WM load condition, thus leading to the over-additive interference effect. However, this interpretation might be challenged by studies that use a Stroop task (Stroop, 1935) to demand inhibitory control while participants perform a WM storage task (Kim, Kim, & Chun, 2005; Park, Kim, & Chun, 2007). These studies reported that the increased (over-additive) interference effect under high WM load only occurred when WM stimuli and Stroop stimuli were of the same modality (e.g., both verbal). In contrast, when stimuli were of different modalities (e.g., spatial-verbal) the interference effect was attenuated under high WM load. This may indicate that stimulus specific interference effects might explain the over-additive effect observed by Lavie and colleagues rather than a depletion of global attentional resources. In contrast, when attentional processes are increased under high WM load there may be even beneficial effects on interference control.

To sum up, whereas the influence of load on WM storage components on single EFs, such as inhibition for example, and the role of attention therein has previously been studied by several authors, the mutual influence of load on WM updating and inhibitory demands has, to the best of our knowledge, not yet been explicitly addressed. Thus, in the current study we focused on the latter research question by using an n-back task that increased load on WM updating with flanker items for simultaneous demands on inhibitory control. This novel task paradigm allowed us to compare the mutual influence of load on these two EFs within one single task, without the additional load on shifting processes as would occur during a dual task.

We hypothesized two possible outcomes for the flanker interference effect under high load on WM updating. Firstly, an increased (i.e., over-additive) flanker interference effect, comparable to the results by Lavie and colleagues, that might be due to a depletion of attentional control under high WM updating load, thus impeding inhibitory control (*depletion hypothesis*). Alternatively, we might find a decreased (i.e., under-additive) flanker interference effect, comparable to the outcomes by Kim and colleagues (Kim et al., 2005), due to a general increased activation of attentional processes under high WM updating load, which facilitates inhibitory control (*facilitation hypothesis*).

In addition to pure behavioral measures, we used electroencephalographic (EEG) and pupil dilation measurement methods that have previously been shown to be sensitive to load on WM updating in n-back tasks and to inhibitory demands in flanker or Stroop tasks. The additional measurement of these parameters provides us with additional information about the influence of load on the task processing. In the following, we will give a brief overview of the measures' sensitivity for updating or inhibition processes and how these measures may allow disentangling between our competing hypotheses.

With respect to behavioral measures, increased load on WM updating due to increased n-back levels normally result in increased reaction times (RTs) and decreased accuracies, that is, number of correct responses (e.g., Jaeggi, Buschkuehl, Perrig, & Meier, 2010). For demands on inhibitory control in a flanker task, increased RTs (and decreased accuracies) are normally observed for incongruent as compared to congruent flanker stimuli (e.g., Erikson, 1995).

With respect to EEG measures, time-frequency representations (TFRs) of oscillatory EEG activity can provide a more detailed picture of the hypothesized shared neuronal networks of attention control with regard to the activation strength and its timing. For example, increased n-back levels typically result in increased EEG theta frequency band power at frontal electrodes as well as decreased EEG alpha band power at parietal electrodes (Gevins, Smith, McEvoy, & Yu, 1997; Gevins & Smith, 2000; Jensen & Tesche, 2002; Pesonen et al., 2007). According to Watter, Geffen, and Geffen (2001), the event-related P300 mean amplitude can serve as an index for the internal distribution of attention when different executive WM functions are required. They observed a decreased P300 amplitude for increased n-back levels. This was interpreted as indicating that attentional processes are more distributed when executive WM updating functions are required during task processing (i.e., in the 1-back or 2-back task) as compared to a purely matching task (i.e., a 0-back task). For flanker tasks, studies describing oscillatory activity (as measured by frequency band power changes) are rather scarce and focus mainly on theta band activity. Nigbur, Ivanova, & Stürmer (2011) observed increased theta power for increased demands on inhibitory control in a flanker task, similar to a study by Hanslmayr, Pastötter, Bäuml, Gruber, Wimber, & Klimesch (2008) for a Stroop task. The P300 amplitude reportedly decreases during incongruent compared to congruent flanker trials (Pratt, Willoughby, & Swick, 2011).

Finally, pupil dilation data might provide a physiological measure of the global load-situation (Beatty & Lucero-Wagoner, 2000). Increased pupil dilation has been observed for higher load during WM updating (Ewing & Fairclough, 2010; Karatekin, Marcus, & Couperus, 2007) as well as for increased demands on inhibitory control (Laeng, Ørbo, Holmlund, & Miozzo, 2011).

Given their sensitivity for updating and inhibition, the behavioral and the physiological measures we used should show dissociable result patterns depending on the relationship between these EFs. If WM updating demands facilitate interference processing (*facilitation hypothesis*), we expected to observe under-additive interference effects under high (compared to low) WM updating load, because the activation of updating processes should be accompanied by improved distractor inhibition. On the contrary, if high WM updating demands deplete resources required for inhibition, then we should find over-additive interactions between updating and inhibition demands in the used measures (*depletion hypothesis*).

Method

Participants

Twenty-two students (M_{age} = 24.64, *SD* = 4.32, 12 females) of the University of Tuebingen participated in the study and received a payment of 8€ per hour. They were all native speakers of German, right-handed (except one) and reported no neurological disorders. All participants had normal or corrected-to-normal visual acuity. The local ethics committee of the Knowledge Media Research Center Tuebingen approved the study. Participants gave their written consent at the beginning of the study.

Stimuli

The letters 'S', 'H', 'C', 'F' were used as stimuli. For each trial, one out of these four letters was randomly chosen and presented centrally on the screen either flanked by the same letters (congruent condition, the same letter appeared three times on the left and on the right sides of the centered target letter, e.g., 'HHH H HHH') or by randomly chosen different letters (incongruent condition, one of the three remaining letters appeared three times on the left and on the right sides of the centered target letter, e.g., 'FFF H FFF'). All letters were presented in gray on black backgrounds in Arial at 25 points font size. Each stimulus was shown for 500 ms, followed by a black screen for 1500 ms (see Figure 1). Thus, one trial lasted 2000 ms.

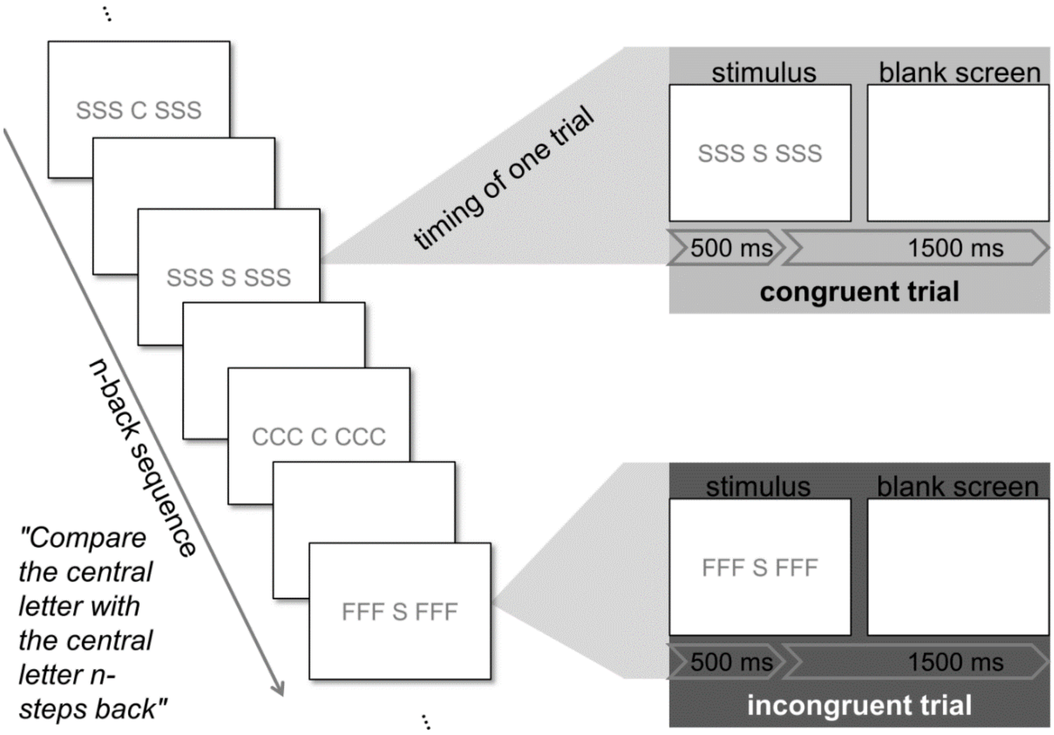


Figure 1. Stimulus sequence and timing of the n-back task with flanker stimuli. Stimuli were originally presented in gray color on black background. For better readability in this publication only we changed the background color to white. Light gray color symbolizes congruent trials, dark gray incongruent trials.

Procedure

Participants performed an n-back task: they saw temporal sequences of the stimuli and indicated via key press ('yes'/'no'-key) whether the central letter of the current trial matched (i.e., was identical to) or mismatched (i.e., was different from) the central letter they had seen in the sequence n-steps back. Henceforth, we will refer to stimuli of the match situation as *n-back targets*, and to stimuli of the mismatch situation as *n-back nontargets*. Participants were instructed to respond as quickly and accurately as possible and to focus only on the centrally presented letter (i.e., to ignore the flanker).

We used three n-back difficulty levels (0-back, 1-back, 2-back). In the 0-back condition, before the stimulus sequence started a randomly chosen letter ('S', 'H', 'C', or 'F') was displayed as n-back target letter. During the following stimuli sequence, each time this letter occurred as central letter, participants had to press the 'yes'-key, in all other cases the 'no'-key. In contrast, in the 1-back and the 2-back condition participants had to base their target/nontarget (i.e., match/mismatch) decision on the particular central letter they saw within the sequence one or two trials before, respectively. On every trial, participants pressed the key 'l' with their right index finger on a standard QWERTZ-keyboard in case of an n-back target, whereas they pressed the key 'd' with their left index finger if the currently presented central letter was not an n-back target. Assignment of 'yes'/'no'-key and left/right index finger was counterbalanced across participants.

The n-back stimuli sequences were presented in blocks. Each block consisted of 154 trials. Half of the trials were targets, half of the trials were nontargets. About one third of the stimuli of each response category were incongruent (e.g., 'FFF H FFF'), two thirds were congruent (e.g., 'FFF F FFF'). The first four trials of each block were always congruent nontargets and were excluded from any further analyses. Stimuli were presented using E-Prime presentation software (E-Prime 2 Professional, Psychology Software Tools, Inc.) with predefined stimuli lists. The trial sequences within the blocks were pseudo-randomized: To avoid attenuation of the interference effect for incongruent stimuli due to conflict adaptation processes (i.e., the so-called 'Gratton effect'; Botvinick, Braver, Barch, Carter, & Cohen, 2001; Davelaar, 2012; Gratton, Coles, & Donchin, 1992), incongruent-incongruent stimuli sequences were excluded in advance during construction of the stimuli lists. To further avoid any Gratton-like effects, congruent trials following incongruent trials were excluded from any further data analyses. Apart from these constraints, the stimulus sequence (i.e., the letters chosen as stimuli as well as their assignment as target/nontarget or congruent/incongruent) was randomly generated for each block and each participant.

To avoid a diminishing flanker effect in the later parts of a block, some randomly chosen stimuli were replaced by stimuli without a central letter (i.e., 10 targets and 10 nontargets per block consisted only of the flankers on both sides of a gap). We instructed participants to remember in these cases the flanker letters of the actual trial for the following comparison and to base their actual target/nontarget judgment on a comparison of the flanker letters with the previous central letter at position 1-back or 2-back, respectively. By means of this instruction we wanted to avoid participants becoming increasingly unaware of the flankers during the course of a block, that is, we wanted to avoid a diminishing flanker effect that would be attributable only to effects of time. We excluded these 'gap'-stimuli and the two immediately following stimuli from any further analyses.

At the beginning of each block, the n-back level that participants' had to perform during the following stimuli sequence was announced (0-back, 1-back, or 2-back). In case of the 0-back, the target letter for this task (randomly chosen out of the four possible letters) was additionally shown. The subsequent stimuli sequence did not differentiate between the n-back load-levels. In accordance with the traditional n-back task design, each block consisted of one n-back level. Each n-back level was presented twice. Thus, participants performed a total of six blocks. The sequence of blocks was randomly assigned for each participant with the constraint that each n-back level was presented once before an n-back level was presented for the second time. One block lasted about five minutes.

At the beginning of the study, participants performed training blocks for each n-back level. Training was repeated until participants reached an accuracy of at least 60 percent correct responses. During training, participants' accuracy was displayed at the end of a block to give them feedback regarding their performance. No feedback was given during the actual task presentation. The total experiment, including EEG preparation, task training, task run, and breaks, took about 2 hours.

Apparatus

The study was run in a quiet room that was dimly lit. Participants sat in a comfortable chair in front of a 22-inch Dell monitor (1680x1050 pixels screen resolution). Below the monitor a 250 Hz SMI remote eye tracking system with infrared-cameras was positioned to record participants' pupil sizes. A chin rest was used to avoid head movements during data recording and to ensure the eyes remained a fixed distance of about 70 cm from the eye tracking device. The eye tracking data were recorded at a sampling rate of 250 Hz (SMI iView X 2.7.13). The eye tracker was calibrated at the beginning and after each break using the built in calibration routines (SMI Experiment Center, 9-point calibration).

EEG was recorded at 27 electrode sites positioned according to the international 10/20 system (Jasper, 1958). The right mastoid served as reference during recording. The ground electrode was positioned at AFz. Three additional electrodes were placed around the eyes for EOG recording. EEG data were recorded (PyCorder 1.0.2) at 500 Hz sampling rate (ActiCHamp, Brainproducts, Inc.) using active electrodes (ActiCap, Brainproducts, Inc.). Impedances were kept below 5 kOhm.

Data Preprocessing and Analysis

Accuracy was calculated for each task condition as the mean percentage of correct response of all trials. Reaction times (RT) were calculated as the mean RT for correct responses in each task condition. Additionally, responses faster than 200 ms were excluded from RT calculation (less than 0.05% of all trials).

Eye tracking data were synchronized with the EEG data and preprocessed using customized Matlab scripts (Matlab 2012b, MathWorks, Inc.; EEGLAB v. 11.0.5.4b, Delorme & Makeig, 2004, with EYE-EEG plugin, Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011). During preprocessing, the eye tracking data were upsampled to 500 Hz to match the sampling rate of the EEG data. Eye blink artifacts (i.e., missing data points) were corrected using linear interpolation. The pupil sizes of the left and the right pupil were averaged. This mean pupil size was averaged over the entire time window from 0 ms to 2000 ms after stimulus onset for each trial. Finally, the mean pupil size values for all trials of each task condition were calculated and used for statistical analyses.

EEG data were preprocessed and analyzed using the same software solutions as described above. During preprocessing the continuous EEG data were filtered (low-pass 40 Hz, high-pass 0.5 Hz, linear finite impulse response filters). EOG artifacts were corrected using independent component analysis (ICA) decomposition. Independent components (ICs) identified as EOG-ICs by visual inspection were rejected. The continuous EEG data were then epoched in stimulus-locked time windows (-500 ms to 2500 ms). Note that the time window we chose for epoching was larger than the length of a single trial. This was done because, for technical reasons, the later time-frequency decomposition of the signal needed some additional data at the trial borders to cover the whole trial (0 - 2000 ms). An automatic artifact removal was performed: Epochs that exceeded ±100 µV were excluded from further analyses (Duncan et al., 2009). No further artifact removal or correction was performed on the EEG data. In the final data set we had in average 56 (*SD* = 11) correctly responded, artifact-free trials per task condition. Finally the EEG data were re-referenced to average reference.

P300 mean amplitude was calculated as the mean power in the typical P300 time window of 250 ms to 500 ms after stimulus onset (Polich, 2007). Statistical analyses were performed for the mean P300 amplitude at electrode Pz.

EEG time-frequency representations (TFRs) give information not only of the activation strength of neuronal network structures (i.e., the frequency band power) but also of their timing (Krause, Pesonen, & Hämäläinen, 2010; Palomäki, Kivikangas, Alafuzoff, Hakala, & Krause, 2012; Pesonen et al., 2007). We calculated TFRs at electrodes Fz and Pz for each task condition and participant separately within a frequency band range from 2 Hz to 32 Hz and a time range from 0 ms to 2000 ms. First, using stepwise fast Fourier transforms (FFTs, 500 ms width moving windows, 10 ms steps), the frequency band power for each TFR data point was calculated for the entire epoch length. Then, the percentage of event-related desynchronization/synchronization (ERD/ERS%; Pfurtscheller & Lopes da Silva, 1999) was calculated for each data point with respect to the mean frequency band power of a global pre-stimulus time interval (-500 ms - 0 ms). Finally, the individual TFRs were averaged over all participants (grand-average). TFR plots were created for the electrodes Fz and Pz (see Figure 4). Pairwise comparisons of consecutive n-back levels and the flanker conflict conditions were performed (EEGLAB bootstrapping statistics, using false-discovery-rate corrections for multiple comparisons).

Beyond their role for comparing different types of load on executive WM functions, we additionally used the TFRs for defining the time-window for calculating the mean frequency band power over time. This specific measure allowed for a more concise comparison with the other load measures that were also averaged over time (i.e., pupil dilation and P300 mean amplitude). EEG frequency band power was calculated using FFTs for the time window of 0 ms to 1000 ms after stimulus onset. This time window showed maximal oscillatory effects for both the updating load and the inhibition load as revealed by visual inspection of the TFR-plots. We used the individual alpha frequency (IAF) band selection account of Klimesch (1999) to calculate the mean frequency band power for individual adjusted frequency band borders of the theta and the upper alpha frequency band.

The ratio of this procedure is to calculate more dynamic frequency band borders that take individual differences into account. The IAF for each participant was calculated at electrode Pz as the frequency point of maximum alpha frequency band power during a rest period (i.e., a 20 seconds EEG recording, within which participants had to relax with eyes open, centrally fixating the screen). The mean IAF for all participants was at 10.29 Hz (*SD* = 0.40). In line with Klimesch (1999), the theta frequency band was then defined individually for each participant as the frequency band of 2 Hz bandwidth that started 6 Hz below the IAF. The upper alpha was defined as the frequency band of 2 Hz bandwidth that started directly above the IAF. Mean frequency band power for IAF adjusted theta and upper alpha frequency bands were calculated for each task condition at the electrodes Fz and Pz, respectively.

We restricted our EEG data analyses to these two electrodes, as frontal-midline electrodes like Fz have been reported to show largest effect sizes for WM load in the theta frequency band whereas parietal-midline electrodes like Pz show largest effect sizes for WM load in the alpha frequency band (e.g., Gevins et al., 1997). P300 amplitude effects have also been reported to be maximally over parietal electrodes (e.g., Polich, 2007), therefore a restriction of this measure to Pz was chosen.

Statistical Analysis

Except for the TFRs, which were statistically analyzed within MATLAB 2012b (MathWorks, Inc.), all statistical analyses were conducted using SPSS Statistics 20.0. (SAS Software, Inc.). For each measure, we computed separate repeated-measures ANOVAs with the factors *n-back level* (0-back, 1-back, 2-back) and *congruency* (flanker interference: congruent vs. incongruent flankers). For the physiological data only artifact-free trials with correct responses were used for data analysis, with additional exclusion of trials that might yield to any Gratton-like effect (see Procedure section above). This procedure resulted in the following amount of trials per factor level: 0-back, congruent, *M* = 54.05, *SD* = 5.24; 0-back, incongruent, *M* = 65.68, *SD* = 7.08; 1-back, congruent, *M* = 52.77, *SD* = 6.68; 1-back, incongruent, *M* = 64.54, *SD* = 9.50; 2-back, congruent, *M* = 44.05, *SD* = 8.89; 2-back, incongruent, *M* = 55.05, *SD* = 11.07. Greenhouse-Geyser corrections were performed on the *p*-values where necessary and additionally epsilon values were given in these cases. For post-hoc pairwise comparisons (t-tests, two-tailed) all *p*-values were Bonferroni corrected for multiple comparisons. Level of significance was set at $\alpha = .05$ for all analyses and partial eta-square is reported as a measure of effect size for the ANOVAs. Cohen's d_z for dependent measures is reported as a measure of effect size for the t-tests.

Results

TFR results are shown in Figure 4, all other results (mean values) in Figure 2. Frequency band power data is given in $10*\log_{10}[\mu V^2/Hz]$. Henceforth, for reasons of readability, we omitted noting the "$10*\log_{10}$". Figure 3 additionally shows the event-related potential (ERP) curves at Pz.

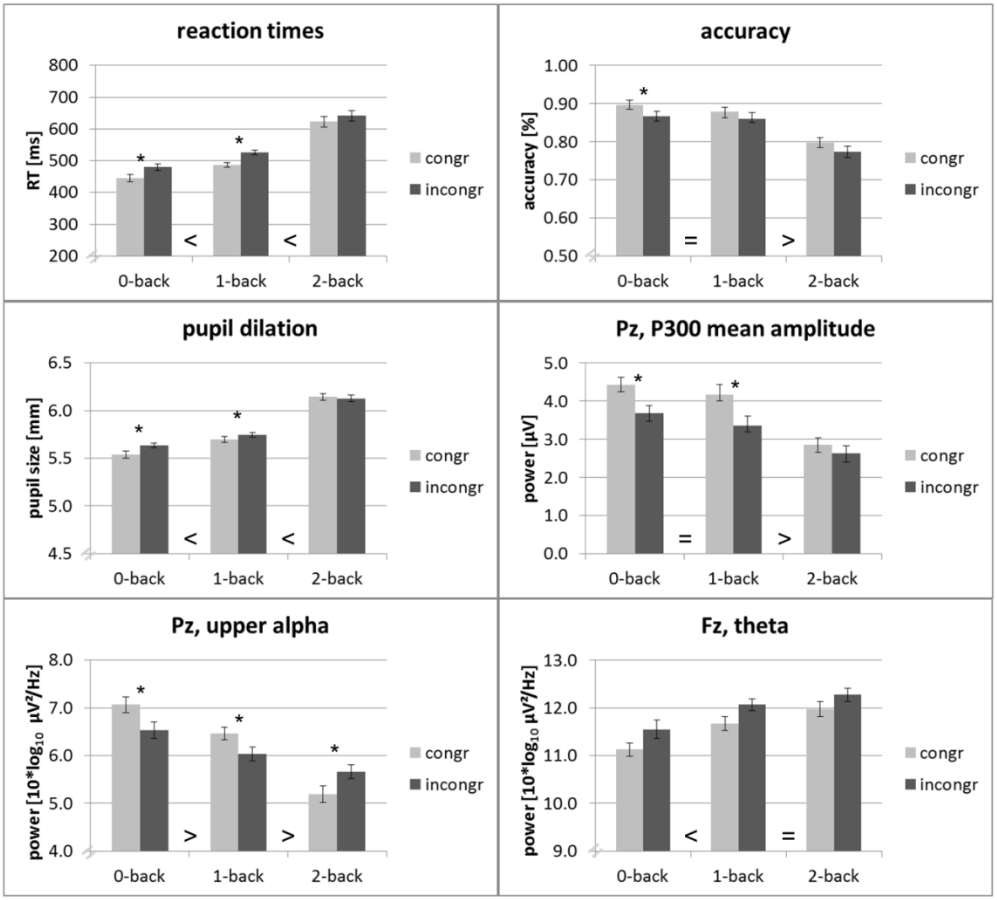


Figure 2. Mean values for reaction times, accuracy, pupil size (indicating pupil dilation), P300 amplitude at electrode Pz, upper alpha power (Pz), and theta power (Fz). Error bars: ± 1 SEM. The *, >, and < mark significant differences ($p < .05$). Light gray color symbolizes congruent trials, dark gray color incongruent trials.

Reaction Times (RTs)

RTs were significantly higher for incongruent flanker trials (549 ms) than for congruent flanker trials (518 ms), indicated by a main effect of congruency, $F(1, 21) = 30.83$, $p = .001$, $\eta_p^2 = .60$. In addition, RTs increased with increasing n-back levels (0-back: 462 ms, 1-back: 506 ms, 2-back: 632 ms; all n-back levels' RTs were significantly different), yielding a main effect of n-back level, $F(2, 42) = 39.55$, $p < .001$, $\varepsilon = .56$, $\eta_p^2 = .65$. These two main effects were modulated by a significant interaction between n-back level and congruency, $F(2, 42) = 3.65$, $p = .047$, $\varepsilon = .78$, $\eta_p^2 = .15$. Post-hoc pairwise comparisons revealed that a significant flanker interference effect could only be observed for the 0-back (35 ms slower reactions for incongruent as compared to congruent flanker trials, $p < .001$, $d_z = -1.44$) and 1-back (39 ms slower reactions for incongruent flanker trials, $p < .001$, $d_z = -1.49$) conditions, but not for the 2-back condition (18 ms slower reactions for incongruent flanker trials, $p = .077$, $d_z = -0.40$). Paired sampled t-tests (two-tailed) were conducted to further assess whether flanker interference effect (i.e., the RT difference in between incongruent and congruent trials) were larger in the 0-back and 1-back conditions than in the 2-back condition. While no significant difference occurred between the 0-back and 1-back condition, $t(21) = 0.73$, $p = .47$, $d_z = -0.16$, the congruency effect was significantly reduced in the 2-back relative to the 1-back condition, $t(21) = 2.27$, $p = .03$, $d_z = 0.48$, and tended to be reduced in the 2-back relative to the 0-back condition, $t(21) = 1.82$, $p = .08$, $d_z = 0.39$. This suggests the congruency effect decreases with increasing WM updating load.

Accuracy

For accuracy, we found significant main effects of congruency, $F(1, 21) = 6.03$, $p = .023$, $\eta_p^2 = .22$ and of n-back level, $F(2, 42) = 17.58$, $p < .001$, $\varepsilon = .62$, $\eta_p^2 = .46$, but no interaction effect ($p = .89$). Overall accuracy for congruent trials was significantly higher (86%) than for incongruent trials (83%). Moreover, accuracy decreased with increasing n-back levels (0-back: 88%, 1-back: 86%, 2-back: 79%). Post-hoc pairwise comparisons revealed that accuracy in the 2-back condition was significantly reduced relative to the 0-back and the 1-back conditions ($p < .001$), with no significant difference occurring between the latter two conditions ($p = .78$).

Pupil Dilation

Results for pupil dilation were in accordance with the RT results. As expected, pupil diameter increased for increasing n-back levels (0-back: 5.59 mm, 1-back: 5.72 mm, 2-back: 6.14 mm), yielding a main effect of n-back level, $F(2, 42) = 56.69$, $p < .001$, $\eta_p^2 = .73$. Additionally, pupil sizes were larger for incongruent flanker trials (5.84 mm) than for congruent flanker trials (5.80 mm), resulting in a main effect of congruency, $F(1, 21) = 20.15$, $p = .001$, $\eta_p^2 = .49$. Moreover, as for RTs we observed a significant interaction between n-back level and congruency, $F(2, 42) = 15.27$, $p < .001$, $\eta_p^2 = .42$. Only in the 0-back and 1-back conditions, we found a significant flanker interference effect on pupil dilation, that is, significantly increased pupil diameter for incongruent as compared to congruent flanker trials (0-back: 0.09 mm difference, $p < .001$, $d_z = -1.08$, 1-back: 0.05 mm difference, $p = .007$, $d_z = -0.63$). For the 2-back condition, we even observed a marginally significant reversed pattern with a slight decrease in pupil dilation (-0.02 mm difference, $p = .052$, $d_z = 0.44$). Further paired-samples t-tests revealed that the interference effect was larger in the 0-back condition relative to the 1-back condition, $t(21) = 2.71$, $p = .01$, $d_z = 0.58$, and the 2-back task condition, $t(21) = 4.78$, $p < .001$, $d_z = 1.02$, as well as in the 1-back relative to the 2-back condition, $t(21) = 3.30$, $p = .003$, $d_z = 0.70$.

P300

Results for the event-related potential curves at electrode Pz are shown in Figure 3. The mean amplitude values are given in figure 2. We found a decrease in mean P300 amplitude with increasing n-back levels (0-back: 4.05 µV, 1-back: 3.76 µV, 2-back: 2.75 µV), indicated by a main effect of n-back level, $F(2, 42) = 12.47$, $p < .001$, $\varepsilon = .71$, $\eta_p^2 = .37$ as well as a decrease in P300 amplitude for the flanker interference, resulting in a main effect congruency, $F(1, 21) = 15.44$, $p = .001$, $\eta_p^2 = .42$. Additionally we found a marginally significant interaction between n-back level and congruency, $F(2, 42) = 2.82$, $p = .07$, $\eta_p^2 = .12$. The pattern of this interaction was the same as for RTs and pupil dilation. Only in the 0-back and 1-back task conditions, P300 amplitude was significantly decreased for incongruent as compared to congruent flanker trials (0-back: -0.74 µV, $p < .001$, $d_z = 0.78$, 1-back: -0.82 µV, $p = .005$, $d_z = 0.67$). For the 2-back task condition, this difference was not significant (2-back: -0.21 µV, $p = .28$, $d_z = 0.24$). While interference effects did not differ between the 0-

back and the 1-back condition, *t*(21) = 0.24, *p* = .81, d_z = 0.05, the interference effect was significantly reduced in the 2-back relative to the 0-back, *t*(21) = 2.10, *p* = .048, d_z = -0.45 and the 1-back condition, *t*(21) = 2.12, *p* = .046, d_z = -0.45.

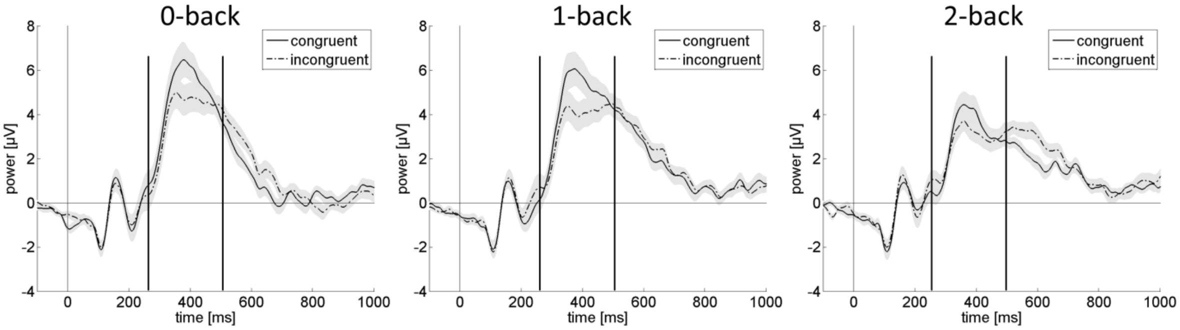


Figure 3. Event-related potential curves at electrode Pz, baseline-related to a 250 ms to 0 ms pre-stimulus time-interval. Solid black curve indicates congruent flanker trials, black dotted curve incongruent flanker trials. From left to right increasing n-back levels. Vertical black lines mark the time window for the calculation of mean P300 amplitude (250 ms to 500 ms after stimulus-onset). Grey shades around the lines are ± 1 SEM.

Time-Frequency Representations (TFRs)

Results for the TFRs are provided in Figure 4. Similar to Pesonen et al. (2007) and Krause et al. (2010), we found for increased n-back levels longer-lasting and more pronounced event-related synchronization (ERS) in the theta frequency range as well as longer-lasting and more pronounced event-related desynchronization (ERD) in the alpha frequency range. The alpha ERD and theta ERS for the flanker conflict were more pronounced within the first 1000 ms after stimulus onset.

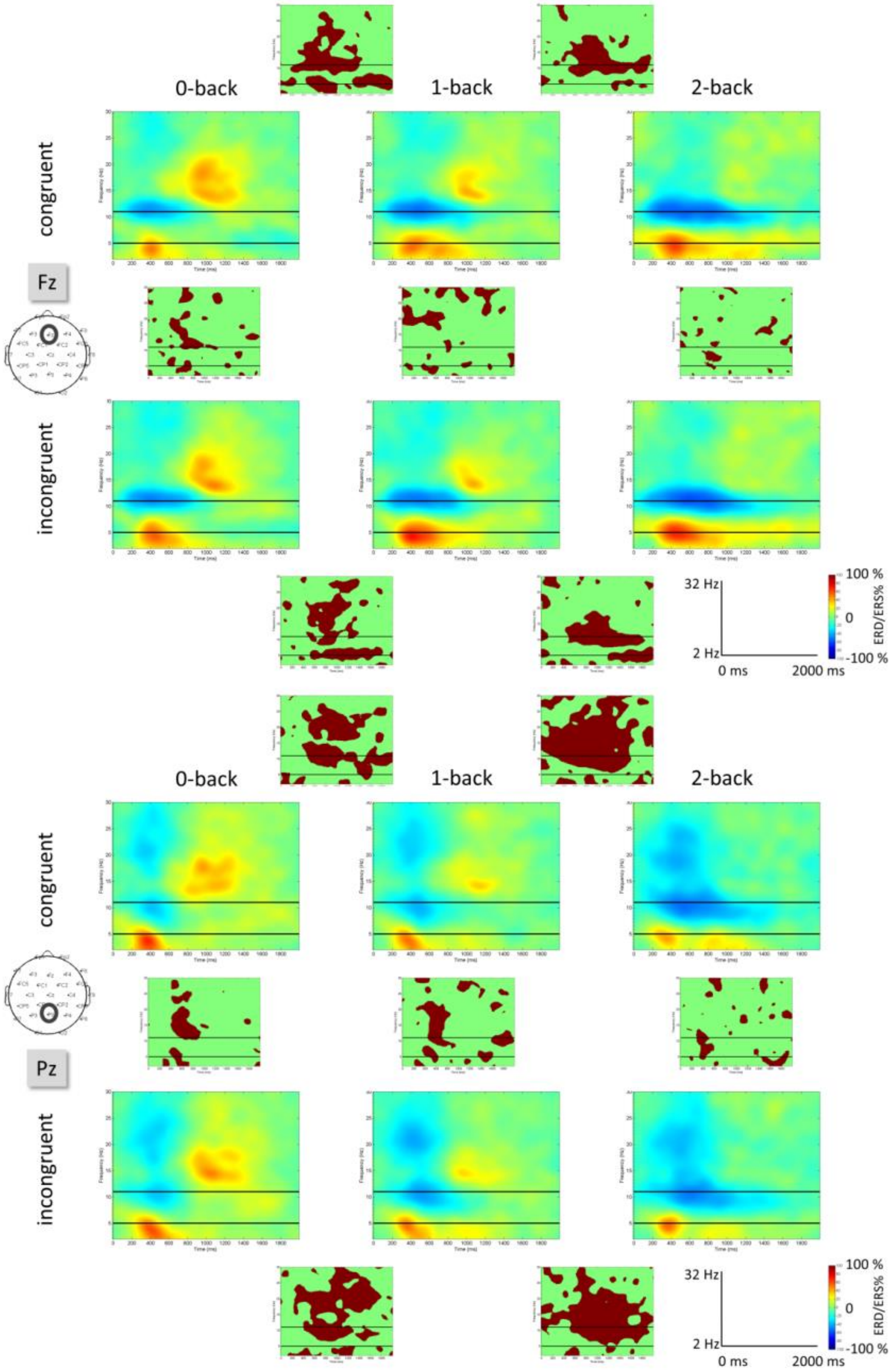


Figure 4. Time-Frequency representations (TFRs) of the n-back levels (increasing from left to right) for congruent and incongruent stimuli at frontal electrode Fz (upper half of the figure) and at parietal electrode Pz (lower half of the figure). Within the Fz and Pz part of the figure, the larger time-frequency plots show stepwise (10 ms bins) the percentage of change in frequency band power (y-axis from 2 Hz to 32 Hz) for the entire trial

length (x-axis from 0 ms to 2000 ms) in relation to a 500 ms pre-stimulus baseline period (i.e., the percentage of event-related desynchronization / synchronization; ERD/ERS%, see Pfurtscheller & Lopes da Silva, 1999, and Pesonen et al., 2007). Blue colors indicate a desynchronization, that is, a reduction of frequency band power in relation to baseline, red colors indicate a synchronization, that is, an increase of frequency band power in relation to baseline. The smaller plots in between give the areas of statistically significant differences (dark red color; $p < .05$) between two conditions (left to right: updating load; upper to lower: inhibition load). Black vertical lines in the TFR plots show the approx. center frequencies of the theta frequency band (5 Hz, lower vertical line) and the upper alpha frequency band (11 Hz, upper vertical line).

Mean Upper Alpha Frequency Band Power

The results for the mean upper alpha frequency band power at electrode Pz are shown in Figure 2. We found a main effect of n-back level, $F(2, 42) = 22.14$, $p < .001$, $\eta_p^2 = .51$, due to a decrease in frequency band power for increasing n-back levels (0-back: 6.80 μV^2/Hz, 1-back: 6.25 μV^2/Hz, 2-back: 5.43 μV^2/Hz). A significant interaction between n-back level and congruency, $F(2, 42) = 8.38$, $p < .001$, $\eta_p^2 = .29$, revealed that the flanker interference effect led to a decrease in upper alpha band power only for the low updating-load conditions (0-back: -0.54 μV^2/Hz, $p = .026$, $d_z = 0.51$, 1-back: -0.43 μV^2/Hz, $p = .023$, $d_z = 0.53$). Comparable to the pupil dilation results, for the 2-back condition we observed the reversed pattern: incongruent flanker trials showed an increase in upper alpha band power instead of the expected decrease (2-back: 0.47 μV^2/Hz, $p = .017$, $d_z = -0.55$). Additional paired-samples t-tests revealed that the flanker interference effects of the 0-back and 1-back task condition were not statistically different, $t(21) = 0.40$; $p = .70$, $d_z = -0.08$, while the flanker interference effect was significantly reduced in the 2-back relative to the 0-back, $t(21) = 3.21$, $p = .004$, $d_z = -0.69$, and 1-back condition, $t(21) = -3.97$, $p = .001$, $d_z = -0.85$.

Mean Theta Frequency Band Power

The results for the mean theta frequency band power at electrode Fz are also shown in Figure 2. As expected, theta frequency band power increased for increasing n-back levels (0-back: 11.34 μV^2/Hz, 1-back: 11.87 μV^2/Hz, 2-back: 12.12 μV^2/Hz), yielding a main effect of n-back level, $F(2, 42) = 7.84$, $p = .001$, $\eta_p^2 = .27$, as well as for incongruent (11.96 μV^2/Hz) as compared to congruent flanker trials (11.59 μV^2/Hz), resulting in a main effect of congruency, $F(1, 21) = 11.81$, $p = .002$, $\eta_p^2 = .36$. We observed no interaction of congruency and n-back level, $F(2, 42) = .10$, $p = .90$, $\eta_p^2 = .01$. Thus, for this measure the flanker interference showed a purely additive effect.

Discussion

In the current study, we investigated the interplay of two EFs, inhibition and updating, manipulated within one single task. We hypothesized that due to common processes of attention control involved in all EFs we might observe interactions between simultaneous load on inhibitory control and WM updating for load-related measures like RTs, EEG theta and alpha frequency band power, P300 amplitude, and pupil dilation. These interactions might be either due to over-additive (*depletion hypothesis*) or under-additive effects (*facilitation hypothesis*).

Our results showed, first, that all measures yielded the expected outcomes for the load on WM updating. With increasing n-back levels, accuracy and mean P300 amplitude at parietal electrodes decreased, whereas RTs and pupil dilation increased (see Watter et al., 2001; Ewing & Fairclough, 2010). Similarly, theta power increased and upper alpha power decreased (see Gevins & Smith, 2000). Additionally we observed similar effects in the TFRs as found by Pesonen et al. (2007) and Krause et al. (2010): Increased n-back levels led to more prolonged and pronounced theta frequency band ERS and alpha frequency band ERD. Thus, the observed effects in behavioral and electrophysiological measures for load on updating are in line with findings of previous studies using the n-back task paradigm, confirming the sensitivity of the used indicators as measures of WM updating load.

Second, all measures were also sensitive to inhibitory demands. As expected, we observed that demands on inhibitory control (i.e., incongruent relative to congruent flanker trials) resulted in increased RTs, theta frequency band power, and pupil dilation as well as in decreased alpha frequency band power and accuracy (see Hanslmayr et al., 2008; Nigbur et al., 2011).

The most important finding with regard to the topic of the current paper was, however, that in contrast to studies by Lavie and colleagues (Lavie et al., 2004; Lavie, 2005, 2010), the flanker interference effect did not increase under high WM updating load as shown by the outcomes of all measures. Moreover, instead of an over-additive effect, we observed an under-additive effect for the flanker interference under high WM updating load on most of our load-related measures. Only for no or low load on updating (i.e., the 0-back and 1-back task conditions), the flanker interference effect led to increased RTs and increased pupil dilation and decreased upper alpha frequency band power and decreased P300 mean amplitude. In contrast, under high load on updating (2-back) we observed no significant flanker interference effect (RT, pupil dilation, P300) or the effect even was reversed (EEG upper alpha frequency band power). Importantly, for these measures we found the flanker interference effect to be significantly reduced in the 2-back compared with the 1-back and 0-back conditions.

Thus, with regard to our two alternative hypotheses of an over-additive or under-additive load effect for simultaneous load on updating and inhibition, the results favor the latter. In line with Miyake et al. (2000), our results indicate that the EFs updating and inhibition might share underlying network structures that serve controlled attention. The high demands of the 2-back task on WM updating might have led to increased levels of controlled attention which in turn facilitated inhibitory control processes and enhanced participants' attentional focus on the task-relevant central letter. This enhanced attentional focus on the central letter may have resulted in a shielding against the distracting flanker letters, leading to the observed reduced flanker interference effects in the 2-back condition.

This explanation is in line with the predictions of the task-engagement/distraction trade-off model by Sörqvist & Rönnberg (2014). In dual task studies investigating auditory distraction under visual WM updating load these authors and others observed that task irrelevant auditory stimuli caused less interference under high as compared to low visual WM updating load (San Miguel, Corral, & Escera, 2008; Sörqvist, Stenfelt, & Rönnberg, 2012). They concluded that the higher task engagement under more difficult task conditions shielded cognition from auditory distraction. Our results might indicate that this shielding from distraction might even work within the same stimulus modality when stimuli that load on WM updating and stimuli that demand inhibitory control are both of the same, visual domain. More importantly however, as we manipulated load on WM updating and demands on inhibitory control within one single task, we avoided additional load on shifting that might confound the results of dual task studies with respect to a specific interpretation of the interplay of updating and inhibition. Thus, we may conclude that our novel task paradigm of an n-back task with flanker stimuli might have identified the specific interplay of WM updating and inhibitory control.

Our *facilitation hypothesis* is supported by the P300 mean amplitude as an indicator of internal attentional distribution (Watter et al., 2001). In line with Watter et al. (2001), we interpret the observed decrease in P300 amplitude for the higher n-back levels and the flanker interference as an indicator of the distribution of attention required by the current task demands. When different EFs are necessary for task performance, attention is thought to be distributed among updating and inhibition functions which resulted in the decrease of the P300 amplitude. The attenuated P300 amplitude reduction for load on inhibitory control in the 2-back task as compared to the 0-back or 1-back task suggests that no additional attentional resources must be attributed to inhibitory control processes in the 2-back condition because attentional resources are already fully focused on the task-relevant central letter due to the severe demands on WM updating.

Upper alpha frequency band power seems to corroborate the *facilitation hypothesis*. Under high WM updating load (2-back) we observed no further decrease of alpha frequency band power for incongruent as compared to congruent flanker stimuli, but rather an increase. This larger decrease of alpha frequency band power for congruent as compared to incongruent flanker stimuli between 1-back and 2-back load level might be puzzling. However, in line with recent interpretations of the functional relevance of alpha oscillatory activity for inhibitory processes and the so called 'inhibition-timing hypothesis' (Klimesch, 2012), we hypothesize that this outcome might indicate strong inhibitory processes under the 2-back load level that led to an increase of alpha oscillatory activity. For example, Händel, Haarmeier, & Jensen (2011) observed that information in a visual hemifield that participants were cued to ignore led to increased alpha frequency band power in the contra-lateral hemisphere. They interpreted this observation in terms of neuronal inhibitory processes triggered by increased alpha activity. Thus, in our study, the severe load on WM updating in the 2-back task condition may have led to strong inhibitory processes with regard to the flanker interference as shown by the increased alpha frequency band power for incongruent compared to congruent flanker stimuli in the 2-back task condition. We speculate that this outcome could potentially reflect overlying effects of oscillatory activity in the alpha frequency band for simultaneous load on updating and inhibition. Clearly future studies are necessary to further elaborate on circumstances under which load on inhibitory control leads to an increase rather than a decrease of alpha frequency band power.

The results for pupil dilation as overall measure of WM load (Beatty & Lucero-Wagoner, 2000) also indicated that under severe load on WM updating inhibitory control processes are facilitated, possibly due to an increased attentional focus on task relevant information. As consequence, incongruent flanker stimuli in the 2-back task condition did not lead to additionally increased pupil diameter relative to congruent stimuli. Pupil dilation data should be interpreted with some caution, although the findings are in line with the outcomes of the RTs, alpha frequency band power, and P300 measures. First, as the inter-trial interval of two seconds used in the current study was relatively short (see, e.g., Laeng et al., 2011, using inter-trial intervals of three seconds in a Stroop task), the pupil dilation data may not be free of carry-over effects. In other words, the magnitude of the pupil dilation on the current trial may be affected by the pupil size in the preceding trial. However, as the sequence of congruent and incongruent flanker trials was random for each n-back load level and each participant, we argue that such effects should affect congruent and incongruent trials equally. Second, we reported absolute pupil diameter values, contrary to studies using single-trial pre-stimulus baseline-corrected pupil diameters (e.g., Laeng et al., 2011; Van Gerven, Paas, Van Merriënboer, & Schmidt, 2004). Importantly, the n-back task paradigm does not allow specifying single-trial pre-stimulus baselines that are free of the cognitive demands of interest in the current study. This is because of the continuous WM demands across trials that are inherent for the n-back task. However, we think that baseline-corrected pupil dilatation data may be mainly important for between-group comparisons (Van Gerven et al., 2004), but it may less be an issue in case of within-subjects designs as in the current study. In within-subject designs, the amount of pupil dilation for cognitive processes has been shown to be rather unaffected by the initial pupil size, that is, the impact of cognitive demands on pupil dilatation appears robust to the pupil diameter in a pre-stimulus baseline (e.g., Beatty & Lucero-Wagoner, 2000; Bradshaw, 1969). Thus, although the pupil results must be interpreted with some caution due to methodological limitations, we think that, nevertheless, the pupil dilatation data provide additional support for our conclusions, given the conformity of the pupil data with other measures.

In conclusion, the pattern of results we consistently found on most load-related measures indicate that severe load on WM updating might have activated processes of controlled attention serving all EFs, thus enhancing inhibitory control. However, the outcomes of the current study are in contrast to studies by Lavie and colleagues (Lavie et al., 2004; Lavie, 2005, 2010). Some conceptual differences might explain the different outcomes. Most important, Lavie and colleagues used dual task paradigms with one task loading on WM storage components and the other demanding inhibitory control. In contrast, we specifically demanded two EFs, namely updating and inhibition within one single task. Thus the over-additive load effect Lavie and colleagues observed might be due to the necessity of dividing controlled attention between executive WM components (i.e., inhibitory control) and WM storage components with additional processes required to handle the overload-situation in the WM storage components. This overload-hypothesis is further corroborated by the fact that the over-additive load effect seems only to be present if the stimuli of the WM storage task and the stimuli of the flanker interference task were of the same category, thereby inducing load on the same, resource-limited storage component (see Kim et al., 2005; Park et al., 2007).

The results of the current study indicate that if EFs are specifically loaded, rather than a depletion of attentional processes (see *depletion hypothesis*) an enhanced activation of attention control might occur, thus leading to a decreased (or even absent) flanker interference effect under high WM updating load (see *facilitation hypothesis*). Naturally, the current study could only serve as a first step in studying the interplay between different EFs when manipulated within one single task. Some open questions may be addressed in more detail in future studies.

Firstly, one might wonder whether increasing WM updating load further beyond the 2-back load level (e.g., 3-back or 4-back) might consistently result in a reduced (or absent) flanker interference effect under all high WM updating load conditions. Given our *facilitation hypothesis*, we would expect to observe an under-additive flanker interference effect for load levels above the 2-back as well. However, for severely n-back overload (e.g., for load levels above 4-back) it might be also plausible to expect a depletion (or collapse) of attentional control processes, thus leading to a reoccurrence (or even strengthening) of the flanker interference effect. This question has to be addressed by future studies.

Secondly, one might wonder whether the interaction we observed between load on WM updating and inhibitory control demands may generalize to other task-combinations. For example, an obvious next study would be to use Stroop stimuli instead of the flanker stimuli to demand inhibitory control. Given such a task, we would expect to observe comparable outcomes as the ones of the current study. Interestingly, we might interpret a recent study by Soutschek and colleagues (Soutschek, Strobach, & Schubert, 2013, experiment 2) as a hint for the generalizability of our assumptions. These authors used the n-back task to manipulate load on WM updating and combined this task with a Stroop task, yet in a dual task setting. Thus, contrary to our current experiment, the task-relevant information in the updating task was not identical with the task-relevant information in the inhibition task. Nevertheless, Soutschek and colleagues also observed a significantly decreased interference effect at the highest (2-back) WM updating load condition. However, as the authors focused on a different research question, they did not explicitly address and discuss these effects. Therefore, further studies on this topic might be desirable.

Finally, one might wonder at which level of the stimulus processing chain the enhanced inhibitory control under severe load on updating takes place. Although this question was not in the primary focus of our current study, we may hypothesize from our results that the task irrelevant stimulus dimension (i.e., the flanker) seems to be processed even in the 2-back task condition at least up to certain neuronal stages. This can be inferred from the purely additive effect that we observed for the theta frequency band power which might indicate that specific aspects of inhibitory control are equally activated under all n-back levels. In line with this reasoning, close connections between specific EFs and frontal oscillatory activity in the theta frequency band range have recently been hypothesized by several authors (e.g., Nigbur, Ivanova, & Stürmer, 2011; Sauseng, Hoppe, Klimesch, Gerloff, & Hummel, 2007; Sauseng, Griesmayr, Freunberger, & Klimesch, 2010; Womelsdorf, Vinck, Leung, & Everling, 2010). According to these authors, theta activity seems to reflect the subordinated coordination of the diverse cognitive processes in WM (Sauseng et al., 2010). Our results might indicate that these subordinated coordination processes are rather unaffected by load on updating.

In sum, our study revealed a close connection of the EFs updating and inhibition which was shown by an interaction between these two EFs on most load-related measures. The under-additive load effect we observed might be explained by commonly activated neuronal network structures of controlled attention. Severe load on WM updating might lead to overall increased processes of controlled attention, thus enhancing inhibitory control and resulting in a decreased flanker interference effect under high WM updating load. Clearly, future studies have to be conducted to overcome the limitations of the current study as discussed above (e.g., concerning the pupil dilation data) and to validate the overall generalizability of our results with respect to a invariable or even decreased flanker interference effect under high WM updating load.

References

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology*, *49*(1), 5–28. doi: 10.1080/713755608

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

Baddeley, A. (2007). *Working memory, thought, and action.* New York, NY, USA: Oxford University Press. doi:10.1093/acprof:oso/9780198528012.001.0001

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29. doi:10.1146/annurev-psych-120710-100422

Baddeley, A., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89. doi:10.1016/S0079-7421(08)60452-1

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. Berndtson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 142–162). Cambridge: Cambridge University Press.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. doi: 10.1037/0033-295X.108.3.624

Bradshaw, J. L. (1969). Background light intensity and the pupillary response in a reaction time task. *Psychonomic Science*, 14(6), 271–272. doi: 10.3758/BF03329118

Collette, F., Van der Linden, M., Laureys, S., Delfiore, G., Degueldre, C., Luxen, A., & Salmon, E. (2005). Exploring the unity and diversity of the neural substrates of executive functioning. *Human Brain Mapping*, *25*(4), 409–423. doi:10.1002/hbm.20118

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A.R.A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42–100. doi:10.1016/j.cogpsych.2004.12.001

Davelaar, E. J. (2012). When the ignored gets bound: Sequential effects in the flanker task. *Frontiers in Psychology*, *3*(552), 1–13. doi:10.3389/fpsyg.2012.00552

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. doi:10.1016/j.jneumeth.2003.10.009

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. doi:10.1146/annurev-psych-113011-143750

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: Analyses and review. *Journal of Experimental Psychology. General*, *140*(4), 552–572. doi:10.1037/a0023885

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, *120*(11), 1883–1908. doi:10.1016/j.clinph.2009.07.045

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. doi:10.1037/a0017891

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23. doi: 10.1111/1467-8721.00160

Eriksen, C. (1995). The flankers task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual Cognition*, *2*(2/3), 101–118. doi:10.1080/13506289508401726

Ewing, K., & Fairclough, S. (2010). The impact of working memory load on psychophysiological measures of mental effort and motivational disposition. In D. de Waard, A. Axelsson, M. Berglund, B. Peters, & C. Weikert (Eds.), *Human factors: A system view of human, technology and organisation.* (pp. 423–433). Maastricht: Shaker Publishing.

Fougnie, D. (2008). The relationship between attention and working memory. In N. B. Johansen (Ed.), *New research on short-term memory* (pp. 1–45). New York, NY, USA: Nova Science Publishers Inc.

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. doi:10.1093/cercor/10.9.829

Gevins, A., Smith, M. E., McEvoy, C. L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, *7*(4), 374–385. doi: 10.1093/cercor/7.4.374

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, *121*(4), 480–506. doi: 10.1037/0096-3445.121.4.480

Händel, B. F., Haarmeier, T., & Jensen, O. (2011). Alpha oscillations correlate with the successful inhibition of unattended stimuli. *Journal of Cognitive Neuroscience*, *23*(9), 2494–2502. doi:10.1162/jocn.2010.21557

Hanslmayr, S., Pastötter, B., Bäuml, K.-H., Gruber, S., Wimber, M., & Klimesch, W. (2008). The electrophysiological dynamics of interference during the Stroop task. *Journal of Cognitive Neuroscience*, *20*(2), 215–225. doi:10.1162/jocn.2008.20020

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. doi:10.1080/09658211003702171

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371–375.

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, *15*(8), 1395–1399. doi:10.1046/j.1460-9568.2002.01975.x

Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., & Koeppe, R. A. (1997). Verbal working memory load affects regional brain activation as measured by PET. *Journal of Cognitive Neuroscience*, *9*(4), 462–475. doi: 10.1162/jocn.1997.9.4.462

Karatekin, C., Marcus, D. J., & Couperus, J. W. (2007). Regulation of cognitive resources during sustained attention and working memory in 10-year-olds and adults. *Psychophysiology*, *44*(1), 128–144. doi:10.1111/j.1469-8986.2006.00477.x

Kim, S., Kim, M., & Chun, M. (2005). Concurrent working memory load can reduce distraction. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(45), 16524–16529. doi: 10.1073/pnas.0505454102

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, *29*(2-3), 169–195. doi: 10.1016/S0165-0173(98)00056-3

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*(12), 606–617. doi:10.1016/j.tics.2012.10.007

Krause, C. M., Pesonen, M., & Hämäläinen, H. (2010). Brain oscillatory 4-30 Hz electroencephalogram responses in adolescents during a visual memory task. *Neuroreport*, *21*(11), 767–771. doi:10.1097/WNR.0b013e32833bfcbb

Laeng, B., Ørbo, M., Holmlund, T., & Miozzo, M. (2011). Pupillary Stroop effects. *Cognitive Processing*, *12*(1), 13–21. doi:10.1007/s10339-010-0370-z

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. doi:10.1016/j.tics.2004.12.004

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, *19*(3), 143–148. doi:10.1177/0963721410370295

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. doi:10.1037/0096-3445.133.3.339

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (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. doi: 10.1006/cogp.1999.0734

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

Nee, D., Brown, J., & Askren, M. (2013). A meta-analysis of executive components of working memory. *Cerebral Cortex*, *23*, 264–282. doi:10.1093/cercor/bhs007

Nigbur, R., Ivanova, G., & Stürmer, B. (2011). Theta power as a marker for cognitive interference. *Clinical Neurophysiology*, *122*(11), 2185–2194. doi:10.1016/j.clinph.2011.03.030

Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation*, *51*, 45–100. doi:10.1016/S0079-7421(09)51002-X

Palomäki, J., Kivikangas, M., Alafuzoff, A., Hakala, T., & Krause, C. M. (2012). Brain oscillatory 4-35 Hz EEG responses during an n-back task with complex visual stimuli. *Neuroscience Letters*, *516*(1), 141–145. doi:10.1016/j.neulet.2012.03.076

Park, S., Kim, M.-S., & Chun, M. M. (2007). Concurrent working memory load can facilitate selective attention: Evidence for specialized load. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1062–1075. doi:10.1037/0096-1523.33.5.1062

Pesonen, M., Hämäläinen, H., & Krause, C. M. (2007). Brain oscillatory 4-30 Hz responses during a visual n-back memory task with varying memory load. *Brain Research*, *1138*, 171–177. doi:10.1016/j.brainres.2006.12.076

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology*, *110*(11), 1842–1857. doi: 10.1016/S1388-2457(99)00141-8

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. doi:10.1016/j.clinph.2007.04.019

Pratt, N., Willoughby, A., & Swick, D. (2011). Effects of working memory load on visual selective attention: Behavioral and electrophysiological evidence. *Frontiers in Human Neuroscience*, *5*(57). doi:10.3389/fnhum.2011.00057

Sanders, A., & Lamers, J. (2002). The Eriksen flanker effect revisited. *Acta Psychologica*, *109*(1), 41–56. doi: 10.1016/S0001-6918(01)00048-8

San Miguel, I., Corral, M.-J., & Escera, C. (2008). When loading working memory reduces distraction: Behavioral and electrophysiological evidence from an auditory-visual distraction paradigm. *Journal of Cognitive Neuroscience*, *20*(7), 1131–1145. doi:10.1162/jocn.2008.20078

Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: A possible function of EEG theta oscillations. *Neuroscience and Biobehavioral Reviews*, *34*(7), 1015–1022. doi:10.1016/j.neubiorev.2009.12.006

Sauseng, P., Hoppe, J., Klimesch, W., Gerloff, C., & Hummel, F. C. (2007). Dissociation of sustained attention from central executive functions: Local activity and interregional connectivity in the theta range. *The European Journal of Neuroscience*, *25*(2), 587–593. doi:10.1111/j.1460-9568.2006.05286.x

Shah, P., & Miyake, A. (1999). Models of working memory: An introduction. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 1–27). New York, NY, USA: Cambridge University Press.

Sörqvist, P., & Rönnberg, J. (2014). Individual differences in distractibility: An update and a model. *PsyCh Journal*, *3*(1), 42–57. doi:10.1002/pchj.47

Sörqvist, P., Stenfelt, S., & Rönnberg, J. (2012). Working memory capacity and visual-verbal cognitive load modulate auditory-sensory gating in the brainstem: Toward a unified view of attention. *Journal of Cognitive Neuroscience*, *11*(24), 2147–2154. doi: 10.1162/jocn_a_00275

Soutschek, A., Müller, H. J., & Schubert, T. (2013). Conflict-specific effects of accessory stimuli on cognitive control in the Stroop task and the Simon task. *Experimental Psychology*, *60*(2), 140–147. doi:10.1027/1618-3169/a000181

Soutschek, A., Strobach, T., & Schubert, T. (2013). Working memory demands modulate cognitive control in the Stroop paradigm. *Psychological Research*, *77*, 333–347. doi:10.1007/s00426-012-0429-9

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*(3736), 652–654. doi: 10.1126/science.153.3736.652

Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643-662. doi: 10.1037/h0054651

Suchy, Y. (2009). Executive functioning: Overview, assessment, and research issues for non-neuropsychologists. *Annals of Behavioral Medicine*, *37*(2), 106–116. doi:10.1007/s12160-009-9097-4

Vandierendonck, A. (2014). Symbiosis of executive and selective attention in working memory. *Frontiers in Human Neuroscience*, *8*(588). doi:10.3389/fnhum.2014.00588

Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, *41*, 167–174. doi:10.1111/j.1469-8986.2003.00148.x

Watter, S., Geffen, G. M., & Geffen, L. B. (2001). The n-back as a dual-task: P300 morphology under divided attention. *Psychophysiology*, *38*(6), 998–1003. doi: 10.1111/1469-8986.3860998

Womelsdorf, T., Vinck, M., Leung, L. S., & Everling, S. (2010). Selective theta-synchronization of choice-relevant information subserves goal-directed behavior. *Frontiers in Human Neuroscience*, *4*(210), 1–13. doi:10.3389/fnhum.2010.00210

2.3 Study 3

Pupil dilation and EEG alpha frequency band power reveal load on executive functions for link-selection processes during text reading

Christian Scharinger[1*], Yvonne Kammerer[1], & Peter Gerjets[1]

[1]Knowledge Media Research Center, Tuebingen, Germany

* Corresponding author E-mail: c.scharinger@iwm-kmrc.de

Abstract

Executive working memory functions play a central role in reading comprehension. In the present research we were interested in additional load imposed on executive functions by link-selection processes during computer-based reading. For obtaining process measures, we used a methodology of concurrent electroencephalographic (EEG) and eye-tracking data recording that allowed us to compare epochs of pure text reading with epochs of hyperlink-like selection processes in an online reading situation. Furthermore, this methodology allowed us to directly compare the two physiological load-measures EEG alpha frequency band power and pupil dilation. We observed increased load on executive functions during hyperlink-like selection processes on both measures in terms of decreased alpha frequency band power and increased pupil dilation. Surprisingly however, the two measures did not correlate. Two additional experiments were conducted that excluded potential perceptual, motor, or structural confounds. In sum, EEG alpha frequency band power and pupil dilation both turned out to be sensitive measures for increased load during hyperlink-like selection processes in online text reading.

Introduction

Imagine you want to gather some information about a certain topic. You may start by reading an introductory article on 'Wikipedia' (http://en.wikipedia.org/), nowadays the standard online encyclopedia in the Internet. While reading the article you will be confronted with hyperlinks that lead to further web pages with additional information. These hyperlinks may be matching your target topic, or they may be only partly relevant or even irrelevant for your current information gathering process [1]. In all cases the hyperlinks will interrupt your current reading process and will call for additional decision processes, that is, you have to decide which links to follow or which links to ignore [2].

Cognitively, these decision processes induced by hyperlinks can be expected to increase load on executive functions (EFs) that are already loaded during normal reading [3]. EFs can be defined as attention-related top-down control processes that are necessary to accomplish complex cognitive tasks that require adaptive behavior [4]. EFs are usually conceptualized to reside within the central-executive component of working memory [5,6]. Often differentiated core EFs are updating, shifting, and inhibition [7,8]. Although being differentiable, these core EFs have been shown to share a common underlying factor that has been attributed to processes of controlled attention [8]. The EF labeled 'updating' is generally defined to incorporate core processes of working memory (WM) functioning, namely the updating, monitoring, and manipulation of WM representations. The EF 'shifting' is defined by processes of shifting between multiple tasks, operations, or mental sets. The EF labeled 'inhibition' in a narrow sense refers to processes of response inhibition (e.g., in a Stroop task, cf. [8]), and more broadly defined refers to general processes of interference control or executive attention [9].

Text reading and comprehension require a number of lower level cognitive processes like letter decoding and word recognition as well as higher level cognitive processes like language and discourse processes, and domain general processes such as WM and EFs [10–12]. WM and EFs are especially required for the comprehension level of text reading [11,13]. According to the influential construction-integration model by Kintsch and colleagues [14,15], text comprehension consists of an iterative sequence of two steps that each refers to a specific level of mental representation. First, in the construction step, a mental model of the propositions of the text (i.e, the textbase) is generated. This textbase that has been constructed purely text-driven is then integrated into a situation model, that is, a mental model of what the text is about [16]. The situation model comprises information from the text as well as inferences made based on the text and prior knowledge [17]. During reading, the situation model has to be continuously updated to integrate the new information [14].

Generally, hyperlinks might affect both steps of the construction-integration cycle during reading. First, the construction step may be affected when hyperlinks interrupt the reading process and demand additional EFs like shifting and inhibition: Readers have to perform a task shift from purely reading to a decision on hyperlink selection. Additionally, inhibitory processes may be required to ignore irrelevant links and focus on relevant links. Second, when readers decide to follow a hyperlink the integration step might also be affected as the text of the following web page has to be integrated into the situation model [2,14]. Traditionally, hypertext research (for a review, see [2]), has studied the entire process of link selection and browsing through the subsequent web pages as a whole, without differentiating between load on EFs imposed by the hyperlink selection processes per se (i.e., affecting the construction step of text comprehension) and the load imposed on EFs through getting disoriented on the following web pages (i.e., affecting the integration step of text comprehension; [1,18–21]).

In the present research we were explicitly interested in load on EFs imposed by link selection processes per se, without inducing additional load due to the retrieval of subsequent hypertext pages and possible additional effects of disorientation. While the load effects of hyperlinks have been stated by several authors (e.g., [22,23]), to the best of our knowledge only one study by Fitzsimmons, Weal, and Drieghe so far has explicitly addressed the influence of links on text reading without the additional effects of load induced by displaying the following hypertext pages [24].

In this study (Experiment 2) by Fitzsimmons and colleagues [24] the authors compared reading of modified Wikipedia articles with texts that were either presented with hyperlinks (blue colored words) or without hyperlinks. Additionally the words chosen as hyperlinks were either high-frequent or low-frequent words. Participants were instructed to simply read through the texts. They had no possibility to actually select the hyperlinks (i.e., to click on the links). During reading, eye-tracking data were recorded. The authors did not find an overall effect of disruption in the reading process due to the hyperlinks. Only low frequent words marked as hyperlinks led to longer fixation durations and a rereading of previous paragraphs. This was interpreted by the authors as indication that hyperlinks may highlight important information for the reader which might be especially of relevance in case of uncommon, difficult concepts. However, as participants in this study did not have to perform selection processes (by clicking on the hyperlinks), the potential influence of actual link selection processes on reading due to a possible increase in load on EFs remained unclear.

Thus, in the current study we wanted to extend this line of hypertext reading research by focusing on the initial selection processes of words marked as hyperlinks. Specifically, we were interested in whether words marked as hyperlinks from which readers had to select appropriate ones, would induce additional load on EFs during the reading process due to the selection processes required. For this purpose we used pupil dilation and EEG alpha frequency band power as measures of load on EFs during reading in a task paradigm that allowed us to record and analyze these physiological measures in a natural reading situation. As will be outlined in the following, both measures, pupil dilation and EEG alpha frequency band power, have been shown to be sensitive for load on EFs when used in highly controlled, low-level tasks, such as working memory or attention tasks. However, only few studies have examined these load-measures in more unconstrained, natural task situations.

Examining low-level tasks it has been shown that the eye pupil dilates if EFs are required, for example, due to demands on the EF 'updating' in an n-back task [25] or the EF 'inhibition' in a Stroop task [26]. However, the eye pupil not only dilates due to specific demands on EFs but also more generally due to load on the cognitive processing system [27,28] as well as due to increased effort [29], or even changes in emotional or motivational states [30]. Recently, a direct connection between pupil dilation and the activity of the locus coeruleus in the brain that is central to the noradrenergic system has been proposed, based on the results of fMRI outcomes [31]. This indicates a close connection between pupil dilation and general states of arousal. Thus, pupil dilation may be seen as a rather overall load measure, including aspects of effort, motivation, arousal, and emotion [30,32–36].

Depending on the environmental lighting condition, the size of the pupils varies between two to nine millimeters [32]. Changes in pupil diameter due to cognitive processing demands are rather small, normally less than one millimeter irrespective of the baseline pupil diameter. Nevertheless, this measure has been proven to be very reliable in low-level tasks [30].

Only few studies have used pupil dilation in rather complex task materials such as unconstrained, free (hypertext) reading situations [37,38]. For example, Di Stasi and colleagues [38] recorded participants' pupil diameter while they had to perform two shopping tasks on a commercial website, either a goal-oriented search task (find and buy a specific object) or an experience-oriented search task (freely browse through the websites and possibly buy objects of own choice). In both task conditions an initial two minutes free exploration of the website without buying objects served as baseline. The authors found that in both task conditions the eye pupils significantly dilated from baseline when participants started one of the tasks. However, in addition, subjective rating scores revealed that the goal-oriented search task was experienced as being more difficult as the browsing task. This subjective difference in difficulty between tasks was not reflected in pupillary results. To conclude, these findings indicate that the sensitivity of pupil dilation as a measure of load in complex task settings is limited. Therefore, part of the research question of the present research was to examine the general sensitivity of pupil dilation as a measure of load on EFs in online reading and hyperlink selection.

EEG alpha frequency band power reflects the strength of EEG oscillatory activity and has been traditionally defined as the frequency range between eight to 13 Hz [39]. For increased demands on cognitive processing, for example, due to load on the EF 'updating' in an n-back task [40] or due to load on the EF 'inhibition' in a Stroop task [41], oscillatory activity has been observed to desynchronize. This event-related desynchronization (ERD) due to cognitive processing load during task performance results in decreased alpha frequency band power as compared to the frequency band power of a baseline interval [42]. The alpha ERD is commonly most pronounced over parietal electrodes [43], but it can also show certain effects of lateralization depending on the task material used (e.g., a left-lateralization for linguistic task material [44]). A synchronization of oscillatory activity in the alpha frequency band range (i.e., an increase in alpha frequency band power) might either reflect the activity of a cortical idling network [45], or, as proposed more recently, the active suppression of those brain networks not required for task performance or of those that are potentially interfering with task-relevant ones [46,47]. The alpha ERD has been associated with processes of attention and semantic memory [48,49]) and generally seems to reflect cognitive processing demands when WM or EFs are demanded ([40,41]). Thus, alpha frequency band power might be particularly suited to capture cognitive processes associated with link selection as described above.

EEG measures have quite recently been started to be applied in the context of unconstraint ('real-world') text reading, mostly using eye fixation-related potentials [50–53]. We sought to extend this line of research by focusing on EEG alpha frequency band power and the use of hypertext material as task domain. Only few studies report EEG alpha frequency band power as measures of load in rather complex, free reading situations (e.g., [54,55]; see [56] for a review in the context of hypermedia research). For example, Antonenko and Niederhauser [54] used the EEG alpha frequency band power to assess the impact of leads (i.e., short previews or descriptions of succeeding websites) on cognitive load during hyperlink selection. As expected, they observed a decrease in alpha frequency band power during link selection (i.e., an alpha ERD). However, this alpha ERD was reduced for those hyperlinks that provided a preview of some initial sentences of the following hyperlinked page via mouse-over. This was interpreted by the authors to reflect reduced cognitive load during the link selection process when leads are given as compared to normal hyperlinks without leads. However, as the task conditions in this study (leads versus no-leads) also varied considerably with regard to perceptual differences, the observed effects might simply go back to perceptual confounds and not to different load-situations per se (for a comprehensive overview of this critical issue in most studies using complex task materials see [57]). By carefully controlling for perceptual confounds, our study was hypothesized to tap deeper into purely load-related effects of link selection processes during text reading.

To sum up, in the present study we sought to further address and extend two currently emerging research directions: (a) a more in-depth analysis of demands on EFs in online reading and hyperlink selection, and (b) the combined recording and analysis of EEG data and pupil dilation data in a complex reading task. To the best of our knowledge a direct comparison of pupil dilation data and EEG alpha frequency band power data as measures of load on EFs has not been conducted before, neither in highly constrained (e.g., working memory) task settings, nor in rather unconstrained, free reading task settings.

As we focused on the influence of initial link selection processes during reading (i.e., during the construction step of text comprehension as described above) without hampering reading comprehension due to following web pages (i.e., the integration step of reading comprehension), we created a rather artificial hypertext reading situation that consisted of one text presented on the screen with words marked as hyperlinks but without any hyperlink functionality (i.e., no further pages could be reached). Furthermore, we simulated the link selection processes by instructing the participants to only click on context-matching words that were designed as links (see Method section for a detailed description of the task material). Although we expected our research account to tap into comparable cognitive effects of link selection during reading that might occur in genuine hypertext reading situations, one should keep in mind that we used a rather artificial hypertext-reading situation with simulated link-selection processes (see the General Discussion section for addressing the validity of our research account).

We hypothesized to observe increased load on EFs during those link selection processes (test condition) in terms of increased pupil dilation (Hypothesis 1a) and decreased alpha frequency band power (Hypothesis 1b) when comparing this test condition to a baseline condition within the same text but consisting of pure text reading. Additionally, we expected that the change of pupil dilation and the change of alpha frequency band power between baseline and test condition would be correlated. As the pupil dilation was expected to increase for increased task demands and the alpha frequency band power was expected to decrease for increased task demands (i.e., an increased alpha ERD) we expected to observe a negative correlation between the two measures (i.e., a larger increase in pupil dilation should be accompanied by larger, yet negatively signed, alpha ERD; Hypothesis 2).

In the following sections we will describe three experiments that were run to address these research questions and to carefully rule out possible alternative explanations due to confounding factors. Experiment 1 most closely simulated a hypertext reading and link selection situation. Experiment 2 and Experiment 3 ruled out possible alternative explanations of the observed results due to perceptual (i.e., word color), motor (i.e., mouse-click), or structural (i.e., sentence difficulty) confounds.

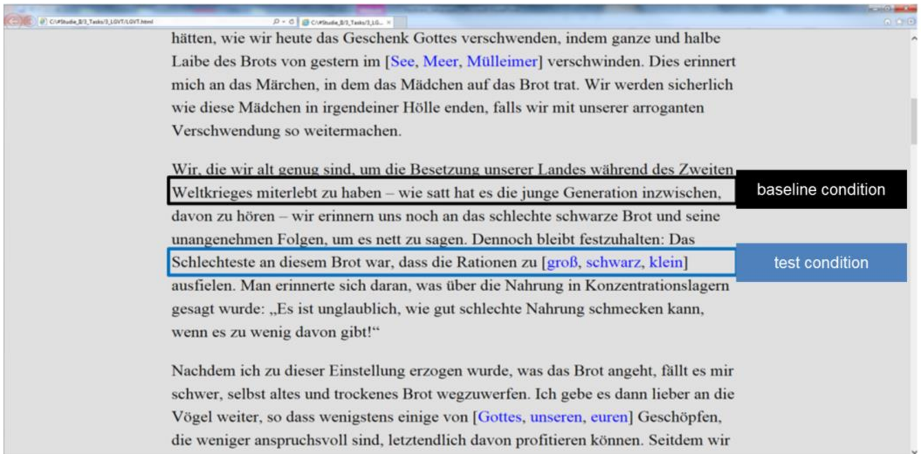
Experiment 1

In the present research we were interested in a specific aspect of load-induction due to the availability of hyperlinks, namely the load on EFs that is induced when text reading is interrupted by words marked as hyperlinks and selection processes have to take place. In Experiment 1 we used a methodology of combined recording and analysis of eye-tracking and EEG data. This methodology allowed us to directly compare the outcomes of two physiological load-measures, pupil dilation and EEG alpha frequency band power, during natural reading of one hypertext page. Two task conditions (baseline and test condition) were implemented within the text. In Experiment 1, parts of the text requiring pure text reading served as baseline condition. Parts of the text requiring additional hyperlink-like selection processes served as test condition.

Method

Participants. Twenty-three university students (mean age = 24.83 years, *SD* = 3.20, 13 females) participated in the study and received a payment of 8 €/h. They were all native speakers of German, right-handed, and reported no neurological disorders. All participants had normal or corrected-to-normal visual acuity. The study was approved by the local ethic committee of the Knowledge Media Research Center Tuebingen. Participants gave their written informed consent at the beginning of the study. None of the participants was familiar beforehand with the task materials we used.

Materials and procedure. Task materials consisted of a text taken out of a German reading-comprehension task (the LGVT, [58]). The text was of standardized difficulty, suitable for testing reading abilities of German high-school students and thus neither over- nor undertaxing a university student sample. The text was 1727 words long. It was presented as one hypertext page in a normal web browser (Microsoft Internet Explorer). Font size was 35 points; spacing was set to 24 points. Font type was Times New Roman, font color black, and background color light gray (see Fig 1 for an exemplary part of the stimulus material used in Experiment 1).



hätten, wie wir heute das Geschenk Gottes verschwenden, indem ganze und halbe Laibe des Brots von gestern im [See, Meer, Mülleimer] verschwinden. Dies erinnert mich an das Märchen, in dem das Mädchen auf das Brot trat. Wir werden sicherlich wie diese Mädchen in irgendeiner Hölle enden, falls wir mit unserer arroganten Verschwendung so weitermachen.

Wir, die wir alt genug sind, um die Besetzung unserer Landes während des Zweiten Weltkrieges miterlebt zu haben – wie satt hat es die junge Generation inzwischen, davon zu hören – wir erinnern uns noch an das schlechte schwarze Brot und seine unangenehmen Folgen, um es nett zu sagen. Dennoch bleibt festzuhalten: Das Schlechteste an diesem Brot war, dass die Rationen zu [groß, schwarz, klein] ausfielen. Man erinnerte sich daran, was über die Nahrung in Konzentrationslagern gesagt wurde: „Es ist unglaublich, wie gut schlechte Nahrung schmecken kann, wenn es zu wenig davon gibt!“

Nachdem ich zu dieser Einstellung erzogen wurde, was das Brot angeht, fällt es mir schwer, selbst altes und trockenes Brot wegzuwerfen. Ich gebe es dann lieber an die Vögel weiter, so dass wenigstens einige von [Gottes, unseren, euren] Geschöpfen, die weniger anspruchsvoll sind, letztendlich davon profitieren können. Seitdem wir

Fig. 1 Exemplary extract of the task materials used in Experiment 1.

Participants were instructed to read through the text quickly yet attentive. As in the standard LGVT task instruction, participants were informed that they would have only four minutes for text reading. We slightly deceived our participants about the actual time constraint: Although we announced a time constraint of four minutes, we let participants read through the whole text in their individual reading speed until they reached the final sentence. By announcing a hypothetical time constraint we wanted to ensure a concentrated, linear text reading process without losing data for the later analysis. A visual inspection of the eye tracking data afterwards confirmed the generally linear reading process. Only when areas of the text were reached that contained the links, participants interrupted shortly the linear reading pattern and some back and forth jumps to nearby words could be observed. A rather shallow text processing in the form of text skimming (e.g., [59]) due to the announced rather short time constraint of four minutes did not occur. This can also be inferred from the rather high reading comprehension scores obtained (see Tables 1 and 2).

At each of 23 different positions in the text three words were set in brackets. One out of the three words matched the context, the two others were distractors. For example, the exemplary text part given in Fig 1 describes the bad food supply situation of the population during World War II, and mentions a certain kind of bad black bread. The sentence in the test condition states: "The worst thing about this bread was that the food rations were too [big, black, small]". Participants had to click on the context-matching word (i.e., "small" in this case). Thus, they had to perform decision and selection processes that we hypothesized to be comparable to initial hyperlink-selection processes. The words in brackets were blue colored to simulate hyperlinks. To provide a visual feedback the blue color changed to red once a link was clicked on. There were no other effects of clicking. The 23 decision items of simulated hyperlinks were in 12 cases nouns, in 7 cases adjectives, in 2 cases verbs, and in 2 cases pronouns. The use of different word forms might resemble the distribution of actual hyperlinks for example in a Wikipedia article where also not only nouns serve as hyperlinks. Furthermore, comparable to real hyperlinks, the 23 decision items required semantic processing. None of the correct items could be inferred by syntactical or morphological processing.

The text out of a German reading comprehension task we used may be regarded as a rather artificial kind of a hypertext page. However, we favored the use of a carefully designed text of standardized difficulty like the LGVT text as task material for two reasons: Firstly, the controlled text ensured that participants' reading comprehension was challenged but not over-taxed, that is, it ensured that the participants were cognitively neither over- nor under-loaded during reading (which otherwise might have confounded our physiological measures). Secondly, the LGVT text provided us with two behavioral measures, reading speed and reading comprehension scores, which we could use to check whether participants had attentively read the text.

The experiment started after the EEG preparation and the calibration of the eye-tracker. Written task instructions were presented as the first page on the screen. Participants reached the LGVT page via a hyperlink at the end of the task instructions. The total duration of the experiment including the technical preparation procedures was about one hour.

Apparatus. The experiment was run in a quiet room that was dimly lit. Participants sat in a comfortable chair in front of a 22-inch Dell monitor (1680x1050 pixels screen resolution) while their EEG and eye-tracking data were recorded. Eye-tracking data were recorded using a 250 Hz SMI (SensoMotoric Instruments) infrared remote eye-tracking system that was positioned below the monitor. A chin rest was used to avoid head movements during data recording and to guarantee a fixed distance of about 70 cm between the eyes and the eye-tracking device. The eye-tracking data were recorded at a sampling rate of 250 Hz (SMI iView X 2.7.13). The eye-tracker was calibrated using the built-in calibration routines (SMI Experiment Center, 9-point calibration) before the first page (written task instructions) appeared on the screen. EEG data were recorded from 27 electrode sites (Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, O2) positioned according to the international 10/20 system [60]. The right mastoid served as reference during recording. Ground electrode was positioned at AFz. Three additional electrodes were placed around the eyes for recording of the electro-occulogram (EOG). EEG data were recorded (PyCorder 1.0.2) at 500 Hz sampling rate (ActiCHamp, Brainproducts, Inc.) using active electrodes (ActiCap, Brainproducts, Inc.). Impedances were kept below 5 kOhm.

Data preprocessing and analysis. During preprocessing, before synchronization with the EEG data, the eye-tracking data were upsampled to 500 Hz to match the sampling rate of the EEG data. Eye-tracking data and EEG data were preprocessed and synchronized using customized Matlab scripts (Matlab 2012b, MathWorks, Inc.; EEGLAB v. 11.0.5.4b, [61], with EYE-EEG plugin, [51]). Eye blink artifacts (missing data points) in the eye-tracking data were corrected using linear interpolation. The continuous EEG data were filtered (low-pass 40 Hz, high-pass 0.5 Hz, linear finite impulse response filters). EOG artifacts were corrected using independent component analysis (ICA) decompositions. Independent components (ICs) identified as EOG-ICs by visual inspection were rejected. EEG data were re-referenced to average reference.

The combined continuous EEG and eye-tracking data were split in epochs of two seconds length for the two task conditions of interest, baseline and test condition. Only first visits of these epochs were considered for analyses. However, a visual inspection of the eye-tracking data showed a rather linear text reading process with no severe view jumps (e.g., regressions to previously read sentence lines). For the baseline condition (pure text reading) the epochs were defined through areas of interest (AOIs) positioned around text lines that lay maximally in between parts of the text where link selection took place. On average 2.9 text lines (range: 1 to 7) were between the test and the baseline conditions. Post-hoc visual inspection of the eye-tracking data ensured that even in cases where the two conditions were separated by only one line, the two conditions were not confounded by each other. For the test condition (link selection) the epochs were defined as ending 500 ms before the mouse click to avoid motor artifacts in the EEG data [62] and to minimize the motor differences between baseline and test condition. In total, 46 data epochs were created, 23 epochs for each task condition. An automatic artifact removal was performed with respect to the EEG data: Epochs that exceeded ±100 µV were excluded from further analyses [63]. By using this criterion, epochs containing severe artifacts (e.g., muscle artifacts) were excluded. No further artifact removal or correction was performed on the EEG data.

For the eye-tracking data, the mean pupil size was calculated by averaging the left and right pupil data. These mean pupil data were further averaged for each of the 2 s epochs and then averaged over all epochs for each task condition. For the EEG data, EEG frequency band power was calculated using fast-fourier transforms (FFTs) for the entire epoch lengths in the alpha frequency band spectrum (8 Hz to 13 Hz). The alpha frequency band power was then averaged individually over all epochs for each task condition.

Results

Behavioral data. Reading comprehension scores were calculated according to the LGVT manual: For the initial four minutes of text reading each correctly selected word counted as +2 points, each wrongly selected word as -1 point. Points were summed up for each participant. Total reading time was defined as the entire time participants read through the LGVT page until they reached the final sentence. Reading comprehension score and total reading time are given in Table 1. These measures show that participants attentively read the text, yet not being cognitively overloaded (i.e., participants' reading comprehension scores are in the upper third of the common LGVT outcomes, a result quite typical for university students).

Physiological data. For each physiological measure (EEG alpha frequency band power at electrode Pz and pupil size), we computed separate one-factorial repeated measures ANOVAs (baseline condition vs. test condition). The results for the physiological variables are shown in Fig. 2. We restricted this analysis to the parietal electrode Pz as alpha frequency band power effects are generally reported to be most pronounced at parietal electrodes with

Pz as representative (e.g., [64,65]). To further explore topographical differences of the alpha frequency band power effects as indicated by the topoplot in Fig 3, an additional 3-way repeated-measures ANOVA was conducted with the factors hemisphere (left / right), electrode site (frontal / parietal) and task condition (baseline / test). For post-hoc pairwise comparisons all p-values were Bonferroni corrected for multiple comparisons. Level of significance was set at $\alpha = .05$ for all analyses and partial eta-square (ηp2) is reported as a measure of effect size.

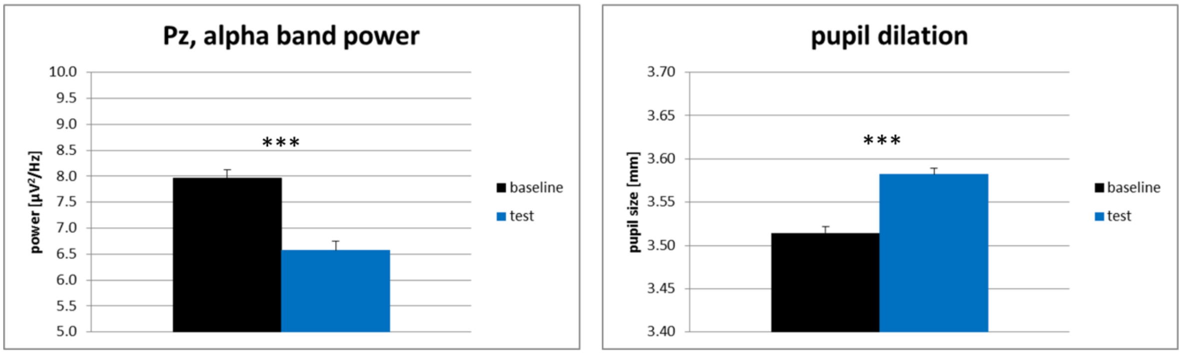


Fig. 2 Mean alpha (8 – 13 Hz) frequency band power at electrode Pz and mean pupil dilation of Experiment 1. Note. *** indicate $p < .001$, black error bars indicate +1 standard error of the mean.

For the pupil dilation data we found a main effect of task condition, $F(1, 22) = 21.85$, $p < .001$, $\eta_p^2 = .50$. In line with hypothesis 1a, pupil sizes in the test condition ($M = 3.58$ mm, $SD = 0.32$) were significantly larger as compared to pupil sizes in the baseline condition ($M = 3.51$ mm, $SD = 0.32$). We also found a main effect of task condition for the alpha frequency band power at electrode Pz, $F(1, 22) = 16.05$, $p = .001$, $\eta_p^2 = .42$. As expected by Hypothesis 1b, the alpha frequency band power in the test condition ($M = 6.58$ µV^2/Hz, $SD = 2.16$) was significantly lower as in the baseline condition ($M = 7.96$ µV^2/Hz, $SD = 2.84$).

The 3-way repeated-measures ANOVA strengthened these results (cf. the topoplot in Fig 3 showing the percent change in alpha frequency band power between baseline and test condition for all electrodes plotted over the scalp). We observed a main effect of hemisphere, $F(1, 22) = 11.32$, $p = .003$, $\eta_p^2 = .34$, with electrodes over the left hemisphere generally showing lower alpha frequency band power values (left: 7.94 µV^2/Hz vs. right: 8.40 µV^2/Hz), and a main effect of task condition, $F(1, 22) = 10.13$, $p = .004$, $\eta_p^2 = .32$, with the test condition showing significantly lower alpha frequency band power values as compared to the baseline condition (test: 7.82 $\mu V^2/Hz$ vs. baseline: 8.40 $\mu V^2/Hz$).

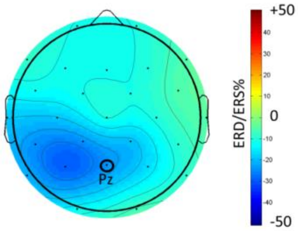


Fig. 3 Topoplot of the percentual changes of alpha (8 – 13 Hz) frequency band power between baseline condition and test condition for Experiment 1. Percentual frequency band power changes (i.e., the event-related desynchronization/synchronization, ERD/ERS%) were calculated after the formula given in [42].

However, these main effects were qualified by a significant interaction between task condition and hemisphere, $F(1, 22) = 20.61$, $p < .001$, $\eta_p^2 = .48$. This interaction was due to the fact that the alpha frequency band power decreased more strongly in the left than in the right hemisphere between baseline and test condition (left: -.88 $\mu V^2/Hz$, $p = .001$, right: -.52 $\mu V^2/Hz$, $p = .028$) and particularly due to the fact that only in the test condition alpha frequency band power between left and right hemisphere differed (test condition: -.64 $\mu V^2/Hz$, $p < .001$, baseline condition: -.29 $\mu V^2/Hz$, $p = .058$). These findings are in line with other studies using linguistic task material that observed left lateralized effects, as will be discussed in the general discussion section of this paper.

Additionally we observed a significant interaction between task condition and electrode site, $F(1, 22) = 15.41$, $p = .001$, $\eta_p^2 = .41$. This interaction resulted from a larger decrease of alpha frequency band power due to test condition at parietal-occipital electrode sites as compared to baseline condition (parietal-occipital: -.91 $\mu V^2/Hz$, $p < .001$, frontal: -.50 $\mu V^2/Hz$, $p = .047$). These findings are in line with typically observed largest effect sizes of alpha frequency band power changes at parietal-occipital electrodes (e.g., [43]).

Correlational data. As we were further interested in whether changes in pupil dilation and changes in EEG alpha frequency band power were measures of comparable sensitivity for changes in load on EFs due to the selection processes in the test condition as compared to the baseline condition, we calculated the mean differences between test and baseline condition (i.e., test condition subtracted from baseline condition) for the mean pupil sizes and the mean alpha frequency band power at electrode Pz (cf., Table 1). We calculated Pearson's correlation coefficients (two-tailed) for these two physiological measures (difference values), as well as the two behavioral measures, that is, the LGVT reading comprehension score and the total reading time. Results of the correlational analyses are given in Table 1. Contrary to Hypothesis 2, we observed no significant correlation between the amount of change in alpha frequency band power and pupil dilation. Interestingly, however, there was a marginally significant negative correlation between participants' LGVT reading comprehension score and the difference value of alpha frequency band power between baseline and test condition. Participants with a higher LGVT reading comprehension score showed a more pronounced decrease in alpha frequency band power (i.e., more negative difference values). Finally, in accordance with the typical outcomes of the LGVT, we additionally found a strong negative correlation between the LGVT reading comprehension scores and the total reading times of the text. Participants who showed higher reading comprehension scores were also faster in reading.

Table 1 Pearsons' correlations coefficients (two-tailed) for the physiological measures (difference in alpha power [in µV^2/Hz] and pupil size [in mm] between baseline and test condition) and the LGVT reading comprehension score and the total reading time [in minutes] of Experiment 1.

Variable	1	2	3	M	SD	N
1. ∆ alpha power				-1.38	1.65	23
2. ∆ pupil size	-.16			0.07	0.07	23
3. reading comprehension	.-41^{+}	.16		21.17	5.62	23
4. total reading time	.16	-.11	-.80***	8.43	1.69	23

Note.$^{+}$ p < .10, * p < .05, *** p < .001.

Discussion

To sum up, in Experiment 1 both physiological measures showed the expected outcomes: When comparing parts of the text that required purely reading (baseline condition) with parts of the text where participants had to interrupt reading and to perform hyperlink-like selection processes (test condition), we observed a significant increase in pupil size as well as a significant decrease of alpha frequency band power. These results are in accordance with our above stated Hypotheses 1a and 1b that the hyperlink-like selection processes in the test condition imposed additional load on EFs as compared to the pure reading situation in the baseline condition.

However, there are two possible alternative explanations that may account for the observed results. As will be outlined in the following the results might go back to motor or perceptual differences between baseline and test condition and not to cognitive processes per se. First, participants' motor activity differed between baseline an test condition. In the baseline condition, participants were reading only whereas in the test condition they had to perform a mouse-click on a word. Although we used a data epoch for data analysis that ended 500 ms before the mouse-click, we cannot fully rule out the possibility that differences in motor activity (or motor preparation processes) have confounded our results and thus serve as an alternative explanation. Second, in the baseline condition all words were black colored whereas in the test condition the words that could be selected were blue-colored and changed their color to red after participants clicked on them. These perceptual differences could also be responsible for the changes in pupil dilation. To rule out these possible effects of the perceptual and motor confounds in Experiment 1, we conducted a second experiment with slightly modified task materials.

Moreover, contrary to Hypothesis 2, we observed no significant correlation between the amount of change in alpha frequency band power and pupil dilation. This was unexpected as separate ANOVAs of pupil dilation data and EEG alpha frequency band power data both showed the expected outcomes, when comparing the two task conditions (baseline and test condition). Therefore we would have expected to find a clear negative correlation between the difference values of the two physiological measures, as the ANOVAs had shown that the pupil size increased significantly from baseline to test condition whereas the alpha frequency band power decreased significantly from baseline to test condition. One potential reason for this unexpected outcome might be the potential perceptual and motor confounds mentioned above, which might have affected the physiological variables differently. For example, one could hypothesize that the different color of the words in the test condition could affect pupil sizes but not alpha frequency band power, thus masking the expected correlation between these two measures.

Finally, we found a trend for a negative correlation between the LGVT reading comprehension score and the difference value of alpha frequency band power between baseline and test condition. Participants with a higher LGVT reading comprehension score showed a more pronounced decrease in alpha frequency band power (i.e., more negative difference values). This might underline the character of alpha frequency band power as a valid measure of essential cognitive processes. If participants were cognitively more engaged, they showed a higher reading comprehension score. In contrast, no such correlation was shown for pupil dilation.

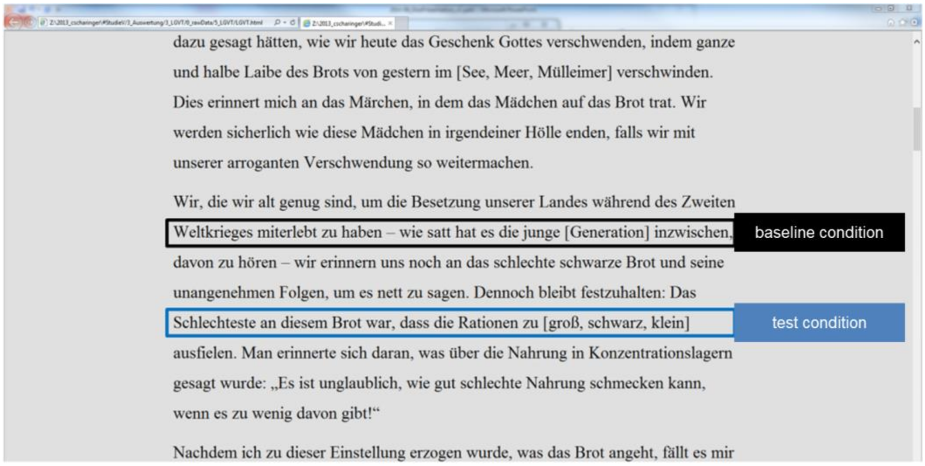
Experiment 2

In Experiment 2 we modified the baseline and test condition of Experiment 1 to avoid possible perceptual-motor differences between them that might have confounded our results in Experiment 1. Thus, in Experiment 2 participants had to perform a mouse-click on a word in the baseline condition in a comparable manner as in the test condition. Furthermore, we changed the color of the words in the test condition to black and removed the color-change to red of words that have been clicked on.

Method

Participants. Twenty additional university students (mean age = 24.90, SD = 3.23, 9 females) participated in Experiment 2. The general subject pool was the same as in Experiment 1 and the same constraints, incentives, and formal procedures were applied. None of the participants had attended Experiment 1 or was familiar with the LGVT task. One participant had to be excluded from data analysis due to technical problems during data acquisition and partly missing data.

Materials and procedure. Task material and presentation was the same as described for Experiment 1 with the following modifications aiming at reducing any motor or perceptual differences between baseline condition and test condition as far as possible: We modified the baseline condition insofar that participants had to click on a single word set in brackets in the baseline condition (instead of purely text reading as in Experiment 1). Furthermore, to avoid any perceptual differences between baseline condition and test condition, the three words in brackets in the test condition were perceptually equalized to the entire text (i.e., black color, no color change when mouse-click was performed). In doing so, we assured in Experiment 2 that any observed difference between test condition and baseline condition should only be due to additional selection processes that were hypothesized to load more on EFs in the test condition as compared to the baseline condition. See Fig 4 for an exemplary part of the task material. As an additional minor modification, we slightly increased the spacing between sentence lines to 28.8 points. This allowed us to more easily define areas of interest that included single sentence lines.



dazu gesagt hätten, wie wir heute das Geschenk Gottes verschwenden, indem ganze und halbe Laibe des Brots von gestern im [See, Meer, Mülleimer] verschwinden. Dies erinnert mich an das Märchen, in dem das Mädchen auf das Brot trat. Wir werden sicherlich wie diese Mädchen in irgendeiner Hölle enden, falls wir mit unserer arroganten Verschwendung so weitermachen.

Wir, die wir alt genug sind, um die Besetzung unserer Landes während des Zweiten Weltkrieges miterlebt zu haben – wie satt hat es die junge [Generation] inzwischen, davon zu hören – wir erinnern uns noch an das schlechte schwarze Brot und seine unangenehmen Folgen, um es nett zu sagen. Dennoch bleibt festzuhalten: Das Schlechteste an diesem Brot war, dass die Rationen zu [groß, schwarz, klein] ausfielen. Man erinnerte sich daran, was über die Nahrung in Konzentrationslagern gesagt wurde: „Es ist unglaublich, wie gut schlechte Nahrung schmecken kann, wenn es zu wenig davon gibt!“

Nachdem ich zu dieser Einstellung erzogen wurde, was das Brot angeht, fällt es mir

Fig. 4 Exemplary extract of the task materials used in Experiment 2.

Data preprocessing and analysis. Data preprocessing and analysis steps were identical to Experiment 1 with the only difference that the epoch alignment in the baseline condition in Experiment 2 was also related to the (now available) mouse-click. This equals the epoch alignment of the test condition (i.e., the 2 s epochs used for data averaging and analysis ended 500 ms before the mouse-click in both task conditions) and minimizes possible non-cognitive differences between baseline and test condition.

Results

Behavioral data. The average LGVT reading comprehension score and the average total reading time are given in Table 2. These behavioral variables were comparable to

Experiment 1, again indicating an attentive yet non-overloaded reading process of the participants.

Physiological data. As in Experiment 1, we found a main effect of task condition for the pupil dilation data, $F(1, 18) = 32.61$, $p < .001$, $\eta_p^2 = .64$. Pupil sizes in the test condition ($M = 3.64$ mm, $SD = 0.36$) were significantly larger as compared to pupil sizes in the baseline condition ($M = 3.57$ mm, $SD = 0.34$), confirming Hypothesis 1a. We also found a main effect of task condition for the alpha frequency band power, $F(1, 18) = 15.00$, $p = .001$, $\eta_p^2 = .45$. As expected by Hypothesis 1b, the alpha frequency band power in the test condition ($M = 8.05$ µV^2/Hz, $SD = 2.03$) was significantly lower as in the baseline condition ($M = 8.87$ µV^2/Hz, $SD = 9.03$). The outcomes of these measures can be seen in Fig 5.

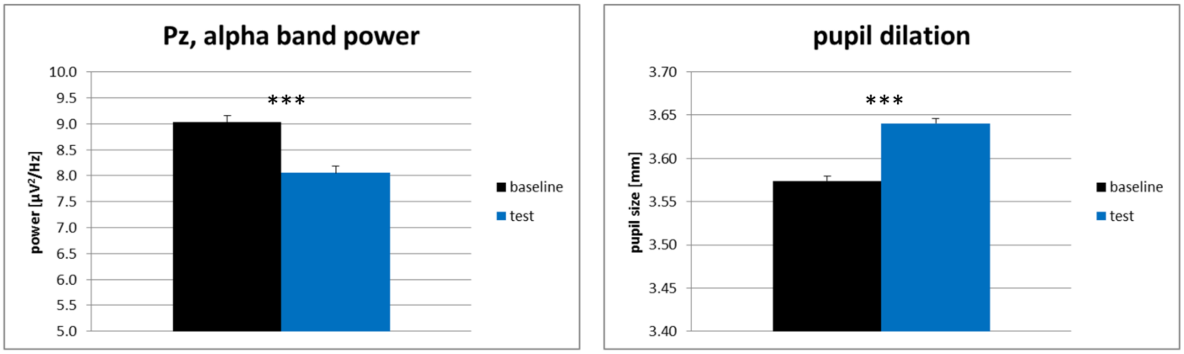


Fig. 5 Mean alpha (8 – 13 Hz) frequency band power at electrode Pz and mean pupil dilation of Experiment 2. Note. *** indicate p < .001, black error bars indicate +1 standard error of the mean.

The 3-way repeated-measures ANOVA revealed a main effect of hemisphere, $F(1, 18) = 8.55$, $p = .009$, $\eta_p^2 = .32$, and a main effect of electrode site, $F(1, 18) = 95.68$, $p < .001$, $\eta_p^2 = .84$. However there was also a significant interaction between hemisphere and electrode site, $F(1, 18) = 20.99$, $p < .001$, $\eta_p^2 = .54$. This interaction was due to the fact that only for parietal-occipital electrode sites alpha frequency band power was in general significantly smaller in the left hemisphere than in the right hemisphere (frontal-left: 7.08 µV^2/Hz vs. frontal-right: 6.89 µV^2/Hz, $p = .09$; parietal-left: 8.42 µV^2/Hz vs. parietal-right: 9.23 µV^2/Hz, $p < .001$). This hemispheric difference did not occur at frontal electrode sites. The significant main effect of electrode site and the interaction between hemisphere and electrode site were not present in Experiment 1. We do not have a concise explanation of this difference between

Experiment 1 and Experiment 2. However, and more importantly, comparable to Experiment 1 we found an interaction between task condition and hemisphere, $F(1, 18) = 8.60$, $p = .009$, $\eta_p^2 = .32$. Although all post-hoc pairwise comparisons were significant ($p < .001$), numerically, alpha frequency band changes for test condition in comparison with baseline condition were larger in the left than in the right hemisphere (left: -.82 $\mu V^2/Hz$, $p < .001$, right: -.65 $\mu V^2/Hz$, $p = .001$), which mirrors the results of Experiment 1. Additionally hemispheric differences were larger in the test condition as compared to the baseline condition (test: -.40 $\mu V^2/Hz$, $p = .006$, baseline: -.22 $\mu V^2/Hz$, $p = .024$). This again indicates, in line with the literature, that the strongest alpha frequency band effects occur at parietal-occipital electrodes (e.g., Gevins et al., 1997) and that left-lateralized effects may be typical for linguistic task material. Fig 6 visualizes the topographic distribution of the alpha frequency band change between baseline and test condition on the scalp.

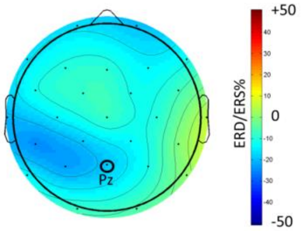


Fig. 6 Topoplot of the percentual changes of alpha (8 – 13 Hz) frequency band power between baseline condition and test condition for Experiment 2. Percentual frequency band power changes (i.e., the event-related desynchronization/synchronization, ERD/ERS%) were calculated after the formula given in [42].

Correlational data. In line with Experiment 1 we found no significant correlations between the difference values of both physiological measures (i.e., the difference value of mean pupil size between baseline and test condition and the difference value of mean alpha frequency band power between baseline and test condition, see Table 2). Also in line with the results of Experiment 1, we found a significant negative correlation between the LGVT reading comprehension score and the difference value of the alpha frequency band power between baseline and test condition. Furthermore, there was a trend for a positive correlation between total reading time and the alpha power difference value. Finally, the observed negative correlation between LGVT reading comprehension scores and total reading times also resemble the results of Experiment 1 (cf. Tables 1 and 2).

Table 2 Pearsons' correlations coefficients (two-tailed) for the physiological measures (difference in alpha power [in μV^2/Hz] and pupil size [in mm] between baseline and test condition) and the LGVT reading comprehension score and the total reading time [in minutes] of Experiment 2.

Variable	1	2	3	M	SD	N
1. Δ alpha power				-0.98	1.10	19
2. Δ pupil size	-.14			0.07	0.05	19
3. reading comprehension	-.50*	-.23		19.84	5.09	19
4. total reading time	.43$^+$	.08	-.84***	8.52	1.90	19

Note.$^+$ p < .10, * p < .05, *** p < .001.

Discussion

To sum up, the outcomes of the physiological measures of Experiment 2 were comparable to the results of Experiment 1. The results strongly support the hypothesis that selection processes in the test condition led to additional load on EFs. This increased load was measurable by increased pupil dilation data (Hypothesis 1a) as well as decreased EEG alpha frequency band power (Hypothesis 1b). As we carefully controlled the task material in Experiment 2 for any motor or perceptual confounds, we can rule out the alternative explanation of our results raised in Experiment 1 based on possible perceptual-motor confounds. However, apart from perceptual-motor confounds a third alternative explanation might be formulated: The parts of the text included in the test condition might per se have happened to be more difficult as compared to the parts of the text included in the baseline condition. To also rule out this alternative explanation, we conducted Experiment 3.

Furthermore, in line with Experiment 1 we found no significant correlations between the difference values of the two physiological measures (i.e., the difference value of mean pupil size between baseline and test condition and the difference value of mean alpha frequency band power between baseline and test condition). As we carefully controlled baseline and test condition for perceptual-motor confounds in Experiment 2, we may rule out any explanation of this non-correlation due to perceptual-motor influences that affected one measure but not the other. We will discuss this unexpected outcome in the general discussion section of this paper.

Finally, also in line with the results of Experiment 1, where we observed a trend for a negative correlation between the LGVT reading comprehension score and the difference value of the alpha frequency band power between baseline and test condition, in Experiment 2 we found a significant negative correlation between these two variables. In Experiment 2 this negative correlation was furthermore accompanied by a trend for a positive correlation between total reading time and the alpha power difference value. As hypothesized above, these results may be interpreted to underline the character of alpha frequency band power as reflecting essential cognitive processes. The more successfully participants performed the task (as indicated by higher LGVT reading comprehension scores and lower total reading times) the more pronounced was the difference of the oscillatory alpha frequency band activity between baseline and test condition (i.e., the more negative were the difference values between baseline and test condition).

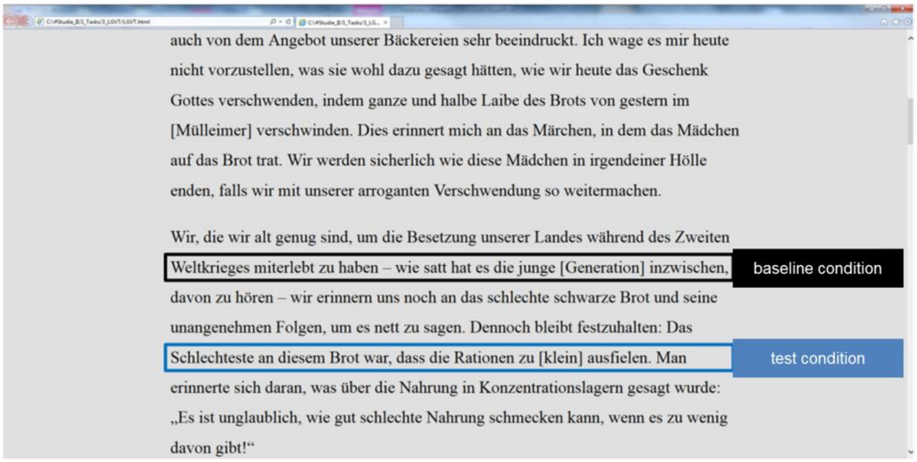
Experiment 3

Experiment 3 was conducted to exclude potential differences between baseline and test condition due to differences in difficulty of the text parts used in these two conditions. Therefore, we modified the test condition of Experiment 2 to resemble the task of the baseline condition (i.e., we eliminated the word selection processes): In the test condition (as well as in the baseline condition) participants had to click on one single word in brackets (cf. Fig 7 for an exemplary part of the task material used in Experiment 3). We expected the textual difficulty in both conditions to be equal, that is, we expected to observe no differences between baseline and test condition in Experiment 3. As pupil dilation data and EEG alpha frequency band power data had been showing similar load-related effects, for reasons of time-efficiency we recorded and analyzed only pupil dilation data in Experiment 3.

Method

Participants. Twenty-four additional university students (mean age = 23.38, *SD* = 2.55, 16 females) participated in Experiment 3. The general subject pool was the same as in Experiment 1 and Experiment 2 and the same constraints, incentives, and formal procedures were applied. None of the participants had attended Experiment 1 or Experiment 2 or was familiar with the LGVT task.

Materials and procedure. Task material and presentation was the same as described for Experiment 2 with the following modification: Instead of presenting three words in brackets in the test condition only the one context-matching word was presented. Thus the task in the test condition was identical to the task in the baseline condition (see Fig. 7 for an exemplary extract of the task materials used). In both conditions participants had to simply click with the mouse cursor on the word in brackets. The baseline condition was identical to Experiment 2.


auch von dem Angebot unserer Bäckereien sehr beeindruckt. Ich wage es mir heute nicht vorzustellen, was sie wohl dazu gesagt hätten, wie wir heute das Geschenk Gottes verschwenden, indem ganze und halbe Laibe des Brots von gestern im [Mülleimer] verschwinden. Dies erinnert mich an das Märchen, in dem das Mädchen auf das Brot trat. Wir werden sicherlich wie diese Mädchen in irgendeiner Hölle enden, falls wir mit unserer arroganten Verschwendung so weitermachen.

Wir, die wir alt genug sind, um die Besetzung unserer Landes während des Zweiten Weltkrieges miterlebt zu haben – wie satt hat es die junge [Generation] inzwischen, davon zu hören – wir erinnern uns noch an das schlechte schwarze Brot und seine unangenehmen Folgen, um es nett zu sagen. Dennoch bleibt festzuhalten: Das Schlechteste an diesem Brot war, dass die Rationen zu [klein] ausfielen. Man erinnerte sich daran, was über die Nahrung in Konzentrationslagern gesagt wurde: „Es ist unglaublich, wie gut schlechte Nahrung schmecken kann, wenn es zu wenig davon gibt!“

Fig. 7 Exemplary extract of the task materials used in Experiment 3.

Data preprocessing and analysis. Data preprocessing and data analysis steps were identical to Experiment 2 for the eye-tracking data. Data epochs of 2 seconds length ending 500 ms before the mouse-clicks in both task conditions were used for data analysis.

Results and Discussion

Behavioral data. Participants had a total mean reading time of about 7.24 minutes (*SD* = 1.67) which was slightly shorter than in Experiment 1 and Experiment 2 (cf. Table 1 and 2). This result is not surprising due to the fact that participants' task in Experiment 3 was less complex as in the two previous experiments. Still, the overall reading time was not severely different from the two other experiments indicating that the participants were thoroughly reading the text, despite the 'easy' task.

Physiological data. We conducted a one-factorial repeated-measure ANOVA for the pupil size data of the two task conditions (baseline versus test condition). As expected we found no significant difference between pupil sizes in the baseline condition (*M* = 3.66 mm, *SD* = 0.35) and the test condition (*M* = 3.67 mm, *SD* = 0.36), *F*(1, 23) = .04, *p* = .85. In line with our expectations, this indicates that the text was of comparable difficulty in the baseline and in the test condition. The results for the pupil dilation are shown in Fig. 8. To conclude, these results can rule out that the observed findings of Experiment 1 and Experiment 2 were based on potentially confounding perceptual (i.e., word color), motor (i.e., mouse-click) or structural (i.e., sentence difficulty) differences between baseline and test condition.

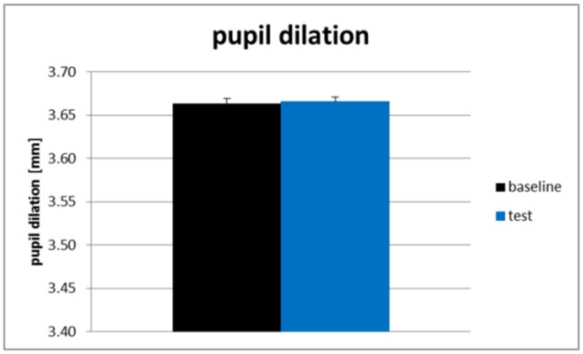


Fig. 8 Mean pupil dilation of Experiment 3. Note. black error bars indicate +1 standard error of the mean.

General discussion

In the present research we aimed at examining increased load on EFs during hyperlink-like selection processes in online text reading. In particular, we were interested in effects of initial link-selection processes without inducing additional load due to subsequent hypertext pages. We used a methodology of combined EEG and eye-tracking data recording and analysis that allowed us to compare two physiological load-measures, namely pupil dilation and EEG alpha frequency band power in the online hypertext-like reading situation.

As expected, both physiological measures were sensitive to increased load on EFs during hyperlink-like selection processes, confirming Hypotheses 1a and 1b. When comparing baseline and test condition, both in Epxeriment 1 and in Experiment 2 we found significantly increased pupil dilation as well as decreased alpha frequency band power for the test condition. These results are in line with studies that manipulated load on EFs in highly controlled working memory or attention tasks and that found increased pupil dilation and decreased alpha frequency band power for increased load conditions (cf., [26,43]). Thus, our study indicates that both measures, pupil dilation and EEG alpha frequency band power, seem to be sufficiently sensitive to also detect changes in load on EFs in an online reading situation.

Experiment 1 most closely simulated a hypertext-reading and hyperlink-selection situation. Participants read a text (baseline condition) and at different parts of the text they had to perform hyperlink-like selection processes (test condition). Experiment 2 and Experiment 3 ruled out possible alternative explanations of the observed results that were based on potentially confounding perceptual (i.e., word color), motor (i.e., mouse-click), or structural (i.e., sentence difficulty) differences between baseline and test condition. Thus, we can conclude that the observed differences we consistently found between baseline and test condition in Experiment 1 and Experiment 2 indeed reflect changes in the load-situation that may be related to increased demands on EFs. These findings extend traditional hypertext research that in general does not distinguish between the load imposed by hyperlink-selection processes per se (as examined in our study) and the load imposed by the additional information of subsequent web pages or the potentially experienced disorientation due to non-linear navigation processes between different pages in a hypertext [2]. Thus, our results can be seen as an empirical validation of the often implicitly made assumption in hypertext research that the presence of links per se might increase the load-situation and therefore might hamper text reading and comprehension (e.g., see [2]).

As the links we used did not lead to any subsequent text page, in our study mainly the construction step of text comprehension might have been affected (i.e., the creation of a textbase). The updating of a situation model of a text (i.e., the integration step of reading comprehension) in contrast might especially be influenced by additional information of subsequent web pages [2,14]. However, although this reasoning might sound plausible, we did not test for effects on reading comprehension by using text comprehension questions afterwards that might have allowed disentangling text comprehension with respect to propositions or inferences (i.e., textbase or situation model construction). Clearly, in future studies the influence of link-selection processes on reading comprehension should be assessed as well.

Some further limitations of the present research have to be addressed. First, as described in the introduction, conceptually it may be justified to assume that link selection during text reading would increase load on EFs. The EF shifting may be loaded because of the task-set shift from purely reading to selection and decision processes when links are encountered. In addition, the EF inhibition may be required when irrelevant links have to be ignored. However, we did not specifically manipulate load on EFs during link selection. In future research such specific manipulation of EFs during link selection might, for example, be done by presenting links of different relevance for a certain information gathering goal. Depending on the relevance of the links, the EF inhibition might be differently loaded, as for example clearly irrelevant links might be easier to ignore than links that are of mixed relevance (see [1] as an exemplary study that manipulated the relevance of links). The research methodology of combined EEG and eye-tracking data analysis might be valuably used in future studies implementing such a fine-grained manipulation of EFs. Yet, our research can serve as an initial step to advance hypertext research by showing increased cognitive demands during initial link-selection processes that might conceptually be linked to load on EFs and that can be assessed by the physiological measures pupil dilation and EEG alpha frequency band power.

Another limitation of the current research to be addressed is the ecological validity of our research paradigm that might be questioned. This is first because we used non-functional links (i.e., links that did not lead to a subsequent web page), second because we presented each time three links in direct sequence, which might be rather uncommon for hypertext reading (but which may be the typical situation readers are confronted with in case of web-search results), and third because of the for hypertexts rather unnatural linear reading situation. With respect to the first two critical aspects, we are confident that the word-selection in our paradigm in principal leads to load-related effects that are comparable to those effects that are essential in genuine hyperlink selection, such as an increased load on EFs due to performing task shifts from purely reading to decision-making, the inhibition of irrelevant words, and the selection of relevant ones. A logical next step would be to use the methodology of a combined EEG and eye-tracking data recording and analysis in a more realistic hypertext reading and browsing situation. As mentioned above, this methodology might also be used to differentiate on a fine-grained level between load imposed by different kinds of hyperlinks, for example task-relevant versus task-irrelevant (but nevertheless interesting) hyperlinks [1,66]. Clearly, such follow-up studies that explicitly manipulate load on different EFs during link selection will be needed to directly test our assumption that link-selection raises demands on EFs.

With respect to the third critical aspect, the task paradigm we used afforded a rather linear text reading (which was confirmed by our visual inspection of participants' eye-tracking data) that might be different from genuine online hypertext reading. Depending on the hypertext-material used, different reading patterns have been observed for online text reading. One classical example are the F-shaped reading patterns that Nielsen and colleagues observed in eye-tracking studies of real hypertext pages (e.g., [67]). Duggan and colleagues [59,68] observed reading patterns of text skimming (i.e., selectively scanning of text parts) when the time to read an online text was limited. Such time-pressure may be the typical situation in online web reading, when a huge amount of web pages addressing a certain topic are available and have to be skimmed for relevance. Reader and Payne ([69], see also [59,68]) observed an online text reading strategy termed satisficing: Parts of the text (or different texts) are skimmed through until the relevant parts of the text (or texts) are reached. These are then read more thoroughly until the individual information gain is reached. Although in the current research we announced a time-limit for text reading, the task instruction afforded our participants to apply a linear text reading strategy. Therefore, one has to keep in mind that other text reading strategies might occur in online text reading when no such task instruction is given. This has to be taken into account in future studies that should use more realistic hypertext materials that additionally might provide more realistic hypertext reading situations with respect to reading strategies than the present research.

Nevertheless, given that an online text adequately meets the readers' information demands, a linear text reading like the one in our current research still might occur [69].

Despite these limitations, the current research showed the general sensitivity of pupil dilation and alpha frequency band power for an increased load-situation during link-selection. Surprisingly, however, the change in pupil dilation and the alpha frequency band power change did not correlate (Experiment 1 and Experiment 2). This outcome was rather unexpected given that the general load effect was observed for both measures. More precisely, as pupil dilation increased and alpha frequency band power decreased for increased load, we expected to observe a negative correlation between the two measures (cf., Hypotheses 2). At least two possible explanations for the observed non-correlation might be hypothesized.

First, the non-significant correlation between the two measures may indicate that they were sensitive to different aspects of load induction during hyperlink-like selection processes. Yet, we may only speculate which different aspects of load induction this may be. Pupil dilation may function as a more global load measure that also includes motivational or emotional aspects of load (e.g., [30,34–36]). This interpretation is corroborated by the non-existent correlation between reading comprehension scores and pupil dilation. In contrast, we observed a negative correlation between LGVT reading comprehension score and EEG alpha frequency band power change. This strengthens the assumption of alpha oscillatory activity being a cognitive correlate: For participants that showed higher reading comprehension scores, we observed a stronger decrease in alpha oscillatory activity. This is in line with current literature reporting a stronger alpha ERD (i.e., decrease in alpha band power) associated with higher semantic memory performance [49]. Generally, as discussed in the introduction, alpha ERD has been related to purely cognitive processes like working memory functioning, attention, and inhibitory control [46,48,49,70,71]. The interpretation of alpha frequency band power as a more sensitive measure of cognitive processes than pupil dilation may be corroborated further by the observation that in both experiments (Experiment 1 and 2) of the present research the alpha frequency band power effect was topographically maximal over left-hemispheric, parietal electrodes. This is in line with studies reporting left-lateralized alpha frequency band power effects for linguistic task material [44,72].

Noteworthy, however, based on our current results, we cannot infer that alpha frequency band power and pupil dilation might be sensitive for different EFs. Rather, the two measures might be sensitive for different aspects of load on EFs: Alpha frequency band power might be sensitive to purely cognitive aspects of load on EFs whereas pupil dilation might additionally be sensitive for emotional or motivational aspects of load on EFs [30,34,35].

Individual differences may be a second explanation for the non-existing correlation observed between the two physiological measures. In some participants pupil dilation might be the 'better' measure to detect a changed load-situation, whereas in other participants EEG alpha frequency band power might be the 'better' one. If this result generally turns out to be true, it has important consequences for the entire research area of 'neuro-ergonomics' [73–75], where researchers try to detect participants' load-situations using physiological measures for evaluating and adapting human-computer interfaces. For this research area our results may suggest to collect several different physiological measures and to select the most sensitive one individually later on. However, as mentioned above, these interpretations and conclusions are somewhat speculative at this point. Clearly, further research should be conducted with combined EEG and eye-tracking data recording in order to directly compare pupil dilation data and EEG frequency band data in diverse task settings and to study more closely their interlinked yet different nature. Although a few other studies have also recorded EEG data and pupil dilation data simultaneously using rather complex task materials, to the best of our knowledge correlations between EEG frequency band measures and pupil dilation measures have never been calculated or reported ([25,76–78]; but see some basic research studies mainly focusing on EEG event-related potentials for comparison with pupil dilation, e.g., [79–82], as referred to below).

To conclude, the present research may serve as an initial step with respect to two currently emerging research directions: (a) a more in-depth analysis of load on EFs in hypertext reading and hyperlink selection, and (b) the combined recording and analysis of EEG frequency band data and pupil dilation (even in 'real-world' tasks) and the exploration of different aspects of load they capture. With respect to (a), we again have to underline that we are well aware of the artificial hypertext situation in the current research as discussed above that may call for conducting additional studies. With respect to (b), we may also suggest further studies of combined EEG frequency band power analysis and pupil dilation. For this research direction using more controlled, low-level tasks may prove to be valuable. Indeed, there are some initial studies of combined EEG data and pupil dilation data analysis in basic research, reporting correlations between pupil diameter and certain EEG event-related potentials, like P300 and N400 [80–82] or between stimulus-evoked pupil dilation and EEG alpha activity [79]. These studies often use single-subject stepwise correlational analysis over the time course of trials, which were beyond the scope of the present paper. However, to the best of our knowledge a direct comparison of EEG frequency band power and pupil dilation data in more complex task situations has not been conducted up to now.

Thus, notwithstanding the limitations of the present work, our data showed for complex task materials and task situations simulating hypertext reading that both pupil dilation and EEG alpha band power overall are sensitive measures to assess the general load-situation during hyperlink-like selection processes, yet, alpha frequency band power might be the more specific measure for cognitive processes (i.e., EFs). The combined recording and analysis of eye-tracking and EEG frequency band data may be a promising methodological account to further study unconstrained, 'real-world' hypertext reading and link selection processes. Hence, this line of research may turn out to be highly relevant for the design and optimization of hypermedia learning environments. For example, future versions of Wikipedia articles may be optimized with regard to the kind and amount of hyperlinks so that additional load on readers' EFs is kept minimal, allowing to allocate more cognitive resources to processes of reading and comprehension. Especially as performance in EFs may be reduced in older adults [83] or impaired in certain populations (e.g., dyslexic or ADHD populations; [84,85]), assessment of load on EFs and the optimization of hypermedia environments with respect to amount and type of links therein causing this load might be valuable and necessary topics of future research. In this vein, individual differences research addressing, for example, the interaction between individuals' WM capacity or general executive function abilities and load during link selection processes might be carried out as well.

References

1. Salmerón L, Cerdán R, Naumann J. How adolescents navigate Wikipedia to answer questions / ¿Cómo navegan los adolescentes en Wikipedia para contestar preguntas? Infanc y Aprendiz. 2015; 1–37. doi:10.1080/02103702.2015.1016750

2. DeStefano D, LeFevre J-A. Cognitive load in hypertext reading: A review. Comput Human Behav. 2007;23: 1616–1641. doi:10.1016/j.chb.2005.08.012

3. Kendeou P, Van Den Broek P, Helder A, Karlsson J. A cognitive view of reading comprehension: Implications for reading difficulties. Learn Disabil Res Pract. 2014;29: 10–16. doi:10.1111/ldrp.12025

4. Diamond A. Executive functions. Annu Rev Psychol. 2013;64: 135–168. doi:10.1146/annurev-psych-113011-143750

5. Baddeley A. Exploring the central executive. Q J Exp Psychol. 1996;49: 5–28. doi:10.1080/713755608

6. Baddeley A. Working memory: looking back and looking forward. Nat Rev Neurosci. 2003;4: 829–839. doi:10.1038/nrn1201

7. Nee DE, Brown JW, Askren MK. A meta-analysis of executive components of working memory. Cereb Cortex. 2013;23: 264–282. doi:10.1093/cercor/bhs007

8. Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. Cogn Psychol. Netherlands: Elsevier Science; 2000;41: 49–100.

9. Engle RW. Working memory capacity as executive attention. Curr Dir Psychol Sci. 2002;11: 19–23. doi:10.1111/1467-8721.00160

10. McNamara DS, Magliano JP. Towards a comprehensive model of comprehension. In: Ross B, editor. The psychology of learning and motivation - advances in research and theory. New York, NY, USA: Elsevier Science; 2009. pp. 297–384. doi:10.1016/S0079-7421(09)51009-2

11. Arrington CN, Kulesz PA, Francis DJ, Fletcher JM, Barnes MA. The contribution of attentional control and working memory to reading comprehension and decoding. Sci Stud Read. 2014;18: 325–346. doi:10.1080/10888438.2014.902461

12. Ericsson KA, Kintsch W. Long-term working memory. Psychol Rev. 1995;102: 211–245. doi:10.1037/0033-295X.102.2.211

13. Sesma HW, Mahone EM, Levine T, Eason SH, Cutting LE. The contribution of executive skills to reading comprehension. Child Neuropsychol. 2009;15: 232–246. doi:10.1080/09297040802220029

14. Kintsch W. The role of knowledge in discourse comprehension: A construction-integration model. Psychol Rev. 1988;95: 163–182. doi:10.1037//0033-295X.95.2.163

15. Kintsch W. The role of knowledge in discourse comprehension: A construction-integration model. Adv Psychol. 1991;79: 107–153. doi:10.1016/S0166-4115(08)61551-4

16. Zwaan RA, Radvansky GA. Situation models in language comprehension and memory. Psychol Bull. 1998;123: 162–185. doi:10.1037/0033-2909.123.2.162

17. Verhoeven L, Perfetti C. Advances in text comprehension: Model, process and development. Appl Cogn Psychol. 2008;22: 293–301. doi:10.1002/acp

18. Amadieu F, van Gog T, Paas F, Tricot A, Mariné C. Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. Learn Instr. 2009;19: 376–386. doi:10.1016/j.learninstruc.2009.02.005

19. Cangoz B, Altun A. The effects of hypertext structure, presentation and instruction types on perceived disorientation and recall performances. Contemp Educ Technol. 2012;3: 81–98.

20. Kim H, Hirtle S. Spatial metaphors and disorientation in hypertext browsing. Behav Inf Technol. 1995;14: 239–250.

21. Scheiter K, Gerjets P. Learner control in hypermedia environments. Educ Psychol Rev. 2007;19: 285–307. doi:10.1007/s10648-007-9046-3

22. Zhu E. Hypermedia Interface Design: The Effects of Number of Links and Granularity of Nodes. J Educ Multimed Hypermedia. 1999;8: 331–358.

23. Carr NG. The shallows: What the Internet is doing to our brains. 1st ed. New York, NY: Norton; 2010.

24. Fitzsimmons G, Weal M, Drieghe D. On measuring the impact of hyperlinks on reading. Proceedings of the 5th Annual ACM Web Science Conference. New York, NY, USA: ACM Press; 2013. pp. 65–74. doi:10.1145/2464464.2464501

25. Ewing K, Fairclough S. The impact of working memory load on psychophysiological measures of mental effort and motivational disposition. In: de Waard D, Axelsson A, Berglund M, Peters B, Weikert C, editors. Human factors: A system view of human,technology and organisation. Maastricht: Shaker Publishing; 2010. pp. 423–433.

26. Laeng B, Ørbo M, Holmlund T, Miozzo M. Pupillary Stroop effects. Cogn Process. 2011;12: 13–21. doi:10.1007/s10339-010-0370-z

27. Kahneman D, Beatty J. Pupil diameter and load on memory. Science. 1966;154: 1583–1585.

28. Van Gerven PWM, Paas F, Van Merriënboer JJG, Schmidt HG. Memory load and the cognitive pupillary response in aging. Psychophysiology. 2004;41: 167–174. doi:10.1111/j.1469-8986.2003.00148.x

29. Piquado T, Isaacowitz D, Wingfield A. Pupillometry as a measure of cognitive effort in younger and older adults. Psychophysiology. 2010;47: 560–569. doi:10.1111/j.1469-8986.2009.00947.x

30. Beatty J, Lucero-Wagoner B. The pupillary system. In: Cacioppo JT, Tassinary LG, Berndtson G, editors. Handbook of psychophysiology. 2nd ed. Cambridge: Cambridge University Press; 2000. pp. 142–162.

31. Laeng B, Sirois S, Gredeback G. Pupillometry: A Window to the Preconscious? Perspect Psychol Sci. 2012;7: 18–27. doi:10.1177/1745691611427305

32. Andreassi JL. Pupillary response and behavior. In: Andreassi JL, editor. Psychophysiology: Human behavior and physiological response. 5th ed. New Jersey: Lawrence Erlbaum; 2007. pp. 289–307.

33. Kahneman D. Attention and effort. Series in experimental psychology. New Jersey: Prentice-Hall; 1973.

34. Granholm E, Steinhauer SR. Pupillometric measures of cognitive and emotional processes. Int J Psychophysiol. 2004;52: 1–6. doi:10.1016/j.ijpsycho.2003.12.001

35. Hess EH, Polt JM. Pupil size as related to interest value of visual stimuli. Science. 1960;132: 349–350.

36. Partala T, Surakka V. Pupil size variation as an indication of affective processing. Int J Hum Comput Stud. 2003;59: 185–198. doi:10.1016/S1071-5819(03)00017-X

37. Iqbal ST, Zheng XS, Bailey BP. Task-evoked pupillary response to mental workload in human-computer interaction. In: Dykstra-Erickson E, Tscheligi M, editors. CHI ’04 Extended Abstracts on Human Factors in Computing Systems. New York, NY: ACM; 2004. pp. 1477–1480. doi:10.1145/985921.986094

38. Di Stasi LL, Antolí A, Gea M, Cañas JJ. A neuroergonomic approach to evaluating mental workload in hypermedia interactions. Int J Ind Ergon. 2011;41: 298–304. doi:10.1016/j.ergon.2011.02.008

39. Krause CM. Brain electric oscillations and cognitive processes. In: Hugdahl K, editor. Neuropsychology and Cognition. Experimental methods in neuropsychology. 21st ed. Boston: Kluwer Academic Publishers Group; 2003. pp. 111–130.

40. Gevins A, Smith ME. Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. Cereb Cortex. 2000;10: 829–839. doi:10.1093/cercor/10.9.829

41. Hanslmayr S, Pastötter B, Bäuml K-H, Gruber S, Wimber M, Klimesch W. The electrophysiological dynamics of interference during the Stroop task. J Cogn Neurosci. 2008;20: 215–225. doi:10.1162/jocn.2008.20020

42. Pfurtscheller G, Lopes da Silva FH. Event-related EEG/MEG synchronization and desynchronization: Basic principles. Clin Neurophysiol. 1999;110: 1842–1857.

43. Gevins A, Smith ME, McEvoy CL, Yu D. High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. Cereb Cortex. 1997;7: 374–385.

44. Spironelli C, Angrilli A. Developmental aspects of language lateralization in delta, theta, alpha and beta EEG bands. Biol Psychol. 2010;85: 258–267. doi:10.1016/j.biopsycho.2010.07.011

45. Pfurtscheller G, Stancák A, Neuper C. Event-related synchronization (ERS) in the alpha band - an electrophysiological correlate of cortical idling: A review. Int J Psychophysiol. 1996;24: 39–46.

46. Klimesch W. Alpha-band oscillations, attention, and controlled access to stored information. Trends Cogn Sci. 2012;16: 606–617. doi:10.1016/j.tics.2012.10.007

47. Palva S, Palva JM. New vistas for alpha-frequency band oscillations. Trends Neurosci. 2007;30: 150–8. doi:10.1016/j.tins.2007.02.001

48. Klimesch W. EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. Brain Res Rev. 1999;29: 169–195.

49. Klimesch W, Doppelmayr M, Hanslmayr S. Upper alpha ERD and absolute power: Their meaning for memory performance. Prog Brain Res. 2006;159: 151–165. doi:10.1016/S0079-6123(06)59010-7

50. Baccino T. Eye movements and concurrent event-related potentials': Eye fixation-related potential investigations in reading. In: Liversedge S, Gilchrist I, Everling S, editors. Oxford handbook of eye movements. Oxford, UK: Oxford University Press; 2011. pp. 857–870.

51. Dimigen O, Sommer W, Hohlfeld A, Jacobs AM, Kliegl R. Coregistration of eye movements and EEG in natural reading: Analyses and review. J Exp Psychol Gen. 2011;140: 552–572. doi:10.1037/a0023885

52. Kliegl R, Dambacher M, Dimigen O, Jacobs AM, Sommer W. Eye movements and brain electric potentials during reading. Psychol Res. 2012;76: 145–158. doi:10.1007/s00426-011-0376-x

53. Frey A, Ionescu G, Lemaire B, López-Orozco F, Baccino T, Guérin-Dugué A. Decision-making in information seeking on texts: An eye-fixation-related potentials investigation. Front Syst Neurosci. 2013;7. doi:10.3389/fnsys.2013.00039

54. Antonenko P, Niederhauser DS. The influence of leads on cognitive load and learning in a hypertext environment. Comput Human Behav. 2010;26: 140–150. doi:10.1016/j.chb.2009.10.014

55. Kretzschmar F, Pleimling D, Hosemann J, Füssel S, Bornkessel-Schlesewsky I, Schlesewsky M. Subjective impressions do not mirror online reading effort: Concurrent EEG-eyetracking evidence from the reading of books and digital media. PLoS One. 2013;8: e56178. doi:10.1371/journal.pone.0056178

56. Antonenko P, Paas F, Grabner R, Gog T. Using electroencephalography to measure cognitive load. Educ Psychol Rev. 2010;22: 425–438. doi:10.1007/s10648-010-9130-y

57. Gerjets P, Walter C, Rosenstiel W, Bogdan M, Zander TO. Cognitive state monitoring and the design of adaptive instruction in digital environments: Lessons learned from cognitive workload assessment using a passive brain-computer interface approach. Front Neurosci. 2014;8. doi:10.3389/fnins.2014.00385

58. Schneider W, Ennemoser M, Schlagmüller M. Lesegeschwindigkeits- und -Verständnistest für die Klassen 6 - 12 (LGVT 6 - 12). In: Hasselhorn M, Marx H, Schneider W, editors. Deutsche Schultests. Göttingen: Hogrefe; 2007.

59. Duggan GB, Payne SJ. Text skimming: The process and effectiveness of foraging through text under time pressure. J Exp Psychol Appl. 2009;15: 228–242. doi:10.1037/A0016995

60. Jasper HH. The ten±twenty electrode system of the International Federation. Electroencephalogr Clin Neurophysiol. 1958;10: 371–375.

61. Delorme A, Makeig S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods. 2004;134: 9–21. doi:10.1016/j.jneumeth.2003.10.009

62. Pfurtscheller G, Zalaudek K, Neuper C. Event-related beta synchronization after wrist, finger and thumb movement. Electroencephalogr Clin Neurophysiol. 1998;109: 154–160. doi:10.1016/S0924-980X(97)00070-2

63. Duncan CC, Barry RJ, Connolly JF, Fischer C, Michie PT, Näätänen R, et al. Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. Clin Neurophysiol. 2009;120: 1883–1908. doi:10.1016/j.clinph.2009.07.045

64. Gevins A, Smith ME. Neurophysiological measures of cognitive workload during human-computer interaction. Theor Issues Ergon Sci. 2003;4: 113–131. doi:10.1080/14639220210159717

65. Smith ME, Gevins A, Brown H, Karnik A, Du R. Monitoring task loading with multivariate EEG measures during complex forms of Human-Computer Interaction. Hum Factors J Hum Factors Ergon Soc. 2001;43: 366–380. doi:10.1518/001872001775898287

66. Scheiter K, Gerjets P, Heise E. Distraction during learning with hypermedia: Difficult tasks help to keep task goals on track. Front Psychol. 2014;5. doi:10.3389/fpsyg.2014.00268

67. Jakob Nielsen. F-shaped pattern for reading web content [Internet]. 2006 [cited 9 Apr 2015]. Available: http://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/

68. Duggan GB, Payne SJ. Skim reading by satisficing: evidence from eye tracking. CHI ’11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, NY: ACM; 2011. pp. 1141–1150.

69. Reader WR, Payne SJ. Texts : Sampling and satisficing. Human-Computer Interact. 2007;22: 263–298.

70. Başar E, Güntekin B. A short review of alpha activity in cognitive processes and in cognitive impairment. Int J Psychophysiol. 2012;86: 25–38. doi:10.1016/j.ijpsycho.2012.07.001

71. Stipacek A, Grabner RH, Neuper C, Fink A, Neubauer AC. Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load. Neurosci Lett. Elsevier; 2003;353: 193–196. doi:10.1016/j.neulet.2003.09.044

72. Strauß A, Kotz S, Scharinger M, Obleser J. Alpha and theta brain oscillations index dissociable processes in spoken word recognition. Neuroimage. 2014;97: 387–395. doi:10.1016/j.neuroimage.2014.04.005

73. Mehta RK, Parasuraman R. Neuroergonomics: A review of applications to physical and cognitive work. Front Hum Neurosci. 2013;7. doi:10.3389/fnhum.2013.00889

74. Parasuraman R, Wilson GF. Putting the brain to work: Neuroergonomics past, present, and future. Hum Factors. SAGE Publications; 2008;50: 468–474. doi:10.1518/001872008X288349.

75. Parasuraman R. Neuroergonomics: Brain, cognition, and performance at work. Curr Dir Psychol Sci. 2011;20: 181–186. doi:10.1177/0963721411409176

76. Brouwer A-M, Hogervorst M, Holewijn M, van Erp JBF. Evidence for effects of task difficulty but not learning on neurophysiological variables associated with effort. Int J Psychophysiol. 2014;93: 242–252. doi:10.1016/j.ijpsycho.2014.05.004

77. John M St., Kobus DA, Morrison JG, Schmorrow D. Overview of the DARPA augmented cognition technical integration experiment. Int J Hum Comput Interact. 2004;17: 131–149.

78. Kruger J, Hefer E, Matthew G. Measuring the impact of subtitles on cognitive load: Eye tracking and dynamic audiovisual texts. Proceedings of the 2013 Conference on Eye Tracking South Africa. 2013. pp. 62–66.

79. Hong L, Walz JM, Sajda P. Your eyes give you away: Prestimulus changes in pupil diameter correlate with poststimulus task-related EEG dynamics. PLoS One. 2014;9. doi:10.1371/journal.pone.0091321

80. Kuipers J-R, Thierry G. N400 amplitude reduction correlates with an increase in pupil size. Front Hum Neurosci. 2010;5: 1–5. doi:10.3389/fnhum.2011.00061

81. Kuipers J-R, Thierry G. ERP-pupil size correlations reveal how bilingualism enhances cognitive flexibility. Cortex. 2013;49: 2853–2860. doi:10.1016/j.cortex.2013.01.012

82. Murphy PR, Robertson IH, Balsters JH, O'Connell RG. Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. Psychophysiology. 2011;48: 1532–1543. doi:10.1111/j.1469-8986.2011.01226.x

83. Fisk JE, Sharp CA. Age-related impairment in executive functioning: updating, inhibition, shifting, and access. J Clin Exp Neuropsychol. 2004;26: 874–890. doi:10.1080/13803390490510680

84. Barkley RA. Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. Psychol Bull. 1997;121: 65–94. doi:10.1037/0033-2909.121.1.65

85. Reiter A, Tucha O, Lange KW. Executive functions in children with dyslexia. Dyslexia. 2005;11: 116–131. doi:10.1002/dys.289

3. General Discussion

The main research goal of the current doctoral thesis was an examination of EEG alpha frequency band power as a global measure of cognitive load which may be grounded in WM functioning and demands on different core EFs like updating, inhibition, and shifting (Diamond, 2013; Miyake et al., 2000). Thus, EEG alpha frequency band power served as primary measure in all three studies. The task settings ranged from basic to applied research. Other measures like the EEG theta and beta frequency band power, the P300, and pupil dilation were used in some studies.

Study 1 compared two classical WM task families, namely an n-back task and an operation span task and a STM digit span task, with respect to EEG correlates (EEG alpha and beta frequency band power, P300). In Study 2 we addressed the interplay of two core EFs, updating and inhibition, by manipulating demands on both within one single WM task, thus avoiding potential perceptual-motor confounds. Furthermore, we were interested in a comparison of EEG alpha frequency band power and pupil dilation as measures of core EFs in a WM task of basic research (Study 2) and as potential measures in a real-world task of hypertext reading and link selection (Study 3).

Study 1 and Study 2 can be considered as basic research that revealed some insights into the interplay of the core EFs and their relation to WM which might be of both, a conceptual as well as a practical value concerning a framework of core EFs. Furthermore, the studies may generally underline the value of using EEG methodologies to examine core EFs. Study 3 can be considered as applied research that mainly addressed the use of EEG alpha frequency band power and pupil dilation to assess different load-situations during hypertext reading (i.e., in comparing text reading and link selection). In the following, I will present and discuss the main outcomes of each study that has been conducted as part of this doctoral thesis, and for each study I will suggest some topics for future research.

3.1 Main Outcomes and Potential Future Research

Concerning the overall research question addressing the role of EEG alpha frequency band power as a measure of cognitive load the three studies conducted as part of this doctoral thesis revealed that a) EEG alpha frequency band power showed an overall sensitivity to demands on different core EFs (Study 1, Study 2), and b) EEG alpha frequency band power could comparably be used in task settings of basic and applied research. Especially, even when potential perceptual-motor confounds were ruled out in a task setting of applied research consisting of link-selection in hypertext reading, EEG alpha frequency band power showed a decrease for increased cognitive load (Study 3). Thus, EEG alpha frequency band power might serve as a global measure of the cognitive load-situation that may be grounded in demands on core EFs. In the following the main outcomes at the level of each individual study will be summarized.

3.1.1 Study 1: The comparison of n-back and WM span tasks

In line with our hypotheses (see Chapter 1.5.1), the outcomes of Study 1 revealed a decreased P300 amplitude for increased load-levels in the n-back and the operation span task but not in the simple digit span task. This indicates that both tasks, the n-back task and the operation span task, share some dual task characteristic, i.e., might comparably demand the EF shifting. Also in line with our hypotheses, overall the TFRs of the n-back and the operation span task were more similar as compared to those of the simple digit span task. Interestingly, the beta frequency band power showed an oscillatory pattern which was comparable to that of the alpha frequency band. This was also the case for the simple digit span task that required no key-pressing (i.e., no motor activity). Thus, the results of Study 1 concerning the beta frequency band power clearly indicate that beta can be regarded to be sensitive to cognitive processing (and not only to motor activity) in WM tasks.

With respect to the long-standing debate concerning the rather weak correlations normally observed between n-back and complex span WM tasks for behavioral performance measures (see, e.g., Jaeggi et al., 2010; Redick & Lindsey, 2013; Schmiedek et al., 2009) and some questioning about the n-back task as genuine WM task (Kane, Conway, Miura, & Colflesh, 2007), Study 1 provides some valuable results to advance the discussion. First, Study 1 showed that performing a verbal n-back task and an operation span task in sum led to comparable EEG patterns. This may be seen to confirm the hypothesis that in both tasks core EFs are comparably active (see 1.3.1 and 1.3.2), thus underlining the assumption that both tasks might be comparably considered as genuine WM tasks. Second, the correlational results of the performance measures in Study 1 showed rather high correlations between the n-back task and the operation span task (and additionally weaker correlations between the n-back task and the digit span task). This was completely in line with the EEG results of the current study and the conceptual considerations with regard to core EFs, but in severe contrast to literature (see, e.g., the meta-analysis by Redick & Lindsey, 2013). One important reason for the rather high correlations we observed in Study 1 might be that we used carefully designed n-back and span tasks with respect to a comparable timing of the single task trials and the overall number of trials. Both factors normally vary severely between complex span and n-back tasks (even in studies that directly compare the both tasks), with the complex span task being used with less trials and longer trial durations (up to self-paced, e.g., Jaeggi et al., 2013, Experiment 1) than the n-back task. Thus, the time-pressure in complex span tasks might normally be less severe than in n-back tasks. As timing is a critical factor with respect to the induced WM load (see Barrouillet, Bernardin, & Camos, 2004; Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007), differences in timing (and amount of trials) may account for the rather weak correlations between n-back and complex span tasks that are normally reported. Our results indicate that if the tasks are controlled for such differences, the correlations between the two tasks are much higher and thus are more in line with theoretical considerations concerning the nature of the both tasks as WM tasks. Clearly, future studies have to be conducted that explicitly address this assumption by directly manipulating time constraints when comparing n-back and complex span WM tasks.

With respect to beta frequency band power, Study 1 might contribute to the recently advanced hypothesis that oscillatory activity in the beta band might not only be associated with motor preparation and motor activity (e.g., Pfurtscheller et al., 1998) but might also reflect cognitive processing (Engel & Fries, 2010; Weiss & Mueller, 2012). We observed an increased beta band ERD for increased load on WM updating in the n-back comparably to Krause and colleagues (Krause et al., 2010; Pesonen et al., 2007). However, as the n-back task contains motor activity (key-presses), the beta band effects in these studies could not assuredly be attributed to cognitive processing but might simply reflect the motor activity. We observed a comparable beta ERD in the operation span task (that also required motor responses in form of key-presses) and more importantly, in the simple digit span task, although this latter task required no motor responses. Thus, our results revealed that comparably to the alpha ERD the beta ERD may reflect load on WM updating. Recently, the role of beta band oscillatory activity and linguistic processing has been put forward (Weiss & Mueller, 2012). Our results may indicate that the beta frequency band might also be important for processes of WM-functioning.

In sum, Study 1 confirmed the conceptually assumptions that the n-back task and the operation span task share some commonalities as genuine WM tasks. The commonalities may be due to comparably demanded core EFs in these tasks, whereas the simple digit span task may mainly demand one single EF, namely WM updating (see 1.3.1 and 1.3.2).

3.1.2 Study 2: The interplay of core EFs

In Study 2 we were interested in the interplay of the core EFs updating and inhibition that we manipulated in one single task consisting of an n-back task (updating) with flanker stimuli (inhibition). We expected to observe an interaction between the n-back levels (i.e., load on WM updating) and the flanker interference (i.e., demands on inhibition). This interaction could be either due to an over-additive or under-additive effect (see 1.5.2). Furthermore, in Study 2 we were interested in the sensitivity of different measures like behavioral data, the P300 amplitude, and alpha and theta frequency band power, for demands on these core EFs. Specifically, we were interested whether EEG alpha frequency band power and pupil dilation would show a comparable pattern of results for load on WM updating and inhibitory demands.

The outcomes of Study 2 revealed first, that all measures that we used were sensitive to load on WM updating as well as demands on inhibitory control. Furthermore, most of the measures consistently showed an interaction between severe load on WM updating and demands on inhibitory control. In the 2-back load level but not in the lower n-back levels, the flanker interference effect was reduced or absent for reaction times, the P300 amplitude, alpha frequency band power, and pupil dilation. This interaction was in line with our *facilitation hypothesis* (see 1.5.2), that is, the assumption that different core EFs may share some common underlying factor which might be attributable to processes of controlled attention (Miyake et al., 2000). Under severe load on WM updating, some common processes of controlled attention might be activated that also serve inhibitory control, and thus facilitate inhibitory control. However, the results of the reduced flanker interference effect under severe load on WM updating are in contrast to observations reported by Lavie and colleagues (Lavie, 2005, 2010; Lavie et al., 2004; see also 1.5.2). These authors reported an increased interference effect under high STM load (i.e., load on WM storage components). The different outcomes of Study 2 and the studies by Lavie and colleagues might be due to the fact that we directly manipulated executive WM load, that is, demands on WM updating and inhibitory demands, whereas Lavie and colleagues manipulated only one EF, namely inhibition in the face of load on STM (i.e., load on WM storage components using a Sternberg task). An additional difference between the studies which might have contributed to the different outcomes was that the manipulation of core EFs was done in a single task paradigm in Study 2, whereas participants in the studies by Lavie and colleagues had to perform two different tasks (a flanker task and a Sternberg task) which were sequentially intermixed.

Taken together, the results of Study 2 underline the closely intertwined character of core EFs. However, the generalizability of these outcomes has to be addressed in future studies. These future studies might use other task designs for manipulating load on WM updating and demands on inhibitory control to exclude any potential task-specificity of our results. Furthermore, other combinations of core EFs (e.g., updating and shifting) might be studied in a task setting comparably to the one we used allowing both EFs to be manipulated within one single task, and thus avoiding potentially confounding perceptual-motor factors.

With respect to the second research question, namely the comparison of different measures for demands on core EFs, Study 2 revealed a comparable pattern of results for pupil dilation and alpha frequency band power. Pupil dilation and alpha frequency band power were sensitive for load on WM updating as well as demands on inhibitory control and both measures also reflected the interaction between the two EFs. Consequently, both measures might be comparably used to study core EFs. In contrast, EEG theta frequency band power increased for both, increased load on WM updating and increased demands on inhibition, but showed no interaction between the two. This might be interpreted to reflect the nature of theta frequency band power as a general correlate of cognitive control (e.g., Nigbur, Ivanova, & Stürmer, 2011; Sauseng et al., 2010). Irrespective of the WM updating load the processes of cognitive control may be active. However, the interpretation of EEG theta frequency band power might be rather speculative at this point, as EEG theta frequency band power in Study 2 was the only physiological measure that showed no interaction between load on WM updating and demands on inhibitory control. Clearly, further studies have to be conducted to examine the robustness of the observed theta frequency band power effects before any valid interpretation can be given.

3.1.3 Study 3: EEG alpha frequency band power and pupil dilation for link selection during hypertext reading

In Study 3 we were interested in the increased load-situation during the initial processes of link selection in hypertext reading that might be attributable to increased demands on core EFs (see 1.3.3 and 1.5.3). As expected, we observed an increased alpha ERD and an increased pupil dilation during link-like selection processes (i.e., the test condition) as compared to a baseline condition. The baseline condition was purely text reading in Study 3 a). Given our assumption that selecting a context matching word out of three response alternatives resemble genuine cognitive processes of hyperlink selection (see 1.5.3), this outcome indicates that already initial processes of link selection during hypertext reading lead to an increased load-situation that can be assessed by pupil dilation or EEG alpha frequency band power ERD and that (at least conceptually, see 1.3.3) might be attributed to core EFs. Study 3 b) replicated these results although a modified baseline was used. In Study 3 b) in the baseline condition participants had to perform a mouse-click on a word, comparably to the test condition, yet without choosing between response alternatives to click on (i.e., baseline and test condition were matched with respect to motor activity related with moving of the mouse cursor and clicking on a word). Additionally, baseline and test conditions were perceptually matched with respect to text color. This way, Study 3 b) eliminated perceptual-motor confounds of Study 3 a) between baseline and test condition thus ruling out the alternative hypothesis that these confounds would have accounted for the observed results in Study 3 a). Finally, Study 3 c) excluded possible context effects (i.e., differences in local text difficulty) that also could be an alternative explanation for the observed different load-situation between baseline and test condition in Study 3 a) and 3 b), that is, when comparing different positions in the text. As Study 3 c) revealed no difference between baseline and test condition when the task in both conditions was equalized to purely mouse-clicking on one word (i.e., no decision and selection processes between response alternatives in the test condition), local text difficulty can also be ruled out as alternative hypothesis for the results of Study 3 a) and 3 b).

In sum, Study 3 showed that both measures, pupil dilation and EEG alpha frequency band power can be applied in a rather real-world, complex task setting like hypertext reading and link selection. Both measures were sensitive to the increased load-situation during link selection in hypertext reading. Surprisingly however, although showing the same patterns of results, the two measures did not correlate with one another. At this point we can only speculate about possible explanations for this rather unexpected outcome. First, as both, the quality of EEG data as well as the quality of eye-tracking data is prune to individual differences (i.e., some participants might show rather noisy EEG data, e.g., because of muscle artefacts, whereas some participants might show rather noisy eye-tracking data, e.g., because of high blink-rates), those technical reasons could account for the non-existent correlation. Furthermore, the rather complex online-reading situation might have affected the EEG data and the pupil dilation data differently in single subjects. However, as we carefully controlled Study 3 for potential perceptual-motor confounds and as we used only artefact-free epochs for data analysis, such technical reasons for the non-existent correlation might be rather ruled out. Another hypothesis concerning the observed non-correlation between pupil dilation and EEG alpha frequency band power ERD might be that pupil dilation also captures additional aspects of a load-situation, like emotional or motivational aspects (e.g., Hess & Polt, 1960; Wykowska, Anderl, Schubö, & Hommel, 2013; see also 1.4.2), whereas alpha frequency band power might capture more purely cognitive aspects of a load-situation (see 1.4.1.1). Thus, these additional aspects of the load-situation might confound the pupil dilation data but not the EEG alpha frequency band power data which may result in the observed outcome of a non-correlation between the two measures.

Clearly, at this point these considerations concerning the observed non-correlation are rather speculative and should be addressed in future studies. These studies might use rather low-level task materials and should be designed as genuine correlational studies (e.g., using larger numbers of subjects). Additionally, to further avoid potential confounds that may selectively affect pupil dilation, auditory task materials might be used, thus avoiding any perceptual influences on pupil dilation.

Another major limitation of Study 3 concerning the task materials has to be addressed in future studies. Although we are quite convinced that the task materials we used triggered comparably processes of decision making, selection, and demands on core EFs that occur during genuine hyperlink-selection in 'real' hypertext reading situations, we are aware that our task materials might be considered as rather artificial. This is mainly because participants in Study 3 simply had to click each time on one context-matching word out of three words. Clearly, future studies have to be conducted using hypertext task materials which more closely resemble genuine hypertext reading with 'true', single hyperlinks to be selected at different positions in the text. Furthermore, the links should provide a 'true' hyperlink-functionality (i.e., lead to subsequent pages) that also missed in Study 3. Noteworthy however, in the current study we were interested in analyzing the initial link-selection processes, that is, without possible confounds due to subsequent web-pages (see 1.5.3). Thus, the non-functional hyperlinks were due to our task paradigm and could not be avoided.

Furthermore, although conceptually the observed increased load-situation during link selection may be attributed to demands on core EFs (see 1.3.3), a direct examination of this assumption still has to be conducted. This could be done in future studies either using a dual task methodology, that is, using secondary tasks that demands specific core EFs (e.g., a flanker task that mainly demands inhibition) while participants read the hypertext and perform the link selections. Such a dual task methodology might allow to selectively test which EF is most strongly demanded during link selection. However, as the dual-task methodology has the risk to negatively interfere with the primary task (i.e., to lead to unexpected confounds), a more suitable methodology might be to specifically manipulate the manner of the hyperlinks in a text to study the induced demands on core EFs. Thus, for example, a comparison between hyperlinks that are highly interesting but not relevant for a current information-gathering task and hyperlinks that are relevant or irrelevant might be conducted. The interesting but irrelevant hyperlinks should lead to increased demands on inhibitory control in comparison to purely irrelevant hyperlinks (i.e., non-interesting and not-relevant hyperlinks).

Despite these obvious limitations discussed above, two major strengths of Study 3 have to be highlighted. First, Study 3 advances traditional hypertext research by using physiological data, and, more importantly, by using a combined EEG-eye-tracking methodology. This methodology allowed us, comparably to the eye-fixation-related potentials methodology (EFRP; e.g., Baccino, 2011; Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011), to use the eye-tracking data to define areas of interest (AOIs) for the EEG frequency band power analysis, that is, to analyze data sections (epochs) of the EEG data that are time-locked to the eye-fixations on specific parts of the text. For example, in Study 3 a) we used this technique to define the text parts for EEG data analysis that contained AOIs of purely text reading. These text parts served as baseline condition in Study 3 a). Noteworthy, in Study 3 b) we defined the EEG data epochs of the baseline condition in a more classical way, namely time-locked to the mouse-event (mouse-click). This was possible as participants in Study 3 b) had to perform a mouse-click in both, baseline and test conditions. Nevertheless, despite these differences between Study 3 a) and 3 b) in defining the baseline condition for the EEG data analysis, the outcomes of Study 3 b) were very similar to Study 3 a). Thus, Study 3 a) and b) may be regarded to proof the feasibility of using AOIs to define EEG data epochs for EEG data analysis even for calculating EEG frequency band power (and not only for the calculation of EFRPs).

Second, in contrast to other studies on hypertext reading using neurophysiological measures (e.g., Antonenko & Niederhauser, 2010; Antonenko et al., 2010; Gerlic et al., 1999), we focused on the load-situation during initial link-selection processes and carefully controlled our task materials for possible confounds like perceptual-motor or content-related differences between the task conditions of interest. Such confounds often hamper the interpretability of studies using complex task materials (e.g., Antonenko & Niederhauser, 2010; see Gerjets et al., 2014 for a discussion of the problems of perceptual-motor confounds in these task-settings especially for the EEG). Thus, in sum, the outcomes of Study 3 may be seen to underline the feasibility and sensitivity of EEG frequency band power and pupil dilation for research on hypertext reading even if confounding factors are ruled out that are often considered as an alternative hypotheses to explain observed effects (Gerjets et al., 2014).

To conclude, Study 3 may serve only as an initial, but important step in studying link-selection processes during hypertext reading and possible demands on core EFs. Follow-up studies using the same methodology of combined EEG-eye-tracking analyses may apply more realistic hypertext materials and may actually manipulate the demands on core EFs within hypertext reading. In the following chapter some general strengths and limitations of the current thesis will be discussed.

3.2 Strengths and Limitations of the Current Thesis

The three studies that were conducted as part of this doctoral thesis span a broad field not only ranging thematically from basic research (Study 1 and Study 2) to applied research (Study 3) but also linking concepts of cognitive science with (neuro-) physiological measurement. Taking such a broad research account and using such diverse measures like EEG frequency band power, P300, and pupil dilation, on the positive side may yield in several new insights on both, a basic level of the single studies (see Chapter 3.1), as well as on an overall level in connecting the single outcomes and the different research areas. Thus, the outcomes of the three studies might provide valuable insights in core EFs and measurement techniques thereof and may serve to stimulate further research (see the following Chapter 3.3).

However, one potential negative side of taking such a global research approach might be the risk of shallowness to a certain extent. Because of limits in time and space (especially within a doctoral thesis), not all aspects of the single different research areas, methodologies, or models may have been presented in an adequate depth. Especially, as stated above, each study we conducted would have called for direct follow-up studies for each single research topic addressing the newly raised questions or potential pitfalls of the single studies. We postponed such studies to the future and instead focused on different aspects of core EFs and physiological measures thereof in diverse task settings. Therefore, instead of digging deep into one specific research question, the current thesis took a multifaceted research approach (resembling the multifaceted nature of EFs) that might serve as an overall, rich source for stimulating future studies in diverse areas of research (see above and Chapter 3.3) that focus on core EFs, physiological measures, and cognitive load.

In the discussion above (Chapter 3.1), some of such possible or necessary follow-up studies have already been presented. Especially the restrictions of Study 3 might have to be addressed in future studies. Thus, the current study might allow only very restricted interpretations with respect to core EFs in hypertext-like task materials and with respect to real-world hypertext reading, as we used a kind of 'simulated' link selection processes without any explicit manipulation of the core EFs therein. Furthermore, it might be questioned whether the task materials that we used (a rather artificial hypertext and link-selection situation) might resemble genuine processes that occur in genuine hypertext reading situations. Thus, although our assumptions concerning the role of core EFs in hypertext reading and link-selection as described in Chapter 1.3.3 may sound plausible, the validity of our assumptions have to be tested further in future studies. These studies might use more genuine hypertext materials with real hyperlink-selection (e.g., Wikipedia articles) and might for example manipulate demands on core EFs like inhibitory control by varying the amount and manner of hyperlinks (e.g., by presenting relevant links, interesting but irrelevant links, and completely irrelevant links; see 3.1.3).

Some further limitations of the current thesis have to be addressed. First, as research on both, EFs as well as EEG oscillatory activity is still a vivid field, with an ultimate definition of a framework of EFs still missing (see Chapter 1.1) and the link of oscillatory activity and specific cognitive processes still being matter of debate (see Chapter 1.4), a concise formulation of a priori hypotheses in the current thesis has been somewhat complicated. Therefore, we choose a research account of formulating specific hypotheses whenever possible and otherwise to present the outcomes rather data-driven if no clear hypotheses could be formulated a priori. For example in Study 1 we had only vaguely expectations concerning commonalities and differences in the TFR patterns of the three different tasks (see 1.5.1). However, as TFRs have not been used before for the comparison of different WM and STM tasks, a rather descriptive presentation of the TFR results seems to be justified.

Second, the EEG analyses we conducted were restricted to EEG frequency band power (and P300) at selected electrodes. The electrodes for the statistical EEG data analysis were chosen in accordance with literature (e.g., Antonenko et al., 2010; Duncan et al., 2009; Gevins & Smith, 2000) to yield maximal effects on each measure. For example, the electrode Pz was chosen to represent oscillatory activity of mid-parietal brain regions that have been previously shown to yield maximal effects on alpha frequency band power and P300 for cognitive processing-load. The restriction to use single electrodes for the main (statistical) data analysis was motivated (and may be justified further) by the overarching goal of the current thesis, namely to examine measures that showed sensitivity to specific cognitive processing like demands on core EFs in basic research (see Study 1 and Study 2) and that may be also used in applied task settings (see Study 3). Especially in applied research it might be wishful to have some few, rather easily (i.e., efficiently) to acquire measures. Despite its advantage as a rather direct measure of brain activity with high time-resolution (see Chapter 1.4), the EEG (at least to date) needs some preparation time to get a good signal-to-noise ratio, that is, a good connection between the single EEG electrodes and the skin. Thus, the reduction to few, most sensitive electrode position may severely decrease the preparation times for EEG data recordings in applied task settings and consequently may help to foster the (efficient) use of the EEG methodology in rather applied research like hypertext reading (see also Chapter 3.3 for some future applications of the EEG methodology in applied task settings).

Apart from the restriction to few, selected electrodes, we limited our EEG data analyses to ERPs (i.e., the P300) and frequency band power for a frequency range between 4 to 30 Hz (theta, alpha, and beta frequency bands). The selection of these measures was also literature-driven (see 1.4.1; e.g., Antonenko et al., 2010; Just et al., 2003; Parasuraman, 2003). The ratio for the chosen measures (i.e., the data analyses methods) was also (like for the selection of the electrodes) to have easily to acquire measures that would allow the application in task settings of both, basic and applied research. Thus, the measures should at least principally allow assessing the individual load-situation based on a few trials. These prerequisites might especially be fulfilled by EEG frequency band power. More advanced data analysis methods to capture oscillatory activity like analyzing phase information and calculating coherence measures (e.g., Sauseng & Klimesch, 2008; Sauseng, Klimesch, Schabus, & Doppelmayr, 2005), or source localization (e.g., Hanslmayr et al., 2008) were beyond the scope of the current thesis. Noteworthy however, such EEG data analysis methods might be used in future studies (at least in basic research) to further increase the understanding of brain oscillatory activity and core EFs.

Third, another limitation of the current thesis might be, as already stated in Chapter 1.1, that we had a rather generic view on core EFs, that is, we did not differentiate possible sub-processes thereunder. Although this view is in line with the prevalent research on core EFs (see Chapter 1.1), a further fractionation of core EFs in possible sub-processes and the analysis thereof using EEG methodologies in future studies may help to identify whether controlled attention is the commonality of different core EFs (as assumed in the current thesis) or whether specific shared sub-processes might constitute commonalities between different core EFs. Thus, for example, updating might be in principal equal to shifting, yet only referring to different WM 'content' (see Oberauer, 2009 and 1.2.4 of the current thesis). Clearly, future research may develop a conceptually more fine-grained framework of core EFs that then may be empirically tested, e.g., using EEG methodologies like the ones proposed in the current thesis.

Finally, to avoid any confusion, we have to address our slightly different focus on the n-back task with respect to the core EFs therein in Study 1 as compared to Study 2. In Study 1 we assumed that the n-back task would incorporate all three core EFs, that is, updating, shifting, and inhibition, comparably to a complex operation span task. In Study 2 we labeled the n-back task as a task that 'mainly' loads on the EF updating. Given our task analysis (see 1.3.1) and the results of Study 1, the definition of the n-back task in Study 2 might be considered as slightly shortened. However, for reasons of brevity and, furthermore, for reasons of conformity with the prevalent literature on the n-back task (e.g., Gevins & Smith, 2000; Gevins et al., 1997; Jonides et al., 1997; Pesonen et al., 2007), it may be justified that we considered the n-back task in Study 2 mainly as representative for WM updating load. Yet, given that the n-back task also task-inherently demands inhibition and shifting, one possible alternative hypothesis for the outcomes of Study 2 might be raised. According to this alternative hypothesis, the observed reduced flanker interference effect under the high, 2-back load-level might be directly due to the already demanded inhibitory control (i.e., processes of inhibition that are incorporated in the n-back) and not only, as stated in 3.1.2, due to the increased activation of some commonly underlying network structure of controlled attention. Therefore, our conclusion of Study 2 might be considered with some caution. Clearly, more studies are needed that manipulate different core EFs in different task settings using the EEG methodologies in order to unravel the exact nature of any underlying common factor.

To conclude, the rather wide-spread research account of the present doctoral thesis might show some limitations at the level of each individual study as discussed above and in Chapter 3.1. However, these limitations may be ruled out by future research. Thus, each current study may serve as a starting point for triggering further research. In sum, the studies showed that (neuro-) physiological research methods like EEG alpha frequency band power or pupil dilation are valuable research accounts to study core EFs in basic and applied task settings (but see the limitations concerning the applied task setting we used as discussed above). Both measures were sensitive to changes in the general load-situation when core EFs are demanded (Study 2). Furthermore, the current thesis showed that the beta frequency band power might be another valuable measure of WM functioning (Study 1). All EEG frequency band power measures were sensitive to demands on different core EFs (Study 2). Yet, none of the EEG frequency band power measures was selectively sensitive to one specific EF. Thus, EEG frequency band power like alpha frequency band power might serve as a rather general measure of the overall cognitive load-situation in task settings of basic and applied research. This overall load-situation might be altered due to demands on core EFs.

In the next chapter potential implications of the main outcomes of the current thesis will be discussed.

3.3 General Implications and Future Research Directions

The outcomes of the present doctoral thesis have several implications for both, basic and applied research. In the following, some main implications will be summarized and some future research directions will be highlighted.

3.3.1 Basic research: WM and EFs

The outcomes of Study 1 confirmed the conceptual similarities of an n-back task and an operation span task which might be attributable to core EFs that are demanded in both WM tasks comparably, and which are less demanded in a STM digit span task. Thus, Study 1 extends purely correlational research accounts that focus on behavioral performance data for comparing WM tasks (e.g., Jaeggi et al., 2010; Redick & Lindsey, 2013), and shows the value of using an (electro-) physiological research approach for examining and comparing different WM tasks. Future studies may take this research approach and compare n-back and WM span tasks using tasks of different modalities (e.g., a spatial n-back task instead of a verbal n-back task, or a reading span task instead of an operation span task). The outcomes of such studies should resemble the results of Study 1 as executive WM processes are hypothesized to be independent of concrete task modalities (Baddeley, 1996). Furthermore, WM and STM tasks beyond the n-back and WM span tasks may be included in the task comparison using the EEG methodology (e.g., Sternberg tasks). Finally, Study 1 demonstrated the importance of carefully designed task paradigms with respect to timing and number of trials for a direct task comparison (see 3.1.3). Future studies interested in the comparison of WM task might explicitly manipulate time constraints or, at least, should keep in mind the importance of matching such constraints.

Study 2 proposed a novel research account to examine core EFs, namely the direct manipulation of core EFs within one single task, thus avoiding potential confounds that are otherwise inevitable when using different tasks for studying core EFs (e.g., perceptual-motor differences or task-set differences between tasks). EEG alpha frequency band power, the P300 ERP, and pupil dilation showed a comparable sensitivity for the interplay of the core EFs updating and inhibition in this study. Consequently, future studies may use the research approach of manipulating different single core EFs within one single task and may also include the manipulation of the EF shifting. This way, the interplay of all core EFs may be studied, and their separable yet intertwined character may be examined further. As there is still debate concerning the exact number and conceptualization of core EFs (see Chapter 1.1 and 3.2), such studies may help clarifying commonalities and differences between different core EFs and thus may advance the development of a general framework of core EFs that grounds on biological underpinnings as identified by physiological measures.

In sum, Study 1 and Study 2 both showed the value of using the EEG for conducting research on WM and core EFs. Thus, future studies on WM and core EFs may also include EEG measures. From the results obtained so far we may conclude that EEG measures (at least the ones we used) are sensitive to changes in the load-situation when core EFs are demanded. However, we did not observe a specific single EEG measure that could be attributed to a specific single core EFs. The P300 may be considered as one measure that might be most closely assigned to one single core EF, namely shifting (see 1.3.1 and 1.4.1.4). Clearly, more research has to be conducted to verify this assumption. EEG frequency band power seems to react more like a global measure of the entire load-situation. EEG frequency band power may not miss when demands on a single core EFs are changed and the overall load-situation is altered, but it may not allow to specifically disentangling different core EFs. Thus, in line with literature on EEG frequency band power (e.g., Krause, 2003; see also 1.4.1) we may conclude that there is no simple one-to-one link between a specific core EF and the oscillatory activity in one specific frequency band.

Pupil dilation may serve as a more easily to acquire, yet comparably sensitive measure for the overall load-situation, as for example, EEG alpha frequency band power (see Study 2 and Study 3). Yet, as Study 3 revealed, the two measures might not be used completely interchangeable. Clearly, more research has to be conducted on the relationship of pupil dilation and EEG (alpha) frequency band power as measures of core EFs. The outcomes of Study 3 may indicate that pupil dilation captures more aspect of the load-situation than EEG alpha frequency band power, for example, aspects that are beyond purely 'cold' cognition, like motivational or emotional aspects. Given that this assumption turns out to be valid, combined analysis of EEG frequency band power data and pupil dilation may be used to disentangle cognitive and emotion-motivational aspects of a load-situation. In doing so, this research methodology may become interesting for the evolving research on core EFs and emotions (e.g., Levens & Gotlib, 2010; Mitchell et al., 2007; Pessoa, 2009).

Finally, as already stated in Chapter 3.2, the outcomes of the current thesis underline the general value of the EEG methodology for studying core EFs and WM. In future studies, more advanced EEG data analyses methods might be taken into account like source localization (e.g., Hanslmayr et al., 2008) or phase-coherence measures (e.g., Sauseng & Klimesch, 2008; Sauseng et al., 2005) which might allow to actually disentangle different core EFs. With respect to EEG frequency band power we may conclude that this measure may serve as a rather global measure of the cognitive load-situation in tasks that demand core EFs.

3.3.2 Applied research: EFs in hypertext reading

Study 3 revealed that physiological measures like EEG alpha frequency band power and pupil dilation can be used in applied research settings like hypertext reading and link selection to assess the global load-situation therein. Importantly, we used in Study 3 highly controlled yet complex task materials and ruled out potential confounds in follow-up experiments (see 1.5.3). Thus, Study 3 advances traditional hypertext and hypermedia research by (a) using physiological measures, (b) at least conceptually focusing on core EFs, and (c) carefully controlling for potentially confounding factors. Future studies may draw upon the outcomes of Study 3 and may adopt the research methodology of combined EEG-eye-tracking data analyses.

In principal there are two main research directions in applied research using physiological measures like EEG alpha frequency band power or pupil dilation with respect to hypertext or, more generally, hypermedia applications, that may benefit from the outcomes of the current thesis. First, physiological measures can be used for the analysis and evaluation of hypermedia applications with respect to the load-situation therein. Second, and more intriguingly, physiological measures may be used to assess the individual load-situation in hypermedia applications online, that is, in single subject, single trial situations. Advanced EEG data analyses methods and the use of machine learning algorithms may allow to automatically detect the instant load-situation of an user (see, e.g., the passive brain-computer interface account proposed by Zander & Kothe, 2011, or the physiological computing account by Allanson & Fairclough, 2004; Fairclough, 2009). The hypermedia application may then be adapted accordingly (e.g., by reducing distracting information like irrelevant hyperlinks that could be faded out, or by inserting relevant hyperlinks which may provide webpages with additional, helpful information).

With respect to the first research area, the analysis and evaluation of hypermedia applications, our results indicate the feasibility and sensitivity of EEG alpha frequency band power and pupil dilation to assess the load-situation therein. Especially the methodology of combined EEG-eye-tracking data recording and analysis (see 3.1.3) may advance future hypermedia research. For example, future studies may use these methodologies to analyze the load-situation during hypertext reading of, for example, Wikipedia articles or webpages presenting search results. These applications may then be evaluated and optimized (see also 1.4 for a discussion for the general advantages of using physiological measures to assess the load-situation in complex, real-world task settings). For example, in case of hypertext webpages, the load-situation therein might be optimized by varying the amount and manner of hyperlinks. Furthermore, the research approach of combined EEG-eye-tracking data recording may be used to identify specific areas in webpages which induce high demands on cognitive processing. In addition, the cognitive integration of different elements on a webpage like textual and pictorial information might be studied using this research methodology. Finally, grounding on our conceptual considerations with respect to core EFs in hypertext reading (see 1.3.3) future research may further address the role of core EFs in hypertext reading. In doing so, traditional models of reading comprehension (e.g., Kintsch & Welsch, 1991; Kintsch, 1988) may be advanced by incorporating core EFs in their conceptual considerations. Focusing on cognitive processes during reading seems generally to be an evolving research direction (see, e.g. Kendeou, Van Den Broek, Helder, & Karlsson, 2014). Given the prevalence of electronic (hyper-) texts we are nowadays faced with (see Chapter 1, Introduction), incorporating hypertext reading and cognitive processes like core EFs that are especially associated with this text category may become an important research direction in the future to advance current models of text comprehension. The outcomes of the current thesis indicate that both, the EEG and pupil dilation can serve as valuable research methodologies for this research direction.

With respect to the second research area, the use of physiological measures for assessing the individual load-situation online (i.e., in a single subject, single trial situation) in order to create load-adaptive hypermedia applications, our results indicate that EEG frequency band power or pupil dilation might be used to assess the global load-situation, but may not be suitable to directly differentiate specific demands on different core EFs. Theoretically, an intriguing idea would be that core EFs may guide specific forms of adaptation. For example, if demands on inhibitory control are detected to be high, distracting information might be reduced. Or, if demands on shifting are detected to be high, some form of attentional guidance might be implemented to help users to focus on the central task. A major prerequisite of such an adaptive system however would be to have measures that are able to differentiate demands on different core EFs. The outcomes of the current thesis indicate that EEG frequency band power alone may not suffice to specifically differentiate demands on core EFs. Although being sensitive to changes of demands on core EFs that alter the global load-situation, we did not observe any specific one-to-one relationship between a single core EFs and a specific EEG frequency band. Thus, more advanced EEG data analysis methods might be necessary to specifically disentangle the demands on single core EFs (see Chapter 3.2). Importantly, the observed non-correlation between pupil dilation and EEG frequency band power in Study 3 of the current thesis implies that especially for research approaches on load-adaptive systems it might be useful to capture different load-measures, as the different load-measures may work differently well for different subjects (see the discussion in 3.1.3).

3.4 Conclusion

In the current doctoral thesis I was interested in research questions concerning cognitive load, its potential origin in WM and demands on core EFs, and measures thereof. As overall research question, we studied EEG alpha frequency band power as a measure of cognitive load which we assumed to originate from demands on core EFs. This was done in task settings of basic and applied research. The outcomes of the three studies we conducted validated EEG alpha frequency band power as a general measure of cognitive load that could be traced back to reflect demands on core EFs. As a prerequisite of this overall research question, we studied the role of core EFs for WM functioning (Study 1 and Study 2) and (with some limitations) for hypertext reading (Study 3). In detail, we studied the relation of core EFs and WM by comparing EEG correlates for different WM tasks (Study 1), the interplay of core EFs within one single WM task (Study 2), and the sensitivity of EEG frequency band power and pupil dilation as measures for demands on core EFs in a WM task (Study 2) and for cognitive load during processes of link selection in hypertext reading (Study 3).

The outcomes of the three studies may serve as starting-points for a variety of follow-up studies also ranging from basic to applied research. They thus may potentially be of relevance for diverse disciplines like cognitive science, neuroscience, or computational science to draw upon. Ideally, the results of the current thesis may initiate (and contribute to) the further development of a thoroughly defined framework of core EFs that is currently missing, and may furthermore advance hypertext research (and, more generally, hypermedia research) by directing future research foci on core EFs therein, and, finally, may have highlighted the use of (electro-) physiological measures as valuable research methodologies for both, basic and applied research areas to study cognitive load.

References

Allanson, J., & Fairclough, S. (2004). A research agenda for physiological computing. *Interacting with Computers*, *16*(5), 857–878. http://doi.org/10.1016/j.intcom.2004.08.001

Allison, B. Z., & Polich, J. (2008). Workload assessment of computer gaming using a single-stimulus event-related potential paradigm. *Biological Psychology*, *77*(3), 277–283. http://doi.org/10.1016/j.biopsycho.2007.10.014

Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review*, *16*(1), 17–42. http://doi.org/10.1007/s11065-006-9002-x

Amadieu, F., van Gog, T., Paas, F., Tricot, A., & Mariné, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction*, *19*(5), 376–386. http://doi.org/10.1016/j.learninstruc.2009.02.005

Andreassi, J. L. (2007). Pupillary response and behavior. In J. L. Andreassi (Ed.), *Psychophysiology: Human behavior and physiological response.* (5th ed., pp. 289–307). Mahwah, NJ: Lawrence Erlbaum.

Antonenko, P., & Niederhauser, D. S. (2010). The influence of leads on cognitive load and learning in a hypertext environment. *Computers in Human Behavior*, *26*(2), 140–150. http://doi.org/10.1016/j.chb.2009.10.014

Antonenko, P., Paas, F., Grabner, R., & Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review*, *22*, 425–438. http://doi.org/10.1007/s10648-010-9130-y

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *Psychology of learning and motivation. Advances in research and theory.* (Vol. 2, pp. 89–195). New York, NY: Academic Press. http://doi.org/10.1016/S0079-7421(08)60422-3

Axmacher, N., Elger, C. E., & Fell, J. (2009). The specific contribution of neuroimaging versus neurophysiological data to understanding cognition. *Behavioural Brain Research, 200*(1), 1–6. http://doi.org/10.1016/j.bbr.2009.01.028

Baccino, T. (2011). Eye movements and concurrent event-related potentials': Eye fixation-related potential investigations in reading. In S. Liversedge, I. Gilchrist, & S. Everling (Eds.), *Oxford handbook of eye movements* (pp. 857–870). Oxford, UK: Oxford University Press.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. http://doi.org/10.1126/science.1736359

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology*, *49*(1), 5–28. http://doi.org/10.1080/713755608

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. http://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. (2002). Fractionating the central executive. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function*. http://doi.org/10.1093/acprof:oso/9780195134971.003.0016

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

Baddeley, A. (2007). *Working memory, thought, and action.* Oxford, UK: Oxford University Press. http://doi.org/10.1093/acprof:oso/9780198528012.001.0001

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63,* 1–29. http://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A. D. (2007). Neuroimaging working memory. In Author (Ed.), *Working memory, thought, and action.* (pp. 211–255). Oxford, England: Oxford University Press.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control*. (pp. 28–61). Cambridge, UK: Cambridge University Press.

Baddeley, A., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89. http://doi.org/10.1016/S0079-7421(08)60452-1

Banich, M. T. (2009). Executive function: The search for an integrated account. *Current Directions in Psychological Science*, *18*, 89–94. http://doi.org/10.1111/j.1467-8721.2009.01615.x

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults’ working memory spans. *Journal of Experimental Psychology. General*, *133*(1), 83–100. http://doi.org/10.1037/0096-3445.133.1.83

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(3), 570–585. http://doi.org/10.1037/0278-7393.33.3.570

Başar, E. (2012). A review of alpha activity in integrative brain function: Fundamental physiology, sensory coding, cognition and pathology. *International Journal of Psychophysiology*, *86*(1), 1–24. http://doi.org/10.1016/j.ijpsycho.2012.07.002

Başar, E., & Güntekin, B. (2012). A short review of alpha activity in cognitive processes and in cognitive impairment. *International Journal of Psychophysiology*, *86*(1), 25–38. http://doi.org/10.1016/j.ijpsycho.2012.07.001

Bastiaansen, M., Mazaheri, A., & Jensen, O. (2012). Beyond ERPs: Oscillatory Neuronal Dynamics. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford Handbook of Event-Related Potential Components* (pp. 31–50). Oxford, UK: Oxford University Press.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276–292. http://doi.org/10.1037/0033-2909.91.2.276

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. Berndtson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 142–162). Cambridge, UK: Cambridge University Press.

Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Archiv Für Psychiatrie and Nervenkrankheiten*, *78*(1), 527– 570.

Best, J., & Miller, P. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641–1660. http://doi.org/10.1111/j.1467-8624.2010.01499.x

Best, J. R., Miller, P. H., & Jones, L. L. (2010). Executive functions after age 5: Changes and correlates. *Developmental Review*, *29*(3), 180–200. http://doi.org/10.1016/j.dr.2009.05.002

Bledowski, C., Kaiser, J., & Rahm, B. (2010). Basic operations in working memory: Contributions from functional imaging studies. *Behavioural Brain Research*, *214*(2), 172–179. http://doi.org/10.1016/j.bbr.2010.05.041

Bledowski, C., Rahm, B., & Rowe, J. B. (2009). What “works” in working memory? Separate systems for selection and updating of critical information. *The Journal of Neuroscience*, *29*(43), 13735–13741. http://doi.org/10.1523/JNEUROSCI.2547-09.2009

Bradshaw, J. L. (1969). Background light intensity and the pupillary reponse in a reaction time task. *Psychonomic Science*, *14*(6), 271–272. http://doi.org/10.3758/BF03329118

Brouwer, A.-M., Hogervorst, M., Holewijn, M., & van Erp, J. B. F. (2014). Evidence for effects of task difficulty but not learning on neurophysiological variables associated with effort. *International Journal of Psychophysiology*, *93*(2), 242–252. http://doi.org/10.1016/j.ijpsycho.2014.05.004

Brown, G. G., Kindermann, S. S., Siegle, G. J., Granholm, E., Wong, E. C., & Buxton, R. B. (1999). Brain activation and pupil response during covert performance of the Stroop color word task. *Journal of the International Neuropsychological Society*, *5*(4), 308–319. http://doi.org/10.1017/S1355617799544020

Brünken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 53–61. http://doi.org/10.1207/s15326985ep3801_7

Brünken, R., Steinbacher, S., Plass, J. L., & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental Psychology*, *49*(2), 109–119. http://doi.org/10.1027//1618-3169.49.2.109

Burke, S. N., & Barnes, C. a. (2006). Neural plasticity in the ageing brain. *Nature Reviews. Neuroscience*, *7*(1), 30–40. http://doi.org/10.1038/nrn1809

Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, *304*, 1926–1929. http://doi.org/10.1126/science.1099745

Cabestrero, R., Crespo, A., & Quirós, P. (2009). Pupillary dilation as an index of task demands. *Perceptual and Motor Skills*, *109*, 664–678. http://doi.org/10.2466/pms.109.3.664-678

Cain, B. (2007). *A review of the mental workload literature*. *Defence research and development Canada Toronto. Human system integration section.* Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA474193

Caldas, A. L., Machado-Pinheiro, W., Souza, L. B., Motta-Ribeiro, G. C., & David, I. A. (2012). The Stroop matching task presents conflict at both the response and nonresponse levels: An event-related potential and electromyography study. *Psychophysiology*, *49*(9), 1215–1224. http://doi.org/10.1111/j.1469-8986.2012.01407.x

Cangoz, B., & Altun, A. (2012). The effects of hypertext structure, presentation and instruction types on perceived disorientation and recall performances. *Contemporary Educational Technology*, *3*(2), 81–98.

Chen, Y.-N., Mitra, S., & Schlaghecken, F. (2008). Sub-processes of working memory in the n-back task: An investigation using ERPs. *Clinical Neurophysiology*, *119*(7), 1546–1559. http://doi.org/10.1016/j.clinph.2008.03.003

Chung, H. J., Weyandt, L. L., & Swentosky, A. (2014). The physiology of executive functioning. In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of executive functioning* (pp. 13–27). New York, NY: Springer.

Collette, F., Hogge, M., Salmon, E., & Van der Linden, M. (2006). Exploration of the neural substrates of executive functioning by functional neuroimaging. *Neuroscience*, *139*(1), 209–221. http://doi.org/10.1016/j.neuroscience.2005.05.035

Collette, F., & Van der Linden, M. (2002). Brain imaging of the central executive component of working memory. *Neuroscience and Biobehavioral Reviews*, *26*(2), 105–125. http://doi.org/10.1016/S0149-7634(01)00063-X

Collette, F., Van der Linden, M., Laureys, S., Delfiore, G., Degueldre, C., Luxen, A., & Salmon, E. (2005). Exploring the unity and diversity of the neural substrates of executive functioning. *Human Brain Mapping*, *25*(4), 409–423. http://doi.org/10.1002/hbm.20118

Conway, A. R. A., 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. http://doi.org/10.3758/BF03196772

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*(2), 163–191. http://doi.org/10.1037/0033-2909.104.2.163

Cowan, N. (1995). *Attention and memory: An integrated framework.* New York, NY: Oxford University Press.

Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control*. (pp. 62–101) Cambridge, UK: Cambridge University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, *24*(1), 87–185. http://doi.org/10.1017/S0140525X01003922

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

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42–100. http://doi.org/10.1016/j.cogpsych.2004.12.001

Dahlin, E., Bäckman, L., Neely, A. S., & Nyberg, L. (2009). Training of the executive component of working memory: Subcortical areas mediate transfer effects. *Restorative Neurology and Neuroscience, 27*, 405–419. http://doi.org/10.3233/RNN-2009-0492

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. http://doi.org/10.1016/S0022-5371(80)90312-6

D'Esposito, M. (2001). Functional neuroimaging of working memory. In R. Cabeza & A. Kingstone (Eds.), *Handbook of functional neuroimaging of cognition*. http://doi.org/10.1037/0894-4105.20.5.497

D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 761–772. http://doi.org/10.1098/rstb.2007.2086

D'Esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., & Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature*, *378*(6554), 279–281. http://doi.org/10.1038/378279a0

D’Esposito, M., & Postle, B. R. (2015). The Cognitive Neuroscience of Working Memory. *Annual Review of Psychology*, *66*(1), 115–142. http://doi.org/10.1146/annurev-psych-010814-015031

DeStefano, D., & LeFevre, J.-A. (2007). Cognitive load in hypertext reading: A review. *Computers in Human Behavior*, *23*(3), 1616–1641. http://doi.org/10.1016/j.chb.2005.08.012

Di Stasi, L. L., Antolí, A., Gea, M., & Cañas, J. J. (2011). A neuroergonomic approach to evaluating mental workload in hypermedia interactions. *International Journal of Industrial Ergonomics*, *41*(3), 298–304. http://doi.org/10.1016/j.ergon.2011.02.008

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. http://doi.org/10.1146/annurev-psych-113011-143750

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: Analyses and review. *Journal of Experimental Psychology. General*, *140*(4), 552–572. http://doi.org/10.1037/a0023885

Doppelmayr, M., Klimesch, W., Stadler, W., Pöllhuber, D., & Heine, C. (2002). EEG alpha power and intelligence. *Intelligence*, *30*, 289–302. http://doi.org/10.1016/S0160-2896(01)00101-5

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., & Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, *120*(11), 1883–1908. http://doi.org/10.1016/j.clinph.2009.07.045

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*(10), 475–483. http://doi.org/10.1016/S0166-2236(00)01633-7

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. http://doi.org/10.1037/a0017891

Einhäuser, W., Koch, C., & Carter, O. L. (2010). Pupil dilation betrays the timing of decisions. *Frontiers in Human Neuroscience*, *4*(18), 1–9. http://doi.org/10.3389/fnhum.2010.00018

Engel, A. K., & Fries, P. (2010). Beta-band oscillations - signalling the status quo? *Current Opinion in Neurobiology*, *20*(2), 156–165. http://doi.org/10.1016/j.conb.2010.02.015

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23. http://doi.org/10.1111/1467-8721.00160

Engle, R. W., & Kane, M. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 44, pp. 145–199). Amsterdam, The Netherlands: Academic Press.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). Cambridge, UK: Cambridge University Press.

Eriksen, C. W. (1995). The flankers task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual Cognition*, *2*(2/3), 101–118. http://doi.org/10.1080/13506289508401726

Ewing, K., & Fairclough, S. (2010). The impact of working memory load on psychophysiological measures of mental effort and motivational disposition. In D. de Waard, A. Axelsson, M. Berglund, B. Peters, & C. Weikert (Eds.), *Human factors: A system view of human,technology and organisation.* (pp. 423–433). Maastricht: Shaker Publishing.

Fairclough, S. H. (2009). Fundamentals of physiological computing. *Interacting with Computers*, *21*(1-2), 133–145. http://doi.org/10.1016/j.intcom.2008.10.011

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology*, *26*(7), 874–890. http://doi.org/10.1080/13803390490510680

Fitzsimmons, G., Drieghe, D., Weal, M. J., & Drieghe, D. (2013). On measuring the impact of hyperlinks on reading. In *ACM Web Science* (pp. 65–74). New York, NY, USA: ACM Press. http://doi.org/10.1145/2464464.2464501

Fougnie, D. (2008). The relationship between attention and working memory. In N. B. Johansen (Ed.), *New research on short-term memory* (pp. 1–45). New York, NY: Nova Science Publishers Inc.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology. General*, *133*(1), 101–135. http://doi.org/10.1037/0096-3445.133.1.101

Fu, S., & Parasuraman, R. (2008). Event-related potentials (ERPs) in neuroergonomics. In R. Parasuraman & M. Rizzo (Eds.), *Neuroergonomics. The brain at work.* (pp. 32–50). New York, NY: Oxford University Press.

Fuster, J. M. (2000). Executive frontal functions. *Experimental Brain Research*, *133*(1), 66–70. http://doi.org/10.1007/s002210000401

Galy, E., Cariou, M., & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types? *International Journal of Psychophysiology*, *83*(3), 269–275. http://doi.org/10.1016/j.ijpsycho.2011.09.023

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*(1), 31–60. http://doi.org/10.1037/0033-2909.134.1.31

Gawron, V. J. (2008). Human workload. In V. J. Gawron (Ed.), *Human performance, workload, and situational awareness measures handbook.* (2nd ed., pp. 87–230). New York, NY: Taylor & Francis.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences, 16*(2), 129–135. http://doi.org/10.1016/j.tics.2011.11.014

Gerjets, P., Walter, C., Rosenstiel, W., Bogdan, M., & Zander, T. O. (2014). Cognitive state monitoring and the design of adaptive instruction in digital environments: Lessons learned from cognitive workload assessment using a passive brain-computer interface approach. *Frontiers in Neuroscience*, *8*(385). http://doi.org/10.3389/fnins.2014.00385

Gerlic, I., & Jaušovec, N. (1999). Multimedia: Differences in cognitive processes observed with EEG. *Educational Technology, Research and Development*, *47*(3), 5–14. http://doi.org/10.1007/BF02299630

Gernsbacher, M. A. (1991). Cognitive processes and mechanisms in language comprehension: The structure building framework. In G. H. Bower (Ed.), *Psychology of learning and motivation. Advances in research and theory.* (Vol. 27, pp. 217–263). San Diego, California: Academic Press. http://doi.org/10.1016/S0079-7421(08)60125-5

Gernsbacher, M. A., & Foertsch, J. a. (1999). Three models of discourse comprehension. In S. Garrod & M. Pickering (Eds.), *Language processing* (pp. 283–299). East Sussex, UK: Psychology Press.

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. http://doi.org/10.1093/cercor/10.9.829

Gevins, A., Smith, M. E., McEvoy, C. L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, *7*(4), 374–385. http://doi.org/10.1093/cercor/7.4.374

Goldinger, S. D., & Papesh, M. H. (2012). Pupil dilation reflects the rreation and retrieval of memories. *Current Directions in Psychological Science*, *21*(2), 90–95. http://doi.org/10.1177/0963721412436811

Goldstein, S., & Naglieri, J. A. (2014). *Handbook of executive functioning*. New York, NY: Springer. http://doi.org/10.1007/978-1-4614-8106-5

Goldstein, S., Naglieri, J. A., Princiotta, D., & Otero, T. M. (2014). Introduction: A history of executive functioning as a theoretical and clinical construct. In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of executive functioning* (pp. 3–12). New York, NY: Springer.

Granholm, E., & Steinhauer, S. R. (2004). Pupillometric measures of cognitive and emotional processes. *International Journal of Psychophysiology*, *52*(1), 1–6. http://doi.org/10.1016/j.ijpsycho.2003.12.001

Gwizdka, J. (2010). Distribution of cognitive load in web search. In *Journal of the American society for information science and technology* (Vol. 61, pp. 2167–2187). New York, NY: ACM. http://doi.org/10.1002/asi.21385

Händel, B. F., Haarmeier, T., & Jensen, O. (2011). Alpha oscillations correlate with the successful inhibition of unattended stimuli. *Journal of Cognitive Neuroscience*, *23*(9), 2494–2502. http://doi.org/10.1162/jocn.2010.21557

Hanslmayr, S., Pastötter, B., Bäuml, K.-H., Gruber, S., Wimber, M., & Klimesch, W. (2008). The electrophysiological dynamics of interference during the Stroop task. *Journal of Cognitive Neuroscience*, *20*(2), 215–225. http://doi.org/10.1162/jocn.2008.20020

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam, NL: North-Holland. http://doi.org/10.1016/s0166-4115(08)62386-9

Hartmann, M., & Fischer, M. H. (2014). Pupillometry: The eyes shed fresh light on the mind. *Current Biology*, *24*(7), 281–282. http://doi.org/10.1016/j.cub.2014.02.028

Hess, E. H., & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science*, *132*(3423), 349–350. http://doi.org/10.1126/science.132.3423.349

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*(3611), 1190–1192. http://doi.org/10.1126/science.143.3611.1190

Hollender, N., Hofmann, C., Deneke, M., & Schmitz, B. (2010). Integrating cognitive load theory and concepts of human–computer interaction. *Computers in Human Behavior*, *26*(6), 1278–1288. http://doi.org/10.1016/j.chb.2010.05.031

Hong, L., Walz, J. M., & Sajda, P. (2014). Your eyes give you away: Prestimulus changes in pupil diameter correlate with poststimulus task-related EEG dynamics. *PloS One*, *9*(3). http://doi.org/10.1371/journal.pone.0091321

Iqbal, S. T., Zheng, X. S., & Bailey, B. P. (2004). Task-evoked pupillary response to mental workload in human-computer interaction. In E. Dykstra-Erickson & M. Tscheligi (Eds.), *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (pp. 1477–1480). New York, NY: ACM. http://doi.org/10.1145/985921.986094

Jacobs, J., Hwang, G., Curran, T., & Kahana, M. J. (2006). EEG oscillations and recognition memory: Theta correlates of memory retrieval and decision making. *NeuroImage*, *32*(2), 978–987. http://doi.org/10.1016/j.neuroimage.2006.02.018

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. http://doi.org/10.1080/09658211003702171

Jainta, S., & Baccino, T. (2010). Analyzing the pupil response due to increased cognitive demand: An independent component analysis study. *International Journal of Psychophysiology*, *77*(1), 1–7. http://doi.org/10.1016/j.ijpsycho.2010.03.008

Jensen, O., Gelfand, J., Kounios, J., & Lisman, J. E. (2002). Oscillations in the alpha band (9-12 Hz) increase with memory load during retention in a short-term memory task. *Cerebral Cortex*, *12*(8), 877–882. http://doi.org/10.1093/cercor/12.8.877

Jong, T. De. (2009). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science*, *38*(2), 105–134. http://doi.org/10.1007/s11251-009-9110-0

Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., & Koeppe, R. A. (1997). Verbal working memory load affects regional brain activation as measured by PET. *Journal of Cognitive Neuroscience*, *9*(4), 462–475. http://doi.org/10.1162/jocn.1997.9.4.462

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*(3), 213–233. http://doi.org/10.1007/s11065-007-9040-z

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology*, *47*(2), 310–339. http://doi.org/10.1037/h0078820

Just, M. A., Carpenter, P. a., & Miyake, A. (2003). Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics Science*, *4*(1), 56–88. http://doi.org/10.1080/14639220210159735

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583–1585. http://doi.org/10.1126/science.154.3756.1583

Kaminski, J., Brzezicka, A., & Wróbel, A. (2011). Short-term memory capacity (7 ± 2) predicted by theta to gamma cycle length ratio. *Neurobiology of Learning and Memory*, *95*, 19–23. http://doi.org/10.1016/j.nlm.2010.10.001

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. http://doi.org/10.1037/0278-7393.33.3.615

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. http://doi.org/10.3758/BF03196323

Kappenman, E. S., & Luck, S. J. (2012). *The Oxford handbook of event-related potential components*. New York, NY: Oxford University Press.

Karatekin, C., Couperus, J. W., & Marcus, D. J. (2004). Attention allocation in the dual-task paradigm as measured through behavioral and psychophysiological responses, *41*, 175–185. http://doi.org/10.1111/j.1469-8986.2003.00147.x

Karatekin, C., Marcus, D. J., & Couperus, J. W. (2007). Regulation of cognitive resources during sustained attention and working memory in 10-year-olds and adults. *Psychophysiology*, *44*(1), 128–144. http://doi.org/10.1111/j.1469-8986.2006.00477.x

Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science*, *12*, 978–990. http://doi.org/10.1111/j.1467-7687.2009.00846.x

Keinrath, C., Wriessnegger, S., Müller-Putz, G. R., & Pfurtscheller, G. (2006). Post-movement beta synchronization after kinesthetic illusion, active and passive movements. *International Journal of Psychophysiology*, *62*(2), 321–327. http://doi.org/10.1016/j.ijpsycho.2006.06.001

Kendeou, P., Van Den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research and Practice*, *29*(1), 10–16. http://doi.org/10.1111/ldrp.12025

Kim, H., & Hirtle, S. (1995). Spatial metaphors and disorientation in hypertext browsing. *Behaviour & Information Technology*, *14*(4), 239–250. http://doi.org/10.1080/01449299508914637

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. http://doi.org/10.1037//0033-295X.95.2.163

Kintsch, W. (1991). The role of knowledge in discourse comprehension: A construction-integration model. In G. E. Stelmach & P. A. Vroon (Eds.), *Advances in Psychology. Text and text processing.* (Vol. 79, pp. 107–153). Amsterdam, The Netherlands: North-Holland. http://doi.org/10.1016/S0166-4115(08)61551-4

Kintsch, W., Patel, V. L., & Ericsson, K. A. (1999). The role of long-term working memory in text comprehension. *Psychologia*, *42*, 186–198. http://doi.org/10.1.1.16.2231

Kintsch, W., & Welsch, D. M. (1991). The contruction-integration model: A framework for studying memory for text. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory* (pp. 367–386). Hillsdale, NJ: Erlbaum.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, *29*(2-3), 169–195. http://doi.org/10.1016/S0165-0173(98)00056-3

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*(12), 606–617. http://doi.org/10.1016/j.tics.2012.10.007

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Research Reviews*, *53*, 63–88. http://doi.org/10.1016/j.brainresrev.2006.06.003

Klimesch, W., Schack, B., & Sauseng, P. (2005). The functional significance of theta and upper alpha oscillations. *Experimental Psychology*, *52*(1), 1–10. http://doi.org/10.1027/1618-3169.52.2.99

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

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, *38*(3), 557–577. http://doi.org/10.1017/S0048577201990559

Krause, C. M. (2003). Brain electric oscillations and cognitive processes. In K. Hugdahl (Ed.), *Neuropsychology and Cognition. Experimental methods in neuropsychology* (21st ed., pp. 111–130). Boston, MA: Kluwer Academic Publishers Group.

Krause, C. M., Pesonen, M., & Hämäläinen, H. (2010). Brain oscillatory 4-30 Hz electroencephalogram responses in adolescents during a visual memory task. *Neuroreport*, *21*(11), 767–771. http://doi.org/10.1097/WNR.0b013e32833bfcbb

Krause, C. M., Sillanmäki, L., Koivisto, M., Saarela, C., Häggqvist, A., Laine, M., & Hämäläinen, H. (2000). The effects of memory load on event-related EEG desynchronization and synchronization. *Clinical Neurophysiology*, *111*(11), 2071–2078. http://doi.org/10.1016/S1388-2457(00)00429-6

Kretzschmar, F., Pleimling, D., Hosemann, J., Füssel, S., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Subjective impressions do not mirror online reading effort: Concurrent EEG-eyetracking evidence from the reading of books and digital media. *PloS ONE*, *8*(2), e56178. http://doi.org/10.1371/journal.pone.0056178

Kuipers, J.-R., & Thierry, G. (2010). N400 amplitude reduction correlates with an increase in pupil size. *Frontiers in Human Neuroscience*, *5*(61), 1–5. http://doi.org/10.3389/fnhum.2011.00061

Kuipers, J.-R., & Thierry, G. (2013). ERP-pupil size correlations reveal how bilingualism enhances cognitive flexibility. *Cortex*, *49*(10), 2853–2860. http://doi.org/10.1016/j.cortex.2013.01.012

Kutas, M., Mccarthy, G., & Donchin, E. (2007). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science*, *197*(4305), 792–795. http://doi.org/10.1126/science.887923

Laeng, B., Ørbo, M., Holmlund, T., & Miozzo, M. (2011). Pupillary Stroop effects. *Cognitive Processing*, *12*(1), 13–21. http://doi.org/10.1007/s10339-010-0370-z

Laeng, B., Sirois, S., & Gredeback, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. http://doi.org/10.1177/1745691611427305

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. http://doi.org/10.1016/j.tics.2004.12.004

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science, 19*(3), 143–148. http://doi.org/10.1177/0963721410370295

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. http://doi.org/10.1037/0096-3445.133.3.339

Levens, S. M., & Gotlib, I. H. (2010). Updating positive and negative stimuli in working memory in depression. *Journal of Experimental Psychology*. *General*, *139*(4), 654–664. http://doi.org/10.1037/a0020283

Linden, D. E. J. (2005). The P300: Where in the brain is it produced and what does it tell us? *The Neuroscientist*, *11*(6), 563–576. http://doi.org/10.1177/1073858405280524

Lisman, J. E., & Jensen, O. (2013). The theta-gamma neural code. *Neuron*, *77*(6), 1002–1016. http://doi.org/10.1016/j.neuron.2013.03.007

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163–203. http://doi.org/10.1037/0033-2909.109.2.163

MacLeod, C. M. (2007). The concept of inhibition in cognition. In D. S. Gorfein & C. M. MacLeod (Eds.), *Inhibition in cognition* (pp. 3–23). Washington, DC: American Psychological Association. http://doi.org/10.1037/11587-000

McEvoy, L. K., Pellouchoud, E., Smith, M. E., & Gevins, A. (2001). Neurophysiological signals of working memory in normal aging. *Brain Research. Cognitive Brain Research, 11*(3), 363–376. http://doi.org/10.1016/S0926-6410(01)00009-X

Mitchell, D. G. V, Nakic, M., Fridberg, D., Kamel, N., Pine, D. S., & Blair, R. J. R. (2007). The impact of processing load on emotion. *Human Brain Mapping Journal*, *34*(3), 1299–1309. http://doi.org/10.1016/j.neuroimage.2006.10.012

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. http://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (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. http://doi.org/10.1006/cogp.1999.0734

Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control.* Cambridge, UK: Cambridge University Press.

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

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review, 18*, 46–60. http://doi.org/10.3758/s13423-010-0034-0

Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'Connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology, 48*(11), 1532–1543. http://doi.org/10.1111/j.1469-8986.2011.01226.x

Nee, D. E., Brown, J. W., & Askren, M. K. (2013). A meta-analysis of executive components of working memory. *Cerebral Cortex, 23*, 264–282. http://doi.org/10.1093/cercor/bhs007

Nigbur, R., Ivanova, G., & Stürmer, B. (2011). Theta power as a marker for cognitive interference. *Clinical Neurophysiology*, *122*(11), 2185–2194. http://doi.org/10.1016/j.clinph.2011.03.030

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, *126*(2), 220–246. http://doi.org/10.1037/0033-2909.126.2.220

Norman, D., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation.* http://doi.org/10.1098/rstb.1996.0132

Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation, 51*, 45–100. http://doi.org/10.1016/S0079-7421(09)51002-X

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

Oliveira, F. T. P., Aula, A., & Russell, D. M. (2009). Discriminating the relevance of web search results with measures of pupil size. In *Proceedings of the 27th international conference on human factors in computing systems - CHI 2009* (pp. 2209–2212). New York, NY: ACM Press. http://doi.org/10.1145/1518701.1519038

Osaka, N., Logie, R. H., & D'Esposito, M. (2007). *The cognitive neuroscience of working memory.* New York, NY: Oxford University Press.

Otero, T. M., & Barker, L. A. (2014). The frontal lobes and executive functioning. In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of executive functioning* (pp. 29–44). New York, NY: Springer.

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. http://doi.org/10.1002/hbm.20131

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*(1), 63–71. http://doi.org/10.1207/S15326985EP3801_8

Palomäki, J., Kivikangas, M., Alafuzoff, A., Hakala, T., & Krause, C. M. (2012). Brain oscillatory 4-35 Hz EEG responses during an n-back task with complex visual stimuli. *Neuroscience Letters*, *516*(1), 141–145. http://doi.org/10.1016/j.neulet.2012.03.076

Palva, S., & Palva, J. M. (2007). New vistas for alpha-frequency band oscillations. *Trends in Neurosciences*, *30*(4), 150–158. http://doi.org/10.1016/j.tins.2007.02.001

Parasuraman, R. (2003). Neuroergonomics: Research and practice. *Theoretical Issues in Ergonomics Science*, *4*(1-2), 5–20. http://doi.org/10.1080/14639220210199753

Parasuraman, R. (2011). Neuroergonomics: Brain, cognition, and performance at work. *Current Directions in Psychological Science*, *20*(3), 181–186. http://doi.org/10.1177/0963721411409176

Parasuraman, R., & Wilson, G. F. (2008). Putting the brain to work: Neuroergonomics past, present, and future. *Human Factors*, *50*(3), 468–474. http://doi.org/10.1518/001872008X288349.

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, *59*(1-2), 185–198. http://doi.org/10.1016/S1071-5819(03)00017-X

Payne, L., & Sekuler, R. (2014). The importance of ignoring: Alpha oscillations protect selectivity. *Current Directions in Psychological Science*, *23*(3), 171–177. http://doi.org/10.1177/0963721414529145

Peavler, W. (1974). Pupil size, information overload, and performance differences. *Psychophysiology*, *11*(5), 559–566. http://doi.org/10.1111/j.1469-8986.1974.tb01114.x

Pesonen, M., Hämäläinen, H., & Krause, C. M. (2007). Brain oscillatory 4-30 Hz responses during a visual n-back memory task with varying memory load. *Brain Research*, *1138*, 171–177. http://doi.org/10.1016/j.brainres.2006.12.076

Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences*, *13*(4), 160–166. http://doi.org/10.1016/j.tics.2009.01.006

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, *21*, 73–89. http://doi.org/10.1146/annurev-neuro-062111-150525

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology*, *110*(11), 1842–1857. http://doi.org/10.1016/S1388-2457(99)00141-8

Pfurtscheller, G., Stancák, A., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band - an electrophysiological correlate of cortical idling: A review. *International Journal of Psychophysiology*, *24*(1-2), 39–46. http://doi.org/10.1016/S0167-8760(96)00066-9

Pfurtscheller, G., Zalaudek, K., & Neuper, C. (1998). Event-related beta synchronization after wrist, finger and thumb movement. *Electroencephalography and Clinical Neurophysiology*, *109*(2), 154–160. http://doi.org/10.1016/S0924-980X(97)00070-2

Plass, J. L., Moreno, R., & Brünken, R. (2010). *Cognitive load theory*. Cambridge, UK: Cambridge University Press.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. http://doi.org/10.1016/j.clinph.2007.04.019

Polich, J. (2012). Neuropsychology of P300. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 1–36). Oxford, UK: Oxford University Press. http://doi.org/10.1093/oxfordhb/9780195374148.013.0089

Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology*, *58*, 1–23. http://doi.org/10.1146/annurev.psych.58.110405.085516

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*(1), 23–38. http://doi.org/10.1016/j.neuroscience.2005.06.005

Privitera, C. M., Renninger, L. W., Carney, T., Klein, S., & Aguilar, M. (2010). Pupil dilation during visual target detection. *Journal of Vision*, *10*(10), 1–14. http://doi.org/10.1167/10.10.3

Rabipour, S., & Raz, A. (2012). Training the brain: Fact and fad in cognitive and behavioral remediation. *Brain and Cognition*, *79*, 159–179. http://doi.org/10.1016/j.bandc.2012.02.006

Redick, T. S., & Lindsey, D. R. B. (2013). Complex span and n-back measures of working memory: A meta-analysis. *Psychonomic Bulletin & Review*, *20*(6), 1102–1113. http://doi.org/10.3758/s13423-013-0453-9

Roux, F., & Uhlhaas, P. J. (2014). Working memory and neural oscillations: Alpha-gamma versus theta-gamma codes for distinct WM information? *Trends in Cognitive Sciences*, *18*(1), 16–25. http://doi.org/10.1016/j.tics.2013.10.010

Royall, D. R., Lauterbach, E. C. E., Cummings, J. J. L., Reeve, A., Rummans, T. A., Kaufer, D. I., & Coffey, C. E. (2002). Executive control function: A review of its promise and challenges for clinical research. A report from the committee on research of the

American Neuropsychiatric Association. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *14*, 377–405. http://doi.org/10.1176/appi.neuropsych.14.4.377

Salmerón, L., Cerdán, R., & Naumann, J. (2015). How adolescents navigate Wikipedia to answer questions / ¿Cómo navegan los adolescentes en Wikipedia para contestar preguntas? *Infancia Y Aprendizaje*, 1–37. http://doi.org/10.1080/02103702.2015.1016750

Salminen, T., Strobach, T., & Schubert, T. (2012). On the impacts of working memory training on executive functioning. *Frontiers in Human Neuroscience*, *6*(166), 1–14. http://doi.org/10.3389/fnhum.2012.00166

Sanders, A., & Lamers, J. (2002). The Eriksen flanker effect revisited. *Acta Psychologica*, *109*(1), 41–56. http://doi.org/10.1016/S0001-6918(01)00048-8

Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: A possible function of EEG theta oscillations. *Neuroscience and Biobehavioral Reviews*, *34*(7), 1015–1022. http://doi.org/10.1016/j.neubiorev.2009.12.006

Sauseng, P., & Klimesch, W. (2008). What does phase information of oscillatory brain activity tell us about cognitive processes? *Neuroscience and Biobehavioral Reviews*, *32*, 1001–1013. http://doi.org/10.1016/j.neubiorev.2008.03.014

Sauseng, P., Klimesch, W., Freunberger, R., Pecherstorfer, T., Hanslmayr, S., & Doppelmayr, M. (2006). Relevance of EEG alpha and theta oscillations during task switching. *Experimental Brain Research*, *170*(3), 295–301. http://doi.org/10.1007/s00221-005-021 l-y

Sauseng, P., Klimesch, W., Schabus, M., & Doppelmayr, M. (2005). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. *International Journal of Psychophysiology*, *57*(2), 97–103. http://doi.org/10.1016/j.ijpsycho.2005.03.018

Scheiter, K., & Gerjets, P. (2007). Learner control in hypermedia environments. *Educational Psychology Review*, *19*(3), 285–307. http://doi.org/10.1007/s10648-007-9046-3

Schmiedek, F., Hildebrandt, A., Lövdén, M., Lindenberger, U., & Wilhelm, O. (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. http://doi.org/10.1037/a0015730

Schüler, A., Scheiter, K., & Genuchten, E. Van. (2011). The role of working memory in multimedia instruction: Is working memory working during learning from text and pictures? *Educational Psychology Review*, *23*, 389–411. http://doi.org/10.1007/s10648-011-9168-5

Schultheis, H., & Jameson, A. (2004). Assessing cognitive load in adaptive hypermedia systems: Physiological and behavioral methods. In P. M. E. De Bra & W. Nejdl (Eds.), *Adaptive hypermedia and adaptive web-based systems. Lecture notes in computer science.* (Vol. 3137, pp. 225–234). Berlin: Springer. http://doi.org/10.1007/b99480

Sevilla, Y., Maldonado, M., & Shalóm, D. E. (2014). Pupillary dynamics reveal computational cost in sentence planning. *Quarterly Journal of Experimental Psychology*, *67*(6), 1041–1052. http://doi.org/10.1080/17470218.2014.911925

Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica*, *50*(3&4), 245–276.

Siegle, G. J., Steinhauer, S. R., & Thase, M. E. (2004). Pupillary assessment and computational modeling of the Stroop task in depression. *International Journal of Psychophysiology*, *52*(1), 63–76. http://doi.org/10.1016/j.ijpsycho.2003.12.010

Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology*, *33*(1), 5–42. http://doi.org/10.1006/cogp.1997.0658

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*, 1657–1661. http://doi.org/10.1126/science.283.5408.1657

Smith, M. E., Gevins, A., Brown, H., Karnik, A., & Du, R. (2001). Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Human Factors*, *43*(3), 366–380. http://doi.org/10.1518/001872001775898287

St Clair-Thompson, H. L. (2007). The effects of cognitive demand upon relationships between working memory and cognitive skills. *Quarterly Journal of Experimental Psychology*, *60*(10), 1378–1388. http://doi.org/10.1080/17470210601025505

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology*, *59*(4), 745–759. http://doi.org/10.1080/17470210500162854

Stipacek, A., Grabner, R. H., Neuper, C., Fink, A., & Neubauer, A. C. (2003). Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load. *Neuroscience Letters*, *353*(3), 193–196. http://doi.org/10.1016/j.neulet.2003.09.044

Strobach, T. (2014). *Executive functions modulated by context, training, and age*. *Habilitation*. Humboldt-Universität zu Berlin.

Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. http://doi.org/10.1037/h0054651

Stuss, D. T. (1992). Biological and psychological development of executive functions. *Brain and Cognition*, *20*, 8–23. http://doi.org/10.1016/0278-2626(92)90059-U

Suchy, Y. (2009). Executive functioning: Overview, assessment, and research issues for non-neuropsychologists. *Annals of Behavioral Medicine*, *37*(2), 106–116. http://doi.org/10.1007/s12160-009-9097-4

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*, 295–312. http://doi.org/10.1016/0959-4752(94)90003-5

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296. http://doi.org/10.1023/a:1022193728205

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://doi.org/10.1037/a0020365

Titz, C., & Karbach, J. (2014). Working memory and executive functions: Effects of training on academic achievement. *Psychological Research*, *78*(537), 1–17. http://doi.org/10.1007/s00426-013-0537-1

Turner, M., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154. http://doi.org/10.1016/0749-596X(89)90040-5

Van Bochove, M. E., Van der Haegen, L., Notebaert, W., & Verguts, T. (2013). Blinking predicts enhanced cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(2), 346–354. http://doi.org/10.3758/s13415-012-0138-2.

Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, *41*(2), 167–174. http://doi.org/10.1111/j.1469-8986.2003.00148.x

Vandierendonck, A. (2014). Symbiosis of executive and selective attention in working memory. *Frontiers in Human Neuroscience*, *8*(588). http://doi.org/10.3389/fnhum.2014.00588

Verhaeghen, P. (2011). Aging and executive control: Reports of a demise greatly exaggerated. *Current Directions in Psychological Science*, *20*, 174–180. http://doi.org/10.1177/0963721411408772

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: A meta-analysis. *Cognitive, Affective & Behavioral Neuroscience*, *3*(4), 255–274. http://doi.org/10.3758/CABN.3.4.255

Watter, S., Geffen, G. M., & Geffen, L. B. (2001). The n-back as a dual-task: P300 morphology under divided attention. *Psychophysiology*, *38*(6), 998–1003. http://doi.org/10.1111/1469-8986.3860998

Weiss, S., & Mueller, H. M. (2012). “Too many betas do not spoil the broth”: The role of beta brain oscillations in language processing. *Frontiers in Psychology*, *3*(201). http://doi.org/10.3389/fpsyg.2012.00201

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. http://doi.org/10.1080/14639220210123806

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*(3), 449–455. http://doi.org/10.1518/001872008X288394

Wickens, C., Kramer, A., Vanasse, L., & Donchin, E. (1983). Performance of concurrent tasks: A psychophysiological analysis of the reciprocity of information-processing resources. *Science*, *221*(4615), 1080–1082. http://doi.org/10.1126/science.6879207

Wykowska, A., Anderl, C., Schubö, A., & Hommel, B. (2013). Motivation modulates visual attention: Evidence from pupillometry. *Frontiers in Psychology*, *4*(59). http://doi.org/10.3389/fpsyg.2013.00059

Yuan, K., Steedle, J., Shavelson, R., Alonzo, A., & Oppezzo, M. (2006). Working memory, fluid intelligence, and science learning. *Educational Research Review*, *1*(2), 83–98. http://doi.org/10.1016/j.edurev.2006.08.005

Zander, T. O., & Kothe, C. (2011). Towards passive BCI: Applying BCI technology for human-machine systems in general. *Journal of Neural Engineering*, *8*(2), 1–10.

Zimmer, H. D. (2008). Visual and spatial working memory: From boxes to networks. *Neuroscience and Biobehavioral Reviews*, *32*(8), 1373–1395. http://doi.org/10.1016/j.neubiorev.2008.05.016

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–185. http://doi.org/10.1037/0033-2909.123.2.162