Cognitive training in younger and older adults: effects on brain and behavior

Buerki, Céline Nadja

UNIVERSITÉ DE GENÈVE
FACULTÉ DE PSYCHOLOGIE ET DES SCIENCES DE L'ÉDUCATION

Section de Psychologie

Sous la direction de la Professeure Anik de Ribaupierre

Cognitive Training in Younger and Older Adults: Effects on Brain and Behavior

Thèse

présentée à la
Faculté de Psychologie et des Sciences de l'Éducation
de l'Université de Genève
pour obtenir le grade de

Docteur en Psychologie

par

Céline Nadja BÜRKI

de

Berne

Thèse No 513

Genève
Septembre 2012

No d'étudiant 00-102-483

UNIVERSITÉ
DE GENÈVE
FACULTÉ DE PSYCHOLOGIE
ET DES SCIENCES DE L'ÉDUCATION

Doctorat en psychologie

Thèse de Céline BUERKI

Intitulée : « Cognitive training in younger and older adults : Effects on brain and behavior »

La Faculté de psychologie et des sciences de l'éducation, sur préavis d'une commission formée par les professeurs : Anik de Ribaupierre, directrice, FPSE, Université de Genève ; Claude-Alain Hauert, FPSE, Université de Genève ; Catherine Ludwig, FPSE, Université de Genève ; Christian Chicherio, FPSE, Université de Genève ; Matthias Kliegel, FPSE, Université de Genève ; Mike Martin, Psychologisches Institut, Universität Zürich

autorise l'impression de la présente thèse, sans prétendre par là émettre d'opinion sur les propositions qui y sont énoncées.

GENEVE, le 21 septembre 2012

Le doyen :

Pascal Zesiger

Thèse No 513
Numéro d'immatriculation : 00.102.483

Uni Mail
40 bd du Pont-d'Arve - CH-1211 Genève 4
Tél. 022 379 90 01 - Fax 022 379 90 20
www.unige.ch/fapse
Doyen-FPSE@unige.ch

ACKNOWLEDGEMENTS

I would like to express my thanks in the first line to my thesis advisor Professor Anik de Ribaupierre who gave me the opportunity to work with her and her team, for the many stimulating discussions, and her support and interest in my work. I am also very grateful to Catherine Ludwig and Christian Chicherio for supporting my thesis with their expertise, for all their help, advice and encouragements over the last few years, and for being in my thesis committee and jury. Professor Claude-Alain Hauert deserves my gratitude for being in my thesis committee and jury, the stimulating exchanges and the insights into the electrophysiological aspect of my work. I would also like to thank Professor Matthias Kliegel and Professor Mike Martin for their interest in my research and for accepting to be members of my jury. An additional thank goes to Matthias Kliegel who gave me the opportunity to work with him and his team during the last few months of my thesis and the stimulating discussions during this time.

I would like to thank Andres Posada, who introduced me into EEG recording and provided support for many issues, from Eprime programming to troubling electrodes. I am also very grateful to Clara James who introduced me to the raw data processing as well as to microstate analyses. I also want to thank the EEG lab meeting members, in particular Joseph Studer, Lore Legrand, Nicolas Burra, and Raoul Jenni, who went through the ups and downs of EEG recording and cleaning with me. I also thank Philippe Golay and Paolo Ghisletta for the introduction into the application of latent growth curve models.

Moreover, I am grateful for the financial support provided by the Hedwig Widmer Stiftung Zürich. I also want to thank all the participants who spent many hours in the lab and without them the present work would not have been possible. Then, my thank goes to the master students Aurélia, Lisa, Yeniffer, and Zahra for their help with the data collection.

Many thanks also go to Judith Dirk, Alexandra Hering, Walter Isler and Ralph Sonderegger, who proof-read parts of this work and to Isabelle Reverte who corrected the French summary.

A big thank you goes to my former and current colleagues, Alexandra, Catherine, Célia, Christelle, Christian, Delphine, Jessica, Judith, Katharina, Manuel, Philippe, Sébastien, and Thierry, as well as to friends I met in Geneva, at the University and beyond, in particular to Anne-Laure, Isabelle, Joana, Joanna, Joseph, Kerstin, Laure, Mathias, Sascha, and to the Coloc rue Necker and satellites, for the multiple ways in which they contributed to this work, from professional discussions to social support. Moreover, I would like to thank my friends in the other parts of Switzerland, who followed my dissertation from a distance, but were nevertheless always present with moral support and encouragement.

Finally, I am deeply grateful to my parents as well as to my brothers Pascal and Lucien, who encouraged me to pursue this doctoral thesis, and who supported me in manifold ways during the last years and especially the last few months.

Most importantly, I am thankful to Gregor, for always being there for me, for the numerous stimulating discussions, for his support, and understanding.

ABSTRACT

Cognitive functions remain plastic in older age. However, they tend to decline with advancing age, even though this does not apply to all functions to the same degree or at the same time. With the present work, we aimed at providing an important cornerstone for the deeper understanding of cognitive aging and the development of means to foster autonomy in older age. We investigated the behavioral and cerebral mechanisms underlying a cognitive training intervention.

The purpose of the present work was first to better understand the behavioral plastic potential in younger and older adults by implementing a working memory (WM) training program. Second, we intended to aliment the current discussion about transfer effects to untrained tasks, whose existence, degree and extent are not clear at present. Third, we aimed at gaining insights into the underlying cerebral plastic mechanisms of training effects and, to our knowledge for the first time, of transfer effects. To this end, we recorded electroencephalogram (EEG) measures at pre-test and post-test on the trained and a near transfer task. By this means, we were finally able to contribute to the debate about the functional meaning of the cerebral changes during cognitive aging, i.e., the frontal over-activation, which naturally occurs in older age and which is still not fully understood. We further aimed at improving several methodological issues that are not systematically respected in training studies to date and do therefore not allow to fully understand the effects of training. First, a crucial ingredient was the inclusion of a younger group which served as an optimal reference group for older adults. Second, in order to provide a control for retest and placebo effects occurring within a training intervention, we included both a no-contact as well as a placebo training group.

We expected to find plastic changes in both age groups during the training, but more pronounced in younger as compared to older adults. We hypothesized similarly to find effects of training in the untrained near transfer tasks measuring WM capacity. Furthermore, we examined whether WM training generates transfer in a fluid intelligence measure, a hypothesis which has gained mixed support. These effects were expected to be found in both WM training groups, but more pronounced in younger as in older adults. On the cerebral level, we hypothesized changes in components of event-related potentials (ERP) which are related to decision-making and attentional resource allocation processes, the N2 and the P3 component. We expected them to decrease in amplitude due to training since we hypothesized that less attentional resources are needed after training to respond to the trained task and the

untrained near transfer task. Additionally, we predicted to find less frontally oriented voltage maps for the N2 and the P3 components. Amplitude in N2 and P3 is usually decreased and less frontally oriented in cognitively lower demanding tasks as compared to higher load tasks.

We implemented a WM training using a verbal *N*-back training procedure during 10 daily sessions of 30 minutes of training in the laboratory. Before and after training, we conducted a large battery of tests in order to evaluate the changes from pre-test to post-test. The battery of cognitive tests included measures of WM, inhibition, processing speed and fluid intelligence. In order to disentangle the training effects from placebo effects, we included an implicit task training in a similar procedure with a duration of 10 sessions about 30 minutes of training per day. Furthermore, we included a no-contact control group, which did not complete a training intervention between the assessment of the cognitive test battery at pre-test and post-test. EEG measures were recorded at pre-test and post-test during the trained verbal task and during a spatial near transfer task.

Results at the behavioral level revealed that the performance of both age groups improved over the 10 WM training sessions. Younger adults exhibited a generally higher training level than older adults and improved faster during training by reaching the asymptote later than older adults. This resulted in a magnified age difference at the end of the training. Further analyses showed that fluid intelligence performance explained individual differences in initial training performance independent of age group. Age group, in turn, accounted for individual differences in the growth curve beyond individual differences in fluid intelligence. As regards training effects, a clear effect of training in the younger and older WM training group was observed for the trained verbal 2-back task as compared to both control groups. The WM load cost, i.e., the difference between the low load 0-back condition and the higher load 2-back condition, also decreased significantly for both WM training groups. Moreover, a near transfer effect was found in the spatial *N*-back task, which was present for the 2-back task as well as for the WM load costs. Similar to the verbal task, the effects were comparable for both age groups. In terms of far transfer effects, there was an effect in the Stroop task, which was also present in the placebo control group. No further transfer effects were found for other WM measures, the fluid intelligence measure or for the processing speed tasks.

The ERP analyses for the verbal 2-back task revealed that both age groups as compared to the control groups showed an activation increase for the N2 and, as expected, an activation decrease for the P3 component. This was reflected by a change from a frontal positivity towards a central negativity and a posterior positivity. Overall, the activation became less distributed over the scalp, tending towards a more selective recruitment in central and posterior regions. These changes were found to describe a reorganization pattern. We observed the inverse pattern for the untrained spatial *N*-back task: The frontal sites showed more positivity in the P3 component after training for the WM training groups whereas the posteriorly oriented maps decreased in presence. This resulted in a redistribution pattern, indicating that similar processes were engaged after training contrary to the verbal task, in which the reorganization pattern showed a change in processes.

The results of the training performance were in line with our predictions, since age differences were increased after training. As regards the findings from training and transfer tasks, we found the predicted gains in the trained and a near transfer task (the spatial *N*-back task), but not in additional WM tasks or the fluid intelligence measure. These findings are in line with a growing body of literature which supports the evidence of preserved, even though attenuated, behavioral plasticity in older age. The small transfer effects in younger as well as in older adults also provided support for the currently mixed body of evidence. Furthermore, transfer effects from WM training to a fluid intelligence measure are currently under debate, since results are even more contradictory. The ERP results revealed that cerebral plasticity was also preserved in older adults and that a similar change in younger and older adults was observed. We were able to link our findings with the scaffolding theory, which states that frontal areas are more activated and therefore attentional resources are preferably recruited in order to respond to the increased demands during learning. When these demands persist, the processes become more efficient which is reflected in decreasing frontal attentional recruitment. The former pattern was found in the near transfer task where frontal recruitment was observed. For the trained task in turn, the latter pattern, that is, decreased frontal recruitment (P3) and a more efficient early process (N2), was found. The results provided additionally evidence regarding the debate concerning the interpretation of age-related cerebral changes. They revealed that a similar change in older adults as compared to younger adults in cerebral processes seems to be beneficial. That is, we were not able to find a differential change due to training in younger and older adults. Thus, we concluded that the frontal over-activation in older as compared to younger adults does not necessarily describe a beneficial functioning.

Taken together, the present dissertation presents evidence that training studies have a great potential to contribute to the understanding of the mechanisms of cognitive aging. Moreover, our findings call for the integration of training studies in more applied contexts by targeting on effects that might transfer into older adults' everyday life.

TABLE OF CONTENTS

INTRODUCTION

> *(...) knowledge about the plasticity of developmental trajectories is essential for improving human welfare. Hence, investigations of age changes in the plasticity of development carry the potential to explain and ameliorate human development.*
>
> (Lindenberger, Li, Lövdén, & Schmiedek, 2007, p. 234)

Life expectancy at birth has increased dramatically over the last century. As a result, the proportion of older adults, i.e., the proportion of people at the age of retirement, within the total population rose considerably. In Europe and North America, this phenomenon will be accentuated in upcoming decades by other demographic developments. In particular, baby boomers are now entering the age of retirement (see e.g., Pruchno, 2012). Related to this rise in the proportion of older people in the total population, society is confronted with new social, psychological and economic challenges, and becomes now increasingly aware of the enormous cost that a senescent population induces for its health care systems (WHO, 2012).

From the point of view of cognitive psychology, an aging population raises a number of questions due to the fact that the cognitive functioning as well as brain structure and function of older adults differ in several ways from that of younger adults. These differences become evident in multiple cognitive dimensions which follow different developmental directions, but in general, a cognitive decline in older age is observed (de Ribaupierre, 2005; de Ribaupierre, Fagot, & Dirk, 2009).

Indeed, as Baltes (1997) argues, neither human biology nor culture is now fully developed for the later phases of life. Therefore, the dynamic interaction between biological and cultural factors leads to an age-related growing incompleteness of the architecture of human ontogeny (Baltes, Lindenberger, & Staudinger, 1998, 2006). It is argued that, as a result of the so-called selection neutrality at older age - i.e., where no evolutionary selection applies to the development - the benefits of evolution decrease with advancing age. In other words, as older people are not under fitness based evolutionary pressure, evolution does not lead to improved biological architecture. This incompleteness leads, in turn, to an increasing demand for social, material and psychological resources compensating for age-related losses.

The efficiency of culture and therefore the provision of compensation are reduced in older age. Culture-based resources seem to be the only way to counter age-related declines so far. It is a central aim of a senescent society to foster the maintenance of health, cognitive functioning and therefore autonomy in older people. Cognitive training interventions, such as the one presented in this study, constitute an important building block to attain these objectives.

This brings us to the concept of plasticity which has been proposed "(…) to overcome limitations of the genome and adapt to the rapidly changing environment" (A. Pascual-Leone et al., 2011, p. 1). Cognitive plasticity is of central interest in cognitive training research and refers to the ability to adapt and change behavior as a response to experience. In a seminal work, Baltes (1987) identified the phenomenon of cognitive plasticity as a way of altering developmental trajectories until old age. He and his collaborators showed that cognitive plasticity is continued in older age but limited compared to younger age (Baltes & Lindenberger, 1988; Kliegl, Smith, & Baltes, 1989). His contribution initiated a still growing body of literature investigating the cognitive plasticity and its potential in cognitive training interventions in old age (for recent reviews see Hertzog, Kramer, Wilson, & Lindenberger, 2009; Martin, Clare, Altgassen, Cameron, & Zehnder, 2011).

Crucial for cognitive training interventions is further its resulting benefits in untrained situations and even in everyday life. The so-called transfer effects refer to gains in untrained tasks that are exclusively caused by the cognitive training intervention (e.g., Willis & Schaie, 2009). However, the growing amount of evidence based on training studies remains inconclusive. Whilst most studies confirmed the plastic changes in younger and older adults, the magnitude or even existence of transfer effects is still unclear, and especially so in older age (H. Li et al., 2011; Melby-Lervåg & Hulme, 2012; Morrison & Chein, 2011; Zehnder, Martin, Altgassen, & Clare, 2009).

The study of brain plasticity, i.e., the adaptation at the cerebral level in reaction to experience, has as well gained growing interest in recent years. Brain plasticity has mainly been investigated by comparing the cerebral organization and structure in experts to novices (e.g., Jäncke, 2009). However, a small number of studies investigated brain changes following a training procedure (e.g., Draganski et al., 2004). Such training studies allow a causal conclusion about the plastic changes, contrary to expert studies. This is an important point to further understand the mechanisms of cerebral plasticity especially in older adults.

Motivated by the apparent discrepancy between society's needs for culture-based resources and the state of scientific knowledge in this respect, we attempt to contribute to a deeper understanding of plasticity in older age by combining the behavioral examination with an analysis at the cerebral level.

Presently, only a few studies have examined training-induced changes at both the behavioral and cerebral level. In particular, there is very little evidence on whether and how cognitive and cerebral changes occur due to a training intervention with respect to age differences in such potential changes. However, the inclusion of a younger group, which is at its most efficient state of functioning, is essential for the interpretation of the changes occurring in older adults. The combined study of behavioral and cerebral plasticity in an age-comparative study is a promising way to better understand the age-related mechanisms in relation with a training intervention.

The combined study of both approaches by including a younger and an older sample in an intervention study is promising for at least two reasons. First, we may investigate not only to what extent a particular training procedure is effective at the behavioral level in older adults as compared to younger adults, but we can also analyze more closely the cerebral processes that are behind the success or failure of the training intervention. Second, the approach allows us to shed light on another important and unresolved issue in neurocognitive aging research - the functional cerebral reorganization process, i.e., the frontal overrecruitment, which naturally occurs in older age, is still not fully understood. Different interpretations are discussed, but today no consensus about the meaning of this reorganization exists. Existing evidence for explaining the functional reorganization is based solely on cross-sectional studies. Therefore, the cerebral approach for investigating age differences in training-induced changes is a promising avenue of research in this respect. By manipulating experimentally the age-related cerebral reorganization, relevant new insights into the interpretation of this reorganization could be gained.

With respect to these aims, we implemented a 10-day working memory training procedure in a sample of younger and older adults. A large battery of tests including electrophysiological brain measures before and after the training intervention was completed. Furthermore, to control for placebo and test-retest effects, we included a passive and, in contrast to many training studies, also an active control group. We considered the inclusion of a no-contact control group as well as an active control group to be essential for deriving reliable conclusions. With this study design, we were able to control for confounding placebo effects, such as social interaction and adaptation to the specifics of the training setting, in addition to retest effects.

The manuscript is organized into the following parts; first, in the *Conceptual Background* section, we provide a literature review and describe the important open issues in detail. Then, we formulate our precise objectives and hypothesis in the section *Aims and*

Hypotheses. This theoretical part is followed by a description of our methodological approach (*Method*). Subsequently, we present the results (*Results*) and discuss them in relation to our hypothesis and finally, we provide conclusions and perspectives for future research (*Discussion*).

CONCEPTUAL BACKGROUND

The conceptual background section aims at providing a review of the relevant literature within the domains of behavioral and cerebral aging and of cognitive and brain plasticity in the context of cognitive training in order to point out the open issues. It is organized in two main sections as follows:

In the initial section, different perspectives on cognitive aging are provided. First, at the behavioral level, classic theories of cognitive aging and lifespan development are introduced. Second, age-related cognitive changes at the structural and functional brain level are addressed. Finally, a review of the more recent research combining both levels of analysis is provided. That is, different hypotheses regarding the functional meaning of age-related brain changes are discussed. At the end of the first section, we come to the conclusion that there is currently still no consensus on how to interpret the cerebral over-activation pattern observed in older adults and whether it is beneficial or not in older age. As a consequence, we proposed to contribute to this open question through the experimental manipulation of the behavioral performance and cerebral functioning by means of a cognitive training intervention. This intervention would allow us to examine age-related differences in plasticity.

The second section covers the relevant literature on cognitive and cerebral plasticity within cognitive training studies. First, we provide deeper insight into cognitive plasticity research with a particular focus on process-based working memory training interventions in younger and older adults. Second, we address the plastic potential of brain activity due to training. Here we focus again on age-comparative and working memory studies in order to gain more insights into age-differences in brain plasticity. Finally, we provide an excursus on the functional neuroimaging method, electroencephalography (EEG) in an event-related design, which we implemented in the present study to measure brain plasticity.

PERSPECTIVES ON COGNITIVE AGING

Cognitive Aging at the Behavioral Level

Complex cognitive abilities like memory, language, perception and attention are used in practically all daily activities and therefore play a central role in our everyday life. Consequently, healthy cognitive functioning constitutes a key ingredient for an individual's well-being and autonomy, which in turn is a central component as well as an outcome of successful aging (Baltes & Baltes, 1990; Rowe & Kahn, 1997; Willis, Jay, Diehl, & Marsiske, 1992). Advanced age, however, is normally accompanied by cognitive declines, which are observed in multiple domains including the aforementioned ones (e.g., Lövdén, Ghisletta, & Lindenberger, 2004).

Multidimensionality and Multidirectionality of Cognition

Cognitive aging does however not reflect a uniform decline but rather results from the different trajectories which cognitive development follows over the entire lifespan (de Ribaupierre, 2005). Baltes (1997) emphasized that cognitive aging is a multidimensional and multidirectional process. Indeed, this was already shown in the 1930s in the first calibration of the Wechsler intelligence scale (Wechsler, 1939, cited by de Ribaupierre, 2005). A dissociation between the two sub-scales was observed: age-related decline was more pronounced in the performance scale than in the verbal scale. This observation has been supported by further psychometric research, which suggests different developmental trajectories for the so-called fluid intelligence and the crystallized intelligence (Horn & Cattell, 1967). These two forms of intelligence have also been described as the *mechanics* and the *pragmatics* of cognition (Baltes et al., 1998). The former depends on basic biological processes and therefore mainly on the integrity of the central nervous system, whereas the latter refers to the culture-related knowledge base accumulated with experience. Cross-sectional studies have shown that fluid intelligence (gf) begins to decrease from middle adulthood on, while crystallized intelligence (gc) increases over the entire lifespan and remains stable or still increases until very late in life (Baltes & Mayer, 1999; Lindenberger & Baltes, 1995). In a more recent large scale study, the non-uniformity of age-related declines has been confirmed (Zimprich et al., 2008). So the demonstrated multidimensionality and multidirectionality of cognitive aging show that there is not only decline but also preservation and even gain in later life.

Even more promising were the results of research that included a longitudinal design. In a well-known large scale study, the Seattle Longitudinal Study (Schaie, 1983), age-related changes in a large battery of cognitive tasks, representing the primary cognitive abilities, were investigated in a sequential study design. By controlling for cohort effects, cognitive decline in advanced age was much less pronounced than generally assumed from cross-sectional studies (Schaie, 1996, 2005). The authors showed that in addition to gc, many gf abilities also remained stable until late in life and only declined at a very old age, between 80 and 90. Furthermore, important individual differences were reported with respect to the amount of abilities which decline. Even though nearly all individuals exhibited at least one declined ability by the age of 74, few individuals showed global decline. By the 80s almost no one showed decline on all abilities and less than 15 percent of the individuals showed more than four declined abilities. These results showed that age-related cognitive decline is present but that there are large inter-individual differences and stability up to a high age.

Mediator Variables between Age and Cognition

Several gf measures have been suggested to act as mediators between age and cognition. The decrease in speed of execution of processing operations was proposed to account for some of the age variance in gf (Salthouse, 1996). Salthouse showed in several studies that, when statistically controlling for processing speed in cognitive tasks, age differences were attenuated in gf performance, mainly in working memory (WM) capacity (Salthouse & Babcock, 1991; Salthouse & Meinz, 1995). WM performance in turn mediated age differences in other fluid tasks (Salthouse, 1991).

Further, inhibitory processes, that is the ability to ignore irrelevant information or so-called distractors, were shown to decrease with age and were proposed as a mediator of age effects in cognition (Hasher & Zacks, 1988). The age-related increase in susceptibility to distractors and the declining inhibitory control was found to have an important impact on WM capacity (Hasher, Zacks, & May, 1999). Inhibition was further investigated together with speed and WM capacity in one study (de Ribaupierre, 2001; de Ribaupierre, Olivier, Leutwyler, & Poget, 1999). The authors showed that speed and inhibition influenced the age-related differences in working memory capacity. In a more recent study, these results have been confirmed: processing speed and inhibition were found to mediate age effects in WM capacity whereas WM capacity accounted for age-related differences in text processing (Borella, Ghisletta, & de Ribaupierre, 2011).

These models reveal that WM appears to play an important role in general cognition and aging processes. WM has been defined in several ways; a commonly adopted view is that it refers to the processes that simultaneously provide temporary storage and manipulation of information for complex cognitive tasks such as language, comprehension, learning, and reasoning (Baddeley, 1986; Baddeley & Hitch, 1974). Its capacity is assumed to be limited; that is, the amount of attentional resources available determines the amount of information that can be simultaneously stored and manipulated (Engle, Kane, & Tuholski, 1999). The amount of attentional resources available has been found to vary across individual and age causing inter-individual differences (Case, 1985; de Ribaupierre, 2001; de Ribaupierre, Fagot, & Lecerf, 2011; Engle & Kane, 2004; Kane, Hambrick, & Conway, 2005; J. Pascual-Leone, 1970; J. Pascual-Leone & Baillargeon, 1994).

Executive functions were also found to mediate the relation between age and general cognitive abilities (Salthouse, Atkinson, & Berish, 2003). The concept of executive functions refers to the mechanisms of control and coordination of cognitive processes and is very closely related to WM processes. The three most important subprocesses are a) shifting between tasks or mental sets, b) updating and monitoring working memory representations, and c) the inhibition of dominant responses (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). However, these three subprocesses were not found to account differently for age-related effects on cognitive functioning (Salthouse et al., 2003).

The Increasing Connection between Cognition and Sensorimotor Functions

The correlation between different cognitive variables has been observed to decrease from childhood to young adulthood and then to increase with advancing age. This phenomenon has been described in the differentiation-dedifferentiation hypothesis (Baltes, Cornelius, Spiro, Nesselroade, & Willis, 1980; Reinert, 1970). It says that the various cognitive abilities become more differentiated during child development until young adulthood, and then again less differentiated and increasingly explainable by a single factor with advancing age.

Observations from the large scale Berlin Aging Study (Baltes & Mayer, 1999) are in line with this hypothesis. An increasingly stronger connection between cognitive, perceptual and motor variables with advancing age was reported. The correlation between sensory and sensorimotor abilities such as vision, hearing, gait, balance, and intellectual abilities in the gf as well as in the gc domain was much larger in old adults than in younger adults (Baltes &

Lindenberger, 1997). The Berlin Aging Study describes a dedifferentiation with advancing age towards a general factor explaining a large amount of individual variance.

In the context of the differentiation-dedifferentiation hypothesis, evidence that inter-individual and intra-individual variability varies with age has been provided (de Ribaupierre, Ghisletta, & Lecerf, 2006; Ghisletta & Lindenberger, 2003). Inter-individual variability refers to differences between individuals whereas intra-individual variability refers to differences within an individual. The intra-individual variability is further divided into inter-task and intra-task variability; that is, the differences in the performance across different tasks by one individual and the differences within one task performed by one individual. The inter-individual variability has been found to be much smaller in younger adults, but larger in children and older adults (de Ribaupierre et al., 2008; Ghisletta, Lecerf, & de Ribaupierre, 2003). The observation was similar for within-task intra-individual variability, whereas the intra-individual variability between tasks is much larger in younger adults as compared to older adults and children. The latter was in line with the differentiation-dedifferentiation hypothesis; that is, inter-task variability, or the difference between abilities, was larger in younger adults than in children and in older adults. This research was able to confirm that there tends to be one general factor explaining inter-individual differences in cognitive performance in children and old people and that there are therefore substantial inter-individual differences.

Two explanations were proposed which are not mutually exclusive but explain the lifespan differentiation-dedifferentiation process from different perspectives. The first is called the common-cause hypothesis (Baltes & Lindenberger, 1997; K. Z. H. Li & Lindenberger, 2002; Lindenberger & Baltes, 1994; Lindenberger & Ghisletta, 2009). The hypothesis suggests that a general neurobiological mechanism regulates the integrity of the brain across a wide range of regions and functional networks. This general mechanism affects all cognitive functions similarly and is more important during childhood, where the development is in progress, and becomes again more important during older age.

The second approach constitutes the aging-induced permeation of sensorimotor functioning with cognition, the so-called cognitive permeation (Lindenberger, Marsiske, & Baltes, 2000; Schäfer, Huxhold, & Lindenberger, 2006). The cognitive permeation hypothesis describes that, with advancing age, more cognitive resources have to be attributed to sensory and sensorimotor functions. In turn, fewer resources are available for intellectual tasks, as the already generally reduced resources in older adults have to be increasingly shared. So the resource overlap and competition between domains increases and compensatory resource allocation trade-offs becomes more frequent (K. Z. H. Li & Lindenberger, 2002). Balance during walking is for example preserved at the expense of the performance in a demanding cognitive task which was simultaneously executed (K. Z. H. Li, Lindenberger, Freund, & Baltes, 2001). This shows that sensory and sensorimotor processes are no longer automatically executed but require increased cognitive resources and turn into more demanding tasks with older age.

The Adaptation of Behavior

The multiple developmental trajectories, the large individual differences and the generally reduced resources older adults have to deal with have been described as a result of the dynamic interaction between biological and cultural (environmental) factors during development (Baltes, 1997; Baltes et al., 1998, 2006). This framework theory from a lifespan perspective demonstrates that the benefit of biology is largest at childhood and younger age and then declines in old age. The need for culture, by contrast, increases with age in order to compensate for the biological inefficiency. Yet the efficacy of culture also decreases with advancing age. The balance between gains and losses has to be adapted over the entire lifespan. However, with increasing age this becomes increasingly challenging (Baltes, 1997).

An effective strategy has been proposed for dealing with the increasing incompleteness of the lifespan development, the so-called selective optimization with compensation (SOC; Baltes, 1997; Baltes, Cornelius, Spiro, Nesselroade, & Willis, 1980). This strategy describes how behavior is adapted in order to achieve the balance between gains and losses. According to SOC, first, the goals have to be selected according to the limited resources available for achieving them. Second, the means for achieving the targets have to be optimized by the application of enhancing methods such as practice, effort, or cultural knowledge. Third, whenever a required means is no longer available due to losses, incompatibility between goals or limiting constraints in time and energy, compensation comes into play. By this strategy, behavior is adapted in order to achieve our goals and maintain autonomy for as long as possible.

A further central concept in the context of SOC is the notion of behavioral plasticity, which refers to the adaptation or the malleability of behavior in reaction to experience (Baltes & Lindenberger, 1988). The amount of plasticity of an individual determines the potential increase of the amount of resources available, promotes its optimization or fosters the efficient use of compensational means to achieve the goals. Given the potential importance of behavioral plasticity, the study of its age-graded differences has been qualified as “a powerful tool for identifying mechanisms of development” (Lindenberger et al., 2007, p. 234).

Summary

To sum up, we have learned that instead of a general cognitive decline there are other patterns observed in aging. First, the process of cognitive aging has been described as a multidimensional and multidirectional process which goes beyond a unitary cognitive decline. Second, despite the impact of the age-related decline in processing speed, WM, inhibition or executive function on the individual differences observed in general cognitive functioning, it has been shown that not all abilities decline at the same time and to the same amount. Third, it has further been shown by a longitudinal study design that there is more preservation in older age than reported from cross-sectional research. Fourth, we have learned that in older age individual differences are substantial. Finally, individuals are able to adapt and modify their behavior by the SOC strategy and by exhibiting behavioral plasticity, also in older age.

Cognitive Aging at the Cerebral Level

As we have learned from the previous section, aging does not have a similar impact on all cognitive domains. The same is true at the cerebral level. Aging does not entail a general deterioration of the cerebral structure and function, but rather exerts a differential effect on the brain describing a balance between gains and losses across the adult lifespan (Ludwig & Chicherio, 2007). Indeed, age-related changes are described not only in terms of decline, but also in terms of preservation and functional reorganization of the mental processes’ neural substrates (Della-Maggiore, Grady, & McIntosh, 2002). Raz (2000) further states that the aging brain is characterized by relative preservations and differential decline not only at the structural level, but also at the functional level. In this section, age-related cerebral changes are reviewed, first at the structural and metabolic levels and then at the functional level. The section concludes with a discussion of the prevailing models and hypotheses addressing the functional meaning of the reported changes.

Structural Changes in the Aging Brain

Postmortem studies were the first means to gain insight into age-related anatomical brain changes. At the macro level, aging was found to be associated with weight and volume losses, as well as with an expansion of the size of the ventricles and sulci (Skullerud, 1985). These studies further documented age effects at the microscopic level, including shrinkage of neuron cell bodies, reduction of synaptic density, and deterioration of myelinated fibers (Raz

& Rodrigue, 2006). More recently, the introduction of magnetic resonance imaging (MRI), a high spatial resolution imaging method, allowed the in-vivo study of age-related differences in the neural substrates, also within longitudinal study designs.

Grey and white matter development

Considering the development of grey and white matter - the two major types of brain tissues - differential age effects are observed. Grey matter refers mainly to the cerebral cortex and consists of neuronal cell bodies, dendrites, short axons, glial cells and blood vessels. White matter consists of myelinated axons and is located subjacent to the cortex. Grey matter volume decreases considerably during early childhood and continues to reduce in a more attenuated way from the first decade onward (Sowell et al., 2003). In contrast, white matter volume increases gradually during childhood, adolescence and adulthood until the fourth decade and then declines. The white matter development describes an inverted U-shaped curve over the lifespan (Fotenos, Snyder, Girton, Morris, & Buckner, 2005; Paus, 2005; Raz et al., 1997).

Several causal hypotheses have been proposed to explain the differential trajectories of grey and white matter during lifespan development. On the one hand, the synaptic and neuronal pruning, a so-called regulatory reduction process, was suggested to be responsible for the loss of grey matter during the first years of life (Paus, 2005). During aging, the observed changes in grey matter are attributed to neuronal atrophy and to the reduction of the dendritic arborization rather than to an effective neural loss, as had been believed (Hof & Morrison, 2004). Concerning white matter, the observed changes during adolescence until the fourth decade could be caused by a progressive myelination of the axons, particularly in the frontal cortex (Paus, 2009). With advancing age, the alteration of the white matter is essentially caused by an accumulation of microlesions and demyelination, which are shown as white matter hyperintensities on MRI scans (Davis et al., 2009; Raz & Rodrigue, 2006; Wozniak & Lim, 2006).

Regional variability

While grey and white matters demonstrate differential developmental trajectories on global measures, another important element is the variability of age effects at the regional level. With respect to aging, observations suggest that the prefrontal regions are more affected than the other cerebral regions in terms of volume loss (Hedden & Gabrieli, 2004; Raz, 2004; Salat et al., 2005). Structures of the medial temporal lobe, in particular the hippocampus and to a lower degree the entorhinal cortex, show age-related volume loss (Raz, Rodrigue, Head, Kennedy, & Acker, 2004). In a longitudinal study run over five years, Raz and colleagues

(2005) found a significant negative correlation between volume and age for the lateral prefrontal cortex, the orbito-frontal cortex, the hippocampus, the caudate and the cerebellum. These correlations were stronger after five years for the prefrontal regions, the hippocampus, the caudate and the cerebellum, which indicates an age-related acceleration of the shrinkage with age in these regions. In contrast, the volumes of the primary visual cortex, the fusiform cortex as well as the inferior parietal regions, were not significantly correlated with age, and the correlations remained stable over five years. A similar regional difference was reported for white matter, with the greatest age-related changes located in the prefrontal cortex and the anterior corpus callosum (Bartzokis et al., 2003).

Metabolic changes

Along with structural changes, metabolic changes have also been reported. Age-related changes were observed in measures of cerebral blood flow, cerebral metabolic rate of glucose and of oxygen use (Slosman, 2001). With respect to neurotransmission, the production and the uptake of many, if not all, neurotransmitters show an age-related reduction (for a review see Rehman & Masson, 2001). However, the most documented findings concern the dopaminergic system. Losses are reported for the dopamine transporter as well as for dopamine D1 and D2 receptors densities (for a review see Bäckman & Nyberg, 2010; Bäckman, Nyberg, Lindenberger, Li, & Farde, 2006; Kaasinen et al., 2000). These changes are mainly observed in the striatum and in the frontal cortex (Suhara et al., 1991) and have been linked to an age-related decline in neuromodulation (S.-C. Li & Sikström, 2002, see next section).

Summary

To sum up, structural and metabolic changes reported with aging seem to be regionally very variable and do not concern the whole brain to the same extent. Associative brain regions like the prefrontal and the medial temporal lobes are more affected by age, whereas primary sensory regions are spared. The former structures cover regions in the anterior part of the brain while the latter are situated in the posterior brain regions. This observation is described by an anterior-posterior gradient of brain changes (Raz & Rodrigue, 2006; Sullivan, Rohlfing, & Pfefferbaum, 2008) which is explained - at least in part - by the progressive myelination during child development and by the decrease in myelin integrity in aging: In a phylogenetic view, the prefrontal regions are the ones which have evolved most recently. At the ontogenetic level the prefrontal regions are the last to myelinate during development. Thus, a *last-in-first-out* hypothesis has been proposed, which means that the regions which developed first are less vulnerable to show the effects of advancing age and are therefore the last to show a decline, and vice versa (Davis et al., 2009; Raz, 2000).

Functional Changes in the Aging Brain

As we have seen in the previous section, age-related changes do not affect all brain structures in the same way. A similar regional variability can be observed at the functional level. Two main patterns of age-related differences in brain activation are usually described and consistently reported. In this section the two models are reported initially, followed by a third and more recent model linking both previously reported patterns. Then three major approaches interpreting these functional age-related changes are discussed.

Models of functional cerebral changes

HAROLD. The first model is derived from the HERA model (Hemispheric Encoding Retrieval Asymmetry; Tulving, Kapur, Craik, Moscovitch, & Hoile, 1994) which was proposed to describe the lateralized neuronal recruitment in younger adults during episodic memory processes. According to the model, the left prefrontal cortex (PFC) engaged in memory encoding processes whereas the right prefrontal regions were occupied with memory retrieval. Studies conducted with older adults failed to report such a lateralization of function: indeed, activity in older adults tends to be less lateralized than in younger adults, irrespective of the episodic memory processes engaged.

These observations were conceptualized as the HAROLD model (Hemispheric Asymmetry Reduction in OLDer adults; Cabeza, 2002 , first proposed by Cabeza et al., 1997). Interestingly, the HAROLD pattern has been found for various other cognitive processes, among them those engaged in WM tasks. Indeed, in young adult’s samples, verbal WM performance is generally associated with left lateralized PFC activation, whereas spatial WM is associated with right lateralized PFC activation (Smith & Jonides, 1997). In older adults however, the functional pattern of lateralization is reduced, corresponding to the HAROLD pattern. Such a pattern is already observed in middle-aged subjects during a number *N*-back task which requires information updating and shifting (Dixit, Gerton, Kohn, Meyer-Lindenberg, & Berman, 2000). Furthermore, the authors demonstrated that the activity of the left dorsolateral PFC and in portions of the left inferior parietal lobule was positively correlated with age (18 – 48 years old). They further reported a positive correlation between age and deactivation of the right dorsolateral PFC and the right parietal lobule, regions that are most robustly activated in younger adults during this task. This supports the fact that older adults rely consistently more on both hemispheres, in conditions in which unilateral recruitment in younger adults is sufficient. A similar age-related symmetry was reported by Reuter-Lorenz and colleagues (2000) in both a letter and a spatial location WM task. The PFC activation was lateralized in the left hemisphere during the verbal task and in the right hemisphere during the spatial task in younger adults. In both tasks, older adults showed an additional contra-lateral PFC activation. No age-related differences were reported in posterior regions.

The HAROLD pattern has been confirmed in a number of functional imaging studies in WM (e.g., Chicherio, 2006; Ludwig, 2005) and across numerous other cognitive domains, including not only episodic memory, but also semantic memory retrieval, perception, and inhibitory control (for a review see Dolcos, Rice, & Cabeza, 2002). Therefore, this pattern cannot simply be attributed to the specific task demands, but rather needs to be considered as an intrinsic characteristic of cognitive aging (Cabeza, Nyberg, & Park, 2005).

PASA. While many of the studies of functional aging initially focused on attention demanding processes, an increasing interest focused on age-differences in less demanding tasks. The second model called PASA, standing for posterior-anterior shift in aging, was derived from this ensemble of studies and was proposed by Davis, Dennis, Daselaar, Fleck, and Cabeza (2008).

PASA describes the age-related reduction in occipital activation concomitant to an increased frontal activation. Grady and colleagues (1994) were amongst the first to describe such a distributed pattern of activity using a face and location matching task. The results revealed that younger adults recruit more occipitotemporal regions than older adults. In contrast, frontal regions showed increased activation in older adults. The authors suggested that older adults recruit anterior regions to compensate for a reduced processing efficiency in posterior regions. The PASA pattern has been replicated within various cognitive domains, and despite its numerous replications, the PASA pattern could be interpreted in terms of age-related differences in task difficulty. To investigate this explanation and specifically address this issue, Davis and colleagues (2008) conducted an experiment in which they controlled for task difficulty across age groups. Younger and older participants received a visual perception task and an episodic memory task for which task accuracy was matched between the two samples. The PASA pattern remained, which led the authors to argue that there seems to be a fundamental difference in task processing between older and younger adults, as has been stated for the HAROLD pattern as well.

However, while the HAROLD pattern is found in tasks engaging higher attentional processes, the PASA pattern is usually observed in less demanding tasks (Dennis & Cabeza, 2008).

CRUNCH. In a more recent model, both patterns of age-related differences are unified and synthesized into a single model. The so-called CRUNCH (Compensation-Related Utilization of Neural Circuits Hypothesis) was proposed by Reuter-Lorenz and Cappell (2008). This model describes the age-related differences in performance and cerebral activation in a continuum from low to high demanding cognitive tasks. The authors predict that in a low load condition older adults will attain the same behavioral performance as younger adults. Regarding cerebral activation however, older adults recruit their frontal resources while younger adults only minimally rely on frontal resources. This pattern corresponds to the PASA model while describing the age-related recruitment of frontal lobes in low demanding tasks. When task demands increase, older adults will maintain the same performance as younger adults for a while, but only by recruiting their frontal lobes even more. Younger adults in turn will exhibit less activation than older adults, even though the frontal activation is now increased. This HAROLD pattern illustrates how older adults recruit frontal lobes increasingly bilaterally when cognitive tasks are demanding. When task demands are further increased, older adults reach sooner limits in their resources; performance as well as activation levels decrease according to the CRUNCH model.

The authors of these models, and specifically the authors of the latter model, argue for a compensational interpretation of the age-related overrecruitment and more diffuse recruitment in frontal lobes. However, alternative interpretations are discussed. These different views are now outlined in the following section.

Interpretations of functional cerebral changes

Research combining these functional models of brain changes with behavioral performance provided evidence for different interpretations of the models. However, to date no consensus regarding the functional significance of the previously reported age-related functional reorganization, i.e., the overrecruitment and the more diffused recruitment of the frontal lobes, exists.

The two main views discussed in recent literature put into question whether a) the aging brain compensates its structural and metabolic decline by a reorganization of the frontal lobes; or whether b) the observed reorganization of the frontal lobes is contradictorily an undesirable outcome of the structural declines, and causes the age-related cognitive decline.

In neuroscientific literature, the former view is called the *compensation approach* whereas the latter is called the *dedifferentiation approach*. The term dedifferentiation originates from the behavioral psychometric lifespan research and describes the increasing correlation between cognitive, sensory and motor task performance with age (see section *The Increasing Connection between Cognition and Sensorimotor Functions*). For the sake of clarity and to avoid confusion with this term in the present work, we use the term *dedifferentiation* only in relation to the phenomenon reported in behavioral data. To describe the non-compensatory interpretation of the functional reorganization, the term *non-selective recruitment* proposed by S.-C. Li and Sikström (2002) is used.

In the present section, both views are discussed by providing evidence linking behavioral performance with brain activity. First, evidence supporting the compensation approach is provided and second, evidence for the non-selective recruitment approach. Third, at the end of this section, both approaches are discussed together by proposing a new view of a co-occurrence of both interpretations.

Compensation approach. One of the most popular studies which provided evidence for the compensation view, by combining behavioral performance and brain activity, was conducted by Cabeza, Anderson, Locantore, and McIntosh (2002). The authors created two groups of older participants, high and low performer groups, based on their respective performance in four memory measures. The old-high group showed no memory performance difference to the young reference groups, whereas the low-old group performed significantly below the younger and the high-old groups. Subsequently the groups underwent a PET-scan during two episodic memory retrieval tasks, a word-pair completion, and a source memory task. The behavioral results showed that the old-low participants' performance was significantly lower than the young and old-high groups' performance, whereas the latter two groups did not show a significantly different performance. As regards the brain activity during the more demanding source memory task, younger and old-low participants showed a right lateralized anterior PFC activity whereas the old-high participants showed a bilateral, that is, right and left anterior PFC activation. These results indicate that older adults who recruit bilateral PFC regions obtain similar behavioral performance to younger adults. The more diffused PFC activation had therefore a compensational function to overcome the age-related decline, whereas the youth-like activation in the low-old was not. This evidence was one of the first which illustrated the HAROLD pattern and provided evidence for the compensation approach.

More recently, another hypothesis has been developed which favors the compensation approach and integrated HAROLD and PASA pattern, the so-called Compensation-Related Utilization of Neural Circuits Hypothesis (CRUNCH; Reuter-Lorenz & Cappell, 2008) mentioned in the previous section. This model was proposed to account for patterns of over-activation and under-activation in older adults. The authors state that the frontal over-activation in older adults is beneficial and allows older adults to perform at similar performance levels to younger adults. Specifically, the over-activation in frontal lobes compensates for the structural decline in the PFC as well as for declines in the medial temporal lobe, for increased noise in perceptual processes as well as for increased general noise. The latter refers to the so-called default network, which is a measure of the brain's activity during nontask periods. The deactivation in younger adults increases when task demands are higher. However, a less adapted deactivation has been observed in older adults when task demands increased (Persson, Lustig, Nelson, & Reuter-Lorenz, 2007). However, as task demands increase, older adults may reach the ceiling of over-activation and show under-activation and decreasing behavioral performance.

Evidence supporting the CRUNCH model was provided for example by Mattay and colleagues (2006). They showed that older adults performed at the same behavioral level as younger adults in a less demanding 0-back task. At the brain level (fMRI) however, older adults showed more bilateral PFC activation during the 0-back task than younger adults. When the difficulty level increased to the 2-back or 3-back level, the accuracy performance in older adults was clearly reduced compared to younger adults. Functional brain activation in contrast decreased in older adults compared to the younger adults. Thus, as the CRUNCH model suggests, older adults used the additional frontal activation to compensate for the age-related cognitive decline and perform at an optimal youth-like level. However, when the task demand increased, the compensation broke down and the performance as well as the frontal activation decreased.

The compensation view has been supported by a number of studies in various cognitive domains by either correlating brain activity with behavioral performance or comparing high versus low performing older adults to younger adults (Daselaar, Rombouts, Veltman, Raaijmakers, & Jonker, 2003; Heuninckx, Wenderoth, & Swinnen, 2008; Rajah & D'Esposito, 2005; Reuter-Lorenz et al., 2000; Reuter-Lorenz, Stanczak, & Miller, 1999; Rypma & D'Esposito, 2000).

Another more recent finding supporting the CRUNCH model is provided by Schneider-Garces et al.(2010). They discussed a fundamental limitation of performance-brain activation correlations, that is, correlation results can be ambiguous and do not indicate which factor is the cause, and which is the consequence. In their study, brain activity was investigated using a continuum from low to high load conditions of the Sternberg task, a verbal WM task. When comparing the low load to the high load condition while correcting for the individual memory span level, the older adults performed at the same level as the younger adults in low load conditions, but their performance decreased in high load conditions. With regard to brain activity, older adults showed overrecruitment at low load conditions and under-recruitment at high load conditions, which is consistent with the CRUNCH model. In sum, this study provided support for the relative utilization of neural networks depending on performance. However, as concerns the compensational view, the conclusions were less strong. The differences in brain activation between younger and older adults could be explained entirely by WM capacity differences. In other words, the results indicated that, given equal objective memory loads, individuals with lower WM capacity were recruiting more brain activation than those with higher WM abilities, regardless of age. Due to the correlational nature of the analysis, the authors could not conclude whether the increased brain activity was used to improve performance, or whether the low memory ability caused the increased activity, and the lower behavioral performance.

An episodic memory study provided further evidence to exclude the possibility that age-related differences in activity result from the simple fact that older adults use different encoding strategies than younger adults (Logan, Sanders, Snyder, Morris, & Buckner, 2002). The authors investigated intentional verbal encoding and showed that older adults under-recruit the frontal lobes. This under-recruitment could be reversed by requiring a deeper elaboration of the verbal material from the study participants. However, as the under-recruitment of frontal regions was diminished in older adults by this strategy, they showed a higher unspecific and non-selective frontal recruitment than younger adults. These findings support the compensation approach: a non-selective recruitment seems to be a characteristic of older adult’s brain activation and this pattern does indeed appear when the same encoding strategy is used, as is also used by younger adults.

Non-selective recruitment approach. The second popular approach for explaining age-related overrecruitment and a more diffused functional distribution of cortical activation is often contrasted with the compensation view. The non-selective recruitment view interprets these activation patterns as a difficulty in recruiting specialized neural resources, the bilateral frontal activation would therefore indicate an undifferentiated processing (e.g., S.-C. Li, 2002; S.-C. Li & Lindenberger, 1999). This view interprets the functional changes as a direct and necessary consequence of neurocognitive impairments in the brain structure and at the metabolic level. It has, in particular, been developed in a neurocomputational model which proposes an explanation of the effect of age-related changes by introducing the neurochemical level (S.-C. Li, Brehmer, Shing, Werkle-Bergner, & Lindenberger, 2006; S.-C. Li & Sikström, 2002). As reported in the previous section, age-related changes in neurotransmitters are mainly found in the dopaminergic system affecting the dopamine transporters and receptors in the PFC.

The neuromodulation model proposed by S.-C. Li and colleagues (2000, 2002) suggests that the age-related attenuation in the dopaminergic system shapes the functional organization of neurocognitive processes. That is, the attenuation leads first to an affected neuronal signal transmission which then leads to an increased signal to noise ratio. This in turn has consequences for the distinctiveness of neurocognitive representations. The representations become less distinctive due to the non-selective or non-specific neuronal recruitment; for example, patterns representing different stimuli are more difficult to differentiate from each other. Therefore different mental representations are less distinct and are more confused with each other. This model has been tested in neural network simulations (S.-C. Li, Lindenberger, & Frensch, 2000). The simulations were able to reproduce many of the generally observed age-related differences patterns: Age-related differences in learning rate, complexity cost, interference susceptibility, intra- and inter-individual variability, and ability dedifferentiation could all be modeled. The neurocomputational model thus illustrates how the attenuation in the dopaminergic system can lead to a non-selective recruitment of PFC regions.

Empirical studies also provided evidence for non-specific recruitment of various brain regions, but particularly in the visual cortex (for a recent review see Goh, 2011). Park and colleagues (2004) showed that older adults' visual ventral regions for specific stimulus categories respond less selectively than in younger adults. For instance, the face regions are more responsive to places in older adults, whereas in younger adults the regions respond selectively.

In another study, the investigation of age-related visual ventral cortex specialization was assessed during a WM task (Payer et al., 2006). The findings confirmed the decreased specialization in visual cortex in older adults compared to younger adults, which also occurred during a more demanding WM task. However, the frontal cortex was simultaneously more bilaterally activated in older adults when compared to younger adults. These results support the non-selective recruitment approach, but at the same time they point towards another view. The authors suggested that the activation in frontal regions would compensate for the non-selective activation in the visual cortex. This co-occurrence of non-selective recruitment and compensation was proposed further in more recent literature and is now reviewed.

Co-occurrence of compensation and non-selective recruitment. More recently, both views have been discussed in a complementary way (e.g., Lövdén, 2010; Ludwig & Chicherio, 2007; Reuter-Lorenz & Park, 2010).

The scaffolding theory of aging and cognition has lately been proposed in order to integrate behavioral and neuroimaging evidence and theories which account for age-related cognitive decline, structural losses and increases in PFC activation (Park & Reuter-Lorenz, 2009; Reuter-Lorenz & Park, 2010). The *scaffolding framework* was initially proposed by Petersen, van Mier, Fiez, and Raichle (1998). It describes the phenomenon during learning and practice when additional brain regions are recruited which serve as a scaffold for the yet inefficient task-specific regions. These scaffolding regions include mainly prefrontal and parietal areas which are associated to attention and cognitive control (Kelly, Foxe, & Garavan, 2006). These brain regions are used to cope with new or increased demands. The scaffolding theory of aging proposes that the aging brain has to deal with age-related neural declines and less efficient or inefficient neuronal networks. The aging brain responds by recruiting and creating additional, alternative neural circuits in order to reinforce declining structures, and maintain a high level of functioning. More precisely, frontal regions compensate for the decreased specificity and non-selective recruitment of posterior regions. However, in this model the non-selective recruitment does not refer to the overrecruitment in frontal cortex as reported in S.-C. Li and colleagues (2006) but rather to losses in specificity of perceptual brain areas such as the ventral visual cortex (Payer et al., 2006).

Park and Reuter-Lorenz call this process compensatory scaffolding. They suggest that the PFC is the most flexible structure in the brain and therefore scaffolding processes would occur mainly in this structure. Similar scaffolding processes are observed during child development, where an increased recruitment of frontal regions is associated with the acquisition of new cognitive skills and the development of specific brain regions (e.g., Church, Coalson, Lugar, Petersen, & Schlaggar, 2008; Klingberg, Forssberg, & Westerberg, 2002a). The scaffolding process is therefore a lifelong process and occurs in situations where an individual has to adapt to environmental challenges.

Recent findings supporting the integration of non-selective activation and compensation are provided by Carp, Gmeindl, and Reuter-Lorenz (2010) who addressed the co-occurrence new view by a multi-voxel pattern analysis of the distinctiveness in the brain's activity. They confirmed a decreased distinctiveness of visual cortical representation in older adults. However, older adults had in contrast a more distinct PFC activation in lower load conditions than younger adults. This was interpreted as a compensational activation in PFC for the less distinct activation in the visual cortex. At higher loads, the pattern reversed and became less distinctive in older adults but more distinctive in younger. The ceiling of compensation was therefore reached in older adults, whereas younger adults begin with an additional recruitment.

However, in a more recent study conducted by the same authors, the findings have not been replicated (Carp, Park, Polk, & Park, 2011). The authors found no evidence that older adults compensate the decreased distinctiveness in visual cortex by engaging additional brain regions. These results are ambiguous and tend to support the non-selective recruitment models of cognitive aging, whilst challenging the compensation approach.

Another example demonstrating the ambiguity in interpretations of functional reorganization was provided by Persson and colleagues (2006). They examined two groups of older adults, one group which showed preserved episodic performance over ten years and another group which declined in episodic performance over the same timeframe. Structural data revealed a decline in hippocampal volume and in anterior corpus callosum integrity for the participants who declined in memory performance only. Functional imaging during episodic memory encoding revealed further an increased activation in left PFC for all participants, but an additional right PFC activation for the memory decliners only. On the one hand, the results can be interpreted in terms of a compensatory response. That is, older adults showing structural declines have to compensate for these changes by an increased frontal recruitment. This would be in line with the compensation approach, which supports the view that structural declines are compensated by PFC over-activation. On the other hand, the PFC over-activation did not yield to a preserved behavioral performance, but was correlated with behavioral decline. This would denote that the structural declines caused the inappropriate over-activation. That would be consistent with the non-selective recruitment approach. However, no clear conclusion can be drawn from these results, as the evidence is based on correlation analyses and therefore no causal explanation can be provided.

Conclusion

In sum there is still no clear and convincing body of evidence which would point towards a unified interpretation of the overrecruitment pattern in older adults. The reviewed evidence from cross-sectional studies remains inconclusive. To date, it is not clear under which conditions the age-related overrecruitment in frontal regions occurs and if it is beneficial to older adults or not.

In the present study we intended to go beyond the cross-sectional study design in which age differences in cognitive performance and cerebral activation patterns were simply compared. For this purpose we aimed at experimentally manipulating the individual's performance and its underlying cerebral organization. Intervention studies have a great potential to provide further insights into this open issue. Training studies come with important validity benefits; by practicing a cognitive task during a certain time and observing individuals' behavior in a longitudinal design, the experimental control is substantially improved. Age differences are no longer captured by a snapshot but can be investigated in a more reliable way, where the individual has the chance to unfold his or her potential. Specifically older adults do not adapt as rapidly as younger adults to new situations and may not always show their optimal performance at the first testing session. As Brehmer and colleagues showed, older adults benefited the most of a strategy instruction session right after the first testing session (Brehmer, Li, Müller, Oertzen, & Lindenberger, 2007).

Pushing the individuals to their limits near asymptotic performance level allows us to gain reliable insights into age-related differences in performance and its plasticity and so into mechanisms of development (Lindenberger et al., 2007). When further combining the training-related changes at the behavioral level to the brain level, a contribution to this debate can be possible. As the change in brain activation from an initial level of performance to a maximum level of performance could be observed, we would be able to better understand the meaning of the age-related differences in the brain organization pattern.

In the following section, the training literature with regard to the behavioral and cerebral level of analysis is introduced.

PLASTICITY AND TRAINING

In this section we introduce the long-standing term of plasticity, which has recently regained its popularity in neurocognitive aging research. In the first part, we introduce and review the significant cognitive plasticity literature in training research. In the second part, we provide an introduction and a literature review of brain plasticity research. Within this review, we will show that in particular WM training interventions offer a great potential for fostering plastic changes at both the behavioral and the cerebral levels of analyses. We close the section by providing an excursus of EEG measures and its sources of individual differences, as we implemented EEG in the present work.

Cognitive Plasticity and Training

Definition of Cognitive Plasticity

In recent years the notion of plasticity has become a widely-used term in many fields of cognitive psychology and behavioral neuroscience. In the present work we use cognitive plasticity in a lifespan developmental view, where it denotes the intra-individual modifiability of behavior and potential for different forms of behavior (Baltes, 1987; Baltes & Lindenberger, 1988). Specifically, plasticity refers to a set of processes which allow an organism to adapt to its rapidly changing environmental demands (de Ribaupierre et al., 2009; A. Pascual-Leone et al., 2011). We adopt further two conceptual definitions of cognitive plasticity. First, we provide the definition proposed by Willis and Schaie (2009) and second we introduce the framework suggested by Lövdén, Bäckman, Lindenberger, Schaefer, and Schmiedek (2010).

In the definition proposed by Willis and Schaie (2009; Willis, Schaie, & Martin, 2009), cognitive plasticity is defined in terms of the individual's cognitive latent potential and the capacity to acquire cognitive abilities to improve cognitive functioning. Willis and colleagues' framework on plasticity is based on the growing body of literature in cognitive interventions in older adults. They support the view that human development is characterized by lifelong modifiability and adaptability. This adaptability encounters several limits according to the phase of development. Their definition of plasticity is rather broad and includes all changes at the neuronal or behavioral level which lead to an adaptation.

Lövdén (2010; Lövdén, Bäckman et al., 2010) defines cognitive plasticity in a more differentiated way. In this framework, plasticity refers to the capacity to change permanently in reaction to a primary change. Such a primary change would cause a mismatch between functional biological supply and environmental demands, and would therefore drive cognitive plasticity. Plasticity occurs, for example in reaction to brain injury, by the processes of restoration and compensation because a mismatch occurred between the organismic supply and the environmental demands.

The authors further distinguish plasticity from variability and flexibility. In their view, variability denotes temporal and reversible fluctuations in behavior whereas plasticity describes a relatively permanent change, from one equilibrium to another. Flexibility is also associated with less permanent change and denotes the capacity to optimize the brain's performance within the current range of performance and functioning. There is a basic difference between flexibility and plasticity: flexibility-based reactions use the existing functional supply whereas plasticity-based reactions result in changes of the functional supply. Plasticity requires brain structure changes and a behavioral flexibility increase. Plasticity can therefore be denoted as a change in supply. According to this framework, plasticity is induced whenever the environmental demands challenge the current range of flexibility. The flexibility range therefore determines at what point a supply-demand mismatch will lead to plastic changes beyond the flexible adaptations. However, we do not aim at differentiating between plasticity and flexibility in the present work.

A supply-demand mismatch can be triggered by a cognitive training procedure. By implementing a cognitive intervention, the environmental demands which are compared to the individual's supply could experimentally be controlled.

Preserved Cognitive Plasticity in Older Adults

Over two decades ago, cognitive plasticity in older adults was addressed in a systematic way in a seminal work by Baltes and Kliegl (Baltes & Kliegl, 1992; Baltes & Lindenberger, 1988; Kliegl et al., 1989). They examined the plasticity of episodic memory by teaching a mnemonic technique, the method of loci. The method of loci is based on a familiar sequence of locations of a mental map which serves as a structure for encoding and retrieving new information. By mental imagery, new information (e.g., words) is associated with a location from the mental map. The associated information is then recalled by mentally revisiting the locations. The authors trained younger and older adults to use this technique in a so-called testing-the-limits procedure. This procedure allowed them identifying three levels of performance: first, the baseline performance, which refers to the performance of an individual under standardized conditions; second, the baseline reserve capacity which refers to what an individual can reach when the conditions of assessment are optimized (e.g., after strategy instruction but without providing extensive practice); third, the developmental reserve capacity, that is the maximal level of performance after a training procedure.

Thirty-eight training sessions were provided in order to let the participants reach asymptote performance and the limits of developmental reserve capacity in learning a list of 30 words (Baltes & Kliegl, 1992). This allowed the authors to examine whether older adults would catch up and obtain the same performance as younger adults with extensive training, or in other words, whether the developmental reserve capacity would differ between groups.

The findings revealed several interesting patterns: First, there were significant age differences in baseline word recall performance before training. Second, average recall performance increased significantly in both age groups over the training sessions and both age groups reached the limits of developmental reserve capacity over the course of the training. Third, age differences in recall performance were magnified after training in favor of the younger adults. The age groups' overlap of the individual performance score distribution after training was reduced and no older adult reached the mean performance of the younger group. Fourth, age turned out to explain increasingly important and unique variance in recall performance as the number of practice sessions increased (Kliegl, Smith, & Baltes, 1990). Finally, the final individual score distribution further revealed that all younger adults improved performance whereas not all older adults did. This finding was later confirmed by the study of Nyberg and colleagues (2003) who used cerebral imaging technique while the method of loci was implemented. They showed that those older adults who did not improve word recall also did not show any brain activation for visual imagery. That is, the mnemonic technique was not implemented at all by these older adults. The generally reduced word recall performance in older adults was however attributed to the generally reduced processing capacity in older adults (this study will be discussed further below).

This initial and systematic age-comparative plasticity research showed that cognitive plasticity is preserved in older age but is nevertheless reduced compared to younger age. Subsequent to this work, a large number of training interventions emerged (for a review see Lustig, Shah, Seidler, & Reuter-Lorenz, 2009).

Transfer of Training Effects

Central in the context of learning and training research is the concept of transfer (Hager & Hasselhorn, 1998; Willis, 2001). This concept has already been discussed in very early research where the "interdependence of mental functions" in relation to training was investigated (Thorndike & Woodworth, 1901, p. 563).

Transfer refers to the generalization of training effects to untrained tasks; in other words, to the transfer of knowledge or abilities to new situations (Willis, 2001; Willis & Schaie, 2009). Generally, two types of transfer are distinguished: near transfer and far transfer (Edwards et al., 2002; Willis, Blieszner, & Baltes, 1981). Near transfer occurs when the demanded abilities in the transfer situation highly overlap with those in the training situation. It can further be distinguished between structurally similar and structurally different near transfer tasks (Perkins & Salomon, 1992). They both share the underlying cognitive processes with the training task, but the former has in addition a common task structure with the trained task whereas the latter shows a different task structure. For instance, a letter comparison task and a pattern comparison are both demanding processing speed performance. They share in addition the task structure and differ only as regards the content material. However, when comparing the letter comparison task to the Digit Symbol-Coding task, they still share common underlying processes but address it in a different task structure.

As concerns far transfer, it qualifies a gain in tasks that share few cognitive processes with the trained task, for example the letter comparison task assessing processing speed and the Reading Span task assessing WM capacity.

Strategy-based Training

Strategy-based training has become again popular in recent years, particularly in aging research, subsequent to the work of Kliegl and Baltes (1989). The main goal of strategy training is to improve episodic memory performance by teaching a specific mnemonic strategy. Many training interventions investigated the effect of specific mnemonic training in older adults (e.g., Bissig & Lustig, 2007; Brehmer et al., 2008; Brooks, Friedman, Pearman, Gray, & Yesavage, 1999; Jennings & Jacoby, 2003; Nyberg et al., 2003; Rebok, Carlson, & Langbaum, 2007). These studies found often large and long lasting effects on the trained task, which were reduced but nevertheless important in older adults as well (for a review see Verhaeghen, Marcoen, & Goossens, 1992). However, strategy training interventions have an important shortcoming as regards transfer effects. Generally they exhibit only very small or no transfer effect at all to untrained situations (Lustig & Flegal, 2008; Zehnder, Martin, Altgassen, & Clare, 2009). The practice of strategies seems to be too specific to the trained task whereas the training of general cognitive processes is more promising (Lustig et al., 2009).

In a large scale intervention study, the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE; Ball et al., 2002) study, these different approaches of training were compared. The ACTIVE study included four groups of participants: two types of strategy training (mnemonic training and reasoning strategy training), a speed-of-processing training group (visual-search and divided attention) as well as a no-contact control group. All three training groups received a 10-session intervention and four sessions of booster training at 11 and 35 month after training. Both the mnemonics and reasoning groups showed improvements on measures that targeted the skills learned which lasted five years after training. The reasoning training group in turn showed improvement that was not restricted to reasoning tasks. Five years post-training, they showed less decline on self-reported Instrumental Activities of Daily Living and on safe driving performance (Willis et al., 2006).

This study demonstrated that despite the large long-lasting effects of strategy training no transfer effects are observed. For process-based training in turn, transfer effects were important and even found in everyday activities.

Process-based Training

Recently, process-based training interventions gained more interest, in particular to remediating age-related decline, since a training of general cognitive processes appeared more promising for fostering transfer effects (Hertzog et al., 2009; Kliegel & Bürki, 2012; Williams & Kemper, 2010). However, training cognitive processes without explicitly training strategies is not a new approach; the training of intellectual abilities has a long history (e.g., Baltes, Kliegl, & Dittmann-Kohli, 1988; Baltes & Lindenberger, 1988; Reisberg, Baron, & Kemler, 1980; Schaie & Willis, 1986; Thorndike & Woodworth, 1901; Willis et al., 1981).

In the present work, we focus on process-based training intervention as it was identified to have a larger potential to foster basic changes in cognitive processes as discussed in the previous section (Lustig et al., 2009). In particular, WM training procedures were found to be very efficient and promising in leading to large transfer effects (Klingberg, 2010; Perrig, Hollenstein, & Oelhafen, 2009; R. J. Sternberg, 2008). We therefore review the literature on WM training studies conducted with younger or older adults. In the last part, we review the few studies that included a sample of both age groups.

Working memory training in younger adults

One of the first studies showing effects from WM training comes from Klingberg, Forssberg, and Westerberg (2002b). They trained children and younger adults with an attention-deficit/hyperactivity disorder (ADHD) on four WM tasks (visuo-spatial rotation task, backward digit span, letter span, go/no-go reaction task) in an adaptive procedure for 26 days over 5 weeks. This led to near transfer effects on a non-trained visuo-spatial WM task and to a far transfer effect on Raven's Progressive Matrices, a fluid intelligence measure.

Two subsequent studies implementing the same WM training procedure, known as Cogmed training (Klingberg et al., 2005; Klingberg et al., 2002b), replicated the findings in a younger healthy sample (Olesen, Westerberg, & Klingberg, 2004; Westerberg & Klingberg, 2007). They reported near transfer to the span board task and digit span task as well as far transfer to the Stroop task and the Raven's Advanced Progressive Matrices. However, their samples were rather small (*N* = 8 and *N* = 3 in the WM training groups, respectively) and no or only a no-contact control groups was included. The findings are therefore to interpret with caution.

More recently, this effect has been replicated in larger samples of younger adults (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008). Ninety-six younger adults (*M* = 26 years) were trained in four different WM training groups. The training groups were matched to no-contact control groups. All groups trained on the same individually adaptive dual *N* -back task, but with different amounts of training sessions (i.e., 8, 12, 17, and 19 days). All participants in the training groups trained daily for 25 minutes, except for the weekends. The dual *N* -back task consisted on the one hand of a visuo-spatial task, for which squares were presented sequentially at eight different locations on a computer screen. On the other hand, the task included an auditory-verbal task, which was presented simultaneously with the visuo-spatial task. It consisted to listen to an auditory sequential presentation of eight consonants. A response was required whenever one of the currently presented stimuli matched the one presented *N* positions back in the sequence. After each block, the level of *N* was adjusted to the individual performance. The results showed large far transfer effects in short versions of the Ravens' Advanced Progressive Matrices and in another fluid intelligence measure. Further a near transfer effect in the Digit Span task was reported; however, no near transfer effect was found in a further WM measure, i.e., the Reading Span task.

In a subsequent study, these finding could be replicated (Jaeggi, Studer-Luethi et al., 2010). The authors trained again younger adults during 20 days on the dual N-back task, but at this time they were compared to a single task training. For the latter training, only the visuo-spatial modality was trained, but in a similar training procedure and adaptivity as the dual task training. Both training groups led to far transfer effects in fluid intelligence measures compared to a no-contact control group. However, again no near transfer effect to the Reading Span task was found. That is somewhat surprising as the transfer from N-back training should be mediated by WM capacity improvements. If that would be so, then the effect should also be found in the Reading Span task.

In another recent study, younger adults were trained during 20 sessions on adaptive verbal and spatial complex WM tasks (Chein & Morrison, 2010). The authors reported a far transfer effect to the Stroop task and to reading comprehension. However, no far transfer effects to a fluid intelligence measure were found, the previous studies could therefore not fully be confirmed by this one.

Working memory training in older adults

One of the first studies which targeted WM was conducted by Mahncke and colleagues (2006), where older adults completed a so called *brain plasticity-based training program*. The training program consisted of several auditory discrimination tasks, an auditory WM task and an auditory episodic memory task. The levels of difficulty were adapted to the participant‘s performance during the training. The training group practiced during a minimum of 40 and a maximum of 50 one-hour-sessions. They were compared to a no-contact and an active control group. The active control group watched DVD-based educational lectures for the same amount of sessions. The overall mean age of the participants was 71 years. The authors observed significant near transfer effects to a global memory score based on six auditorily presented cognition tasks, which assessed short-term memory, WM and episodic memory performance. Maintenance of transfer effects in the three-month follow-up testing was demonstrated for only one short-term memory task; the digit-span forward task. Even though it is impossible to separate the effect of the WM task training from the other process-training task, there is a remarkable transfer effect after the training program. However, the maintenance of transfer effects was rather small.

Buschkuehl and colleagues (2008) conducted a WM training study with older adults (*M*=80 years) over 23 sessions. The training included WM tasks and speed tasks. The WM tasks were a visual short term memory task and two versions of a picture span task. These three tasks were adaptive according to the participants' performance. Speed training tasks consisted of two choice reaction time tasks. As a control group they included a physical training group which exercised at a low cardiovascular engagement and trained to the same amount of sessions and time as the cognitive training group. The WM training group improved performance in all three trained tasks significantly. A near transfer effect was found in a merged forward and backward block span task score and a far transfer in a visual free recall task. No transfer effects were found in the digit-span task and verbal free recall task, and in addition, no maintenance one year after the training was observed. This study showed that older adults can exhibit reliable transfer effects, also in an active control group design, even though they remain here rather sparse.

Striking transfer effects have been reported by Borella, Carretti, Riboldi, and Beni (2010). They investigated the effect of WM training in older adults (65-75 years). The authors trained the participants during three sessions lasting 60 minutes each, on an adaptive categorization WM span task. During this task the participants had to listen to several lists of words and tap their hand on the table whenever the word was an animal noun. At the end of a series of word lists, they were asked to recall the last noun of each list. The training group was compared to an active control group which completed questionnaires about autobiographic memory and well-being. The analyses of the training and transfer tasks revealed important results: There was a substantial gain in the trained task for the WM training group. Further, large near and far transfer effects and maintenance (over eight months) effects for the WM training group were found with nearly all tasks used for investigating training effects. That is, in tasks assessing performance in short-term memory, WM, processing speed, inhibition and fluid intelligence.

Transfer effects from WM training in older adults were further recently addressed by Richmond, Morrison, Chein, and Olson (2011). They were also interested in transfer effects to self reported everyday cognitive functioning. The authors conducted a verbal and a spatial complex WM span training over 20 sessions in older adults between 60 and 80 years old. The training group was compared to an active control group which completed 20 sessions of cognitively engaging but low WM load quizzes. The results revealed an improvement in a WM task and in an episodic memory task. When excluding the participants without improvement during WM training from analysis, more far transfer effects for episodic memory measures were observed. The everyday functioning self-report revealed a significant improvement in everyday attention. This last reviewed study confirmed the prior findings that older adults exhibit near and far transfer effects.

Age differences in working memory training

In a further study the training was conducted on WM tasks only. Dahlin, Nyberg, Bäckman, and Neely (2008) trained younger (*M* = 24 years) and older (*M*= 68 years) adults on five adaptive WM tasks which all engaged updating processes. The participants trained over five weeks during 15 sessions for 45 minutes each. They were compared to a no-contact control group. The authors found a significant training gain for both age groups compared to the no-contact control group which performed this task only twice, at pre-test and at post-test. Younger adults improved more on the trained task than the older adults. However, transfer effects were limited to younger adults who showed a near transfer effect to a very similar task (number 3-back) as the ones trained (updating colors, numbers etc.). There was further a far transfer effect in younger adults in an episodic memory task which required an updating of memory representations. The training gain and the near transfer effect were maintained some 18 months after training.

S.-C. Li and colleagues (2008) also trained younger (20-30 years old) and older (70-80 years old) adults on a spatial WM updating and shifting task during 45 sessions. The training task however was not adaptive to the individual's performance, contrary to the training methods in the two previous studies. Participants were compared to a no-contact control group as well. Both age groups improved significantly in the trained tasks, with a trend towards more improvement in older adults. Near transfer effects were found to three very similar tasks to the trained tasks. Contrary to the author's expectation the transfer effects were similar in younger and older adults. These effects maintained over three months. No evidence for far transfer to complex span tasks was found. These results are contradictory to the results reported by Dahlin and colleagues (2008), who only discovered transfer effects in younger adults.

In the more recent COGITO study by Schmiedek, Lövdén, and Lindenberger (2010) younger (20 - 31 years old) and older adults (65 - 80 years old) were trained during 100 sessions. They practiced on twelve different tasks engaging WM (updating and shifting), episodic memory or processing speed. The task difficulty was adapted to the participant's pre-test performance at the beginning of the training; whereas during training no adaptive procedure was used. They were compared to a no-contact control group. Instead of analyzing the individual transfer test separately, the authors used a latent difference score model approach, where transfer effects were examined on latent cognitive ability factors. The results revealed a near transfer effect to a latent factor of WM ability in both age groups. The tasks constituting this latent WM factor required the same processes as the practiced task, but were based on different content material. Far transfer effects to the latent factor of fluid intelligence and episodic memory was limited to younger adults. Transfer to the latent factor of processing speed tasks did not reach a significant level.

This study also showed that transfer effects remain possible in older age and can even be similar to younger adults. However, here the range of transfer was limited in older adults compared to younger adults. The authors discuss several limits of the study like the lack of an active control group and the limited adaptivity of the training procedure. They further suggest that there should be more time allotted to WM training, as their study showed that WM training seems to be more efficient than others. Finally, the authors proposed that future research should investigate the neural mechanisms underlying transfer effects, in order to better understand the changes in neural system efficiency.

Summary and discussion

With respect to younger adults, WM training appears to have a great potential to exhibit near and far transfer. However, the far transfer effect to fluid intelligence measures seems not to be consistently confirmed.

As regards older adults, considerable transfer effects were reported as well, even though the pattern was less clear than in younger adults. However, only one study reported very broad near and far transfer effects (Borella et al., 2010). These results are indeed surprising considering that to date no such complete transfer effects are reported. The authors argue that their training task engaged multiple processes including encoding, maintenance and inhibition of information, task-switching and shifting attention. On the one hand, when compared to simple updating training tasks (e.g., Dahlin, Stigsdotter Neely, Larsson, Bäckman, & Nyberg, 2008), there were certainly more processes engaged by the training of Borella and colleagues (2010). On the other hand, Schmiedek and colleagues (2010) for example trained three different processes on twelve different tasks. However, no such large extent of transfer effects was manifested. Borella and colleagues argue that they used an adaptive training protocol which kept the task challenging. Borella and colleagues also suggested that the shorter training regime (three sessions only) and the human-human interaction during the training sessions might have had a positive effect on motivation and effort recruitment during training, which results in a beneficial training effect.

In sum, significant gains in the trained task were reported in all training studies. Transfer effects for younger and for older adults were reported in nearly all presented studies. However, when the transfer effects were compared to the younger training group, older adults exhibited in general less transfer effects. Overall, the body of evidence appears still sparse with regards to the comparison between younger and older adults. Moreover, there exists a considerable diversity of training procedures and experimental designs which leads to heterogeneous results whereof no clear conclusions can be drawn to date. As Martin and colleagues (2011) suggested, more standardized study protocols are needed, in order to maximize comparability of studies and the possibility of data pooling. To fully understand the potential of a WM training procedure and its limits in transfer effects in older adults, more research is evidently needed.

Brain Plasticity and Training

Brain plasticity refers to the processes which enable the nervous system to modify structure and function in order to adapt the behavior to the rapidly changing environmental demands (e.g., Gazzaniga, 2004). These processes involve the strengthening, weakening, pruning, or adding of synaptic connections, and the promotion of neurogenesis all over the lifespan (A. Pascual-Leone et al., 2011).

One of the main goals of the current work was the investigation of age-related differences in training-induced cerebral plasticity.Here we review the few respective studies in more detail. We first introduce structural brain plasticity, and then, functional brain plasticity with a focus on WM training including younger or older adults separately. This review provides on the one hand evidence that brain plasticity in brain structure as well as in brain function cannot solely be observed in comparative expertise studies, but also after training interventions. On the other hand, these studies provide evidence that age-related functional decline in frontal regions is not an inevitable process of aging, but can be altered and reversed with training.

Structural Brain Plasticity

Plasticity at the structural brain level has mainly been investigated by comparing the cortical organization of experts versus laymen. Differences have for example been observed in Braille readers (Sterr et al., 1998), where the somatosensory representation of the fingers is topographically disordered in Braille readers compared to the control group. Structural differences in grey and white matter has been reported in musicians when compared to laymen (Jäncke, 2009; Oechslin, Imfeld, Loenneker, Meyer, & Jäncke, 2010). Further, increased hippocampal grey matter volume has been found in taxi drivers compared to laymen controls or bus drivers (Maguire et al., 2000; Maguire et al., 2003; Maguire, Woollett, & Spiers, 2006). However, these observations have to be confirmed by an experimental study, as the cerebral predisposition and practice are confounded. No answer can be given to the question of whether the brain’s organization is the result or the cause of expertise.

Training-induced plasticity

The literature is scarce as concerns training studies which include younger and older adults by using brain structural measures. And even more as concerns cognitive training. Here we review a study reporting structural changes with motor skill learning training and a second study investigating induced changes from WM training - to our knowledge the first study in this domain.

Juggling Training. Draganski and colleagues (2004; Driemeyer, Boyke, Gaser, Büchel, & May, 2008) provided one of the first studies in which brain plasticity was experimentally manipulated. They used random assignment and trained younger adults to juggle for three months. Structural MRI scans were conducted before and after juggling training and compared to a control group. The training group showed structural changes after the three months of training, which partially persisted for another three months after the end of the juggling training. This evidence showed that a structural cerebral change can be provoked by training as well, mediated by the process of brain plasticity.

In an extension of this study, Boyke, Driemeyer, Gaser, Büchel, and May (2008) used the same procedure and trained an older sample (M = 60 years) during three months to juggle. An MRI scan was carried out before and after the three months of training and again three months after the juggling training finished. Results revealed that there was a training-related transient increase in grey matter volume, mainly in the middle temporal area of the visual cortex. This increase, however, returned to the initial baseline level in the third scan. This area is associated with processing and storage of complex visual motion. Further, the authors compared the younger and older adults directly. Results showed that older adults did by far not achieve the same level in juggling as the younger adults did. Nevertheless, both age groups showed an increase in grey matter volume in the middle temporal area, but it was somewhat less pronounced in the older adults. The grey matter expansion observed at post-test was decreased at the three-month-follow-up stage for both age groups. Additionally, the authors found a significant correlation between grey matter change and juggling performance, but only in the younger adults. This study confirmed the preserved structural plasticity in older age in grey matter.

WM Training. In another more recent study, a subsample from the COGITO longitudinal study was used (see above section Working Memory Training Studies; Schmiedek et al., 2010). Lövdén, Bodammer and colleagues (2010) trained younger (20 - 31 years old) and older adults (65 - 80 years old) during 100 sessions in different tasks: engaging WM (updating and shifting processes), episodic memory and processing speed. The task difficulty was adapted to the participant’s pre-test performance at the beginning of the training, whereas during training no adaptive procedure was used. In addition to the behavioral measurement, MRI scans (including diffusion tensor imaging (DTI) measures) were assessed. The training group was compared to a no-contact control group which performed the imaging sessions only. The authors were interested in white matter changes mainly in the anterior corpus callosum. These white matter tracts connect the prefrontal cortices and show pronounced age-related changes.

The results revealed behavioral training improvements in both age groups, but more pronounced ones in younger adults. The DTI results showed increased white-matter microstructure in both younger and older training groups to the same amount. Changes were most important in the anterior part of the corpus callosum. However, no significant correlations were obtained between white matter changes and behavioral performance improvements. The authors explained the lack of the behavior-brain correlation with the small number of participants (*n*= 20 for younger; *n*=20 for older). According to the authors this study showed that experience-dependent plasticity in white matter microstructure was also present in later adulthood. The observed structural changes showed that the myelination processes occur beyond the maturational myelination and continue until older age.

Since these results showed a similar brain structure change in both age groups, they contradict the findings from juggling training, which showed less cerebral change in older adults than in younger adults. However, the training regimens taxed different domains so that the results may not be directly comparable. Nevertheless, these age-comparative studies showed that white matter as well as grey matter is still plastic in older adults.

Functional Brain Plasticity

Functional brain plasticity has also been studied in experts, for instance in musicians, where functional analyses were able to show that musical experts showed a different processing of musical stimuli compared to musical laymen (James, Britz, Vuilleumier, Hauert, & Michel, 2008). Practice-related functional brain plasticity has largely been investigated in motor processes as well, in its development and in motor skill learning (e.g., Hauert, 2004; Karni & Ungerleider, 1995; Ungerleider, Doyon, & Karni, 2002).

Kelly and Garavan (Kelly et al., 2006; Kelly & Garavan, 2005) distinguished between four patterns of functional changes associated with practice or patterns of functional plasticity: Activation decrease, activation increase, functional reorganization as well as redistribution of regional activations. Activation decreases in reponse to a task, which results in a decreased spatial extent of activation or a reduction of activation, are associated with an increased neural efficiency reflected by the more specific use of neuronal networks and a decreased number of neurons firing strongly (Poldrack, 2000). Activation increases correspond to increases in strength of activation or to activation expansions which reflect the recruitment of additional cortical units. Further, Kelly and Garavan describe two patterns which combine activation decreases and increases at the same time. They distinguish between activation redistribution and true activation reorganization. The activation reorganization reflects a qualitative change in the activation pattern which implies a shift in the cognitive processes underlying task performance (Poldrack, 2000). This refers to a process switch, for example from one strategy to another, which results in a coordinated decrease and increase of activation in separate brain regions. The activation redistribution in contrast, refers to a quantitative change in the activation pattern. That is, the same brain regions are activated before and after practice, but the levels of activation within those regions have changed. This implies that the cognitive processes engaged did not fundamentally change by practice.

The redistribution pattern often reflects a decreased demand on attentional processes and an increased demand in task-specific brain regions with practice. Kelly and Garavan linked this pattern with the scaffolding framework we have previously introduced (Petersen et al., 1998). It describes the processes when frontal brain regions are recruited in order to serve as a scaffold for task-specific regions to cope with the new task demands. Once the task-specific processes are more efficient, the scaffolding network is no longer needed and decreases. This decrease is known as pruning or neurophysiological trimming (Ramsey, Jansma, Jager, Van Raalten, & Kahn, 2004) of attentional processes which occurs with practice.

This scaffolding process has been proposed to integrate the non-selective recruitment approach with the compensation approach for explaining the frontal overrecruitment in older adults (Goh & Park, 2009; Park & Reuter-Lorenz, 2009; Reuter-Lorenz & Park, 2010). The frontal scaffolds compensate for the inefficient and non-selectively recruited task-specific processes in older adults.

As regards the qualitative interpretation of these patterns of functional changes in relation with the frontal over-activation pattern in older adults, we propose the following interpretations. When activation decrease occurs in frontal regions, it may be qualified as a decrease in non-selective recruitment. In contrast, when an activation increase occurs in frontal regions the compensational activity might be increased due to training. However, these interpretations are only possible if the performance at the behavioral level effectively increased due to training. The change pattern has to be beneficial and therefore underlie an improved performance.

Furthermore, both the reorganization and the redistribution pattern could reflect less compensational recruitment whenever the frontal activation decreases and posterior and task-specific regions increase in activity. This would be the interpretation according to the scaffolding hypothesis. However, these patterns could equally reflect a decrease in non-selective recruitment towards a more task-specific recruitment in posterior regions. To disentangle these confounded interpretations in older adults, it is therefore indispensable to include a younger group which serves as a reference. For younger adults, we assume that no non-selective recruitment is observed in contrast to older adults. If the training-induced changes lead to increased age differences after training, we could interpret the overrecruitment as compensational. That is, if the frontal overrecruitment pattern persists through training and leads to improved performance in older adults, age differences would not decrease with training. The age-specific activation pattern would be needed for the improved performance and can be interpreted as beneficial and therefore compensational.

In contrast, whenever training brings reduced age differences, we assume that the overrecruitment pattern before training was not beneficial. In other words, when the older adults' activation pattern after training tends to be similar to the younger one, we assume that an age-common pattern would be most beneficial and underlie the improved performance. The frontal overrecruitment would in this case not underlie an improved performance and therefore be interpreted as a non-selective recruitment.

Again, these interpretations require the assumption that with training the activation pattern would be optimized and underlies an improved behavioral performance. The pattern after training would in that case be more beneficial than the one before training.

In recent years an increasing number of studies emerged aiming at investigating the potential and limits of functional brain plasticity with a particular focus on older age. Brain plasticity has been studied in connection to physical training as well as with several kinds of cognitive training interventions such as process-based training or strategy training (for reviews see Greenwood & Parasuraman, 2010; Kramer & Erickson, 2007; Lustig et al., 2009; Willis & Schaie, 2009). The focus in the following review is first on brain plasticity associated with WM training, first in younger adults and then in older adults. Finally, we focus on age-comparative training studies, where we review findings from a strategy- based training, a dual-task training and finally from a WM training.

Plasticity due to working memory training in younger adults

Recent reviews of the WM training-induced effects in younger and older adults revealed no clear or consistent pattern of cerebral changes (Buschkuehl, Jaeggi, & Jonides, 2012; Klingberg, 2010). Evidence for all of the four different brain plasticity patterns proposed by Kelly and Garavan (2005) was reported.

In one of the first studies investigating functional brain plasticity, younger adults were trained daily for five weeks (Olesen et al., 2004). The training was adaptive to the individual performance and included several visuo-spatial WM tasks such as spatial rotation tasks from the commercial Cogmed training program (Klingberg et al., 2005). The task used during scanning is probably very similar to the trained tasks, but no further information is provided. During the scanned task after training, an activation increase in middle frontal gyrus and superior and inferior parietal cortices during a visuo-spatial WM task was found. At the behavioral level, transfer effects to the Stroop task, the span board task and the digit span task were reported. The results showed an increased pattern of the brain activity in WM related frontal regions. However, the training effects were not controlled by the inclusion of a control group.

Westerberg and Klingberg (2007) also found increased activity after a training intervention in three younger adults, who trained between four to six days a week for five weeks on three WM tasks. The WM training was adaptive and included a visuo-spatial WM task, a backwards digit span task and a letter span task (Cogmed; Klingberg et al., 2005). After training, the middle and inferior frontal gyrus exhibited increased activity in a very similar task to the trained visuo-spatial WM task. There were no additional brain regions involved but a small increase in the extent of the activated area was observed. The authors reported further near and far transfer effects to a nontrained visuo-spatial WM span task and the Raven's Advanced Progressive Matrices. The trained participants were compared to a no-contact control group.

Hempel and colleagues (2004) trained younger adults on a visuo-spatial *N*-back task which was randomly presented in a 0-, 1- or 2-back condition. The participants trained twice a day during four weeks and were scanned before training, after two weeks and after the four weeks of training. After two week of training, they observed increased activity in the right inferior frontal gyrus and the right intraparietal sulcus. At the end of training however, the same regions showed decreased activation. The authors suggested that training-related cerebral activation changes follow an inverse U-shaped quadratic function. The reported frontal activity increases were comparable to the activity increases found in the two previous studies. However, Hempel and colleagues were able to show in addition, that with extensive practice on a single task, brain activity decreases. Even though the former two studies trained to almost the same duration, they included different tasks, so their training was probably not that extensive for one single task as in Hempel and colleagues (2004). The scaffolding framework could explain this pattern, since the frontal activation was increased during the first sessions of training in order to serve as scaffold for the new demand. After continued training, the scaffold was probably no longer used and decreased as a consequence. However, the authors did not show a clear redistribution of activation, so the results cannot fully be explained by the scaffolding process. Since no control groups were included in the study, the training-induced pattern is therefore at least confounded with test-retest effects.

Another study which reported activation decreases also trained the younger adults on an *N*-back paradigm (Schneiders, Opitz, Krick, & Mecklinger, 2011). The authors compared a visual *N*-back task training with black- and white pattern stimuli to an auditory one with bird voice stimuli. They further included a no-contact control group. The *N*-back training was adaptive and was implemented in ten sessions of 50 minutes. A 2-back condition of the trained visual task was used during scanning, which was therefore a training gain task for the visual training group and a structurally similar near transfer task for the auditory group. A more pronounced behavioral gain during training was observed for the visual training group compared to the auditory training group. For the task during scanning, there was as expected more gain from training for the visual training group. On the brain level, a decreased activity was found in the right middle frontal gyrus after the visual training only. Further, both training conditions led to decreased activation in the superior right middle frontal gyrus and the posterior parietal region as compared to the no-contact control group. The authors concluded that there was on the one hand an increased neural efficiency of visual processes after the visual training only. On the other hand, there was an increased neural efficiency of general control processes after both modalities of *N*-back training. One could qualify the latter changes as task-modality-unspecific changes in WM processes, since previous studies also reported changes to WM training in these regions. However, as these effects were present in both training groups compared to a no-contact control group, the training effects cannot be attributed to the specificity of *N*-back training alone. Another placebo training group should have been included.

Plasticity due to working memory training in older adults

A further study, conducted with older adults (60-70 years old), also reported activation decreases after training (Brehmer et al., 2011). One group received an adaptive WM training whereas another group received low-level non adaptive training and served as control group. The participants were trained during five weeks including 25 sessions on four visuo-spatial and three verbal WM tasks (Cogmed; Klingberg et al., 2005). The WM task during scanning was again very similar to one of the trained visuo-spatial WM tasks. The authors reported a near transfer effect to the span board task and far transfer effects to a sustained attention task and an episodic memory task. However, performance in the task during scanning did not show any training-related improvement as the task exhibited a ceiling effect similar to previous studies (Olesen et al., 2004; Westerberg & Klingberg, 2007). On the brain level, they observed on the one hand intervention-general effects, i.e., in both training groups, an activity decrease in the inferior frontal, anterior cingulate and inferior parietal cortices as well as in occipital and hippocampus regions. On the other hand, intervention-specific activation decreases for the adaptive training group were reported in dorsolateral prefrontal cortex (DLPFC) and in superior temporal and occipital regions. Similar to the findings from Schneiders and colleagues (2011), they found training-level specific and training-level independent activity changes. The authors conclude that the training-related decreases indicate increased neural efficiency which was more pronounced in the adaptive training group. However, the intervention-general effects were confounded with test-retest effects, since a no-contact control group was not included.

Age differences in training-induced plasticity

To our knowledge one PET study and four MRI studies exist to date which investigated age-related differences in changes after cognitive training and one study which investigated changes after motor-skill-learning training.

Strategy-based training. The first study combining both the behavioral and the brain level of analysis in younger (*M*=26, *SD*=2.6) and older adults (*M* = 69, *SD*=2.5) within a training program was reported by Nyberg and colleagues (2003). They conducted a PET study to investigate the neural underpinning of the acquisition and the use of a mnemonic strategy, the method of loci which implied visualizing and memorizing locations and associating words to these locations (for more details see above in the section *Seminal Cognitive Plasticity Research in Older Adults*). They submitted the participants to a training procedure including three phases: a pre-test phase, a loci acquisition phase and a loci utilization phase. In each phase they underwent several PET scans and practice sessions. The mnemonic training was not adaptive to the participant's performance, but additional training was provided when a minimum of correct responses was not reached after the acquisition phase.

The behavioral results showed the classic pattern of training-related age differences in episodic memory performance: both age groups improved memory performance, but the difference between the age groups was magnified due to a larger improvement in younger adults than in older adults. After the data acquisition they divided the older training sample into a subset of older adults not benefiting from the mnemonic training (unimproved group) and a subset improving memory performance from pre-test to post-test (facilitated group).

They showed that the magnified age difference after training was mainly due to the performance of the considerable number of unimproved older adults. The improved older adults sample showed therefore the same amount of gain from pre-test to post-test as younger adults.

As concerns the PET results, the acquisition of the loci structure, i.e., learning to visualize and memorize locations, was overall associated with an increased activation in the bilateral parietal and medial parietal cortex, as well as with the left hippocampal region. No age group differences were found. As regards the activation during the loci mnemonic use comparing pre-test versus post-test, overall analysis revealed activation in the left occipito-parietal cortex and the left DLPFC. Further analyses comparing the subgroups in older adults and the younger adults revealed the following group differences: The younger and the older facilitated groups showed activation in the left occipito-parietal cortex, but not the unimproved older group. The left DLPFC, in contrast, was specific to the younger group only; hence there was no increased frontal activation in both older groups.

In order to exclude the possibility that older adults would already have shown more frontal activity before training, the pre-test scan was contrasted to a non-demanding baseline task scan. They showed that no age-differences and thus no compensatory and HAROLD-like activation pattern was observed in the older adults before training commenced.

The parietal activation was associated with encoding of spatial information. This activation was also present in the older unimproved group during the loci acquisition but not during the loci use. The results therefore suggested that the unimproved older adults had difficulties with the implementation of the loci method, but not in acquiring it.

Furthermore, only the improved older and the younger group showed an increased occipito-parietal activation after training, which was associated with the use of visual imagery. According to the authors, these results suggest that the older unimproved group did not use the loci mnemonic for word encoding. In contrast, neither of the older groups showed an increased frontal activation during loci use after training. The authors interpreted this as a sign of the generally reduced processing capacity in older adults. The left dorsal frontal cortex is linked with processes like the integration of information in WM, or the creation of an organizational structure. As these processes and their underlying frontal activation are usually reported as reduced in older adults, Nyberg and colleagues (2003) suggested that the lack of frontal activation in older adults was due to a general age-related deficit in basic cognitive resources. However, older facilitated adults achieved a change in occipito-parietal activation, which shows that not all older adults do show a general failure to engage cognitive resources in an appropriate way.

This study showed that age differences were magnified in behavior as well as in cerebral functioning when considering the entire older group. However, when comparing only the facilitated older group to the younger group, the behavioral improvement was no longer magnified, but the cerebral functioning was still different. It is not completely obvious how the facilitated older adults were able to improve their performance to the same amount as the younger, but with limited processing resources and without recruiting additional frontal resources. The authors did not provide an explanation relative to this issue. These results can be compared to the results of Hempel and colleagues (2004), where younger adults showed an increase in activation during learning which decreased in the course of further training. Similarly in this study, the activation was increased during learning and then decreased or no longer present at the same sites during successful strategy use. However, this plastic pattern has to be qualified as activation increase with training, since activation was still increased at post-test compared to pre-test.

The results seem in addition contradictory to the generally observed over-activation pattern in frontal regions in older adults. They generally recruit more and less specialized frontal areas, but show a decreased activation in occipital regions (see section Models of Functional Cerebral Changes; Cabeza et al., 2004; Payer et al., 2006). In particular Cabeza and colleagues (2002) showed that high-performing older adults seem to compensate their age-related neural deficits by recruiting bilaterally and increasingly the frontal areas in order to perform at the same level as the younger adults. According to Cabeza and colleagues (2002) the low-performing older adults recruited a similar network as young adults, but used it inefficiently. Thus the interpretation of Nyberg and colleagues' results is not clear; in older adults they found neither a generally increased frontal activation before training, nor a change in frontal activity after the training. In contrast, the younger adults increased frontal activity with the mnemonic training. So these results are neither accounting for the compensation view nor for the non-selective recruitment hypothesis. However, no control group was included in this study, which would have at least controlled for test-retest effects.

Dual-Task Training. Erickson and colleagues conducted a single and dual task training during five 1-hour sessions with younger (19-32 years old) and older (55-76 years old) adults (Erickson et al., 2007a, 2007b). The single tasks required the participants either to make a decision about the color of an *X* (green or yellow) or about the letter presented (*B* or *C*). In the dual-task condition, a response was required for both conditions at the same time.

Feedback adapted to their response times after each block was provided. The training groups were compared to a no-contact control group.

The results revealed that the younger and the older training groups decreased reaction time and increased accuracy under both conditions to the same amount, compared to the control group. Concerning the neuroimaging data, the overall results showed significant signal changes in ventral prefrontal cortex (VLPFC) and DLPFC in both hemispheres in the training groups only. In the left VLPFC the older training group showed increased activation after training compared to the control group. The younger group in contrast did not change signal activation in left VLPFC, but age-differences in activation were less pronounced after training for the training groups. The activation in right VLPFC in turn was decreased after training in both training groups and age-differences were less pronounced. Thus, as concerns the VLPFC, there was an increased asymmetry in activation in both age groups and reduced age-differences in activation in both hemispheres after training.

As regards the DLPFC, older adults showed reduced activation in both hemispheres after training whereas the younger training group showed increased activation. The changes were such that age difference no longer existed after the training. These findings from the DLPFC were in line with the findings from VLPFC. Furthermore, all activation changes were correlated with improved performance in the dual-task condition and were therefore relevant for behavioral changes.

The authors associated the activation change in the left VLPFC to the use of verbal strategies during dual-task performance. As this VLPFC increase occurred in older adults only, the authors suggested that older participants made more use of this strategy with training as compared to younger adults. The activation reduction in the right VLPFC in turn was associated with the more efficient use or its reduced need and was independent of age. The VLPFC results revealed that both age groups showed a similar training-related increase in lateralization indicating that they responded in a similar way to the training. However, the authors suggested that the increased asymmetry in the VLPFC regions in older adults conflicts with the compensation approach of reduced asymmetry (Cabeza, 2002). The DLPFC results also revealed decreased age-differences after training, which was inconsistent with the compensation hypothesis. The authors hypothesized that older adults may have learned to use these regions, which are related to cognitive control and task coordination, in a more efficient manner; whereas younger adults may have learned to more extensively recruit these regions in order to reduce demands in other regions.

The changes in younger adults could be qualified as an activation reorganization pattern. The changes in DLPFC could be considered as an activation in a new brain area for younger adults, since it was very close to zero at pre-test. However, no information was provided whether the difference at pre-test was significant from zero or not. For older adults, we propose to label the change an *activation redistribution*, as there was increase and decrease on the same cerebral regions.

WM Training. In a further study addressing training-induced cerebral changes in younger and older adults, Dahlin, Neely et al.(2008) added fMRI measures at pre-test and post-test to their updating training procedure (see above section *Working Memory Training Effects in Older Adults*). Their aim was to find activation overlaps between the trained and the transfer task. During MRI scans, the participants conducted the training task, i.e., the letter updating task and two untrained tasks, i.e., the 3-back task and the Stroop task. For the trained letter memory task, a significant increase in the left striatum activation was observed in both age groups. However, only the younger adults already showed significant striatum activation at pre-test, whereas the older adults showed no activation before training but only after training. Further and in younger adults only, there was an additional decrease in frontoparietal activity and a striatum activation overlap between the trained updating task and the untrained 3-back task. No activation overlaps were found for the Stroop task.

These findings were consistent with the behavioral findings, younger adults showed a near transfer effect into the 3-back task whereas older adults did not exhibit transfer effects. The striatum was therefore found to mediate transfer effects and regulating updating processes through dopaminergic regulations.

The changes in younger adults can be considered as an activation redistribution pattern, since increases in subcortical activation occurred combined with decreases in frontal regions. The results were in line with the scaffolding hypothesis, as the frontal attentional recruitment activation decreased but task-specific activity, i.e., the striatal updating related activity, increased. This has been confirmed by the activation overlap in the striatal regions in another untrained updating task.

For older adults, the pattern corresponds to an activation increase with training, but in task-specific regions only. This is in line with Nyberg and colleagues' study where similarly only a change in task-specific regions was reported but no change in frontal regions. This contradicts however the findings from Erickson and colleague's study, as they reported frontal changes in both age groups.

Even though no transfer effects were observed in older adults, the effect of training at the cerebral level seems to be partially similar in both age groups.

Summary and discussion

In sum, all four training-induced plasticity patterns were observed: Activation increases were reported in frontal areas from Olesen and colleagues and Westerberg and colleagues whereas task-specific posterior or subcortical increases were reported by Nyberg and colleagues and by Dahlin and colleagues for older adults. A frontal activation increase and then a decrease were reported from Hempel and colleagues. Frontal activation deceases were showed by Schneiders and colleagues as well as by Brehmer and colleagues, whereby the latter in older adults. Then, an activation redistribution in frontal regions was found by Erickson and colleagues in older adults and in frontal and subcortical regions by Dahlin and colleagues in younger adults. Finally, Erickson and colleagues reported a reorganization in frontal areas for younger adults. So for younger adults, all four patterns were reported and frontal areas were involved in all changes. However, for older adults all but a reorganization pattern was found but frontal changes were not always implied. Subsuming these findings, neither an age-specific nor a WM training-specific pattern could be detected.

A clear interpretation with respect to the issue of the over-activation in older adult was not possible either. The results from Brehmer and colleagues indicated less non-selective recruitment after training, as frontal activity was decreased and we therefore assume that the frontal activity was not beneficial before training. However, the reported activation increases from the study of Nyberg and colleagues as well as Dahlin and colleagues involved task-specific regions but no change in frontal activation in older adults. This could reflect that the frontal overrecruitment was beneficial as it was not changed by training. However, as already discussed, the former study entails an unclear pattern before training. Hence, as concerns the Dahlin and colleagues' findings older adults still required the similar frontal activity as before training, in contrast to the younger adults who decreased frontal activity. The frontal activation in older adults could therefore be interpreted as compensational since it did not change with training. In contrast, the redistribution pattern in older adults found in Erickon and colleagues' study showed that the frontal activity was reduced to the level of younger adults after training. This indicated that a youth-like frontal activity is more beneficial and would therefore support the hypothesis of a non-selective recruitment rather than the compensation hypothesis.

Even though the brain plasticity pattern is inconclusive, we were able to draw several important conclusions and a caveat from this body of evidence.

First, there is a potential for training-induced functional cerebral plasticity in both younger and older adults. Second, the changes seem to occur after different types of intervention, after WM training as much as after dual-task training as well as episodic memory training.

Third, as concerns the age-comparative studies, there is no clear pattern of change. The results were contradictory, since the age-related differences in brain activity were increased by training in Nyberg and colleagues' study as well as in Dahlin and colleagues' study whereas they were decreased in Erickson and colleagues' study. More research to understand these changes is therefore needed.

Fourth, older adults showed task-specific changes similar to younger adults. But younger adults showed in addition training-induced changes in frontal regions which were less extensive in older adults.

Finally, to our knowledge, only two studies (Dahlin, Stigsdotter Neely et al., 2008; Schneiders et al., 2011) investigated cerebral activity in an untrained transfer task to date. The study of transfer effects at the cerebral level would potentially be very instructive to better understand the mechanisms of training-induced brain plasticity.

However, there is an important caveat to this literature review; all the reported studies demonstrated a lack in their methodological design. The control for retest and placebo effects was in all studies insufficient, since no control group, or either a no-contact or an active control training group was included, but none of the studies included both. This does not allow assuming that the observed effects are proper to the training intervention; no reliable conclusions can therefore be drawn.

Generally, the body of evidence from older adults and even more from age-comparative training research is scarce and heterogeneous and does not yet provide a conclusive pattern. More research from well-designed studies implementing the process-based training approach is needed.

Excursus: Differential Effects on ERP Components

In this section we provide an introduction to the cerebral measures implemented in the present study. In order to measure plastic changes at the brain level, we chose the EEG method. EEG allows the recording of voltage fluctuations at the scalp, which reflect the summation of post-synaptic potentials in cortical pyramidal neurons (Pizzagalli, 2007). Event-related potentials (ERP) are extracted when the EEG signal is averaged over several epochs that are time-locked to the onset of a stimulus. The ERP is a time series of the scalp voltage in microvolt (μV) over time and contains the information directly related to the processing of the stimulus (Handy, 2005). Within an ERP, different waveforms or components are distinguished, which show either positive or negative voltage amplitude during a particular time window (latency) after stimulus onset at particular electrode sites. They result from specific voltage maps, that is, a specific polarity distribution over the scalp, which occurs in the course of these time windows. An ERP component can be defined as the reflection of the brain activity on the scalp in particular brain regions over time (Michel, Koenig, Brandeis, Gianotti, & Wackermann, 2009).

We first introduce the advantages of the use of an ERP approach with EEG, followed by the ERP correlates of updating and WM processes. Finally we discuss several sources of individual differences in these ERP measures: age, cognitive load, practice, and task performance.

Advantages of EEG in Cognitive Aging Research

The reasons for the choice of electrophysiological ERP measures and against hemodynamic imaging techniques were, besides its availability in the laboratory of the FPSE, manifold: First, compared to hemodynamic methods such as fMRI or PET, EEG provides a high temporal resolution of the neuronal signal in the millisecond range. In contrast, hemodynamic signals lag the corresponding neuronal signal by several seconds. This is in particular a problem when comparing younger and older adults, as the signal lag may be larger and more variable for older adults such that the BOLD signal may differ in quality between age groups (Fabiani & Gratton, 2005). Second, the cardiovascular response can be altered in older adults as compared to younger adults and therefore change the BOLD response without necessarily reflecting a proportional change in the underlying neuronal signal (D'Esposito, Zarahn, Aguirre, & Rypma, 1999). Third, BOLD measures integrate information about the time spent on the task. Older adults do generally process slower than younger adults and spend therefore more time on the task, which can alter the hemodynamic response. So the hemodynamic measures could potentially confound the time on task with the magnitude of activation (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). Finally, as proposed by Zöllig and Eschen (2009), EEG in an ERP approach is particularly recommenced in age-comparative research. ERP allows to analyze the correct, false and non-responses separately, that is, the comparison of successful and unsuccessful trials. This is a great advantage as compared to BOLD measures, in which these types of responses are confounded. This point is indeed important for research comparing younger and older adults, as there are usually significant age-related accuracy differences in addition to the reaction times differences.

ERP Components Related to Working Memory and Updating Processes

In the present study we investigated plastic changes in *N*-back tasks and its underlying ERP correlates, where WM processes and in particular updating processes are demanded. The components of interest in the *N*-back paradigm were the N2b (aka N200b) and the P3b (aka P300b; Friedman, 2008). The N2b is a negative deflection occurring at approximately 200-300 ms post-stimulus and has its maximum negative peak at central-frontal electrode sites. The P3b is a positive deflection occurring at approximately 400 - 600 ms post-stimulus and shows its maximum positive peak at central-parietal electrode sites. In the following we will use the terms N2 for N2b and P3 for P3b.

These components are observed in tasks where participants allocate and focus attention to make a task-relevant decision. The N2 component is associated with early mismatch detection, decision processes and response selection which control the behavioral response (Folstein & Van Petten, 2008; Gajewski, Stoerig, & Falkenstein, 2008; Ritter, Simson, Vaughan, & Friedman, 1979; van Veen & Carter, 2002). The P3 component is associated with an attention driven comparison process which evaluates the representation of the previous stimuli in WM, is related to updating processes and more generally to resource allocation capacity (Donchin & Coles, 1988; Picton, 1992; Polich, 2007; Watter, Geffen, & Geffen, 2001). These components are widely used within the *N*-back paradigm (e.g., Chen & Mitra, 2009b; Chen, Mitra, & Schlaghecken, 2008; Daffner et al., 2011; Missonnier et al., 2004; Segalowitz, Wintink, & Cudmore, 2001; Shucard et al., 2009; Wintink, Segalowitz, & Cudmore, 2001).

Age effects

The N2 as well as the P3 component undergo age-related changes in amplitude, latency and scalp topography; here we focus on changes during adulthood.

The N2 component is characterized by an age-related increase in latency (Friedman, 2008), but evidence is less clear with respect to its amplitude. Several studies reported amplitude decreases in older adults compared to younger (Anderer, Semlitsch, & Saletu, 1996; Missonnier et al., 2004; Mueller, Brehmer, von Oertzen, Li, & Lindenberger, 2008), whereas others reported amplitude increases with age (Friedman & Simpson, 1994). The N2 topography was reported to show a more equipotential amplitude distribution over the entire scalp in older adults as compared to the anterior negativity in younger adults (Anderer et al., 1996; Daffner et al., 2010). This reflects an age-associated decrease in anterior negativity.

As regards the P3 topography, consistent age-related changes are generally reported: latency is increased, amplitude is decreased and the voltage topography is more frontally oriented in older adults compared to younger adults (Fabiani, Friedman, & Cheng, 1998; Fabiani & Gratton, 2005; Ford, Roth, Mohs, Hopkins, & Kopell, 1979; Friedman, 2008; Polich, 1996; Rossini, Rossi, Babiloni, & Polich, 2007). The P3 topography is therefore shifted to anterior areas such that the positive peak has its maximum no longer at centro-parietal sites but more on frontal sites.

These age-related topographical changes have been interpreted as reflecting attentional changes in older adults, who would continue to recruit frontal processes (Friedman, 2003).

Effects of cognitive load, practice and performance

The P3 amplitude is sensible to cognitive load, that is, amplitude decreases as a function of increasing task demands (Watter et al., 2001). This was reported within an *N*-back paradigm where cognitive load was gradually increased from 0-back to 2-back. The amplitude decreased in turn gradually as a function of load in younger and older adults (Daffner et al., 2011; Segalowitz et al., 2001). Latency was not modulated by task load.

P3 amplitude was further found to decrease with practice. A hyperfrontality of P3 was observed within the first blocks of an *N*-back task which then decreased within the subsequent task blocks (Segalowitz et al., 2001). The authors interpreted this initial hyperfrontality as an increased attentional allocation before habituation to the task. However the hyperfrontality was observed to be maintained during a longer period in older adults compared to younger adults (Wintink et al., 2001). They associated this maintained hyperfrontality to a reduced prefrontal adaptive functioning. A more frontally oriented P3 topography was also found to be associated with low performance on standardized neuropsychological tests in older adults (Fabiani et al., 1998).

Individual differences were further found when comparing older high versus older low performers. Low performers generally exhibited more P3 hyperfrontality as compared to high performers who in contrast showed a youth-like and less widespread brain topography (Duarte, Ranganath, Trujillo, & Knight, 2006; Lorenzo-López, Amenedo, Pazo-Álvarez, & Cadaveira, 2007). However, the inverse pattern has also been found where older high performers exhibited a more frontally distributed P3 as compared to older low performers and younger adults (Riis et al., 2008). In a more recent study, younger and older high performers were found to increase P3 as well as N2 frontality as a function of increasing task load

(Daffner et al., 2011). Low performers in contrast decreased hyperfrontality with increasing task load. The pattern was therefore different for low and high performers but showed the same trend for both age groups. Despite Daffner and colleagues (2011), no such effects were investigated as regards the N2 component.

In sum, there seem to be in general an age-related P3 hyperfrontality which is mainly present in low performers at least at low cognitive demands.

SUMMARY AND CONCLUSIONS

In the first part of the conceptual background section, we learned from cognitive aging and lifespan research that despite cognitive decline there is a preservation of cognitive functions, and that compensational mechanisms as well as behavioral plasticity are present also in older age. Similar observations were reported concerning age-related structural and metabolic brain changes. They are similarly characterized by decline and preservation when integrating various cerebral regions. The age-related reorganization patterns at the functional brain level are observed to be mainly characterized by a frontally oriented over-activation in older adults as compared to younger adults during nearly all cognitive tasks. With respect to this pattern, several leading attempts to interpret its meaning by linking behavioral and cerebral evidence were discussed. We concluded that a consensus regarding interpretation and the question of whether this pattern is beneficial to older age or not, is still lacking to date. We closed with the statement that age-comparative training studies combining behavior and brain measures may have a large potential to bring light to this issue.

In the second main section, the notions of cognitive and brain plasticity were introduced. Cognitive plasticity, referring to the malleability of performance at the behavioral level, is preserved in older adults, as has been shown in strategy-based as well as process-based training studies. Here we focused on WM training studies since they targeted a key cognitive ability and provided promising results. By reviewing the literature, we discovered that younger and older adults were able to learn and modify their training performance. Nevertheless, cognitive plasticity seemed reduced in older adults as compared to younger adults. As regards transfer effects, the results were ambiguous: Transfer effects to untrained tasks after training were generally reported, but in most studies to a small extent. Furthermore, it was not possible to determine what kind of WM training task would lead to transfer in which tasks. Near and far transfer effects were reported, but no systematic pattern was able to be observed. The only systematic finding was the generally reduced transfer effects in older adults compared with younger adults, even though this did not apply for all studies.

We were able to draw similar conclusions from brain plasticity research, where the training-related cerebral malleability was investigated. The brain remains plastic until old age, even though to a smaller extent in older than in younger adults. However, the general training-induced pattern of change again remained inconclusive, with studies reporting patterns of activation increases, decreases, redistribution and reorganization in both younger and older adults. Further, the brain changes underlying transfer effect have hardly ever been investigated. Research in WM training comparing younger and older adults and including behavioral and brain measures still seems to be very scarce.

We believe that the potential of well-designed age-comparative neurocognitive training studies, like the one presented here, to better understand the mechanisms of aging is very large indeed.

AIMS AND HYPOTHESES

AIMS OF THE THESIS

The main purpose of the present work was to further explore age differences in cognitive and cerebral plastic change. We followed a promising avenue by designing a study in which we implemented ten sessions of a process-based training paradigm applicable among younger as well as older adults, and included behavioral and cerebral training gain and transfer measures before and after training. Based on the present literature we determined and assessed the following three postulates:

a) Cognitive plasticity remains preserved in older adults, even though it is attenuated compared to younger adults (e.g., Kliegl et al., 1989; Schmiedek et al., 2010). However, its extent in younger and older adults during WM training and in the effects in training gain as well as transfer effects, needs further evidence to be fully understood (Martin et al., 2011; Shipstead, Redick, & Engle, 2010). For that purpose we implemented an adapted version of a promising WM training procedure, the *N*-back training (Chicherio, 2006; Jaeggi et al., 2008; Jaeggi, Studer-Luethi et al., 2010; Ludwig et al., 2008). We trained younger and older adults on the verbal *N*-back task and investigated, first, age-differences in and predictors of learning of the working memory task; second, age differences in the impact of WM training on the trained WM task; third, its impact on near transfer tasks, which addressed similar cognitive processes as the trained task (WM); and finally its impact on far transfer tasks, which addressed different cognitive processes to the ones trained.

b) Brain plasticity is preserved in older adults (Dahlin, Bäckman, Neely, & Nyberg, 2009). However, results regarding the pattern of WM training-induced changes in younger and older adults are even less conclusive than the ones from cognitive plasticity research (Brehmer et al., 2011; Buschkuehl et al., 2012; Erickson et al., 2007a, 2007b). Further research is needed to understand these patterns of change. For this purpose we included cerebral measures (EEG) before and after the training intervention for the trained task and for a near transfer task.

c) The interpretation of the age-related cerebral over-activation pattern is still under debate (Cabeza, 2002; S.-C. Li et al., 2006; Reuter-Lorenz & Park, 2010). Therefore, the examination of age-related differences in intra-individual behavioral and cerebral plastic change in a study design that includes longitudinal data should provide a valuable contribution to the debate. To this end we combined the aims a) and b).

DESIGN OF THE STUDY

Here we introduce the study design we chose for the present work. We have detected several shortcomings in previous training study designs, which were addressed by the present design.

Training Paradigm

Process-based training compared to strategy-based training has been identified as having a great potential to lead to transfer effects in untrained tasks (Lustig et al., 2009). We chose a WM training paradigm as it targets a central cognitive process used in all complex tasks, which is affected by age and responsible for large individual differences (de Ribaupierre, Fagot, & Lecerf, 2011; Engle et al., 1999). Further, WM training studies have shown promising effects in younger adults (Jaeggi et al., 2008; Jaeggi, Studer-Luethi et al., 2010) as well as in older adults (Borella et al., 2010; Buschkuehl et al., 2008; Mahncke et al., 2006).

We further considered that an adaptive training paradigm is more effective for fostering transfer effects than training at a general and non-individually adapted level (Buschkuehl, 2007; Kliegel & Bürki, 2012). This approach has been applied in several WM training studies (e.g., Dahlin, Nyberg et al., 2008; Olesen et al., 2004; Schneiders et al., 2011), but in other studies the difficulty level was only individually adapted at the beginning of the training, but no longer during training (100 training sessions: Schmiedek et al., 2010; Shing, Schmiedek, Lövdén, & Lindenberger, 2012).

As a cognitive training intervention we therefore implemented an adaptation of the continuously and individually adaptive *N*-back task training from Jaeggi and colleagues (2008). We used a single visuo-verbal version of the *N*-back task (Chicherio, 2006; Ludwig et al., 2008). To our knowledge, this training has never been carried out with older adults before. The training procedure was adaptive to the participant's performance within a training session, i.e., the accuracy was evaluated at each block, and the subsequent difficulty level was then adapted accordingly.

Sample

To date, there have been few studies that explore cognitive plasticity in an age-comparative way, and even fewer investigations in which younger and older adults are directly compared with regard to cerebral plasticity. However, the combination of a young and an older age group in one training study offers great potential for further insights into the mechanisms of cognitive aging (Lindenberger et al., 2007). For this purpose we included a younger (18-38 years old) and an older (60-84 years old) group.

Investigation of Training and Transfer Effects in Behavior and Brain

In order to investigate behavioral training effects as well as generalized effects, we included a large battery of cognitive tests before and after the training. This cognitive test battery was conducted with all experimental groups in order to identify training and transfer effects. First, we included two conditions of the trained task, a verbal 0-back and a verbal 2-back task in order to assess direct gains from training. Then we included a near transfer task which was structurally similar to the trained one (spatial 0-back and 2-back task). Two near transfer tasks which were not structurally similar but assessed WM processes were also added. The first one was a number updating task which should capture updating processes as well, as in the trained task. Transfer from updating training to an updating task has been shown in a previous study in younger adults (Dahlin, Nyberg et al., 2008). The second near transfer task was a Reading Span task which assessed WM capacity in general. Transfer effects from *N*-back training to Reading Span task have been shown by Schmiedek and colleagues (2010) in younger and older adults. Further, we included several far transfer tasks, first a gf measure, the Raven's Progressive Matrices, to which transfer effects have been reported after *N*-back training, in a single as well as in a dual task training paradigm (Jaeggi et al., 2008; Jaeggi, Studer-Luethi et al., 2010). However, the debate concerning transfer effects to gf measures has recently increased in amplitude (Redick et al., in press). Finally, an inhibition task as well as a processing speed task were included, in order to cover all important cognitive processes which are potentially affected by age. We included a gc measure as a control variable at pre-test only, since we did not expect transfer effects in gc performance.

With respect to the performance at pre-test in these cognitive tasks, we hypothesized that age differences would be observed for all cognitive tasks in favor of the younger adults with the exception of the crystallized measure, where older adults would in turn outperform younger adults (e.g., Baltes & Mayer, 1999).

The combination of several levels of analyses has been proposed as a very promising avenue (Lindenberger et al., 2007). In addition to the behavioral measures, we recorded EEG during the trained task as well as during a near transfer task. The investigation of WM training-induced cerebral change during a transfer task has, to our knowledge, only been reported once in literature so far (Dahlin, Stigsdotter Neely et al., 2008).

As regards the EEG measures, the ERP components of interest in relation to WM performance and updating processes were the N2 and the P3 components (Daffner et al., 2010; Folstein & Van Petten, 2008; Friedman, 2008; Ritter et al., 1979). These two components were investigated before and after the training in the verbal as well as the spatial *N*-back task. As concerns the specific patterns of functional cerebral plasticity associated with practice, we were able to distinguish activation increase and activation decrease within an ERP component. We associated an increased amplitude in the ERP component to an increase pattern and so to a larger recruitment of cortical areas. By analogy, a decreased ERP amplitude was associated to a decrease pattern and to an increased neural efficiency.

In order to distinguish reorganization from redistribution as further proposed by Kelly and colleagues (2006; Kelly & Garavan, 2005), we had to investigate the polarity distribution over the whole scalp, since these patterns can better be examined from voltage maps than from single electrode sites. To this end, another analysis which considers all electrode sites at once was used. We implemented the so-called spatio-temporal microstate segmentation which determines several stable topography maps which occur in a specified time frame of the ERP. Each of this stable topography maps represents a *functional microstate* of information processing in the brain (Lehmann, 1987; Michel, Seeck, & Landis, 1999).

Due to this analysis, we were now able to distinguish between a potential reorganization in information processing steps and a redistribution of activation. Whenever the microstate maps changed from pre-test to post-test we assumed that reorganization took place. We suggested this interpretation according to Michel and colleagues (2009) who qualified a change in voltage map as a change in underlying processes. In turn, when for instance the duration of presence for a map increased or decreased, but no new map appeared or no old one disappeared, we assumed that no new processes were engaged and that a redistribution was observed. This is in line with the distinction that Kelly and Garavan (2005) proposed for reorganization versus redistribution pattern. The former reflects a qualitative change in underlying processes whereas the latter represents a quantitative redistribution of activity by increases and decreases in similar brain regions, but no change in underlying processes. A redistribution pattern should therefore not change the voltage map.

With regard to the ERP components at pre-test in the verbal and spatial tasks we hypothesized that in the low level conditions, the 0-back tasks, older adults would exhibit a smaller P3 amplitude than younger adults. Concerning the scalp topography, a more frontally distributed positivity for the N2 as well as the P3 component would also be found in older adults compared to younger adults (Friedman, 2008; Friedman & Simpson, 1994; Segalowitz et al., 2001). Furthermore, similar effects are expected in the higher level conditions, the 2-back tasks. Age differences would however be more pronounced, that is, older adults would exhibit a more frontally distributed N2 and P3 component compared to younger adults (Daffner et al., 2011; Fjell, Rosquist, & Walhovd, 2009; Polich, 2007). No task specific lateralization was expected (Chen & Mitra, 2009a, 2009b).

Control for Confounding Variables

Control for confounding placebo and retest effects from training has been completely or partially absent in previous training studies. Several studies included a no-contact control group which was tested at pre-test and at post-test without any additional session in-between (Dahlin, Nyberg et al., 2008; S.-C. Li et al., 2008; Schmiedek et al., 2010; Westerberg & Klingberg, 2007). Other studies compared training effects to an active control group which performed either a low load version of the training task (Brehmer et al., 2011) or another task not targeting WM, for example physical training (Buschkuehl et al., 2008), during the same amount of training time as the WM training group. A few studies did not include any control group (Hempel et al., 2004; Olesen et al., 2004). However, to our knowledge, only one study included both, a no-contact control groups as well as an active control group (Mahncke et al., 2006).

In the present study we included two control groups, i.e., a no-contact control group as well as an active control group. The purpose of the former group was to control for test-retest effects which occur when the same task is conducted twice; i.e., to control for practice effects and familiarization with the task and the testing setting. The latter group, the active control group, enabled the control for placebo effects. It is well known that training can enhance several non-specific side-effects, so-called placebo effects, such as changes in achievement motivation or habituation to the experimental setting, which result in decreased anxiety (Goldberg, Keefe, Goldman, Robinson, & Harvey, 2010). The term Hawthorne Effect has been used to describe the fact that participant's behavior is potentially affected by their involvement within a closely monitored study (Green & Bavelier, 2008; Shipstead et al., 2010).

The active control group we included in the present study conducted a control training during the same number of sessions. In order to provide a similar training setting and a similar amount of training time but no constant or increasing demand of attentional resources, we implemented an implicit memory training program.

Generally, implicit memory tasks like the serial reaction time task (SRTT) seem to be relatively unaffected by age compared to explicit memory tasks (Nissen & Bullemer, 1987; Parkin, 1993). In a study comparing younger and older adults in implicit motor serial learning, no age difference was found with regard to the total amount of learning. However, older adults showed a slower learning rate (Daselaar et al., 2003). The authors also observed similar brain network activations in fMRI during the SRTT in both age groups. Hence, age differences were found neither at behavioral level nor at brain level in motor sequence learning.

In a recent literature review, several forms of implicit learning were nevertheless found to be affected by age (Howard & Howard, in press). A first important element in implicit learning was the predictability of the implicit information. A sequence can either be probabilistic or deterministic; whereas the former describes learning about probabilities in series, the latter describes sequences where each trial is fully determined by the preceding trials. No age-differences are reported in the latter, the deterministic SRTT (Gaillard, Destrebecqz, Michiels, & Cleeremans, 2009; Howard & Howard, 1992). The complexity of the sequence structure was another important factor that mediated age effects in implicit learning (Howard et al., 2004; Rieckmann & Bäckman, 2009). Age differences increased when the information to be integrated increased from preceding trials, enabling predictions regarding the current trial.

Overall, these findings indicate that deterministic implicit learning and performance in lower order implicit learning tasks are preserved with age and engage the same brain regions in both younger and older adults (Daselaar et al., 2003). For our study, we therefore used a deterministic and lower order SRTT in a training procedure for the active control group.

In summary, with the present study, we aimed to implement an improved training design compared to previous training studies. We introduced and combined the following essential methodological features in a single study: a) the use of an adaptive process-based WM training procedure, b) the inclusion of a younger and an older sample in one study, c) the combination of behavioral and cerebral measures in one study, with the inclusion of cerebral measures in a near transfer task in addition to the ones in the trained task, d) the inclusion of two control groups: no-contact group and an active control group.

General Hypotheses

Hypotheses at the Behavioral Level

According to the present literature we formulated the following hypotheses for the behavioral results. Hypotheses 2 – 4 are formulated with regard to the WM training groups and have to be understood as effects hypothesized for the WM training groups as compared to both control groups for which we did not expect a significant change from pre-test to post-test.

1. With regard to the WM training performance we hypothesized that:
    a) there would be an increase in performance level over the ten training sessions in both younger and older adults. Age differences would persist throughout the training, but become more pronounced by the end of training. This would lead to magnified age differences with training (Baltes & Kliegl, 1992; Brehmer, Westerberg, & Bäckman, 2012);
    b) both age groups would reach asymptote performance during ten sessions of WM training (Jaeggi, Studer-Luethi et al., 2010; S.-C. Li et al., 2008; Schneiders et al., 2011);
    c) age would become a unique predictor of training performance by the end of training beyond the assessed cognitive variables (Kliegl et al., 1990).
2. With respect to WM training gain we hypothesized that:
    a) there would be a training gain in the *verbal N-back task* performance in both age groups, but more pronounced in younger adults.
3. Concerning the near transfer effects we hypothesized that:
    a) there would be a near transfer effect to the performance in the structurally similar *spatial N-back task* for both age groups, but more pronounced in younger adults (Brehmer et al., 2012; Dahlin, Nyberg et al., 2008; S.-C. Li et al., 2008);
    b) there would be a near transfer effect to the performance in the structurally different *Updating task* for both age groups, but more pronounced or even only present in younger adults (Brehmer et al., 2012; Dahlin, Nyberg et al., 2008);
    c) there would be a near transfer effect to the performance in the *Reading Span task* for both age groups, but more pronounced or even exclusively present in younger adults (Brehmer et al., 2012; Schmiedek et al., 2010).

4. With regard to far transfer effects we hypothesized that:

a) there would be a far transfer effect to fluid intelligence performance in the *Raven's Progressive Matrices* in younger adults and probably also in older adults. However, this effect has never been investigated in older adults (Jaeggi et al., 2008; Jaeggi, Studer-Luethi et al., 2010; Schmiedek et al., 2010);

b) no specific transfer effects are expected for the performance in the inhibition and processing speed tasks; they were included for exploratory reasons and to examine the comparability of the different experimental groups.

Hypotheses at the Cerebral Level

For the EEG we formulated the following hypotheses according to the current literature. However, given the scarce and inconclusive body of evidence, precise predictions with respect to the specific pattern of the plastic cerebral change and its age-related differences were difficult to make. Therefore the hypotheses about training-induced changes were of exploratory nature.

Here again, all hypotheses are formulated for the WM training groups with regard to changes due to training and have to be understood as compared to both control groups for which we hypothesized that no significant changes from pre-test to post-test would be observed.

1. With respect to changes in ERP components in the WM training gain, the verbal *N*-back task, we hypothesized that:

a) for younger adults, there would be a pattern of decrease in the N2 as well as in the P3 component that would be reflected by less pronounced amplitudes. This would be in line with a high performers' profile but at a decreased cognitive load (Daffner et al., 2011). As regards microstates, we expected that a reorganization would take place towards a less frontal recruitment, that is, a change towards a more efficient and posteriorly distributed map, as generally training-induced frontal changes were observed for younger adults (Dahlin, Stigsdotter Neely et al., 2008; Duarte et al., 2006; Erickson et al., 2007b; Lorenzo-López et al., 2007; Nyberg et al., 2003);

b) for older adults, there would similarly be a decrease in N2 and P3 amplitude. But as regards microstates, a pattern of redistribution which would be reflected by an increased presence of a less frontally and less widely distributed voltage map would be observed (Daffner et al., 2011; Daffner et al., 2010; Duarte et al., 2006;

Erickson et al., 2007b; Lorenzo-López et al., 2007; Missonnier et al., 2004; Nagel et al., 2009; Riis et al., 2008).

c) there would consequently be a differential change between younger and older adults: A reorganization in older adults and a redistribution in younger adults, as younger adults were generally found to exhibit more plastic change. This does however not imply that age differences would be magnified (Erickson et al., 2007b). Since we expected a decrease in frontal activation in both age groups, we would show that frontal overrecruitment is not beneficial in older age which would be in line with the non-selective recruitment approach (S.-C. Li et al., 2006).

2. As concerns ERP components in the near transfer task, the spatial *N*-back task, we hypothesized that:

a) an equal change pattern like for the verbal 2-back task in both age groups would be found. But we expected it to be slightly less pronounced than for the trained verbal task (Schneiders et al., 2011).

b) as regards the microstate maps, predictions were even more difficult to make as compared to the trained task. From the scarce body of literature we would also expect similar changes as in the trained task, since the processes overlap between the trained and the structurally similar near transfer task appears large (Dahlin, Stigsdotter Neely et al., 2008).

METHOD

The method section compasses a detailed description of the used methodology in the present study. In the first part, we delineate the behavioral method, describing in turn the participants' characteristics, the general procedure, all the administered tasks, the specific hypothesis regarding the operationalization, and finally the implemented analyses. In the second part, we describe the EEG method. EEG was conducted with a subsample of participants. Here also, the participants' characteristics, the adapted general procedure and the material, the specific hypothesis, and the respective ERP analyses are described. In the final part, we provide the hypothesis and the analyses considering the link between behavioral and EEG results.

BEHAVIORAL METHOD

Participants

A total of 278 individuals were contacted for the study, whereof 130 individuals, 64 younger and 66 older adults, met the inclusion criteria, were available and agreed for participation in the study. They were pseudo-randomly assigned to one of the three experimental groups with the purpose of balancing gender, age and education across the groups. Before testing, a telephone interview was conducted in order to check the inclusion criteria concerning age and French language skills. Younger adults between 18 and 38 years old were tested and older adults had to be 60 years or older. Only native French speakers or participants being fluent in French for at least five years were included. All participants had normal or corrected-to-normal vision. Younger adults were nearly all undergraduate students of the University of Geneva recruited by advertisements posted in the buildings of the University of Geneva. Older participants were recruited from the University of the Third Age, from senior citizen centers, and from other retirement clubs in the canton of Geneva.

The present study was approved by the ethics committee of the Faculty of Psychology and Educational Sciences of the University of Geneva. All participants signed a written informed consent for participation. Several control tests were performed such as the Freiburg Visual Acuity & Contrast Test (FrACT; Bach, 1996) for visual acuity control, where higher scores reflect better vision. For further control tests, the participants completed a self-rating health questionnaire and a crystallized intelligence (gc) test, the French adaptation of the Mill Hill Vocabulary Scale (Deltour, 1993; J. Raven, Raven, & Court, 1998).

Table 1. Demographic sample characteristics.

	Younger adults			Older adults		
Variable	WM (n = 22)	Implicit (n = 20)	No-contact (n = 21)	WM (n = 22)	Implicit (n = 20)	No-contact (n = 23)
Gender	73	70	71	82	70	70
Age ***						
M	24.68	24.35	25.52	67.64	67.70	68.61
SD	5.26	5.18	4.54	4.73	4.96	6.54
Range	18- 38	19-38	18-35	61-81	61-78	60-84
Health *						
M	2.04	2.04	2.00	2.25	2.07	2.24
SD	0.45	0.43	0.26	0.31	0.50	0.38
Range	1.29-2.86	1.14-2.86	1.57-2.57	1.86-2.86	1.29-3.00	1.57-2.86
FrACT *						
M	0.75	0.74	0.80	0.63	0.68	0.71
SD	0.21	0.22	0.28	0.20	0.21	0.22
Range	0.54-1.18	0.49-1.18	0.05-1.18	0.29-1.18	0.41-1.13	0.42-1.18
Education						
M	14.86	14.65	15.57	15.00	15.20	13.52
SD	2.03	1.81	2.38	2.51	2.76	1.71
Range	12-19	10-19	12-21	12-22	11-22	11-18
Diploma ***	91	85	90	59	55	30
Mill Hill ***						
M	34.59	35.15	33.76	39.32	38.00	38.30
SD	5.26	5.10	4.37	2.63	4.40	3.30
Range	21-43	17-42	18-40	34-43	28-44	32-43
Raven ***						
M	36.23	34.85	35.57	27.05	30.30	25.17
SD	5.82	6.12	5.48	7.55	6.87	6.10
Range	24-45	24-46	22-43	9-36	14-42	14-36

Note. Significant differences between younger and older adults: *** $p < .001$. * $p < .05$. Gender values reflect percentage of women in each group. Age and education values reflect years. Health scores reflect average response to seven questions on a scale from 1 = excellent health to 5 = very bad health. FrACT = Freiburg Visual Acuity & Contrast Test. Diploma values reflect percentage of participants who hold a university-entrance diploma. Mill Hill scores represent total correct responses out of 44 words. Raven score represents total correct responses out of 48 at pre-test (for test description see below). Range = minimum – maximum value.

One older participant dropped out after the pre-test session and one younger was excluded due to insufficient language skills; hence, the final sample consisted of a total of 128 participants with 63 young and 65 older adults. However, several subjects had to be excluded for the Reading Span task and the Stroop task: For the reading span task one younger and three older participants were excluded due to a bad semantic judgment at pre-test and/or post-test. In the Stroop task, one older person was color-blind and not able to distinguish reliably between red and green. This participant was excluded from the Stroop task analysis.

The experimental groups were compared separately for each age group regarding the characteristics listed in Table 1Table 1. With one exception, there were no significant differences between groups. For the older group the total number of years of education differed significantly between experimental groups ($F(2, 62) = 3.38$, $p = .04$). The older no-contact control group had significantly less years of education compared to its trained peers.

As regards the comparison of the whole group of younger adults to the whole group of older adults, analyses revealed several expected differences: Despite mean age, there were age-related differences in the health questionnaire score, the visual acuity score (FrACT), the diploma score, the Mill Hill score as well as the Raven score. For the health score, younger adults ($M = 2.02$, $SD = 0.39$) exhibited a lower score which indicates that younger adults evaluated their subjective health status better than older adults did ($M = 2.18$, $SD = 0.41$; $t(126) = -2.14$, $p = .034$). Also for the visual acuity score, younger adults exhibited a better score ($M = 0.76$, $SD = 0.23$) than older adults ($M = 0.67$, $SD = 0.21$; $t(126) = 2.31$, $p = .023$). As regards the percentage of university-entrance diploma holders, there were more holders in younger adults (89%) than in older adults (48%; $\chi^2 = 24.94$, $p < .001$). Results concerning age differences in the Mill Hill and the Raven scores are provided in the results section in the section *Age Differences in the Cognitive Tasks*.

General Procedure

All participants were tested individually in a quiet laboratory in the building of Uni Mail of the University of Geneva. The general procedure is illustrated in Figure 1; the EEG recording part will be described in the *Electrophysiological Method* section. During the pre-test session, the participants underwent a battery of cognitive tasks assessing WM, fluid intelligence (gf), inhibition as well as perceptual-motor and processing speed. All tasks and all blocks within one task were administered in the same fixed order across participants (see Appendix A, Table A1). At post-test, after the training period, a similar session was administered. One session lasted about 120 to 150 minutes. Between pre-test and post-test, participants in the training groups performed 10 training sessions distributed over two to four weeks. The participants were trained with either a WM task or an implicit sequence learning task. The training sessions lasted about 20 to 30 minutes per day and were administered individually in the same laboratory at Uni Mail. The control groups were not contacted during two to four weeks between pre-test and post-test.

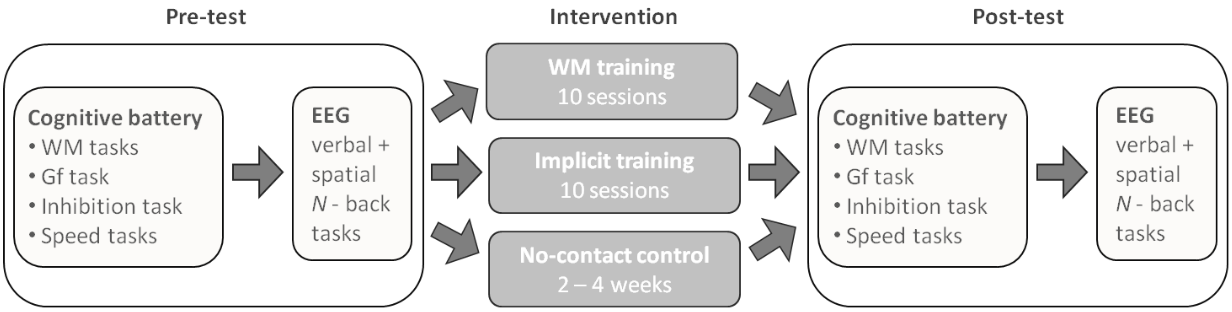

Figure 1. General procedure of the training study. For detailed administration order of the tasks at pre-test and post-test, see Appendix A.

All computerized tasks were programmed with E-prime 1.1 software (Psychology Software Tools, Inc., Pittsburgh, PA, USA). The scripts were run on a Fujitsu Siemens personal computer and displayed on a 15 inch video graphics array color computer monitor. According to the task demands, participants responded using the keyboard, the E-Prime response box or the E-Prime microphone. The distance between the participant and the screen was kept at approximately 60 cm.

A pilot study has been conducted for the present study with 20 younger and 20 older participants divided into a WM training and a no-contact control group each. We conducted the pilot study to assess the efficacy of the WM training and the transfer effects. These data were kept in the final sample. Some modifications of the protocol were made subsequent to the pilot study: On the one hand, two tests were excluded, one test out of two tests assessing inhibition and one out of two tests measuring processing speed. They were excluded because no transfer effect was found in the pilot study neither for inhibition tasks nor for processing speed tasks. On the other hand, a new updating task with the purpose of replicating the findings from Dahlin, Nyberg and colleagues (2008) was included. Therefore the *N* for the Updating task is smaller than for the other tasks and corresponds to the one of the EEG subsample.

Material

Training Tasks

Working memory training

The WM training group performed an adaptive verbal *N*-back task training. The dual *N*-back task training used by the research group of the University of Berne was adapted for this research (Jaeggi et al., 2008). A single task condition with verbal stimuli based on Chicherio (2006) and Ludwig and colleagues (2008) was used.

Description. The task consisted in judging whether the current letter matches the letter *N* positions back in a sequence of letters presented one by one (see Figure 2, left panel). A response was required for each stimulus. The stimuli (twelve consonants) and their presentation time were the same as in the verbal 2-back task examined at pre-test and post-test described below. The letter sequences were presented in blocks of 30 stimuli. Each sequence was randomly created and enclosed ten targets (matched items) and 20 non targets (non match items). The inter-block delay was determined by the participant and each training session contained 15 blocks.

Reaction times and accuracy for targets and non targets were registered. The level of difficulty was varied by adapting the load in each block according to the participant's performance reached in the preceding block. Response accuracy in all items was used as criterion. When the participant's accuracy was below 70 percent in one block, the load in the next block was decreased by one *N*-level. If accuracy was 90 percent or more, the *N* increased by one level. The load level remained unchanged whenever the percent of correct responses was between 70 and 90 percent. At the end of each block, a feedback on accuracy was provided to the participant, along with the level of the block to be presented subsequently. The training started with a block of 1-back level and from the fourth session on with a block of 2-back level. The training was conducted over ten sessions which were distributed over two to four weeks.

At the end of the ten training sessions, a questionnaire about strategy use was provided. We asked the participant whether they implemented one or several strategies during training. The participants had to describe them and judge their usefulness during training (see Appendix A).

Score. The mean *N*-level over the 15 blocks calculated for each session separately was used as dependent measure.

Implicit sequence learning training

The control training group was trained with an implicit sequence learning task. This training served as a placebo training in order to control for the confounding variables resulting from a training setting. The goal of this training was that participants recruit as little attentional resources as possible during the training. Therefore, a simple implicit sequence learning task was chosen. The objective was that the sequence should be implicitly learned and the participants should respond by using less attentional control in the course of the training.

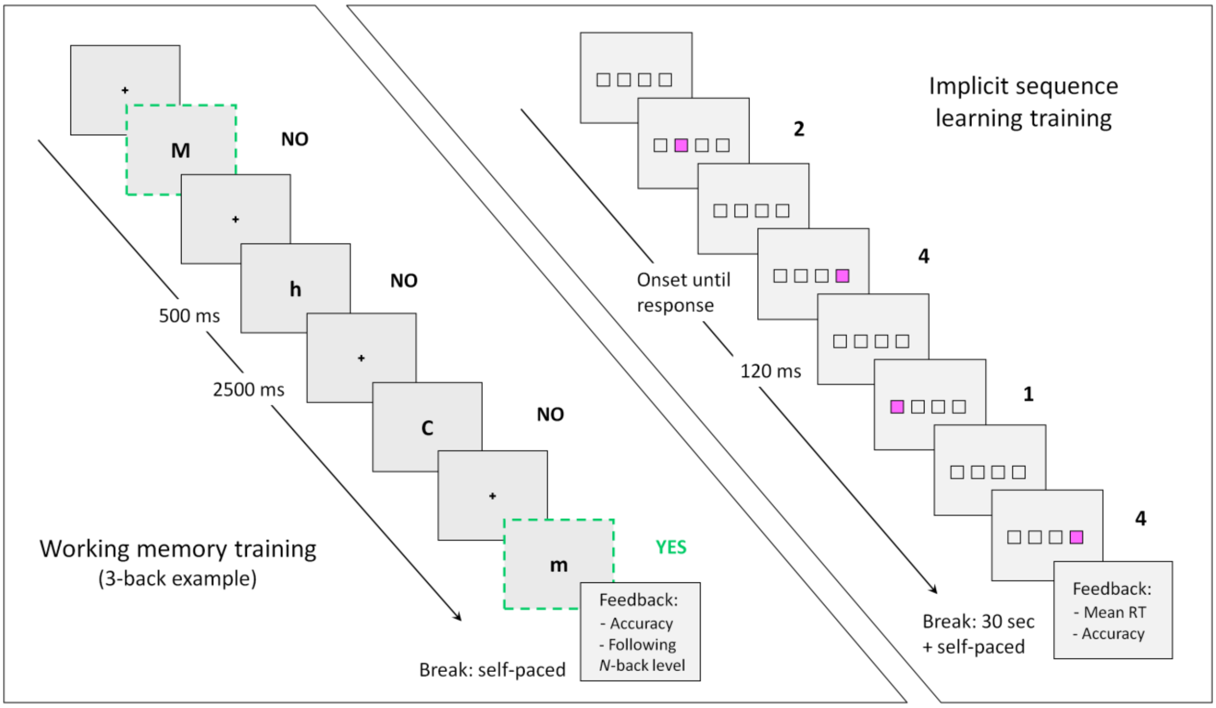

Figure 2. Illustration of the training tasks. Working memory training (left panel); illustration of a 3-back example in the adaptive *N*-back task training. Implicit sequence learning training (right panel); illustration of a part of the trial sequence.

Description. The task consisted of four light grey squares with a black border presented horizontally aligned in the center of the screen on a light grey background (see Figure 2, on the right). One stimulus consisted of one pink and three grey squares, the four possible stimuli were used. They were presented in such a way that one square after the other changed the color to pink. The participants had to respond as fast and as accurate as possible by pressing the key which matched the square colored in pink. Four fingers of the right hand had to be placed on the numeric keypad such that the index finger is positioned on the key 1 and the little finger on the key *Enter* with the middle and ring finger on key 2 and 3, respectively. Left-handed participants used the four fingers of the left hand in the inversed order. The key 1 was associated to the first square on the left, key 2 to the second square from the left, key 3 to the third and the *Enter* key to the square on the right border. The stimulus was presented until the participant's response which was followed by an inter-stimulus interval (ISI) of 120 ms and the following stimulus. The ISI facilitated the distinction of two stimuli which were presented consecutively at the same position.

The stimuli were presented in a determined second order sequence where one trial is determined by the two preceding trials (Reed & Johnson, 1994; Ungerleider et al., 2002). The four stimuli are presented four times which results in a sequence of 16 trials. The 16-trials sequence was repeated six times without interruption in one block resulting in 96 trials per block. The starting point in the sequence in each block was defined randomly in order to avoid explicit learning of the sequence. Consequently, the first two trials in each block could not be learned but determined the following trials. Between each block a rest of a minimum of 30 seconds was demanded in order to relax the hand. At the same time, the participant received a feedback containing the mean reaction time and the percentage of the correct responses from the preceding block.

Similar to the WM training, 15 blocks per training session were provided. The participant trained during ten sessions distributed over two to four weeks.

Score. The dependent variable used from this training was the mean correct reaction time for each session. Reaction times were recorded beginning with the stimulus onset. With advanced training time, however, the participants learned the series and were no longer waiting for the next stimulus to appear on the screen. Therefore, with advancing training time the 120 ms of the ISI became part of the reaction time, however, it was not taken into account for reaction time recording. This is the reason why especially the young adults show very fast reaction times by the end of the training.

Pre-test and Post-test Tasks

Working memory

WM performance was evaluated by three paradigms: the N-back task in a visuo-verbal and a visuo-spatial version (Braver et al., 1997; Owen et al., 1999), the number updating task (Carretti, Cornoldi, & Pelegrina, 2007) which was introduced after the pilot study and the Reading Span task (Daneman & Carpenter, 1980; de Ribaupierre & Bailleux, 1995; de Ribaupierre, Lecerf, Leutwyler, & Poget, 1997; Delaloye, Ludwig, Borella, Chicherio, & de Ribaupierre, 2008). The first two paradigms require maintenance and updating of verbal, spatial or numerical information, whereas the latter requires to simultaneously process the semantic content of sentences while storing a series of words. In this research, the verbal *N*-back task was used to investigate gain from WM training. The spatial *N*-back task was included to explore a near transfer effect from WM training to a structurally similar task. Further, another updating task was included in order to investigate near transfer effect to a different task structure. Finally, the Reading Span task was used to quantify a far transfer effect from WM training to a structurally different task measuring WM capacity.

***N*-back Task.** The *N*-back paradigm was first introduced in the late fifties to study the interaction of WM load and age in the WM performance of younger and older adults (Kirchner, 1958). In 1966, Sternberg introduced the *N*-back paradigm to investigate response latency differences generated by variation in WM load. Subsequently, the *N*-back paradigm has been widely used and by the introduction of neuroimaging techniques it became a popular experimental paradigm in functional neuroimaging studies (e.g., Cohen et al., 1997; Gevins et al., 1996; Jonides et al., 1997; for a review see Owen, McMillan, Laird, & Bullmore, 2005). The task has been established in the verbal as well as in the spatial domain (e.g., McEvoy, Smith, & Gevins, 1998). The *N*-back task can be considered as a measure of WM performance in accordance with different theories and concepts of WM (Baddeley & Hitch, 1974; Engle et al., 1999; J. Pascual-Leone & Baillargeon, 1994). The *N*-back task requires both, the control of encoding and retrieval of information and the active maintenance, that is keeping information available online (e.g., Cohen et al., 1997). In the model of executive functions proposed by Miyake, Friedman, Emerson, Witzki, and Howerte (2000), the *N*-back task has been suggested to measure updating and monitoring processes of WM representations. That is, updating requires monitoring incoming information for relevance to the task and revising the items held in WM by replacing old, no longer relevant information by new and relevant information (Morris & Jones, 1990). Further, the updating function measured by the *N*-back task has been associated to activation in the dorsolateral prefrontal cortex (Jonides & Smith, 1997).

Description. In the present study an N-back task in the verbal as well as in the spatial domain was used, presented visually in a computerized version. Two load levels, a 0-back and a 2-back condition, were used in each domain. In the 0-back condition, participants were required to detect an initially defined stimulus in a sequence of stimuli. The 0-back condition served as a control condition as it doesn't require updating processes, but monitoring of the sequence of stimuli and maintenance of the target item. The 2-back conditions consisted in deciding whether each stimulus matched the one presented two stimuli back in the sequence (see Figure 3). This condition requires updating and monitoring processes.

Description verbal task. For the verbal task, sequences of upper and lower case letters presented one by one were used (Chicherio, 2006; Ludwig, 2005; Ludwig et al., 2008). As stimuli, thirteen different characters in the 0-back condition and twelve in the 2-back condition were used, selected according to the following criteria: They had to be monosyllabic consonant in the French pronunciation (no vowels; no *W* or *Y*), not sharing identical patterns with any other consonant in lower or upper case (no L or I) and be frequent as initial letter in the French language (above a frequency of 1000, according to the Brulex reference database, Content, Mousty, & Radeau, 1990). The selected stimuli were *B, C, D, F, H, M, N, P, R, S, T,* and *V*. In order to prevent a direct visual match, their corresponding lowercase letters were used for half of the stimuli. In addition, the consonant *X* was inserted as the target stimulus for the 0-back condition. In each condition 6 blocks of 36 trials each were presented resulting in a total of 216 trials per condition. According to the procedure proposed by Braver and colleagues (1997), one-third of the trials were targets (match trials) and two-third were non targets (non match trials). The proportion was maintained within each block (12 targets and 24 non targets). In the 2-back condition, each consonant was equally presented as a non target, as a non target prime and as a target. In both conditions, the presentation of the stimuli was fixed in a pseudo-random sequence following several criteria: More than two consecutive stimuli of the same consonant and an alphabetical order of the stimuli were avoided. No target was presented in the first trial in the 0-back or in the third trial in the 2-back condition. All consonants were displayed in upper and lower case print in order to avoid visual strategies and force the participant to process the stimuli at the phonological level. The stimuli were typed in 125 pts in black bold Arial font and presented centered on a light grey background. The six blocks within each condition were given in a constant order with breaks between each block.

Description spatial task. The material for the spatial condition was adapted from Owen and colleagues (1999). Configurations of six black circles on a light grey background were used (Chicherio, 2006). The circles were arranged randomly in 6 locations out of 49 (7 x 7) possible positions. Twelve configurations of six circles were selected according to the following criteria: no more than three circles in a straight line and no configuration resembling a drawing easy to process verbally. Each configuration was presented in one block, six configurations in the 0-back condition and six in the 2-back condition. For each configuration six stimuli were created such that each circle was colored once in pink whereas the other circles stayed colored in black. In the 0-back condition one stimulus per block was defined as the target stimulus. In the 2-back condition the participants had to decide whether each configuration matches the one presented two configurations back in the sequence. Like in the verbal task, in each block one third of the stimuli were targets and two third were non targets. The six blocks within each condition were given in a constant order with breaks between the blocks.

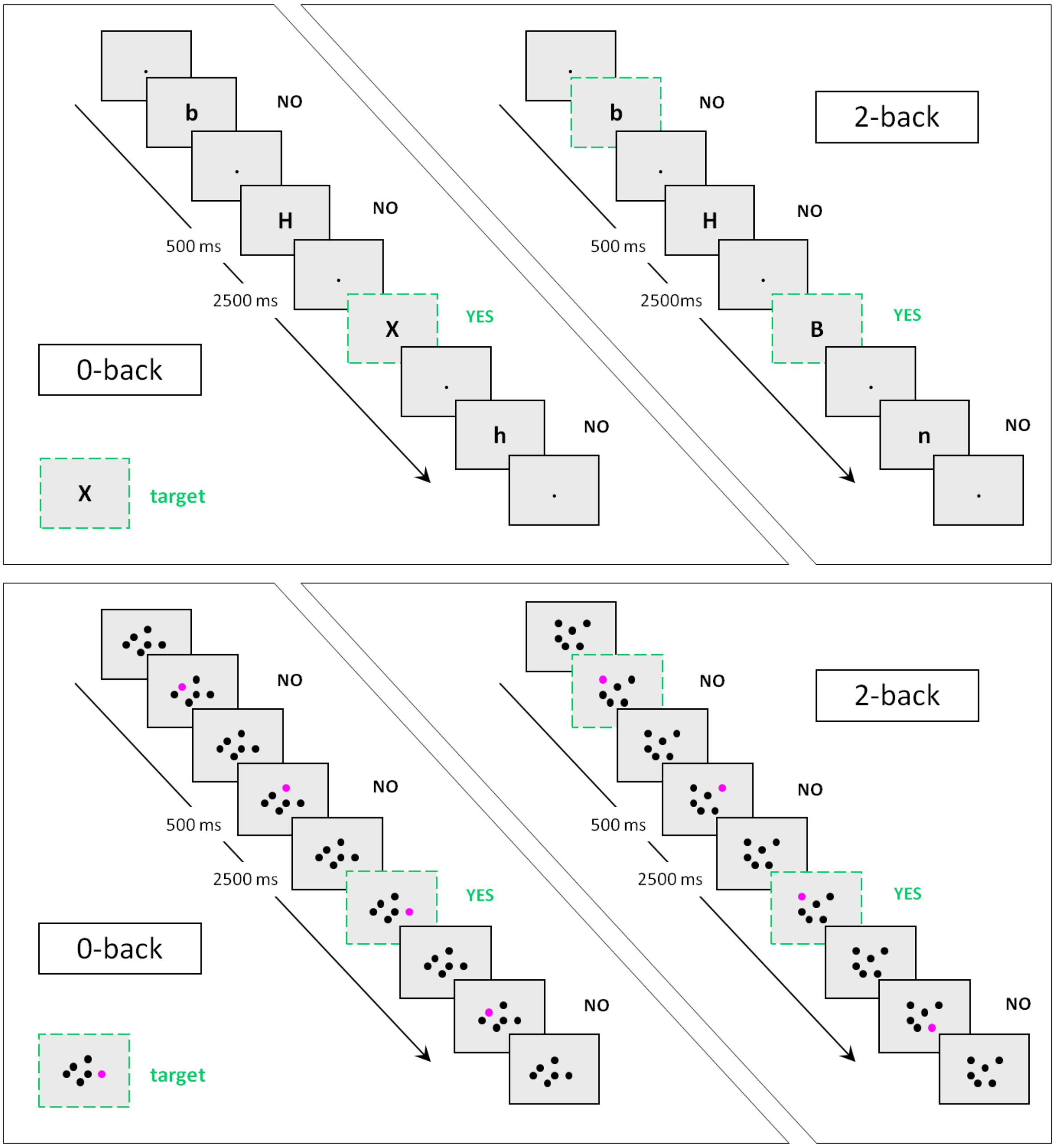

Figure 3. *N*-back task. Illustration of the trial sequences in the 0-back conditions (left panels) and in the 2-back conditions (right panels) of the verbal and spatial tasks (upper and lower panel, respectively).

In the verbal as well as in the spatial *N*-back task, the stimuli were displayed for 500 milliseconds (ms) followed by an ISI of 2500 ms, resulting in a stimulus onset asynchrony (SOA) fixed to 3000 ms (Braver et al., 1997; Smith & Jonides, 1999). The timing of the trial sequences is illustrated in Figure 3. During the ISI a fixation point was presented in the verbal task, and in the spatial task the same configuration as the stimulus but with black circles only was displayed.

A response was required for each trial by a button press with the index finger for targets (yes) or the middle finger for non targets (no) of the dominant hand. The response window was identical to the SOA. Participants were told to respond as fast and as accurately as possible. Response time and accuracy were recorded while very fast responses below 200 ms were discarded as anticipations and excluded from the analysis.

Scores. The scores used for the verbal and the spatial 2-back tasks were the proportion of correct responses. The mean correct response time was also calculated; however, it was of limited interest as the WM training was aimed at primarily improving accuracy and not response time. Further, a *d* prime (*d'*) score was calculated in order to quantify the quality of the discrimination between targets and non targets. *d'* is calculated from hit rate and false alarm rate as follows: *d'* = *Z*(hit rate) - *Z*(false alarm rate) (Macmillan & Creelman, 1990). The Z transformation was the calculation of the inverse of the standard normal cumulative distribution function for the rates (probabilities). Whenever the false alarm rate equaled zero, a correction was applied using the resulting value of the calculation 1/(2 x maximum value non-targets); (for more details see Macmillan & Creelman, 1990). In order to investigate the WM load costs, differences between 2-back scores and the low load 0-back scores (no updating control condition) were calculated. That is, 0-back minus 2-back for accuracy and *d'* scores and 2-back minus 0-back for response times was calculated.

Updating Task. The Updating task was introduced after the pilot study, the *N* for the Updating task is therefore smaller than for the other tasks and corresponds to the one of the EEG subsample. The Updating task was translated and adapted from Carretti and colleagues (2007).

Description. The task required maintaining and updating numerical information. The computerized task consisted of 13 series of 10 two-digit numbers. The presented numbers ranged from 24 to 95 and were presented one after the other in the centre of the screen. At the end of each series, the participants had to recall the three smallest numbers in the correct order of presentation. Thirteen series were presented, six of the series were composed of numbers close in size (e.g., 34, 46), seven of numbers distant in size (e.g., 28, 73). Further, two levels of difficulty in terms of numbers of updating processes were used: for each distance size group, six of the series included five numbers that initially had to be considered as relevant, but later had to be updated (numbers close in size; five updating processes: e.g., 50 39 53 49 42 47 56 45 44 55). The other seven series contained only three numbers that had to be updated later (numbers distant in size; three updating processes: e.g., 58 34 52 64 28 46 79 30 72 61). Each number was presented for two seconds, with an ISI of one second. Within the

ISI after 500 ms a mask (i.e., ##) appeared on the screen for 500 ms. After the presentation of each series, the participants had to use the keyboard to type the three smallest numbers respecting the correct order of presentation. The task started with two series for practice in which participants received feedback about accuracy. Then the participants responded to the 13 test series.

Scores. The scores considered for this task were first the total number of correctly recalled numbers, independent of the correct serial order, divided by the total number of series (13). Second, the proportion of entire correctly recalled series was evaluated, again without considering the order of presentation. The first score varies between 0 and 3 whereas the latter score varies between 0 and 1.

Reading Span Task. The Reading Span task was initially introduced by Daneman and Carpenter (1980) in a study of individual differences in WM capacity and reading comprehension. In the present study, a French adaptation of the task was used (de Ribaupierre & Bailleux, 1995; de Ribaupierre & Ludwig, 2003; Delaloye et al., 2008).

Description. The Reading Span task required simultaneous processing of the semantic content of sentences while maintaining a series of words. The task was preceded by a semantic judgment task in order to assess the semantic processing time without a simultaneous maintenance of verbal material. Sentences were presented sequentially on a computer screen in white characters on a black background.

In the semantic judgment task, the participants were instructed to decide whether the presented sentence is semantically plausible or not. Responses were given by pressing either the *yes* or the *no* key tagged with stickers on the keyboard. The index fingers of both hands were used for key press and the *yes* button was on the side of the dominant hand. Twenty sentences were given in total, whereof four were given as study trials. The Reading Span task followed immediately and consisted of series of two to five individually presented sentences. In addition to the semantic judgment by key press, participants had to actively maintain the last word of each sentence. At the end of each series of sentences, a white triangle was displayed and the participants were asked to recall the words orally in the correct order of presentation. The Reading Span task consisted of two two-sentence series for study phase and of four series of each item length (2 - 5) resulting in a total item number of 16.

In both tasks all sentences remained on the screen until the participant responded by key press. The presentation of the sentences and the series of sentences were administered in identical order for all the participants. A parallel version was constructed according to the same criteria in order to avoid familiarization with sentences and words. The order of the two versions was counterbalanced among participants. The reliability of version 1 and version 2 for the mean number of words correctly recalled was good with a Pearson's correlation coefficient of $r = .74$ ($p < .001$). Further, the score did not differ between versions neither at pre-test nor at post-test.

Scores. Accuracy and reaction time for the semantic judgment were recorded by computer and the recalled words were recorded by the experimenter. The accuracy proportion for the semantic judgment in the Reading Span task was calculated in order to control the compliance with task instructions. That is, to check whether the participants were processing the sentences semantically and were not pressing the response keys by chance. Only subjects obtaining an accuracy proportion of 85 percent or above were included (Delaloye et al., 2008).

One dependent variable was the mean number of words correctly recalled per series. It was calculated by dividing the total number of correctly recalled words by the total number of series (16) regardless of whether the entire series was completely recalled. This score varies between 0 and 3.5.

The second dependent variable was the proportion of entirely correct series for which all words of one series are required, independent of the correct serial order, and no supplementary word was allowed. The latter score varies between 0 and 1. Previous analyses showed that there was no significant difference between a score which took the serial order of the recalled words into account and a score which did not (de Ribaupierre et al., 1997).

In addition, the total number of intrusions committed in the Reading Span task as well as the difference in response times between single and dual task conditions was calculated. Intrusions were incorrectly recalled words, which appeared in previous sentences (Robert, Borella, Fagot, Lecerf, & de Ribaupierre, 2009).

Fluid intelligence

For the evaluation and replication of the results of Jaeggi and colleagues (Jaeggi et al., 2008; Jaeggi, Studer-Luethi et al., 2010), a far transfer effect on a fluid intelligence measure, Raven's Progressive Matrices were used (J. Raven, Raven, & Court, 1998, updated 2003).

Raven's Progressive Matrices.

Description. The task consisted of matrices of figures and patterns composed of abstract shapes and lines in which one element is missing. The participants had to complete the matrices by visual analogy. The correct pattern out of six or eight response alternatives had to be selected for each problem.

In order to avoid a ceiling effect in younger and a bottom effect in older adults, both versions, i.e., the Raven's Standard Progressive Matrices (SPM) and the Raven's Advanced Progressive Matrices (APM) were used (J. C. Raven, 1958, 1962). Odd and even trials of each version were separated and administered as different forms at pre and post-test as described in (Jaeggi, Studer-Luethi et al., 2010). Participants answered first to 30 standard trials and then to 18 advanced trials. The trials were presented in increasing difficulty within each version of the test.

The trials were presented on paper bound in a folder, and participants had to note the response on a separate sheet. For the total of 48 trials the time was restricted to 30 minutes. The time restriction was determined according to the recommendation in the test manuals that is, 20 minutes for the SPM and 40 minutes for the APM, summed-up and divided by two as the tests were split into two forms.

Reliability for the two versions was good; Pearson's correlation coefficient for the total number of correct responses of the odd items version and the even items version was at $r = .87$ ($p < .001$). However, the two versions yielded significantly different scores at pre-test ($M_{odd} = 33.2$ (7.5), $M_{even} = 29.57$(7.36); $t(126)= 2.67$, $p <.01$) and at post-test ($M_{odd} = 34.57$ (6.58), $M_{even} = 31.26$ (8.54); $t(126)= 2.67$, $p <.05$), with higher scores for the odd version compared to the even one. As the two versions were fully counterbalanced between pre-test and post-test and within each experimental group, a bias of the level difference between the two versions can be excluded.

Scores. The scores used in this task were the total number of correct responses and the proportion of correct trials over the done trials as in (Ackerman, Kanfer, & Calderwood, 2010).

Inhibition

To measure inhibitory control capacity, the Color Stroop task (Stroop, 1935) was used. It was used to quantify the capacity to inhibit or to resist a dominant response .

Color Stroop Task. The task was originally described by Stroop and was later adapted from Spieler, Balota, and Faust (1996) for trial by trial computerized assessment (Ludwig, 2005; Ludwig, Borella, Tettamanti, & de Ribaupierre, 2010).

Description. The Color Stroop task assessed the restraint function of inhibition (Hasher et al., 1999) by requiring participants to name the color in which words are printed while ignoring the word's meaning. The trials consist of color words *(vert, jaune, rouge, and bleu*), neutral words (*grave, fort, plein, and neuf*) and colored rectangles printed in different colors (green, yellow, red, and blue). There were four conditions: congruent (e.g., *bleu* printed in blue, correct answer is blue), incongruent (e.g., *bleu* printed in green, correct answer is green), neutral (e.g., *fort* printed in red, correct answer is red) and a control condition (e.g., yellow rectangle printed in yellow, correct answer is yellow).

The task started with a practice block of eight trials followed by a testing phase with 5 blocks of 28 trials each. The number of congruent, incongruent, neutral and control trials was counterbalanced within each block. Each item was individually displayed on a black screen and remained until the participants gave their response. A fixation point was presented during 1000 ms prior to the item and an empty screen during 800 ms followed the item. In between blocks, participants could decide to take breaks.

The participants were required to name the color of the ink of the presented word or rectangle. The reaction times were recorded using the E-Prime microphone while the experimenter noted the accuracy of the answers in a protocol. Fast and therefore implausible responses below 200 ms were excluded.

***Scores*.** The dependent variables were the mean reaction times for correct trials calculated separately for each condition. An interference ratio on reaction times using the following equation was calculated: (incongruent trials - control trials)/control trials. For the facilitation ratio the same equation was used, but replacing the incongruent trial time by the congruent one. As a further score the proportion of intrusions was used. That is, the number of errors committed on incongruent trials (e.g., when the answer was blue for the word *bleu* printed in red) over all valid incongruent trials (Borella, Delaloye, Lecerf, Renaud, & de Ribaupierre, 2009; Ludwig, 2005).

Processing Speed

In order to test processing speed abilities, three different processing speed tasks were performed.

Letter and Pattern Comparison Tasks. The Letter Comparison Task was developed by Salthouse and Babcock (1991) in order to assess the verbal information processing speed (de Ribaupierre et al., 2011).

***Description*.** The participants were required to determine as fast and as accurate as possible, whether two series of three, six or nine consonants were identical or not. The pairs of series were presented in rows on two pages of paper with 21 trials each. The participants had to note on the respective row between the two to be compared series the letter *O* for yes (oui) or the letter *N* for no (non). Three practice trials were provided before the 42 test trials were presented.

For the spatial version, the Pattern Comparison Task consisted of a series of two patterns which had to be compared. The patterns were constructed out of three, six or nine lines and alike the verbal task, the answers had to be noted between the patterns. After the three practice trials, two pages of 30 pairs of patterns each were presented.

***Scores*.** The scores used were the average time spent for each page and the sum of errors committed on both pages for each task separately.

Simple Reaction Time Task. The perceptual-motor speed was evaluated with the computerized Simple Reaction Time Task (SRT; de Ribaupierre et al., 2011; Lecerf, de Ribaupierre, & Neimer, 2008) adapted from Hultsch, MacDonald, Hunter, Levy-Bencheton, and Strauss (2000).

***Description*.** The participants were asked to press a button on the E-prime button box as quickly as possible using their dominant hand when a white cross appeared in one of five different positions on a black screen. Prior to the stimulus, a white fixation point was presented with a varying ISI of 500, 800, 1100, 1400, or 1700 ms. The ISI were equally distributed across blocks. The order of the trials within a block was randomized except for two constraints: No more than two consecutive trials belonged to the same position and no more than two consecutive trials were presented with the same ISI. Six practice trials preceded 5 blocks of 24 trials each for a total of 120 testing trials.

***Score*.** Response latency was recorded for the delay between the apparition of the cross and the participant’s response. Extremely fast responses below 150 ms were excluded. The dependent variable was the mean reaction time across the remaining items. We additionaly calculated the intra-individual standard deviation for the purpose of exploring intra-individual variability.

Analyses

To address the above introduced hypotheses, the analyses described in the following sections were conducted. The statistical package SPSS for Windows was used for data analysis when not indicated differently.

Training Performance

The aim of analyzing the training data was to gain insight into age effects on training level and progression as well as on their predictors during the training.

ANOVA

First, age effects in training performance and its interaction with training sessions were investigated. To this end a mixed between-within subjects (repeated measures) analysis of variance (ANOVA) was conducted for both training conditions. The factors were Age Group (younger, older) as between-subjects factor and Training Session (session 1 to 10) as within-subjects factor. The same ANOVA was repeated with the initial and the last training session only, in order to verify if the interaction still persisted. Whenever variance homogeneity assumption was violated, checked with the Mauchly's test of sphericity, results with Greenhouse-Geisser adjusted degrees of freedom were reported. Significant interactions were followed up by pairwise comparisons with adjusted significance level using Bonferroni corrections to avoid family-wise alpha error inflation by multiple comparisons. For the ANOVAs, partial eta squared values (η^2_p) were reported as effect size indices. An alpha level of .05 for all statistical tests and an alpha level of .10 for marginally significant effects were used.

Regression analyses

Then, the questions whether cognitive abilities would predict *N*-back level at the different training sessions was addressed and whether these variables would explain age differences during training.

Five cognitive predictor variables were used, representing measures of gf, WM, interference, speed, and gc performance at pre-test: Raven score, Reading Span performance, interference ratio of the Stroop task, mean reaction time of the SRT task, and Mill Hill score, respectively. The predictor variables were the same measures as the ones used in training gain and transfer effect analyses.

To investigate the explained variance of the predictors, separate hierarchical multiple regression analyses were conducted for each training session. All cognitive variables were entered in a first step and age group in a second step. If R^2 was still significantly modified when adding age group as predictor, it implies that the cognitive variables were not sufficient to account for all age differences and a specific age effect was observed.

Latent growth curve model

To confirm the results from regression analyses, a latent growth curve model (LGCM) analyses, a standard analysis frequently applied to such longitudinal learning data, was introduced. The aim of the analyses was to model individual growth curves which allow examining intra-individual change in learning over time as well as inter-individual variability in intra-individual change (Preacher, Wichman, MacCallum, & Briggs, 2008).

The present sample size however hardly reached the required statistical power for the reliable use of the model fit indexes; this type of analyses was therefore introduced for additional exploratory purpose. Especially the Root Mean Square Error of Approximation (RMSEA) tends to overreject models at small sample sizes (Hu & Bentler, 1999). Therefore χ^2 -test, Standardized Root Mean Square Residual (SRMR), Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) were reported for absolute model fit evaluation. Values of the SRMR smaller than .08, of the CFI above .90 and of the TLI above .95 indicate adequate model fit.

Analyses were conducted in Mplus (Muthén & Muthén, 2004) using the maximum likelihood estimator. An unspecified model with a random intercept and a shape factor as slope was fitted (Duncan, Duncan, & Strycker, 2006; Tisak & Meredith, 1990). This specification allowed exploring the shape of the growth function without defining a priori the slope factor. The factor loadings were therefore estimated from data by constraining the first and the last loadings to 0 and 1, respectively, and estimating the remaining eight loadings. The estimated loadings revealed the shape of the longitudinal trend.

For the sake of parsimony, the variances of the ten training sessions were fixed on the same estimated value. Further, a linear slope model was tested against the unspecified one in order to test the hypothesis about linear growth. However, regarding χ^2 value the unspecified model ($\chi^2(51, n = 44) = 86.62$, $p = .001$, SRMR= .04, CFI = .95, TLI = 0.96) revealed a significantly better model fit than the linear one ($\chi^2(59, n = 44) = 113.62$, $p < .000$, SRMR= .04, CFI = .92, TLI = 0.94). The unspecified model was therefore retained for further analyses.

In a next step, the effect of age group on intercept and slope was examined by introducing it as a dummy variable predictor into the model. Next, the effect of cognitive predictors of individual differences in gf, WM, inhibition and speed performance were included into the model. All cognitive predictors were again measures from pre-test assessment and were entered as grand mean centered variables. The LGCM analysis was conducted for the WM training only, because no adequate model fit was identified for the implicit sequence learning training data.

Age Differences in Cognitive Tasks

Before investigating the effects of training, we compared the age groups on task performance at pre-test, in order to control for the classic age differences in cognitive variables. We conducted independent two-sample *t*-tests comparing younger and older adults on each cognitive variable separately.

Training Gain and Transfer Effects

To investigate age effects and to separate true training gains from confounding placebo effects and test-retest effects, all experimental groups were compared on the trained verbal 2-back task and on the transfer tasks.

Gain scores

Gain scores were calculated in order to reduce the ANOVA by the factor time of assessment (pre-test vs post-test). They were defined as the absolute difference between pre-test and post-test. We calculated post-test minus pre-test performance for accuracy data and pre-test minus post-test performance for correct response times.

ANOVA

The gain scores of the respective dependent variable were submitted to a 2 x 3 two-way ANOVA with the two between-subjects factors Age Group (young, older) and Training Group (WM training, implicit training, no-contact control).

As the total number of years of education was significantly lower in the older no-contact control group compared to the older training groups, education was included as a covariable in all analyses. Partial η^2_p values were reported as effect size for the two-way ANCOVAs. Moreover, it was assumed that participants had similar performance levels at pre-test between training groups within each age group. This assumption was overall not or only marginally violated.

High versus low performers

To further consider individual difference in initial performance level, two groups of participants were created in order to compare high performers (above median) to low performers (below median) at pre-test. Level effects were investigated by an Age Group x Training Group x Pre-test Level (low, high) ANOVA for those dependent variables where a significant training effect was found.

Cohen’s d

Finally, in order to estimate the relevance of the training effects, Cohen’s *d* effect size (Cohen, 1988) was calculated. To this end, the difference between pre-test and post-test was divided by the pooled standard deviation for both test occasions for each training and age group separately. This allowed the comparison of the training effects between experimental groups and different transfer tasks by a measure which is independent from raw scores. At the same time this score provides a measure of the practical relevance of the training effect. Effect sizes higher than $d = 0.8$ are considered as large effects, scores about $d = 0.5$ are medium effect sizes and scores smaller than $d = 0.2$ are conceived as small effects.

Specific Hypotheses

In the present study, we distinguished three different categories of change or gains in performance which formed our focus of interest: first, changes in training performance from the first through the last training session, entitled *training performance*; second, the gain in task performance (verbal *N*-back task) that was assessed at pre-test and post-test and was trained in the WM training group but not in the control groups, headed *training gain*; third, gains on entirely untrained cognitive tasks that were assessed at pre-test and post-test, entitled *transfer effects*. In the below sections, first specific hypotheses are formulated separately for each category of change according to the theoretical general hypotheses, and then, specific hypotheses are formulated which link these three categories.

Training Performance

First, we expected that the WM training task scores would significantly increase over the course of training and that there would be significant age differences showing that younger perform on a higher level than older. As concerns the interaction, training task scores would increase at a higher rate for younger adults than for older adults. Age differences should therefore be increased at the end of the training. Similar effects were expected for the ANOVA including the initial and the last trainings session only. Age differences would be magnified with trainings (Baltes & Kliegl, 1992; Brehmer et al., 2012).

As concerns the implicit sequence training, we expected age differences in reaction times and a significant decrease in reaction times over the course of trainings. We further expected a significant interaction as younger adults would learn faster and therefore reach the asymptote earlier than older adults. However, as proposed in previous studies, the total amount of learning should not differ between younger and older adults (Daselaar et al., 2003). Therefore, the ANOVA including only the initial and the last training session should not show a significant Age x Session interaction. Older and younger adults would show the same amount of training gain.

Second, we expected that the learning curves of both age groups would attain the asymptote over the course of the 10 sessions of WM training (Jaeggi, Buschkuehl, Perrig, &

Meier, 2010). This would be reflected in decreased significant differences in the training level between training sessions in the second half of the training.

Finally, according to the first hypothesis, the impact of age as a predictor of training performance should increase with advanced training for the WM training group. It was additionally expected that age would explain more of the variance than the cognitive predictors with advancing WM training as shown by Kliegl and colleagues (1990).

The first three hypotheses were assumed for both training programs. However, the last one was only hypothesized for the WM training as implicit learning is less affected by age.

Age Differences in Cognitive Tasks

We expected age differences in all cognitive variables as constantly has been reported in the literature (Salthouse, 1991; Schaie, Willis, Jay, & Chipuer, 1989): As concerns the WM tasks, that is, the *N*-back tasks, the Updating task and the Reading Span task, we expected less accurate responses in older than in younger adults (Chicherio, 2006; Delaloye et al., 2008; Kirchner, 1958; Verhaeghen & Basak, 2005). We expected the *N*-back WM load costs to be larger for older adults than for younger adults. Regarding response times in the *N*-back tasks, we expected significantly slower response times in older than in younger adults (Chicherio, 2006). As regards the Color Stroop task, we expected that older adults would be more susceptible to interference and therefore show a significantly larger interference ratio (Ludwig et al., 2010). For the speed measures, that is, the SRT task as well as the letter and pattern comparison tasks, we expected significantly slower response times for older adults than for younger (de Ribaupierre et al., 2011). For the Raven's Progressive Matrices task, we also expected less correct responses in older than in younger adults whereas in the Mill Hill scale assessing crystallized intelligence, we expected the inverse patter; that is more correct responses for older than for younger adults (de Ribaupierre et al., 2008; Horn & Cattell, 1967).

Training Gain

Concerning the trained verbal 2-back task, we expected first that there would be a significant training gain for older and younger adults of the WM training group only (main effect intervention). According to previous findings and the hypothesis regarding training performance, it was expected that the younger WM-trained adults would improve more in accuracy performance in the verbal 2-back task than older WM-trained adults (interaction Age x Intervention; Brehmer et al., 2012; Dahlin, Nyberg et al., 2008).

Concerning the WM load costs, that is, the difference in accuracy between 0-back and 2-back, we also expected that younger adults would decrease to a greater amount than older adults. The same was expected for *d* prime scores.

However, it was not clear whether a participant would show a large improvement when the initial performance level is low or when it is high. Therefore, a final exploratory analysis was introduced which compared the low and high performance level participants per age group.

Transfer Effects

The same hypotheses are proposed for the near transfer task, the spatial 2-back task, for accuracy in the 2-back task, WM load cost in accuracy and *d* prime. The hypotheses were that there would be, first, significant gains in the WM training group only (main effect intervention) and second, more gain effects in younger than in older adults for the WM training group (interaction Age x Intervention; Dahlin, Nyberg et al., 2008; S.-C. Li et al., 2008). Furthermore, the initial level effect was also explored.

Concerning the remaining cognitive tasks, we further hypothesized that a significant transfer effect would be observed for WM performance measured with Updating task and Reading Span task. According to some literature there should also be a transfer effect into the fluid intelligence measure - Raven's Progressive Matrices (Jaeggi et al., 2008). Again, a main effect of intervention group in favor of the WM training group and an interaction with age in favor of the younger adults was expected. Concerning the tasks of speed and inhibition performance, no significant gains were expected.

Linking Training Performance, Training Gain and Transfer Effects

Concerning the WM training group, we expected that training performance would be correlated with the training gain and transfer effects. The more the participants improve from session one through session ten, the more gains and transfer effects are expected.

ELECTROPHYSIOLOGICAL METHOD

In this section we will introduce the EEG method we implemented for a subsample of the behavioral population in order to gain insight into the cerebral changes due to WM training.

Participants

A subsample of 76 individuals (39 younger and 37 older adults) out of the above described sample conducted the *N*-back tasks at pre-test and post-test during EEG data recording in an additional session. Besides the above mentioned inclusion criteria (age between 18 - 38 and 60 or older and good French language skills), several additional inclusion criteria were examined by telephone interview before testing. Only right-handed participants were included, for whom a handedness score was calculated using the Edinburgh Handedness Inventory (Oldfield, 1971). The inventory includes twelve questions about hand preferences for several activities. The score is obtained by the following computation: (right - left)/(right+left) x 100; where right and left reflect the total score of hand preference respectively. The participants of the EEG subsample received financial compensation of approximately 10.- CHF per hour.

Participants were excluded if they reported a history of sustained head injury or neurological or psychiatric diseases. Moreover, subjects with regular use of stimulants, β-blockers, or psychotropic drugs were excluded. Three younger and one older participant were excluded due to a large number of artifacts in EEG data and therefore to an insufficient quantity of trials for averaging. The final EEG subsample consisted of 72 participants, 36 younger and 36 older adults. The group characteristics of the final EEG subsample are presented in Table 2. The three experimental groups did not differ on any sample characteristic examined separately per age group.

General Procedure and Material

With the EEG subsample, we implemented the same procedure as with the sample without EEG measures but added two specific sessions for EEG recording; one at pre-test and one at post-test (see Appendix A, Table A1). The first session at pre-test contained the same battery of cognitive tasks as described above, but without the *N*-back tasks. The participants performed the verbal and spatial *N*-back tasks in a second session at pre-test, while EEG data were recorded. Two similar sessions were administered at post-test after the training intervention. Prior to each EEG recording, some preliminary tests and preparations of the participants were conducted, aiming at enhancing the quality of the EEG measurement. These preliminary preparations and tests are now described, and subsequently details on the specifics of the testing environment are provided.

Table 2. Demographic sample characteristics for the EEG subsample.

	Younger adults			Older adults		
Variable	WM (n = 12)	Implicit (n = 12)	No-contact (n = 12)	WM (n = 12)	Implicit (n = 12)	No-contact (n = 12)
Gender	83	83	75	67	67	67
Age ***						
M	23.33	23.17	24.75	67.42	67.92	68.25
SD	3.78	4.84	4.37	5.47	4.36	6.23
Range	18-31	19-35	18-32	62-81	62-78	60-77
Health						
M	1.93	2.12	2.00	2.21	1.98	2.31
SD	0.40	0.51	0.27	0.31	0.50	0.38
Range	1.29-2.57	1.14-2.86	1.71-2.57	1.86-2.86	1.29-2.71	1.71-2.86
FrACT						
M	0.73	0.64	0.84	0.65	0.68	0.81
SD	0.23	0.16	0.18	0.21	0.23	0.23
Range	0.54-1.18	0.49-1.05	0.61-1.17	0.40-1.18	0.41-1.13	0.41-1.18
Oldfield						
M	83.65	88.89	84.72	87.5	83.3	86.51
SD	12.02	10.86	19.41	12.56	20.64	23.07
Range	66.67-100	66.67-100	33.33-100	66.67-100	41.18-100	33.33-100
Education						
M	14.58	14.67	15.33	16.17	14.83	14.08
SD	1.73	0.99	2.77	2.76,	2.21	2.03
Range	12-17	12-16	12-21	12-22	12-19	12-18
Diploma ***	92	92	92	75	50	42
Mill Hill ***						
M	36	34.17	33.17	39.5	38.33	40
SD	4.09	5.77	5.37	2.58	4.89	2.17
Range	29-43	17-39	18-40	35-43	28-44	37-43
Raven ***						
M	35.92	34.17	35.25	30.75	31.00	27.33
SD	6.43	4.24	6.31	(4.09,	4.18	5.14
Range	24-45	27-42	22-43	23-36)	26-39	20-36

Note. Significant differences between younger and older adults: *** $p < .001$. Gender values reflect percentage of women in each group. Age and education values reflect years. Health scores reflect average response to seven questions on a scale from 1 = excellent health to 5 = very bad health. FrACT = Freiburg Visual Acuity & Contrast Test. Oldfield values reflect percentage right-handedness. Diploma values reflect percentage of participants who hold a university-entrance diploma. Mill Hill scores represent total correct responses out of 44 words. Raven score represents total correct responses out of 48 at pre-test. Range = minimum – maximum value.

Preliminary Tests and Preparation of Participants

Before starting with an EEG session, the participants had to indicate the number of hours of sleep during the previous night as well as food, coffee, and tea intake since getting up in the morning. They were also asked to indicate medication intake during the past 24 hours. This information was asked to be prepared for the eventuality that EEG recording would yield an abnormal pattern. Further, we informed participants about causes of artifacts that potentially occur while EEG recording, such as eye blinks and other muscular contractions, and they were encouraged to reduce them. Prior to the implementation of the *N*-back tasks, several control measures were recorded: The participants were asked to close their eyes for one minute, followed by a minute of focusing a point in the center of the screen. Three beep sounds marked the end of each minute. This procedure was repeated once. Then, the participants had to focus a red point in the center of a black and white checkerboard, again for one minute. The checkerboard switched the black and white squares each 500 ms. These measures served as a final check for ill-behaved electrodes and as a control for alpha waves during the closed eyes sequences.

Testing Environment

Participants were tested individually in the laboratory of psychophysiology in the building Uni Mail of the University of Geneva. Including EEG material preparation and hair washing, one EEG session lasted about 120 to 150 minutes. Participants were seated in an armchair in a darkened and sound- attenuated room. The distance to the screen was approximately 90 cm. The scripts were run on a Dell computer and displayed on a 19 inch VGA color computer monitor. The room was monitored by a camera and equipped with an interphone in order to allow the experimenter to stay in contact with the participant during data recording. EEG was continuously recorded with an Active-Two BioSemi 64-channel Ag-AgCl active electrode acquisition system in conjunction with ActiView BioSemi software (BioSemi Active-Two, V.O.F., Amsterdam, the Netherlands). Data was digitized at a sampling rate of 2048 Hz in a bandwidth filter of 0-417 Hz. All electrode offsets were held between ± 20 microvolt (μV).

Analyses

Raw Data Pre-Processing

After data acquisition, EEG raw data was processed using the software Brain Vision Analyzer (Brain Products GmBH). First, they were offline down-sampled to 256 Hz and band-pass filtered between 0.1 and 30 Hz (24 dB/oct) with a notch filter of 50 Hz (Luck, 2005). Further, the signal was referenced to the average of all electrodes and a global DC detrend correction was applied (Hennighausen, Heil, & Rösler, 1993; Picton et al., 2000). Event-related potential (ERP) epochs were extracted from 200 ms pre-stimulus to 1000 ms post-stimulus. Epochs containing artifacts with a threshold of ± 75 µV were removed. Data covering substantial eye movement artifacts in the epochs were corrected by means of an independent component analysis before artifact rejection (Jung et al., 2000). Only components consistent with topographies for eye blinks and horizontal eye movements were removed. Channels containing substantial noise were interpolated using a spherical spline interpolation (Perrin, Pernier, Bertrand, & Echallier, 1989). A 200 ms pre-stimulus baseline correction was applied. We only retained artifact free trials of correct responses. Target and non target trials were averaged separately for each task. Data files which did not reach a minimum of 25 epochs per average were eliminated at pre-test and post-test (Chen et al., 2008). However, in order to preserve as much information as possible, files with at least 16 valid epochs were retained (Angel, Fay, Bouazzaoui, Baudouin, & Isingrini, 2010; Friedman & Johnson, 2000). Additionally, the files were only kept whenever the visual inspection of individual scalp topography ascertained proper ERP patterns. After this data pre-processing procedure, the data set consisted of one file for younger adults and ten files for older adults with 16 - 24 epochs.

Classic Event-Related Potential Analysis

In the following sections, several methodological ERP issues are discussed. First, different methods for extracting ERP measures are introduced. Second, the implementation of these measures in the present study is specified. Third, an important issue is the choice of the electrode sites and their assignment to so-called regions of interest (ROIs) that will be analyzed further. Then, we determined how to compute a gain measure based on the ERP measures; this procedure is similar to the gain measures derived from behavioral data. Finally, the statistical analyses conducted are introduced.

Methods for extracting ERP measures

The two main dependent variables analyzed from ERP components are the amplitude in µV and its latency in ms. As concerns the amplitude, two strategies for extracting a measure for the component’s amplitude out of the stream of signals are commonly used (Handy, 2005; Luck, 2005). The first strategy, the so-called peak-amplitude measure, relies on the maximum amplitude within a particular time window. The window is determined relative to stimulus onset based on commonly used time windows for specific components and post-hoc visual amplitude inspection. The second strategy, the so-called mean amplitude measure, also analyzes the stream of signals within a pre-determined window. However, it calculates the mean voltage instead of the maximum voltage. Even though both measures are widely applied, recent literature recommends using the mean amplitude measure for the following two reasons. First, the peak amplitude measures are less reliable as there is a considerable inter-and intra-individual variability in ERP waveforms (Picton et al., 2000). Therefore, it is possible that the maximum amplitude occurs at the edge of the determined time window, leading to a biased measure. To avoid this problem, a local peak amplitude detection can be used which is defined as the maximum value that is surrounded on both sides by smaller values (for illustration see Figure 4). It is, however, questionable from a theoretical point of view, whether a component occurs at different latencies on different electrode sites (Picton & Stuss, 1980). Second, the various electrodes are not signed the same way. As an example, frontal electrodes show negative peaks during the N2 component whereas occipital electrodes show positive ones. Therefore, several electrodes show a negative peak while others necessarily show a positive peak. For that reason it is impossible to extract a reliable negative peak for all electrodes within a given time window. The recommended alternative is the use of a representative reference channel at which the component's latency is defined (Picton et al., 2000). The amplitude measure is then taken for all electrodes at the reference electrode latency. However, the local peak amplitude with or without the use of a reference channel is artificially increased when the noise level is high. This occurs especially when averaging over a small number of trials. Further, ERP components can have no clear peak and therefore not providing a definitive point at which to measure peak amplitude (Handy, 2005). For these reasons, the mean amplitude measure is less polluted by the noise level of the data, such that it is highly recommended to rely on this method. Regarding the latency measure, it can either be defined as the average of individual peak latencies or as the peak latency of a defined reference channel. The problems with the peak detection described above are also relevant in this case, but no good alternatives are available.

Implementation of amplitude and latency measures in the present study

In the present study, the amplitude as well as latency measures were extracted for the N2 and the P3 component. In each case, a time window had to be determined prior the extraction of the respective measure. Furthermore, for the peak amplitude measure, it had to be indicated whether negative or positive peaks were searched, that is, for local minima or maxima. These issues are now discussed in turn.

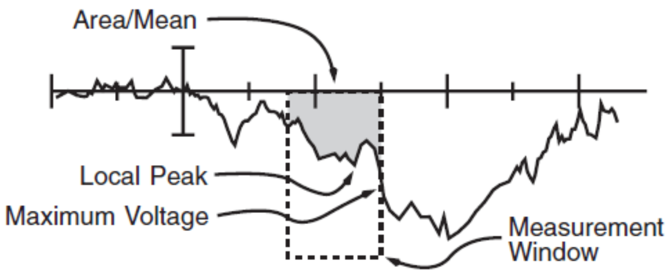

Figure 4. ERP waveform with illustration of different measurements: mean amplitude, local peak amplitude and maximum peak amplitude (called voltage in the figure) in a defined measurement window (Luck, 2005; Chapter 6).

Implementation of the amplitude measures. For amplitude measures we used the mean amplitude as recommended by the literature (see above; e.g., Handy, 2005). The time windows were identified by visual inspection of grand average brain topographies and waveforms. The windows were set separately for younger and older adults as well as for verbal and spatial conditions. For younger adults, the N2 mean amplitude was calculated between 180-300 ms for the verbal tasks and between 210-300 ms for the spatial tasks. The P3 mean amplitude for younger adults was calculated for both task conditions - verbal and spatial - between 300-550 ms after stimulus onset. For older adults, the windows were set with a longer delay after stimulus onset compared to the delay for younger adults. For the verbal tasks, the N2 mean amplitude was averaged between 220-350 ms and for spatial tasks between 240-350 ms. For verbal and spatial tasks, the P3 component was averaged between 350-600 ms after stimulus onset (see Table 3 for time windows summary). The same time windows were used for pre-test and post-test data.

Implementation of the latency measures. For the latency measure, a local peak detection on a reference electrode was used. The N2 component latency was determined as the local negative peak for the reference electrode Cz within a 175-350 ms window after stimulus onset. For the P3 component, the local positive peak on the Pz electrode within 275-650 ms after stimulus onset was measured (see Table 3). The time windows were defined based on the method by Daffner and colleagues (2011) who also investigated ERP from *N*-back tasks comparing younger and older adults. In the present study however, the N2 window was slightly enlarged after visual inspection of the grand average waveforms; it begins at 175 ms instead of 210 ms. The end of the window is kept at a latency of 350 ms as suggested by Daffner and colleagues (2011). Reference channels were specified as the most representative channel of the component by visual inspection of grand average brain topography. The same time windows were used for pre-test and post-test data as well as for both *N*-back conditions and age groups.

Table 3. Specified time windows in ms after stimulus onset for the components N2 and P3 for mean amplitude measure and local peak latency detection.

	Mean Amplitude				Local Peak Latency
	Younger adults		Older adults		
Component	Verbal	Spatial	Verbal	Spatial	*(Reference electrode)*
N2	180 – 300	210 – 300	220 – 350	240 – 350	175 – 350 *(Cz)*
P3	300 – 550		350 – 600		275 – 650 *(Pz)*

Location of the electrodes and assignment to ROIs

For component analyses, single electrode site measures are commonly used, as proposed in several ERP studies using the N-back paradigm (e.g., Daffner et al., 2011; McEvoy et al., 1998; Missonnier et al., 2004). The chosen electrodes were located in order to cover regions where N2 and P3 components showed maximum peaks. For the present study, the electrode sites were chosen following the mentioned existing ERP studies. We also computed several regions of interest (ROIs) by averaging a selected electrode plus 2 adjacent ones. This was possible since a higher number of electrodes was recorded in the present study (64) as compared to previous *N*-back studies (32; e.g., Daffner et al., 2011). Furthermore, the use of ROIs has several advantages (Dien & Santuzzi, 2005). ROIs represent descriptive units and allow reducing the number of levels per factor, which improves the ease of result interpretation. Another advantage is that it reduces variability and attenuates the potential bias which can be introduced by a single noisy electrode. To this end, nine ROIs were defined according to the proposed electrode locations by averaging three electrodes for each ROI (see Figure 5). The ROIs were then named in relation to their laterality position on the x-axis (left, central, right) and their anterior - posterior position on the y-axis (frontal, central, parietal/occipital; see Figure 5). The ROI averages were calculated for amplitude and latency measures of each component separately.

ERP scores

In analogy to the analysis of behavioral data, we calculated gain scores for ERP data. Thereby, we proceeded as follows. To calculate the gain score, the difference between pre-test and post-test measures was extracted for both components separately. If post-test measures were larger than pre-test measures, positive gain scores resulted. Note that for the N2 component, an increased amplitude is reflected by a more negative amplitude. The gains were calculated as pre-test minus post-test. The following example for the N2 case illustrates this. Suppose that the N2 pre-test amplitude reaches -1 µV and the corresponding post-test amplitude is - 2 µV. In this example, the gain score would be -1 µV - (- 2 µV) = 1 µV. For the P3 component in contrast, because it is a positive amplitude component, the gain is calculated as post-test minus pre-test measures. Suppose that the P3 pre-test amplitude is 3 µV and the corresponding post-test amplitude is 4µV. In this case, the gain score would be 4 µV - 3 µV = 1 µV.

The gain score for latency measures was calculated as pre-test minus post-test measures for both, the N2 and the P3 component. This resulted in a positive gain score in ms whenever latency was decreased at post-test compared to pre-test.

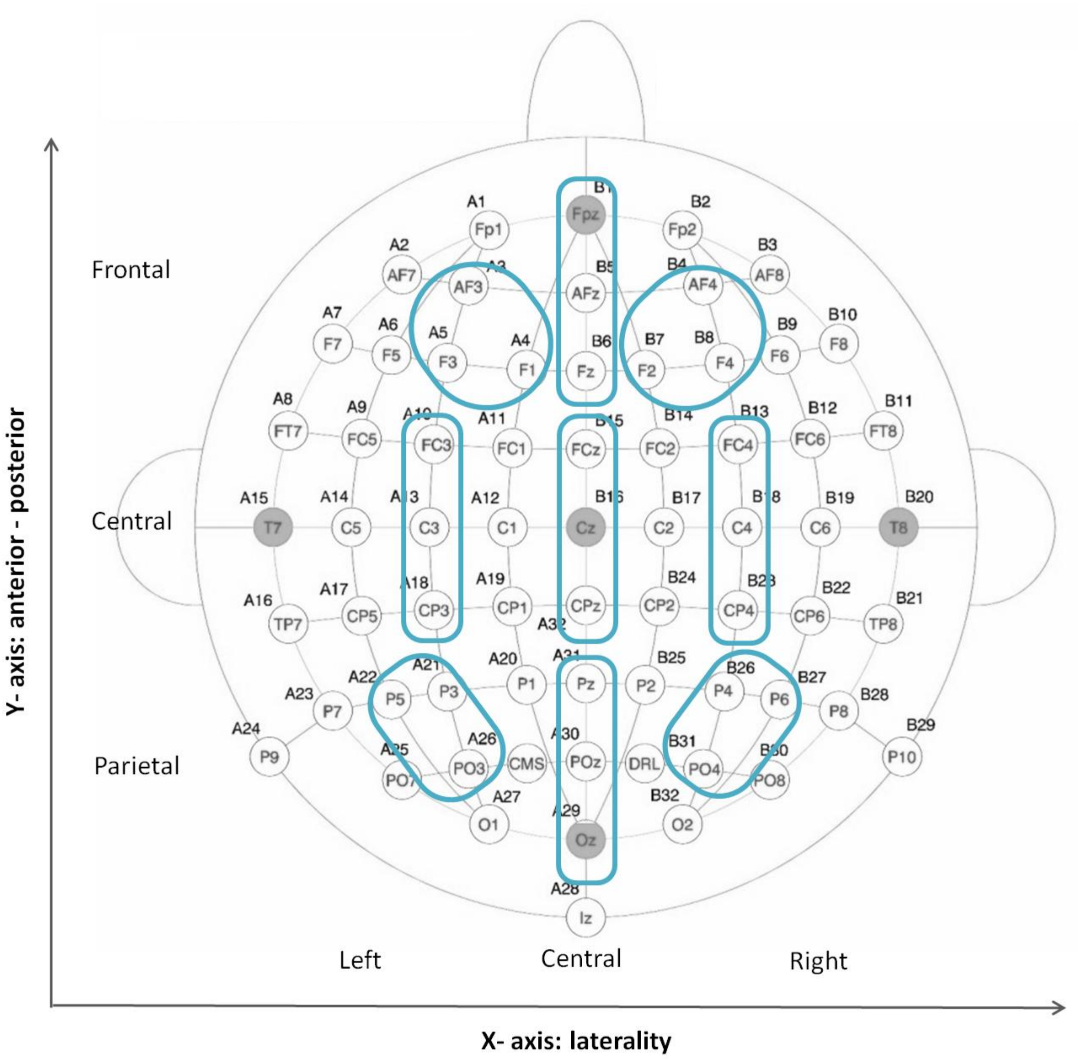

Figure 5. Biosemi montage for 64 electrodes illustrating the nine ROIs created for analyses.

Statistical analyses

A repeated measures ANOVA with Age Group (younger, older) and Training Group (WM training, implicit training, no-contact control) as between-subjects factors and X-axis ROI (left, central, right) and Y-axis ROI (frontal, central, parietal) as within-subjects factors was conducted for each measure. In analogy to the analysis of behavioral data, Greenhouse-Geisser and Bonferroni adjustments were applied if necessary.

Spatio-Temporal Microstate Segmentation

In a second step, a determination of functionally relevant time periods was applied by means of the spatio-temporal microstate segmentation of the ERPs (Brunet & Murray, 2011; Michel et al., 2004; Michel et al., 1999). The analysis was performed using the Cartool software programmed by Denis Brunet from the Functional Brain Mapping Laboratory in Geneva (www.brainmapping.unige.ch/cartool; supported by the Center for Biomedical Imaging of Geneva and Lausanne, Switzerland).

This method allowed the examination of spatial variations of the scalp topography over time, between conditions and groups (Murray, Brunet, & Michel, 2008). The advantage of this method was that all 64 electrodes were considered at once such that the measure was not dependent from one single electrode. The analysis reduces the ERP data to an optimal number of stable topography maps, of which each represents a *functional microstate* of information processing in the brain (Lehmann, 1987; Michel et al., 2009). These periods of stable microstates were identified in the data set with a cluster analysis.

The analysis consisted of several steps: First, for hypothesis generation, we conducted a *K*-means clustering analyses at the group level on the grand average ERPs. This microstate segmentation was conducted for each *N*-back condition separately, including for each condition all grand average files from pre- and post-test as well as from all age and training groups. We used 300 random trials and allowed 1 to 20 clusters. The analyses were run in a time frame of 0 to 600 ms after stimulus presentation. Further, two microstate maps were merged, when they correlated above .88 and maps smaller than 20 ms were rejected (Fingelkurts, 2006).

Then, the number of microstate maps that best explained the data across groups and session was determined by several indicators of goodness-of-fit: the absolute minimum in the Cross Validation (CV) value, a high Krzanowski-Lai (KL) value as well as a large global explained variance (GEV). Then, a time frame of interest was determined where microstate maps differed between session and/or groups. These microstate maps were subsequently fitted to the ERPs of each participant on the basis of spatial correlation. More precisely, for each time point of the individual ERP, the scalp topography was compared to each microstate map present in the grand average in the determined time frame. The time point is then labeled according to the microstate map that correlated best.

Finally, several dependent measures were provided from this fitting procedure: The most important ones were the duration of a given map and the GEV of a given map. The former indicates the number of data points at which the microstate map was present. Cartool reports the duration in time frames, an arbitrary time unit, for which in our analyses 1 time frame corresponded to 4 ms. The latter dependent variable, GEV, indicates how much variance was explained by the map and represents the squared spatial correlation in a score between 0 and 1.

Scores and statistical analyses

Similarly to the previous analyses, we first calculated gain scores for the duration and the GEV for each map. Then, a 2 x 3 two-way ANOVA with the two between-subjects factors Age Group (young, older) and Training Group (WM training, implicit training, no-contact control) was conducted.

Specific Hypotheses

Training Gain

According to the theoretical hypotheses we expected for the ERP components a decreased amplitude for the N2 as well as the P3 component. This would indicate that the N2 was more equipotential and the P3 less frontally distributed after training (Daffner et al., 2011; Daffner et al., 2010; Duarte et al., 2006; Erickson et al., 2007b; Lorenzo-López et al., 2007; Missonnier et al., 2004; Nagel et al., 2009; Riis et al., 2008). However, decreases were expected to be larger for younger adults than for older adults. As regards latency, it would decrease for younger as well as older adults probably to the same amount.

With respect to microstate maps, we expected to observe a new map for younger adults after training, reflecting a reorganization pattern. For older adults, we expected no reorganization, but a redistribution pattern which would be observed in a longer duration and larger GEV for maps which showed less frontally oriented positivity and in analogy a decrease for frontally oriented maps (see *Investigation of Training and Transfer Effects in Behavior and Brain* on page 61 for the operationalization of the reorganization and the redistribution pattern).

Transfer Effect

We expected similar decrease patterns for the spatial 2-back task for the ERP components as well as for the microstate maps. However, changes would in general be less pronounced than for the trained verbal 2-back task.

RESULTS

Behavioral Results

In the behavioral results section, the performance in both training tasks is analyzed as a first step. As a second step, age differences at pre-test are reported. In a third step, the differential effects of training condition and age group in the trained verbal 2-back task, so-called training gain, and in the untrained transfer tasks, so-called transfer effects, are reported. These analyses were conducted on the whole sample of N = 128 when not indicated differently. As a last step, training gain and transfer effects are linked to the training performance for the WM training group.

Training Performance

Working Memory Training

Figure 6 shows the individual training curves for each participant in thin lines (for descriptive statistics see Table 4). The N-back levels of the 15 training blocks were averaged within each individual for each training session. The dashed bold lines show the group means for younger adults in light green and for older adults in dark blue color.

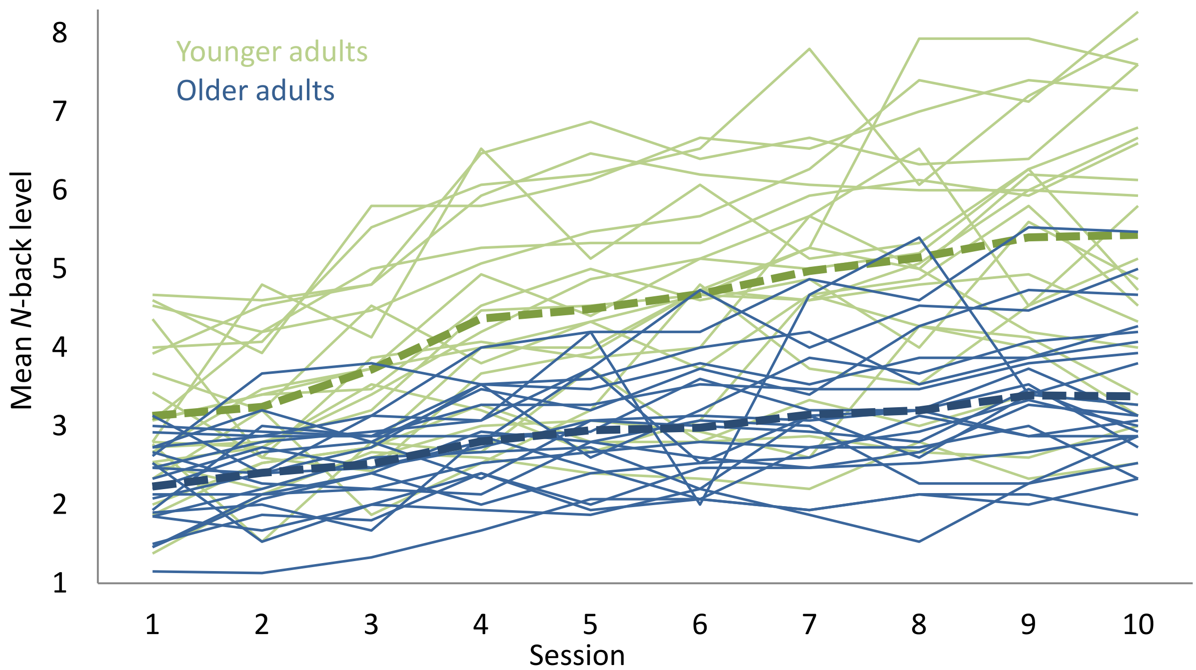

Figure 6. Working Memory Training. Individual training curves of the averaged N-back level as a function of session; younger adults in light green, older adults in dark blue, bold dashed lines represent the mean of each age group.

ANOVA

In order to investigate age effects on training level and training progression a repeated measures ANOVA was conducted. The ANOVA revealed a significant effect of age ($F(1, 42) = 31.13$, $p < .001$, $\eta^2_p = .43$) and session ($F(4.41, 185.24) = 54.7$, $p < .001$, $\eta^2_p = .57$) as well as a significant Age x Session interaction ($F(4.41, 185.24) = 6.62$, $p < .001$, $\eta^2_p = .14$). Younger adults (*M* = 4.46, *SD* = 0.2) exhibited a higher mean level *N*-back level than younger adults (*M* = 2.9, *SD* = 0.2). For older adults, pairwise comparisons revealed that the initial session was significantly different from the 5th to the 10th session and that the last session was significantly different from the 1st to the 3rd session, but no longer from the 4th session on. For younger adults however, the first session was significantly different from the 3rd session on until the 10th, and the last session was significantly different from the 1st through the 6th session (for pairwise comparison results see Appendix B, Table B1). These results showed that the older adults attained the asymptote earlier than the younger adults. Age differences persisted all along the ten training sessions.

Table 4. Mean, standard deviation and range (min. – max.) of the average *N*-back level per training session.

	Young (*n* = 22)			Older (*n* = 22)		
Session	*M*	*SD*	Range	*M*	*SD*	Range
1	3.13	0.92	1.38 - 4.67	2.23	0.55	1.15 - 3.13
2	3.25	0.90	1.53 - 4.80	2.41	0.59	1.13 - 3.67
3	3.72	1.06	1.87 - 5.80	2.52	0.56	1.33 - 3.80
4	4.37	1.26	2.53 - 6.53	2.81	0.60	1.67 - 4.00
5	4.49	1.26	2.40 - 6.87	2.95	0.69	1.87 - 4.20
6	4.68	1.31	2.33 - 6.67	2.98	0.79	2.00 - 4.73
7	4.97	1.39	2.20 - 7.80	3.15	0.82	1.87 - 4.87
8	5.15	1.45	2.67 - 7.93	3.20	0.93	1.53 - 5.40
9	5.40	1.53	2.33 - 7.93	3.39	0.87	2.00 - 5.53
10	5.43	1.78	2.53 - 8.27	3.38	0.93	1.87 - 5.47

Regression analyses

Hierarchical multiple regression analyses were conducted in order to disentangle the amount of variance explained by the cognitive variables from the amount of variance explained by age. The five cognitive variables were entered in a first step and age group in a second step into the analysis, which was conducted for each session separately. The results are summarized in Table 5. Regarding the R^2 the cognitive predictors explained all along the ten training sessions an important part of the variance in *N*-back level.

Adding age group as a predictor into the model did not explain more variance in the first three sessions. However, from the 4th session on, age group accounted for a significant amount of variance beyond the cognitive variables. As regards the different cognitive variables, gf significantly predicted the average *N*-back level from the initial through the last sessions. Gc however did not predict training performance at any training session. The results were mixed for WM, interference, and speed. Speed performance predicted eight out of the ten training sessions significantly or marginal significantly. WM performance only marginally predicted the initial, the 5th and the last training session. Interference however seemed to get more important during training by marginally predicting the last four training sessions.

In sum, these analyses showed that age did not significantly explain more of the performance variance in the initial training sessions than the cognitive predictors. But when training advanced, age group became a significant predictor of training performance beyond the inter-individual differences in cognitive abilities. However, it has to be noted that gf ($r(44) = -.57$, $p < .000$), Speed ($r(44) = .38$, $p = .011$) and gc ($r(44) = .50$, $p = .001$) performance were significantly correlated with age group before training, whereas reading span ($r(44) = .04$, $p = .795$) and interference ($r(44) = .18$, $p = .256$) performance was not.

Table 5. Hierarchical multiple regression analyses of *N*-back level on cognitive abilities and age group per training session.

	Session									
Measure	1	2	3	4	5	6	7	8	9	10
Predictors										
gf	.46**	.50**	.43**	.45**	.47**	.59**	.54**	.46**	.44**	.41**
WM	.24+	.20	.19	.18	.23+	.15	.16	.13	.21	.26+
Int.	-.07	-.03	-.23*	-.09	-.18	-.11	-.24*	-.20+	-.20+	-.22+
Speed	-.33**	-.19	-.38**	-.33**	-.23+	-.10	-.31**	-.40**	-.25+	-.30*
gc	-.05	-.01	-.13	-.18	-.13	-.08	-.05	-.09	-.13	-.14
R^2	.57**	.46**	.61**	.55**	.53**	.50**	.65**	.61**	.51**	.55**
ΔR^2 age	.03	.04+	.02	.06*	.07*	.10**	.06*	.05*	.09**	.06*

Note. Values for predictors are standardized regression coefficients. Predictors were entered in the first step; age group was entered in the second step. Int. = Interference. ** $p < .01$. * $p < .05$. + $p < .10$.

Latent growth curve model

To underpin these results, a latent growth curve was modeled that aimed at examining intra-individual change in learning over time as well as inter-individual variability in intra-individual change. The retained unspecified LGCM on the ten training sessions revealed a significant intercept of 2.69 (*SE* = 0.13, p < .001) and a significant slope of 1.76 (*SE* = 0.18, p < .001; for model selection procedure see Method section; no figure for the basic model is provided). Slope and intercept were positively correlated (*r* = .56, cov = .44, p = .003). This correlation suggested that those individuals who started at a higher level of performance (intercept) gained more in cognitive performance during the training.

A conditional model with age group as predictor further indicated that being in the one or the other age group significantly predicted initial level and growth curve: For older adults the initial level was 0.86 (std. coefficient = -.57, p < .000) *N*-back levels lower and slope was 1.24 (std. coefficient = -.59, p < .000) *N*-back levels lower than for younger adults. Age group accounted for 33% of the total variance in intercept and for 35% of the total variance for slope. Slope and intercept were now just marginally correlated (*r* = .33, cov = .17, p = .088).

Further, to examine the potential modifiers of training level and growth curve despite age, LGCM analyses were repeated with all cognitive predictors separately while controlling for age group. Speed, gf and the spatial *N*-back measures significantly predicted intercept when controlling for age. Reading span and the verbal *N*-back measures also significantly predicted slope. There was no significant effect of gc or interference.

In all these models age group still significantly predicted slope and intercept with one exception: gf performance was a significant predictor of intercept whereas age group significantly predicted slope only (see Figure 7). That is, with every additional point reached in the Raven score, the intercept increased by 0.05 *N*-back levels (std. coefficient = .57, p < .000) while age group predicted the difference between groups of 0.95 *N*-back levels in the growth curve (std. coefficient = -.46, p = .006; lower score for older adults). These results suggested that high gf ability predicted a high initial training performance while age group predicted the development during training. Belonging to the younger group predicted a higher development during training but did not predict the initial level. Together, the predictors explained 55% of the variance in intercept and 39% in slope. The additional explained variance proper to gf after controlling for age effects was 22% for intercept and 4% for slope. Slope and intercept were not significantly correlated in this model (*r* = .25, cov = .1, p = .244). The fit indexes did not change substantially by introducing any reported predictor into the model; fitting procedures suggested therefore that the models were still tenable.

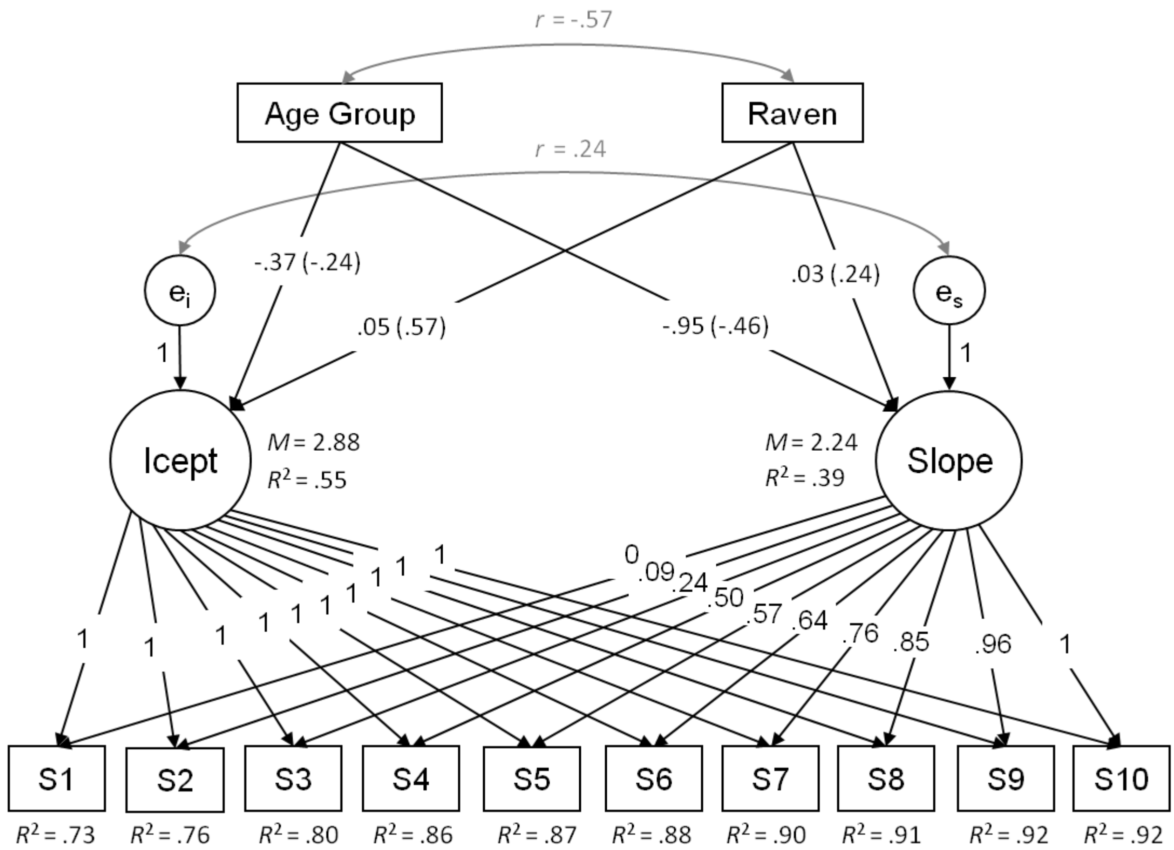

Figure 7. Path diagram for the conditional unspecified growth curve model with Age Group and Raven score (at pre-test) as predictors. Observed variables in rectangles (S= session), latent variables in circles (e=error; Icept = intercept). Regression paths are unstandardized coefficients; standardized coefficients are in parentheses.

Strategies used during training

We further explored the strategies used during the WM training. Participants were free to report the strategy, no strategies were proposed. We created nine post-hoc groups of strategies: the use of 1) memory or concentration, 2) intuition, 3) verbal rehearsal of the letters, 4) counting the letters, 5) retaining the series of letters (according to the required *N*), 6) retaining the series of letters and comparing and replacing the newly occurring letters, 7) updating of the series, 8) assorting the letters in small groups, 9) visualization of the series of letters. Two strategies were significantly more often used (or at least reported) in younger adults than in older adults: a) assorting the letters of the series (strategy 8; $t(42) = 2.45$, $p = .019$) and b) retaining the series of letters until the required *N*, comparing this series with the newly appearing letters in the task and replacing them when a new letter appeared (strategy 6; $t(42) = 2.49$, $p = .017$). These were two very efficient strategies which were useful at higher *N*-back levels. Older adults in contrast reported marginally more often that they tried to focus and use concentration and memory without further specifying the strategy (strategy 1; *t*(42) = -1.9, *p* = .064).

Implicit Sequence Learning Training

As regards the implicit sequence learning training, Figure 8 shows the individual training curves for each participant (for descriptive statistics see Table 6). Correct reaction times were averaged for each training session. The dashed bold lines show the group means for younger adults in light green and for older adults in dark blue. It has to be noted that with advanced training sessions, the participants learned the series implicitly and were no longer waiting for the next stimulus to appear on the screen. Therefore, the ISI of 120 ms became a part of the reaction time but was not included in the measure since reaction times were measured from stimulus onset on. This resulted in very fast reaction times by the end of the training.

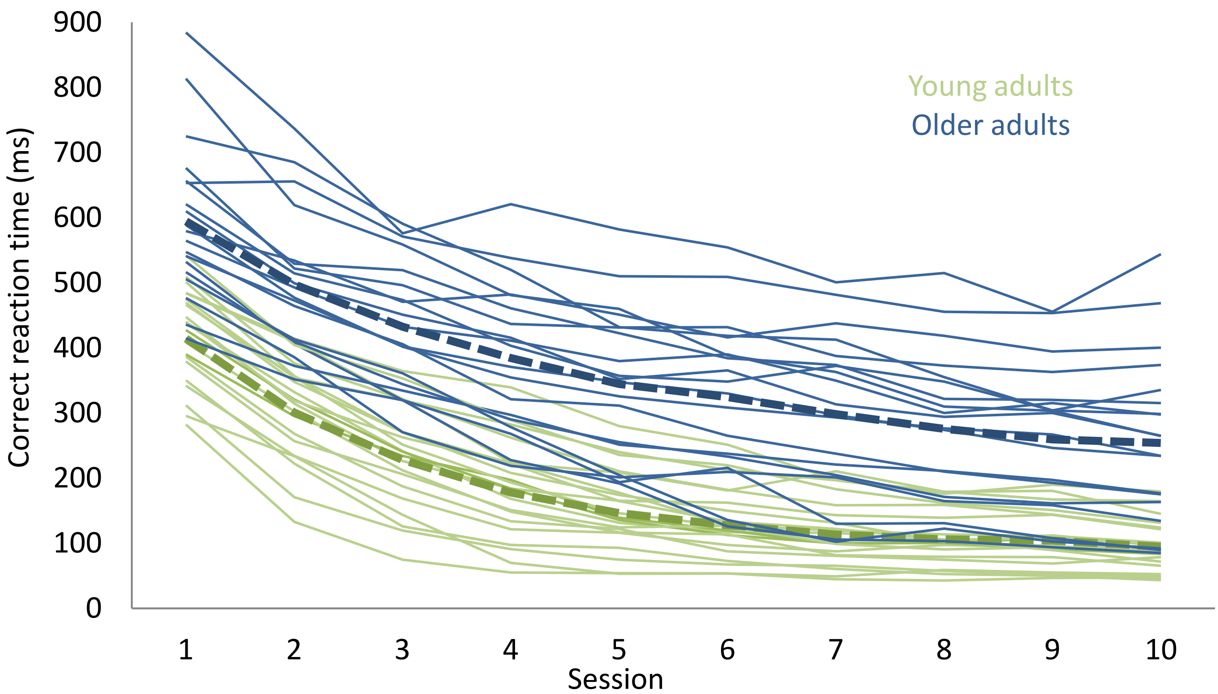

Figure 8. Implicit Sequence Learning Training. Individual training curves of the averaged correct reaction time as a function of session; younger adults in light green, older adults in dark blue, bold dashed lines represent the mean of each age group.

ANOVA

A repeated measures ANOVA was conducted in order to analyze age effects on training level and training progression. The ANOVA revealed a significant effect of age ($F(1, 38) = 46.06$, $p < .001$, $\eta^2_p = .55$) and session ($F(2.05, 77.78) = 508.99$, $p < .001$, $\eta^2_p = .93$) and a significant Age x Session interaction (F(2.05, 77.78) = 3.73, p = .024, η^2_p = .09). The post-hoc comparisons revealed that all sessions were significantly different from each other with the exception of the 9th and 10th session for older adults and from the 7th to the 10th session for younger adults (for pairwise comparison results see Appendix, B Table B2). Younger adults reached the asymptote therefore earlier than older adults. Age differences however persisted through the 10 training sessions.

In order to verify whether the initial and the final age difference changed, we conducted a Training session x Age ANOVA. It revealed that the interaction was no longer significant (F(1, 38) = 0.83, p = .369, η^2_p = .02) when considering only session 1 and 10 and that therefore no magnification of age differences was present.

Table 6. Mean, standard deviation and range (min. – max.) of the average reaction time in ms per training session.

	Young (n = 20)			Older (n = 20)		
Session	*M*	*SD*	Range	*M*	*SD*	Range
1	414	70	282 - 509	594	118	415 - 884
2	301	75	132 - 414	497	107	352 - 737
3	228	76	74 - 364	433	97	271 - 591
4	178	72	55 - 339	384	111	219 - 621
5	146	59	53 - 280	345	114	192 - 582
6	127	52	53 - 251	324	117	126 - 554
7	113	48	45 - 211	298	121	102 - 501
8	106	42	43 - 179	276	116	103 - 515
9	104	43	47 - 190	260	111	94 - 456
10	95	38	43 - 180	254	128	85 - 544

Regression analyses

A hierarchical multiple regression analysis was also conducted for this training task in order to disentangle the amount of variance explained by the cognitive variables from the amount of variance explained by age. The five cognitive variables were entered in a first step and age group in a second step into the analysis, which was conducted for each session separately. The analyses revealed that the role of cognitive predictors and age did not change in the course of the training (see Table 7). Age group accounted for a significant part of the variance in reaction times in all sessions beyond the cognitive variables. Concerning the five cognitive predictors, there was also no considerable change over training sessions in accounting for variance with the exception of interference. All cognitive variables showed a significant or marginally significant beta value at all training sessions, with the exception of interference whose beta value was no longer significant from the second session on. These analyses show that age significantly accounted for performance variability from the first session on until the last one.

Table 7. Hierarchical multiple regression analyses of correct reaction times on cognitive abilities and age group per training session.

	Session									
Measure	1	2	3	4	5	6	7	8	9	10
Predictors										
gf	-.33*	-.34*	-.29+	-.31*	-.33*	-.27+	-.29+	-.29+	-.27+	-.29+
WM	-.27+	-.30*	-.28+	-.27+	-.29*	-.28+	-.28+	-.28+	-.30*	-.33*
Int.	.24*	.19	.17	.13	.16	.17	.16	.13	.12	.10
Speed	.29+	.23	.28+	.32*	.31*	.34*	.28+	.29+	.30+	.29+
gc	.27*	.33*	.34*	.31*	.30*	.32*	.34*	.34*	.34*	.29*
R^2	.55**	.52**	.50**	.47**	.49**	.54**	.43**	.43**	.43**	.51**
Δ R^2 age	.12**	.18**	.22**	.19**	.18**	.17**	.16**	.16**	.14**	.13**

Note. Values for predictors are standardized regression coefficients. Predictors were entered in the first step; age group was entered in the second step. Int. = Interference. ** $p < .01$. * $p < .05$. + $p < .10$.

Age Differences in the Cognitive Tasks

Differences between the younger and the older group with all intervention groups confounded were analyzed for the pre-test scores. The results showed the expected age differences between younger and older adults in nearly all the assessed cognitive tasks (Table 8). Age differences for the 0-back tasks were significant for response times but not for accuracy measures. This was due to a ceiling effect in accuracy measures which led to censored data. All the 2-back task and WM load cost measures in contrast showed significant age differences, that is, less accurate and slower responses as well as more WM load costs for older adults than for younger adults in line with the literature (Chicherio, 2006). Only the d’ measures did not always reach more than marginal significance. As regards the other cognitive tasks, all expected age-differences in accuracy and latency were found with the exception of the Reading Span task. Younger and older adults showed the same average number of correct words, which was not in line with the literature, since older adults usually recall less words than younger adults (e.g., de Ribaupierre et al., 2011).

Table 8. *T*-tests for the difference between younger and older adults on the cognitive task scores at pre-test.

	Younger *M (SD)*	Older *M (SD)*	*t*	*p*
0-back tasks				
Verbal accuracy	.97 (.03)	.97 (.03)	-0.43	.665
D prime	4.12 (0.67)	4.22 (0.63)	-0.91	.365
RT (ms)	460 (65)	533 (69)	-6.19	< **.001**
Spatial accuracy	.96 (.04)	.95 (.04)	1.46	.147
D prime	3.78 (0.71)	3.71 (0.86)	0.44	.660
RT (ms)	505 (112)	661 (170)	-6.11	< **.001**
2-back tasks				
Verbal accuracy	.91 (.06)	.86 (.11)	3.25	**.002**
D prime	3.02 (0.81)	2.73 (1.12)	1.66	.099
RT (ms)	630 (167)	865 (260)	-6.05	< **.001**
Spatial accuracy	.92 (.07)	.86 (.12)	3.35	**.001**
D prime	3.13 (0.81)	2.78 (0.98)	2.20	**.029**
RT (ms)	697 (182)	1051 (275)	-8.57	< **.001**
N-back WM load costs				
Verbal accuracy	.06 (.06)	.11 (.10)	-3.62	< **.001**
D prime	1.10 (0.86)	1.49 (0.99)	-2.40	**.018**
RT (ms)	171 (141)	332 (237)	-4.66	< **.001**
Spatial accuracy	.04 (.06)	.09 (.11)	-3.00	**.003**
D prime	0.65 (0.79)	0.94 (1.05)	-1.75	.082
RT (ms)	192 (127)	390 (218)	-6.25	< **.001**
Updating accuracy	.80 (.12)	.74 (.12)	2.06	**.043**
Reading Span (average correct words)	2.74 (0.46)	2.76 (0.42)	-0.33	.740
Color Stroop (interference ratio)	.28 (.13)	.33 (.13)	-2.26	**.026**
Letter comparison RT (sec)	61.20 (17.15)	80.85 (17.96)	-6.32	< **.001**
Pattern comparison RT(sec)	50.98 (12.76)	73.07 (18.78)	-7.76	< **.001**
SRT task (ms)	291 (46)	326 (48)	-4.26	< **.001**
Raven (total correct responses)	35.57 (5.74)	27.38 (7.07)	7.18	< **.001**
Mill Hill (total correct responses)	34.49 (4.88)	38.55 (3.47)	-5.44	< **.001**

Note. Accuracy scores indicate the proportion of correct responses. RT = response time. *N*-back WM load cost = Difference between 0-back and 2-back condition. df = 73 for the updating task; df = 122 for the Reading Span task; df = 125 for the Stroop task; df = 126 for the remaining tasks.

When consulting the norms for the Reading Span task based on a French sample of younger and older adults (Delaloye et al., 2008), younger and highly educated adults got a *z* score of -0.5 for the average of 2.75 correct words. The same average of correct words for older adults corresponds to a *z* score of 0.12 for highly educated adults between 65 and 69 years old (we only considered the standard scores for women, since 70 percent or more of the sample were female). The *z* scores indicated that both age groups were within the normal range (z score = +/- 1.65) and deviated less than a standard deviation from the mean (z score = +/- 1). However, the younger adults were slightly more deviant from the mean (*z* score = 0) than the older adults.

These results indicate that the sample showed the classic age difference patterns and can therefore be considered as a good representation of the population.

Training Gain

Verbal 2-back Task

In order to investigate age and training effects in the trained verbal 2-back task, a 2 (Age Group) x 3 (Training Group) two-way ANOVA on the difference between pre-test and post-test score, the so-called gain score or training gain, was conducted. The ANOVA on the gain score of the proportion of correct responses yielded a significant effect of age ($F(1, 121) = 5.64$, $p = .019$, $\eta^2_p = .05$) and training ($F(2, 121) = 10.66$, $p < .001$, $\eta^2_p = .15$), but no significant Age x Training interaction ($F(2, 121) = 1.13$, $p = .273$, $\eta^2_p = .02$). Pairwise comparisons revealed that younger adults (*M* = .03, *SD* = .06) had a significantly lower gain score than older adults (*M* = .06, *SD* = .08; see Appendix C, Table C1 and Table C3 for descriptive statistics of the raw data and gain scores, respectively). The average gain score was significantly higher for the WM training group (*M* = .08, *SD* = .09) than for the implicit training group (*M* = .02, *SD* = .05) and the no-contact control group (*M* = .02, *SD* = .06). The latter two however did not differ significantly (see Figure 9).

Further, gain scores for d’ measures were analyzed (see Appendix C, Tables C2 and C3for descriptive statistics). Results showed a significant effect of training (F(2, 121) = 18.79, p < .001, η2p = .24), but no significant age effect or interaction (F(1, 121) = 0.64, p = .425, η2p = .01 and F(2, 121) = 0.31, p = .733, η2p = .01, respectively). The WM training group (M = 1.17, SD = 0.9) had a significantly higher gain score than both control groups (M = 0.27, SD = 0.71 for implicit and M = 0.3, SD = 0.68 for no-contact group). In summary, these analyses revealed that the WM training group improved performance significantly in the trained task and that no differences between age groups were found concerning this improvement in the WM training group.

Analysis for correct reaction times did not yield any significant result.

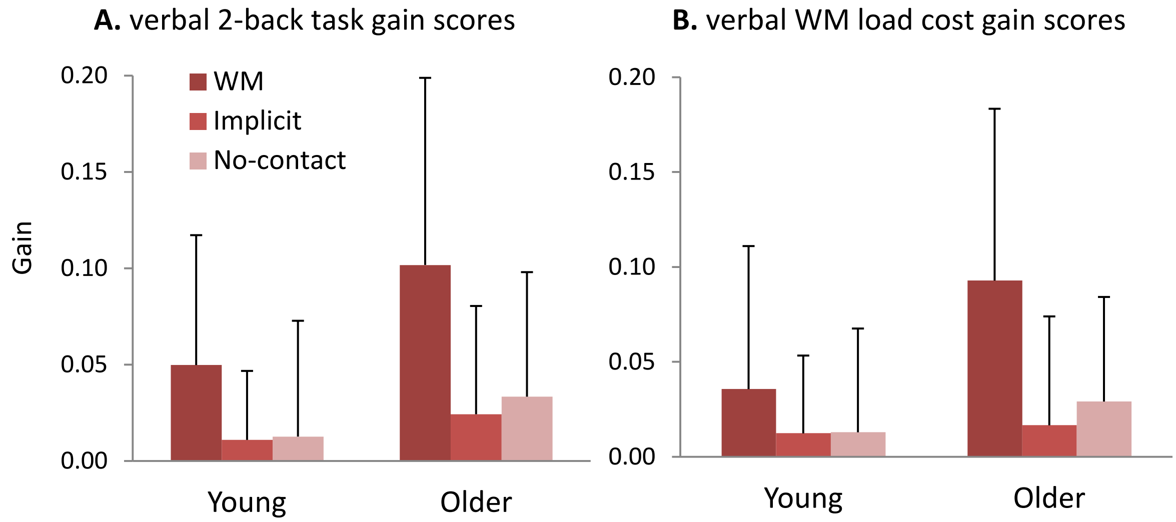

Figure 9. Verbal *N*-back task. A. Gain score means (post-test - pre-test) and standard deviations of the proportion of correct responses for the 2-back task and B. gain score means (pre-test - post-test) and standard deviations of the WM load costs, that is, the difference between 0-back minus 2-back task for the proportion of correct responses.

WM Load Costs

With the aim of analyzing the WM load costs, the difference between 0-back and 2-back accuracy was calculated. We expected that the WM load costs would decrease with training and calculated therefore the WM load costs gain score as pre-test minus post-test. The ANOVA for WM load cost gain scores on proportion of correct responses revealed a significant effect of age ($F(1, 121) = 4.62$, $p = .034$, $\eta^2_p = .04$) and training ($F(2, 121) = 7.6$, $p = .001$, $\eta^2_p = .11$) but no significant interaction ($F(2, 121) = 2.06$, $p = .132$, $\eta^2_p = .03$; see Figure 9 and Appendix C, Table C3 for descriptive statistics). Pairwise comparisons revealed that the gain score was significantly smaller for younger adults (*M* = .02, *SD* = .06) than for older adults (*M* = .05, *SD* = .08). They further showed that the WM load cost gain score was significantly larger for the WM training group (*M* = 0.06, *SD* = 0.09) than the other training groups (*M* = 0.02, *SD* = 0.05 for implicit and *M* = 0.02, *SD* = 0.06 for no-contact group). The analyses on *d*' gain scores showed a main effect of training only ($F(2, 121) = 3.22$, $p = .044$, $\eta^2_p = .05$). Pairwise comparisons indicated that the gain scores of the WM training group (*M* = 0.64, *SD* = 0.9) were significantly larger than both control groups (*M* = 0.25, *SD* = 0.75; *M* = 0.28, *SD* = 0.66 for implicit and no-contact group, respectively). WM load cost gain scores on correct response time did not reveal significant results. To sum up, these results revealed that WM load cost decreased significantly with training for the WM training group, which was true for both age groups. Furthermore, these results showed the same pattern as the previous

2-back analyses. That is, an effect of training was found for *d'* and the proportion of correct responses, but an effect of age was found for the proportion of correct responses only.

High versus Low Performers

To gain insight into the effect of individual difference in initial performance level, further analyses were conducted. A median split comparison of low versus high performers on the verbal 2-back task at pre-test was implemented (see Appendix C, Table C4 for descriptive statistics of the raw data). This comparison was added as the between-subjects factor performance level in the ANOVA. The analyses of proportion of correct responses gain scores revealed an effect of age ($F(1, 115) = 8.76$, $p = .004$, $\eta^2_p = .07$), training ($F(2, 115) = 13.74$, $p < .001$, $\eta^2_p = .219$) and performance level ($F(1, 115) = 28.58$, $p < .001$, $\eta^2_p = .20$). Pairwise comparisons indicated that gain scores were larger for the WM group compared to both control groups; larger for the older adults compared to the younger adults and larger for the low level participants compared to the high level participants. Furthermore, the Training x Performance Level interaction was significant ($F(2, 115) = 7.15$, $p = .001$, $\eta^2_p = .11$) and revealed that gain scores were significantly higher for the WM group as compared to the other training groups in low level participants only. Moreover, only in the WM group high and low level participants had significantly different gain scores. These results suggest that only low level participants, independent of their age group, improved significantly with a WM training procedure.

Cohen's d

Cohen's *d* effect sizes are displayed in Figure 10 for all dependent variables. They confirm the results reported above: For accuracy and *d'* the effect sizes were large to very large for the WM training groups. For accuracy data the effect sizes were even larger for older adults than for younger adults which confirmed the trends in pairwise comparisons for the interaction.

To sum up the training gain results, the older and younger WM training groups improved their performance in the trained 2-back task and showed decrease in the load costs to the same extent, compared to both control groups regarding proportion of correct responses and *d'* performance. However, additional analyses considering the initial level revealed that only low level participants improved their performance significantly. Furthermore, the constant lack of Age x Training interactions could be due to a potential ceiling effect bias. Whenever the task is too easy for one or both age groups at post-test or at both sessions, an interaction with age can potentially be suppressed. As a rule of thumb, a ceiling effect is assumed when the sum of mean and standard deviation of the score exceeds the maximum possible score (proportion of 1.0 in our case). This is true for almost all groups at least at post-test (see raw data scores Appendix C). In order to exclude this possible bias, accuracy data were reanalyzed in different exploratory ways; first, by excluding potential ceiling participants from the sample and then by a Tobit model which takes into account the censored nature of data (see Appendix D). The results of these additional analyses however did not show any change in results and confirmed the above reported findings. There was no significant Age x Training interaction but the initial level still had a significant effect on the gain score. Thus, these additional analyses suggested that results seemed not to be biased by ceiling effects and that gains were similar in both WM training age groups.

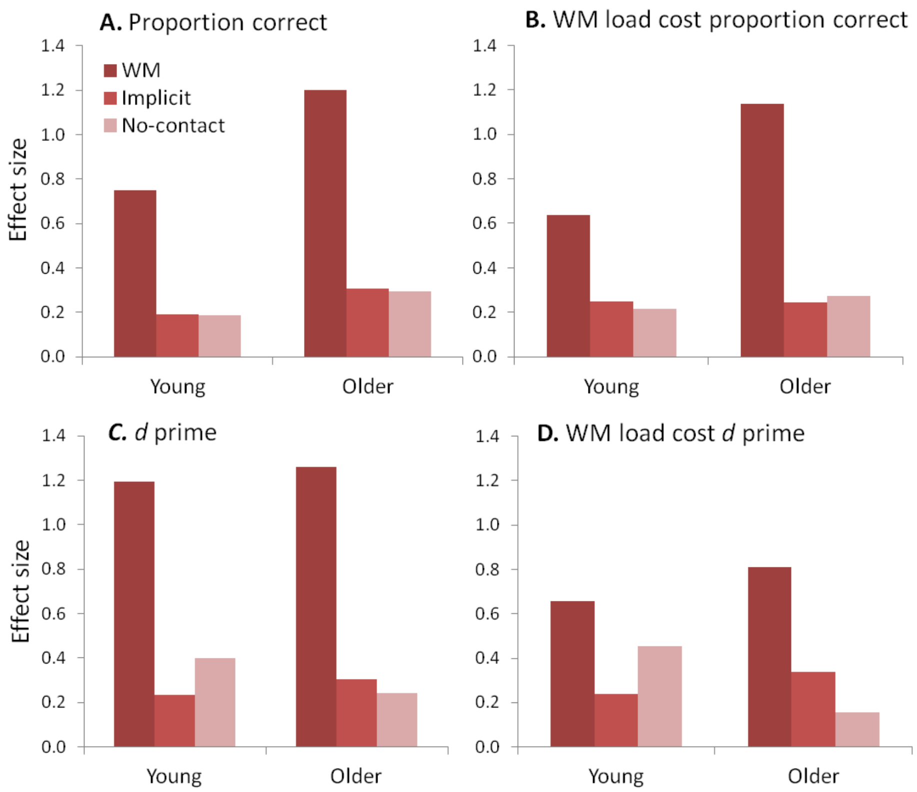

Figure 10. Verbal 2-back task. Effect sizes (Cohen's *d*) for A. the proportion of correct responses, B. the difference between 2-back and 0-back for correct responses (WM load cost), C. *d'* and D. WM load costs for *d'*.

Transfer Effects

In the following sections, the results of all transfer tasks that were conducted at pre-test and post-test are reported. For the spatial *N*-back task, the figures are showing gain scores and effects size results for different dependent variables (see Figures 11 and 12). For the remaining transfer tasks, one figure summarizing the effect sizes for all tasks is provided (see Figure 13; see Appendix E, Tables E1 and E2 for descriptive results).

Working Memory

Spatial N-back task

With the purpose of investigating age and training effects in the untrained spatial 2-back task, a 2 (Age Group) x 3 (Training Group) two-way ANOVA on the gain scores was conducted. The analysis on the gain score of the proportion of correct responses yielded a significant effect of age ($F(1, 121) = 8.78$, $p = .004$, $\eta^2_p = .07$) and training ($F(2, 121) = 3.28$, $p = .041$, $\eta^2_p = .05$), but no significant interaction Age x Training ($F(2, 121) = 1.42$, $p = .247$, $\eta^2_p = .02$). Younger adults (*M* = .004, *SD* = .06) had significantly lower gain scores than older adults (*M* = .05, *SD* = .09; see Appendix C, Tables C1 and C3 for descriptive statistics). Pairwise comparisons revealed that the WM training group had a marginal significantly higher average gain score (*M* = .05, *SE* = .09) than the implicit training group (*M* = .01, *SE* = .04) and the no-contact control group (*M* = .01, *SE* = .09; see Figure 11). The latter two groups did not differ significantly.

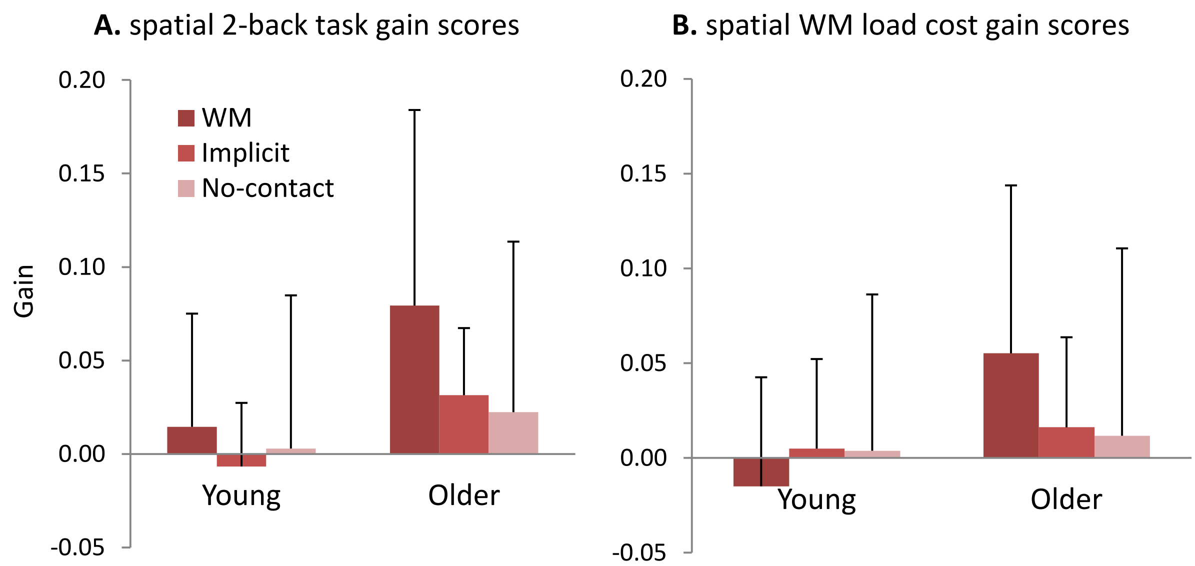

Figure 11. Spatial *N*-back task. A. Gain score means (post-test - pre-test) and standard deviations of the proportion of correct responses for the 2-back task and B. gain score means (pre-test - post-test) and standard deviations of the WM load costs, that is, the difference between 0-back minus 2-back task for the proportion of correct responses.

As concerns the *d'* gain score, there was only a significant main effect of age ($F(1, 121) = 6.03$, $p = .015$, $\eta^2_p = .05$) whereas the main effect of training ($F(2, 121) = 1.51$, $p =$

.224, $\eta^2_p = .03$) and the interaction ($F(2, 121) = 1.07$, $p = .346$, $\eta^2_p = .02$) were not significant. Older adults ($M = 0.46$, $SD = 0.7$) showed overall more gain for *d'* than younger adults ($M = 0.15$, $SD = 0.74$).

Regarding mean correct response times, the two-way ANOVA on gain score revealed no significant results similar to the verbal 2-back task.

WM load costs. The WM load cost gain score was also analyzed here. The results of the ANOVA conducted with WM load cost on proportion of correct responses revealed a significant effect of age ($F(1, 121) = 4.91$, $p = .029$, $\eta^2_p = .04$) and a marginally significant Age x Training interaction ($F(2, 121) = 2.44$, $p = .092$, $\eta^2_p = .04$) but no significant training effect ($F(2, 121) = 0.35$, $p = .709$, $\eta^2_p = .01$; see Appendix C, Table C3 for descriptive statistics). Pairwise comparisons revealed that younger adults ($M = -.002$, $SD = .06$) had a significantly lower WM load cost gain score than older adults ($M = .03$, $SD = .08$). They further showed that younger ($M = -.02$, $SD = 0.06$) and older adults ($M = 0.06$, $SD = 0.08$) were only significantly different from each other in the WM training group. However, the pairwise comparisons further showed that both WM training age groups were not significantly different from both control groups. Analyses with *d*' on the difference between 0-back and 2-back tasks showed a marginally significant main effect of training ($F(2, 121) = 2.39$, $p = .096$, $\eta^2_p = .04$). Pairwise comparisons however did not indicate significant differences between groups. No significant effects resulted for correct reaction times. These results confirm the previous analyses in that *d*' does not show any transfer effect but the proportion of correct responses did.

High versus Low Performers. Further, the initial performance level effect and its influence on the gain score were explored. The samples were divided via median split in low and high level participants with respect to the spatial 2-back performance at pre-test. The analyses of proportion of correct responses gain scores revealed an effect of age ($F(1, 115) = 10.57$, $p = .002$, $\eta^2_p = .08$), training ($F(2, 115) = 3.74$, $p = .027$, $\eta^2_p = .06$) and performance level ($F(1, 115) = 11.79$, $p = .001$, $\eta^2_p = .09$). Pairwise comparisons indicated that the WM group, the older and the low level participants had significantly higher gain scores. Further, the interaction Age x Performance Level ($F(1, 115) = 5.38$, $p = .022$, $\eta^2_p = .05$) and the triple interaction Age x Performance Level x Training ($F(2, 115) = 2.36$, $p = .099$, $\eta^2_p = .04$) were (marginally) significant. Pairwise comparisons indicated that only the older WM training group of the low level participants was significantly different from both control groups. These results suggest that only older low level participants improved significantly with a WM training procedure. The same trend was found when the analyses were conducted for WM load costs and *d* prime gain scores.

Cohen's *d*. Furthermore, effect sizes for the different variables were calculated (see Figure 12). They were all large to very large for older adults with the exception of WM load costs. For younger adults however, they were small or even negative for WM load costs.

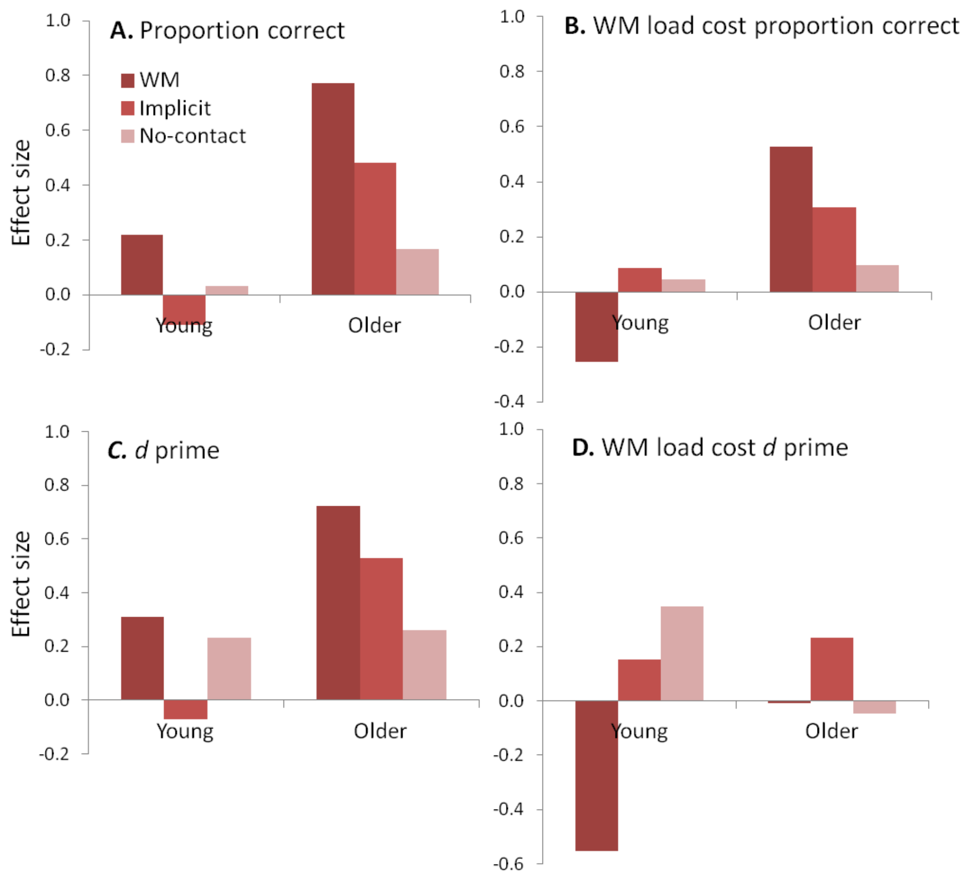

Figure 12. Spatial 2-back task. Effect sizes (Cohen's *d*) for A. the proportion of correct responses, B. the difference between 2-back and 0-back for correct responses (WM load cost), C. *d'* and D. WM load costs for *d'*.

To sum up, as seen for the trained verbal 2-back task, the older and younger WM training groups improved performance in the spatial transfer task to the same extent compared to both control groups. However, there was an age-difference trend indicating more gain in older adults after WM training, which was confirmed by taking into account the initial performance level. No transfer effect was observed for *d'*.

As for the verbal task, the same procedure to address the ceiling effects in the spatial 2-back task was applied (see Appendix D). However, the results of the additional analyses did not show any change in results and confirmed the above reported findings: There was no significant Age x Training interaction but the initial level still had a significant effect on the gain score.

Updating task

The sample for the Updating task contained only the participants of the EEG subsample (see below), it is therefore smaller with a total *N* = 75. As dependent variables the gain scores of the average number of correct responses per series and the proportion of correct series were used. The ANOVA revealed no significant results for the first score, therefore only the results from the latter one are reported: there was a marginally significant age effect ($F(1, 68) = 3.41$, $p = .069$, $\eta^2_p = .05$), but no significant training effect ($F(2, 68) = 0.31$, $p = .733$, $\eta^2_p = .01$) or interaction ($F(2, 68) = 1.62$, $p = .205$, $\eta^2_p = .05$) for the proportion of correct recalled series. Pairwise comparisons revealed that younger ($M = .07$, $SD = .19$) showed overall more gain than older adults, who even demonstrated a slightly negative change ($M = -.01$, $SD = .18$).

Reading Span task

The sample was *N* = 124 (out of *N* = 128) for the Reading Span task after exclusion of participants who did not reach at least 85% correct responses in the semantic judgment task at pre-test and/or post-test ($n = 4$). Like for the updating task, the dependent variables were the mean number of correct words recalled per series and the proportion of correct series recalled. Both gains scores of the dependent variables were analyzed; however, the analyses revealed no significant results for any of the gain scores. Only the results from the mean number of correct words are reported. The ANOVA revealed no significant differences between age groups ($F(1, 117) = 1.83$, $p = .179$, $\eta^2_p = .02$) or training groups ($F(2, 117) = 0.71$, $p = .493$, $\eta^2_p = .01$) and no significant interaction ($F(2, 117) = 0.05$, $p = .95$, $\eta^2_p = .00$) in gain score. Controlling for mother tongue by adding it as a covariate did not change the results.

Fluid Intelligence

Raven's Progressive Matrices

The dependent variable was the total number of correct responses in the Raven's Progressive Matrices. The ANOVA revealed no significant effects on the gain score; no significant effect of age ($F(1, 121) = 0.37$, $p = .543$, $\eta^2_p = .00$) or training ($F(2, 121) = 0.30$, $p = .739$, $\eta^2_p = .01$) nor a significant interaction ($F(2, 121) = 0.23$, $p = .797$, $\eta^2_p = .00$) was found.

Inhibition

Color Stroop task

The total sample for the Color Stroop task was N = 127, as one older participant had to be excluded due to daltonism. The analyses of the gain scores revealed for one particular score, the interference ratio, a marginally significant result. The effect of training reached marginal significance ($F(2, 120) = 2.56$, $p = .082$, $\eta^2_p = .04$), but neither the effect of age ($F(1, 120) = 0.25$, $p = .617$, $\eta^2_p = .00$) nor the interaction did ($F(2, 120) = 0.11$, $p = .894$, $\eta^2_p = .00$). Pairwise comparisons showed that the WM training group (M = 0.06, SD = 0.10) had a significantly larger gain score, and did therefore show a larger decrease in interference ratio, than the no-contact control group (M = 0.01, SD = 0.12). However, the implicit training group (M = 0.04, SD = 0.08) had a gain score value between the other training groups and did therefore not show any significant difference to either of the groups. These results are reflected by the effect size values (see Figure 13) where the WM training group as well as the implicit training group reached nearly medium effect sizes, at least for younger adults.

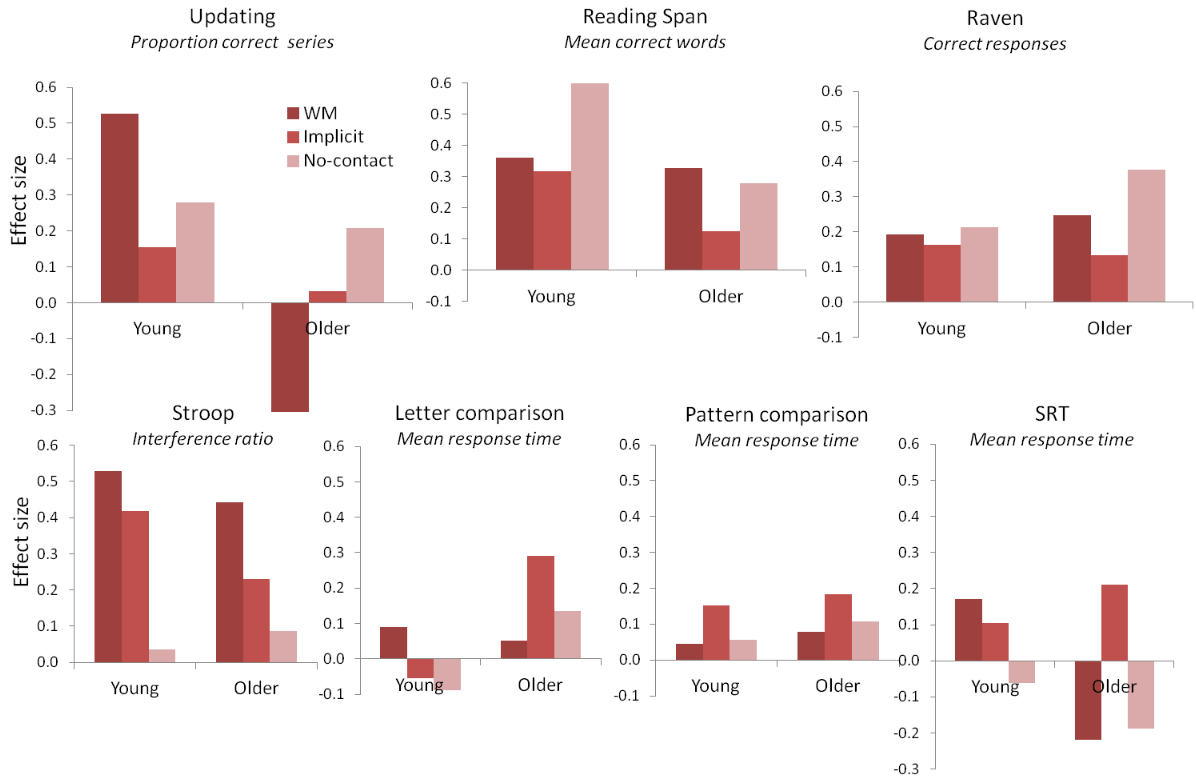

Figure 13. Effect sizes (Cohen's *d*) for pre-test and post-test tasks.

Processing Speed

Letter and pattern comparison tasks

For these two tasks as well, no significant effects of training on gain scores were found. Here, only the findings for mean reaction times are reported. For the letter comparison task, the age effect (F(1, 121) = 2.16, p = .145, η^2_p = .02), training effect (F(2, 121) = 0.35, p = .709, η^2_p = .01) and interaction (F(2, 121) = 1.18, p = .310, η^2_p = .02) were not significant. As concerns the pattern comparison task, the effects of age (F(1, 121) = 0.34, p = .559, η^2_p = .00), training (F(2, 121) = 0.23, p = .797, η^2_p = .00) and the interaction (F(2, 121) = 0.02, p = .979, η^2_p = .00) were not significant.

Simple reaction time task

The mean reaction time and the intra-individual standard deviation of the Simple Reaction Time task did not significantly change from pre-test to post-test. The effect of age (F(1, 121) = 0.90, p = .345, η^2_p = .01) and training (F(2, 121) = 1.29, p = .280, η^2_p = .02) as well as the interaction (F(2, 121) = 1.01, p = .368, η^2_p = .02) for the mean reaction time were not significant.

Linking Training Performance, Training Gain and Transfer Effects

Finally, gain scores were linked with training improvements for the WM training group in one analysis. The correlations were conducted in the previously reported LGCM. The raw scores of the verbal and spatial 2-back tasks and the gain scores for all transfer tasks were correlated with intercept and slope values. The correlations were examined in three different models: first, in a model without controlling for any predictor, second, in a model controlling intercept and slope for the predictor age group, and third, in a model controlling intercept and slope for the predictors age group and fluid intelligence (gf) together.

The results revealed several interesting patterns (see Table 9 for a summary). As concerns the correlations of the raw scores for the 2-back tasks, they were all positively correlated with intercept, but only the verbal pre-test score was significantly correlated with slope. When controlling for age group the correlation with slope became marginally significant for the verbal post-test score as well. However, when controlling for age together with gf, the correlations with slope were no longer significant, whereas the ones with intercept remained stable. This indicates that the correlations between slope and raw performances were to a great amount mediated by gf performance. In sum, these correlations show that the initial level of training, but not the training improvement, was relied to 2-back updating performance at pre-test and/or post-test.

Table 9. Correlations between WM training intercept and slope and gain scores for the non-conditional model, the conditional model with age group and the conditional model with age group and Raven score (gf).

	Non cond. model		Cond. model *age*		Cond. model *age and gf*	
	Icept	Slope	Icept	Slope	Icept	Slope
Raw Scores 2-back task						
Verbal pre-test	.67**	.41*	.63**	.31+	.36*	.18
post-test	.33*	.23	.43*	.30+	.41*	.27
Spatial pre-test	.57**	.27	.47**	.10	.16	-.05
post-test	.52**	.23	.62**	.28	.53**	.20
Gain Scores						
N-back tasks						
Verbal 2-back	-.57**	-.33+	-.47**	-.18	-.18	-.04
Verbal WM load cost	-.55**	-.44*	-.44*	-.31+	-.19	-.19
Spatial 2-back	-.28+	-.14	-.09	.09	.21	.21
Spatial WM load cost	-.31+	-.09	-.08	.20	.10	.29+
Updating	.44*	.71**	.12	.45*	.04	.42*
Reading Span	.00	.17	-.07	.12	-.14	.11
Raven	-.09	.12	-.07	.19	.20	.30+
Stroop Interference Ratio	-.14	-.04	-.21	-.09	-.23	-.09
Letter comparison RT	.39*	.11	.44*	.10	.40*	.05
Pattern comparison RT	-.11	.02	-.12	.04	.03	.11
SRT task	.03	.37*	-.11	.30+	-.07	.33+

Note. Con. = Conditional. Icept = Intercept. RT = response time. SRT = simple reaction time. WM load cost = difference 2-back - 0-back. ** $p < .01$. * $p < .05$. + $p < .10$. $n = 44$; $n = 24$ for updating task.

A different pattern was observed as concerns gain scores. First, regarding the non conditional model, it is notable that all *N*-back gain scores were negatively correlated with intercept and/or slope, whereas the remaining transfer task gain scores, that showed a significant correlation, were positively correlated with intercept and/or slope. This means that a high gain score in the *N*-back task was correlated with a small initial training level and a flat growth curve. However, when correlations were examined in the third model, which was controlling for age and fluid intelligence together, the correlation with spatial *N*-back gain scores became positive. The correlation between WM load cost (difference between 2-back and 0-back condition) and slope even reached marginal significance. This means that the gradient of the growth curve was positively associated with gain score beyond the effect of age and fluid intelligence. This could be interpreted as a transfer effect, which implies that those who improved more during training showed more gain in a near transfer task.

As regards the correlations with the verbal *N*-back gain scores in the third model, they were no longer significant but still negative. This could be due to the ceiling effect which did not allow the high performer to improve from pre-test to post-test as they already showed a high pre-test score. The high performers had therefore a low gain score but probably a high initial training level and a steep growth curve. This bias was in part confirmed when considering the raw 2-back scores, where high performance at pre-test and post-test was associated with a high initial training level.

In terms of the remaining transfer tasks, the Updating task, the letter comparison task and the simple reaction time task (SRT) gain scores showed a significant positive correlation with intercept and/or slope in the first model (see Table 9). These correlations were still present, although attenuated for Updating task and SRT, in the second and the third model with control for age and age together with gf respectively. The correlation with intercept revealed that the letter comparison gain score was high for participants who showed a high initial training level. The correlation with slope indicated that the participants, who showed a steep training curve, also showed a high gain score in the Updating and SRT task, beyond the effect of age and fluid intelligence. Furthermore, there was a marginally significant positive correlation between Raven task and slope in the third model, which was not observed in the former two models. This means that when controlling for age and fluid intelligence effects, participants who reached a steep training curve also attained a higher gain score in the Raven task. This indicates that the association between Raven gain score and the growth curve during training was masked by the effects of age and the fluid intelligence baseline performance.

Overall, these results showed that, when taking into account the effect of age group and fluid intelligence level, the gradient of improvement over the ten training sessions was associated with several gain scores. Concretely, there was a positive association with the gain scores in spatial 2-back WM load cost, Updating task, Raven task and SRT task. These findings were more or less in line with the hypothesized transfer effects. The hypotheses predicted that there would be a significant transfer effect in tasks engaging updating processes and in the fluid intelligence task. The association of the training improvement with the gain scores for the Updating task and the spatial *N*-back WM load cost as well as with the Raven gain score were therefore in line with the hypotheses. These effects were probably masked by inter-individual differences in fluid intelligence and by age effect. However, the finding that speed gain score, but not Reading Span, would be influenced by training was not in line with the hypotheses.

Summary

To sum up the findings, the results of the training procedures revealed that participants improved performance over the ten training sessions in both training tasks. However, younger adults had a higher initial training level than older adults. Moreover, younger adults improved faster during training and reached the asymptote earlier than older adults. This was true for both training tasks.

Further, hierarchical regression analyses revealed that in the first three sessions of the WM training, age did not or only marginally explain more of the variance in training data than the cognitive abilities. However, with advancing training, age did explain a significant part of the training performance variance beyond the cognitive variables. This indicates that in the first part of the training, individual differences in cognitive abilities explained the training performance. The assessed cognitive abilities already accounted for all age differences in training performance. With a growing number of training sessions however, age explained an additional part of variance in training performance beyond the assessed cognitive variables. These findings were confirmed by the LGCM analysis. It showed that fluid intelligence was a significant predictor of the initial training performance, but not of the training improvement. Age group, in contrast, was a significant predictor of the improvement during training, but not of the initial training performance. Here again, fluid intelligence variability accounted for differences in initial training performance independent of age effects, whereas the age group accounted for individual differences in growth curve.

As regards the implicit training, age also accounted for an additional part of variance beyond the cognitive variables in the training performance. This did not change in course of the training and was observed for all training sessions. However, it is possible that age was still confounded with implicit memory performance after controlling for the cognitive variables, since no implicit memory task has been included.

As regards the training gain in the verbal *N*-back task, analyses showed a clear effect of training. Younger and older WM training groups showed a significantly larger gain score for accuracy in the trained task than both control groups. Similar results were found for the spatial *N*-back task, the near transfer task, even though the effects were less pronounced as no significant gains were found for *d'*.

However, no significant interaction with age was found for both *N*-back tasks; neither for the training gain nor for the near transfer task. This indicates that age groups improved to a similar extent from pre-test to post-test. Nevertheless, data contained a potential ceiling effect bias, for which additional analyses were conducted. These additional analyses did not confirm a ceiling effect since the results did not change when taking it into account.

Differences between the low and high performers in the *N*-back tasks were assessed by dividing the groups via median split according to pre-test performance into low and high performer groups. These analyses revealed that only the low performers showed a higher gain score than both control groups. In the spatial *N*-back task, results even indicated that only the older low performers showed higher gain scores than both control groups.

Furthermore, there was a marginally significant effect of training in the Stroop task. The training groups showed a significantly larger gain score than the no-contact control group, but did not differ from the implicit training group. For all the other pre-test and post-test tasks, analyses did not reveal any significant effect.

As regards the Cohen's *d* effect sizes for the verbal *N*-back task, they were large (d = 0.8) to very large (up to d = 1.2) for younger and older WM training groups, while they were in general slightly larger for older adults. The effect sizes for the spatial *N*-back task in older adults were medium (d = 0.5) to large for the WM training group with the exception of *d'* WM load costs, for which the effect size was zero. However, those for the implicit control group reached up to medium values too, which led to small effect size differences between WM training and implicit training group. The effect sizes for the younger WM training group in the spatial *N*-back task were small (d = 0.2) or even negative for the WM load costs.

As concerns the remaining transfer tasks, effect sizes were in general rather small for all groups, especially in the Raven task and in the speed tasks. Regarding the Updating task, the younger WM training group reached medium effect size whereas the older WM training group showed a negative score. For the Reading Span and the Stroop tasks, both WM training groups showed nearly medium effect sizes. However, the control groups also reached considerable effect sizes, which led to small effect size differences between training conditions.

In summary, these effect sizes showed that training effects were largest for the trained task. The effect sizes for the WM training group were considerable in the spatial *N*-back task for older adults and in the Updating task for younger adults, that is, in two near transfer tasks.

Finally, the training intercept and slope values for the WM training group were linked to all gain scores. Analyses revealed that there was a positive relationship between training improvement and gain scores after controlling for age effects and fluid intelligence differences. This emerging relationship was found for spatial WM load costs, the Updating task, the Raven task and the SRT task. A transfer effect had been hypothesized for all these tasks, except for the SRT task.

ELECTROPHYSIOLOGICAL RESULTS

In the following section, we report the results of the EEG data derived at pre-test and post-test. The EEG subsample is smaller than the sample analyzed in the previous section with a total $N = 72$ (for the sample description see section *Electrophysiological Method*).

First, we examined the training gain and transfer effects at the behavioral level for the EEG subsample. This allowed us to conclude whether the EEG subsample showed in general the same trends in training effects as the whole sample. Then, we analyzed the classic ERP data. Finally, we conducted the microstate segmentation analyses.Within each type of analysis, first, we report age differences at pre-test, then, training gains on the verbal *N*-back tasks and finally, transfer effects to the spatial *N*-back task.

Behavioral Results for the EEG Subsample

As for the whole sample, we conducted for the EEG subsample a 2 (Age Group) x 3 (Training Group) two-way ANOVA with the gain scores for all cognitive tasks. However, we did not include the covariate Education, as there was no difference between groups with respect to this or any other independent variable.

Verbal N-back Task

The ANOVA on the 2-back gain score of the proportion of correct responses yielded a marginally significant main effect of training ($F(2, 66) = 3.09$, $p = .052$, $\eta^2_p = .09$), but no significant main effect of age ($F(1, 66) = 0.56$, $p =.457$, $\eta^2_p = .01$) nor an interaction ($F(2, 66) = 1.01$, $p = .367$, $\eta^2_p = .03$). Pairwise comparisons revealed that the average gain score was significantly higher for the WM training group ($M = .05$, $SD = .07$) than for the implicit training group ($M = .01$, $SD = .03$) but did not differ significantly from the score of the no-contact control group ($M = .03$, $SD = .05$).

Results for *d'* gain scores showed a significant effect of training ($F(2, 66) = 8.84$, $p < .001$, $\eta^2_p = .21$), but no significant age effect or interaction ($F(1, 66) = 0.01$, $p = .934$, $\eta^2_p = .00$ and $F(2, 66) = 0.75$, $p = .474$, $\eta^2_p = .02$, respectively). The WM training group ($M = 1.01$, $SD = 1.0$) had a significantly higher gain score than the implicit control ($M = 0.13$, $SD = 0.55$) and a marginally significant higher gain score than the no-contact group ($M = 0.56$, $SD = 0.52$).

Analysis for correct reaction times did not yield any significant result as obtained for the entire sample as well.

As regards the WM load costs, the analyses also revealed a similar pattern in results as the analyses for the whole sample. The ANOVA for WM load cost gain scores on proportion of correct responses revealed a marginally significant interaction ($F(2, 66) = 2.87$, $p = .064$, $\eta^2_p = .08$), but no significant main effect of age ($F(1, 66) = 0.31$, $p = .579$, $\eta^2_p = .01$) and training ($F(2, 66) = 2.3$, $p = .114$, $\eta^2_p = .06$). Pairwise comparisons revealed that older adults with WM training (*M* = .06, *SD* = .07) had a significantly larger WM load cost difference than younger WM training adults (*M* = .01, *SD* = .07). They further showed that only in older adults the WM training group showed a significantly larger WM load cost gain than both control groups (*M* = 0.00, *SD* = 0.04 for the implicit group; *M* = 0.01, *SD* = 0.06 for the no-contact group).

The analyses on *d*’ WM load cost gain scores showed a significant main effect of training ($F(2, 66) = 4.09$, $p = .021$, $\eta^2_p = .11$) but no significant age effect or interaction ($F(1, 66) = 0.28$, $p = .600$, $\eta^2_p = .00$ and $F(2, 66) = 1.52$, $p = .226$, $\eta^2_p = .04$, respectively). Pairwise comparisons indicated that the WM training group (*M* = 0.55, *SD* = 0.9) showed a significantly higher WM load cost difference than the implicit training group (*M* = 0.00, *SD* = 0.53) and a similar gain score like the no-contact control group (*M* = 0.43, *SD* = 0.61).

WM load cost gain scores on correct response times did also reveal a significant main effect of training ($F(2, 66) = 3.95$, $p = .024$, $\eta^2_p = .11$) but no significant age effect or interaction ($F(1, 66) = 0.07$, $p = .795$, $\eta^2_p = .00$ and $F(2, 66) = 0.88$, $p = .420$, $\eta^2_p = .03$, respectively). Pairwise comparisons indicated that the WM training group (*M* = 107, *SD* = 189) showed a significantly higher WM load cost difference than both control groups (*M* = 25, *SD* = 58 for the implicit group; *M* = 12, *SD* = 87 for the no-contact group).

To sum up, the EEG subsample showed similar training gain effects compared to the whole sample with some slight exceptions. The differences were not always significant to both control groups, but at least to one. This could be due to the reduced *N* in the subsample. However, in general, the same results for training gains can be assumed.

Spatial N-back Task

The ANOVA on the spatial 2-back gain score of the proportion of correct responses yielded a significant main effect of age ($F(1, 66) = 4.97$, $p = .029$, $\eta^2_p = .07$), but no significant main effect of training ($F(2, 66) = 1.93$, $p =.153$, $\eta^2_p = .06$) nor an interaction ($F(2, 66) = 1.58$, $p = .213$, $\eta^2_p = .05$). Pairwise comparisons revealed that the average gain score was significantly higher for older adults (*M* = .03, *SD* = .05) than for younger adults (*M* = .01, *SD* = .03). When consulting the pairwise comparisons of the non significant interaction, they indicated that there was a trend towards a marginally significantly larger gain score in the WM training group than in the control groups, but only for older adults.

Results for *d'* and correct response time gain scores did not show significant effects.

The ANOVA for WM load cost gain score on proportion of correct responses revealed a marginally significant interaction ($F(2, 66) = 3.19$, $p = .047$, $\eta^2_p = .09$), but no significant main effect of age ($F(1, 66) = 1.20$, $p = .278$, $\eta^2_p = .02$) and training ($F(2, 66) = 0.00$, $p = .998$, $\eta^2_p = .00$). Pairwise comparisons revealed that older adults with WM training (*M* = .03, *SD* = .04) had a significantly larger WM load cost difference than younger WM training adults (*M* = -.02, *SD* = .06). WM load cost gain scores on *d*' and correct response times did not reveal significant effects.

To sum up, the EEG subsample showed less near transfer effect than the entire sample. However, in general there was a trend towards the same results but it did not reach significance, probably due to the reduced *N* in the subsample.

Other Transfer Effects

No transfer effects were observed for the remaining cognitive tasks in the EEG subsample. This is consistent with the findings for the whole sample.

Classic Event-Related Potentials Analyses

Age Differences in ERP Measures

With regards to age differences in the amplitude and latency measures, we observed the classic age-related pattern (Fabiani et al., 1998; Fabiani & Gratton, 2005; Ford et al., 1979; Friedman, 2008; Polich, 1996; Rossini et al., 2007). First, younger adults showed in general shorter latencies for both the N2 and the P3 component in all task conditions as compared to older adults (see Table 10). Similar were the results for the amplitude measure: Younger adults exhibited larger negative amplitudes for the N2 component at the central frontal ROI and larger positive amplitudes for the P3 component at the central parietal ROI in all task conditions compared to older adults.

Table 10. *T*-tests for the difference between younger and older adults on the ERP component measures at pre-test.

	Younger *M (SD)*	Older *M (SD)*	*t*	*p*
0-back tasks				
Verbal N2 amplitude	-2.65 (2.82)	-0.19 (2.02)	-4.25	< **.001**
N2 latency	242 (38)	280 (35)	-4.43	< **.001**
P3 amplitude	5.76 (3.57)	1.56 (2.2)	6.01	< **.001**
P3 latency	382 (51)	502 (75)	-7.91	< **.001**
Spatial N2 amplitude	-2.36 (2.64)	-0.9 (2.05)	-2.62	**.011**
N2 latency	247 (41)	273 (39)	-2.60	**.011**
P3 amplitude	4.83 (3.26)	2.15 (1.83)	4.30	< **.001**
P3 latency	375 (65)	458 (88)	-4.48	< **.001**
2-back tasks				
Verbal N2 amplitude	-1.4 (2.35)	-0.19 (1.62)	-2.54	**.013**
N2 latency	236 (40)	275 (38)	-4.25	< **.001**
P3 amplitude	4.56 (3.04)	1.04 (1.51)	6.24	< **.001**
P3 latency	407 (91)	448 (106)	-1.79	.078
Spatial N2 amplitude	-1.18 (2.14)	-0.04 (1.41)	-2.67	**.009**
N2 latency	245 (42)	283 (51)	-3.48	**.001**
P3 amplitude	4.13 (2.55)	1.88 (1.71)	4.40	< **.001**
P3 latency	406 (82)	472 (93)	-3.22	**.002**

Note. Amplitude measures in µV; latency measures in ms. The N2 amplitude measure reflects the mean amplitude for the central frontal ROI; the P3 amplitude measure reflects the mean amplitude for the central parietal ROI. df = 70.

Verbal N-back Task

The N2 component

The repeated ANOVA on mean amplitude gain scores for the verbal 2-back task revealed a significant effect of age group, training group, X-axis ROI (laterality) and Y-axis (anterior-posterior) ROI. Further, the interactions X-axis x Y-axis, X-axis x Age x Training Group, and Training Group x X-axis x Y-axis reached significance (see Figure 14, 15, and 16 for ERP of the WM training group, the implicit training group, and the no-contact control group, black arrows indicate significant N2 changes compared to both control groups; see Table 11 for statistics and Appendix F, Table F1 for descriptive values).

Regarding the main effects, pairwise comparisons revealed that in general younger adults marginally and the WM training group clearly showed significantly higher gain scores than older adults and both control groups, respectively. Concerning the ROIs, the central ROIs on the X-axis showed a higher gain score than the right ROIs and the parietal Y-axis ROIs showed significantly smaller gain scores than the central and frontal ROIs.

Pairwise comparisons of the significant interactions showed that gain scores for the WM training group are significantly larger on the left-central and central-central ROI compared to both control groups. The N2 became significantly larger (more negative) on left-central and central-central sites compared to the central-frontal and left-frontal sites. Those central sites exhibited also a larger gain score compared to parietal sites, however, the control groups also showed this kind of difference. Regarding the interaction with age group, younger adults showed larger gain scores for the ROIs on the right side whereas older adults showed larger gain scores for the ROIs on the left side compared to both control groups.

Table 11. N2 component: Results of the repeated measures ANOVA Age x Training Group x X-axis x Y-axis for verbal 2-back amplitude gain scores and the WM load cost gain scores (0-back minus 2-back gain scores).

	Verbal 2-back			WM load cost		
Source of Variance	*df*	*F*	η^2_p	*df*	*F*	η^2_p
	Between subjects					
Age (A)	1	3.78 +	.06	1	0.12	.00
Training (T)	2	6.64 **	.17	2	4.38 **	.12
A x T	2	0.01	.00	2	0.03	.00
Error	66			66		
	Within subjects					
X-axis	2 (1.6)	3.41 *	.05	2 (1.8)	3.88 *	.06
Y-axis	2 (1.3)	9.15 **	.12	2 (1.5)	1.66	.02
X-axis x A	2 (1.6)	1.41	.02	2 (1.8)	1.6	.02
X-axis x T	4 (3.2)	0.44	.01	4 (3.6)	1.18	.03
X-axis x Y-axis	4	3.44 *	.18	4	2.40 +	.13
Y-axis x A	2 (1.3)	0.78	.01	2 (1.4)	0.17	.00
Y-axis x T	4 (2.6)	1.19	.32	4 (2.8)	1.15	.03
X-axis x A x T	4 (3.2)	4.11 **	.11	4 (3.6)	5.00 **	.13
Y-axis x A x T	4 (2.6)	0.57	.02	4 (2.8)	0.15	.00
X-axis x Y-axis x A	4	1.37	.08	4	1.35	.08
X-axis x Y-axis x T	8	2.13 *	.12	8	1.48	.09
X-axis x Y-axis x A x T	8	1.55	.09	8	1.00	.06
Error (X-axis)	132 (105.8)			132 (119.6)		
Error (Y-axis)	132 (85.1)			132 (91.6)		
Error (X-axis x Y-axis)	128			128		

Note. *df* in parentheses represent Greenhouse-Geisser corrected values. *** $p < .001$. ** $p < .01$. * $p < .05$. + $p < .10$.

Concerning the latency measure taken at the local peak on the Cz electrode, no significant effects were observed.

In order to investigate the WM load cost gain score, the gain score for the difference between verbal 0-back and 2-back task measures were calculated. The repeated ANOVA on mean amplitude gain scores revealed a significant effect of training group and X-axis ROI. Further, the interactions X-axis x Y-axis and X-axis x Age x Training Group reached significance (see Table 11 for statistical results and Appendix F, Figures 1, 2, and 3 for ERP waveforms of the 0-back task). Regarding the main effects, pairwise comparisons revealed that in general the WM training group showed significantly higher gain scores than both control groups and the central ROIs on the X-axis showed a higher gain score than the right ROIs. Regarding the interaction with age, pairwise comparisons of the significant interactions revealed that gain scores for the younger WM training group were significantly larger for the ROIs on the right side whereas older adults showed larger gain scores for the ROIs on the left and the central side compared to both control groups.

As regards the differences on latency gain score measures taken at the local peak on the Cz electrode, a significant age effect ($F(1, 66) = 4.92$, $p = .03$, $\eta^2_p = .07$) appeared. Younger adults showed over all higher gain scores than older adults.

The P3 component

The repeated ANOVA on mean amplitude gain scores for the verbal 2-back task revealed a significant effect of age group, training group, X-axis ROI and Y-axis ROI. Further, the interactions X-axis x Y-axis and Age x Training Group x X-axis reached significance (see Figure 14, 15, and 16 for ERP waveforms, white arrows indicate significant P3 changes compared to the control groups; see Appendix F, Table F2 for descriptive values and Table 12 for statistical results). Regarding the main effects, pairwise comparisons revealed that in general younger adults and the WM training group showed significantly larger negative gain scores compared to the respective other groups. That is, the amplitude decreased more with training in these groups compared to the other groups. Pairwise comparisons of the significant interaction with age group showed that younger adults had significantly decreased P3 amplitude on the right side ROIs whereas older adults decreased significantly on P3 at the ROIs on the left side compared to both control groups.

Regarding the latency gain score measures taken at the local peak on the Pz electrode, a significant age effect ($F(1, 66) = 4.6$, $p = .036$, $\eta^2_p = .07$) was observed. Younger adults showed over all higher gain scores than older adults.

Regarding the latency gain score measures taken at the local peak on the Pz electrode, a significant age effect (F(1, 66) = 4.6, p = .036, η^2_p = .07) was observed. Younger adults showed over all higher gain scores than older adults.

Table 12. P3 component: Results of the repeated measures ANOVA Age x Training Group x X-axis x Y-axis for verbal 2-back amplitude gain scores and the WM load cost gain scores (0-back minus 2-back gain scores).

	Verbal 2-back			WM load cost		
Source of Variance	*df*	*F*	η^2_p	*df*	*F*	η^2_p
	Between subjects					
Age (A)	1	6.14 **	.09	1	0.01	.00
Training (T)	2	5.72 **	.15	2	1.04	.03
A x T	2	0.68	.02	2	0.95	.03
Error	66			66		
	Within subjects					
X-axis	2 (1.8)	2.69 +	.04	2 (1.9)	2.55 +	.04
Y-axis	2 (1.5)	2.51	.04	2 (1.6)	0.16	.00
X-axis x A	2 (1.8)	0.58	.01	2 (1.9)	2.47 +	.04
X-axis x T	4 (3.6)	1.00	.03	4 (3.8)	0.66	.02
X-axis x Y-axis	4	3.00 *	.16	4	1.51	.09
Y-axis x A	2 (1.5)	0.19	.00	2 (1.6)	0.21	.00
Y-axis x T	4 (3.0)	0.73	.02	4 (3.2)	0.48	.01
X-axis x A x T	4 (3.6)	2.50 *	.07	4 (3.8)	2.22 +	.06
Y-axis x A x T	4 (3.0)	0.93	.03	4 (3.2)	1.14	.03
X-axis x Y-axis x A	4	1.01	.06	4	0.88	.05
X-axis x Y-axis x T	8	1.46	.08	8	0.78	.05
X-axis x Y-axis x A x T	8	0.45	.03	8	1.40	.08
Error (X-axis)	132 (118)			132 (126.8)		
Error (Y-axis)	132 (97.4)			132 (106.8)		
Error (X-axis x Y-axis)	128			128		

Note. *df* in parentheses represent Greenhouse-Geisser corrected values. *** $p < .001$. ** $p < .01$. * $p < .05$. + $p < .10$.

The repeated ANOVA on mean amplitude for the WM load cost gain scores revealed marginally significant effects of X-axis, X-axis x Age, and X-axis x Age x Training (see Table 12 for statistical). Regarding the main effect X-axis, the central ROIs showed significantly larger gain scores than the left ROIs. Pairwise comparisons for the X-axis x Age interaction revealed that this effect was larger for older adults than for younger adults. For the

X-axis x Age x Training interaction, pairwise comparisons showed that age differences were significant in the WM training group, but not in both control groups. Younger adults showed more gain on right ROIs whereas older adults showed more gain on central ROIs.

As regards latency gain score measures taken at the local peak on the Pz electrode, a significant age effect ($F(1, 66) = 12.99$, $p = .001$, $\eta^2_p = .16$) and interaction Age x Training group ($F(2, 66) = 3022$, $p = .046$, $\eta^2_p = .09$) was observed. Younger adults showed overall higher gain scores than older adults. Age difference was however not significant for the implicit training group, but for WM training and no-contact control groups.

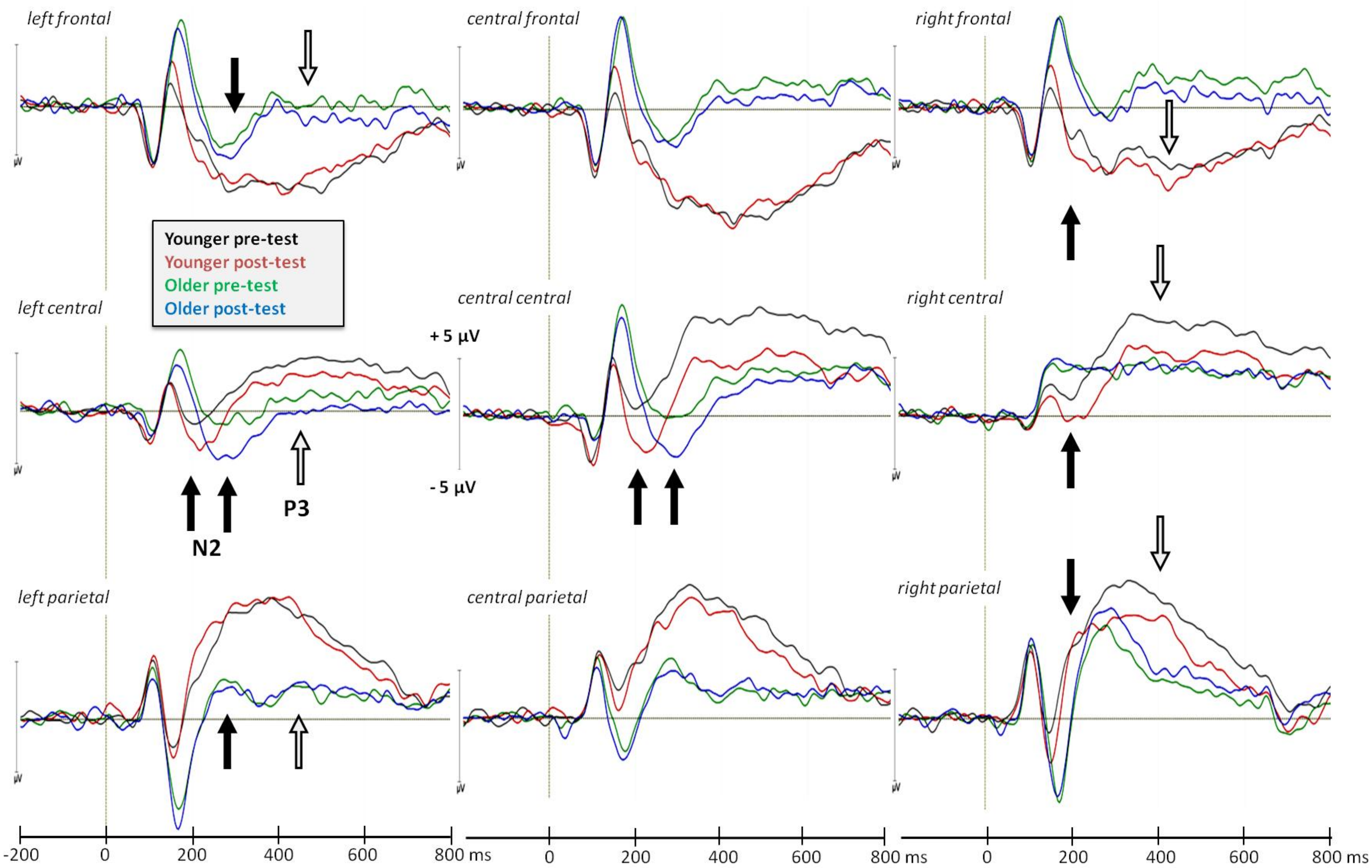

Figure 14. **WM training group**: ERPs for the verbal 2-back task at the nine ROIs for younger and older adults at pre-test and post-test; significant changes from pre- to post-test compared to the control groups are indicated with black arrows for the N2 component and with white arrows for the P3 component.

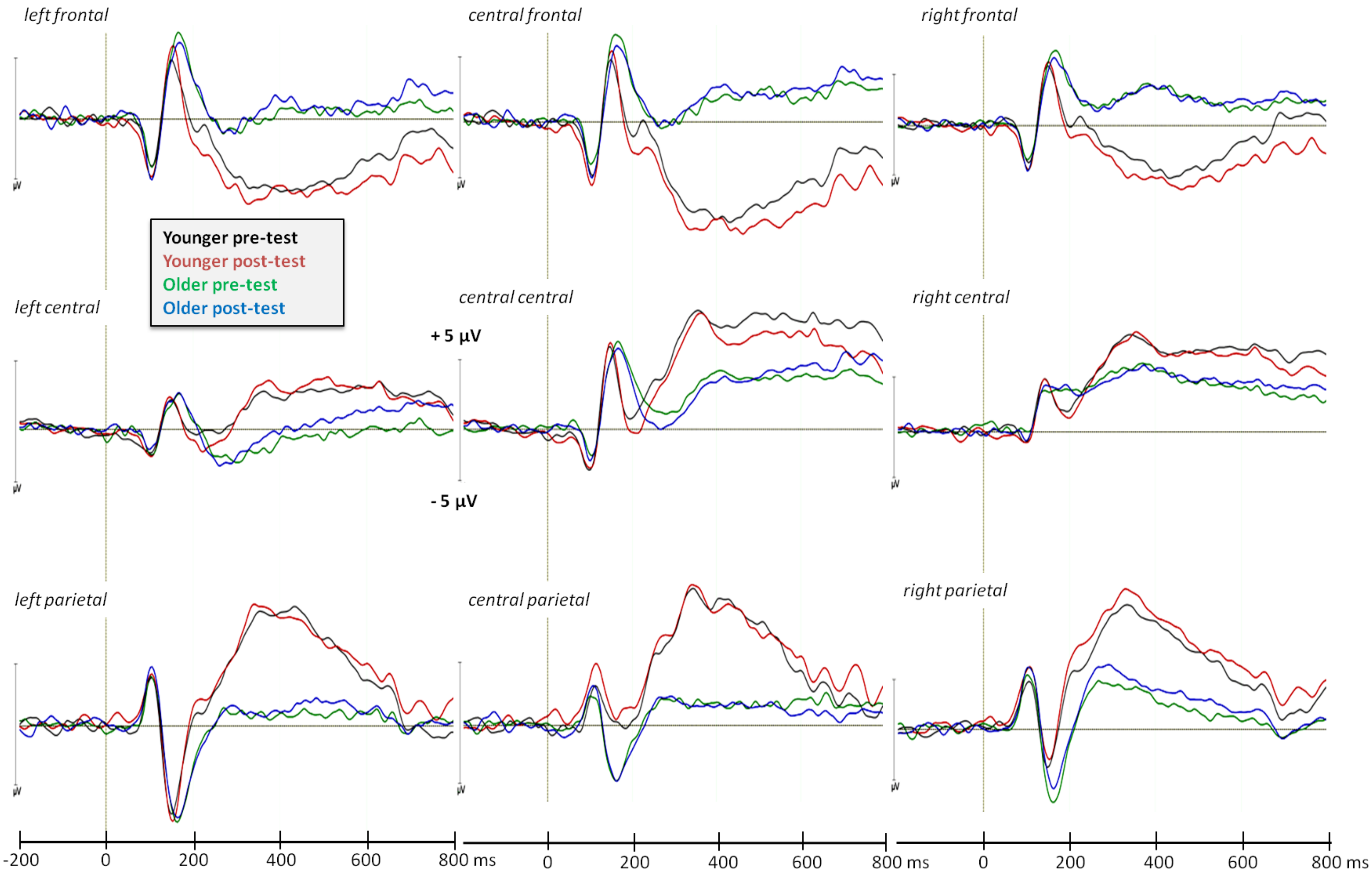

Figure 15. **Implicit training group**: ERPs for the verbal 2-back task at the nine ROIs for younger and older adults at pre-test and post-test.

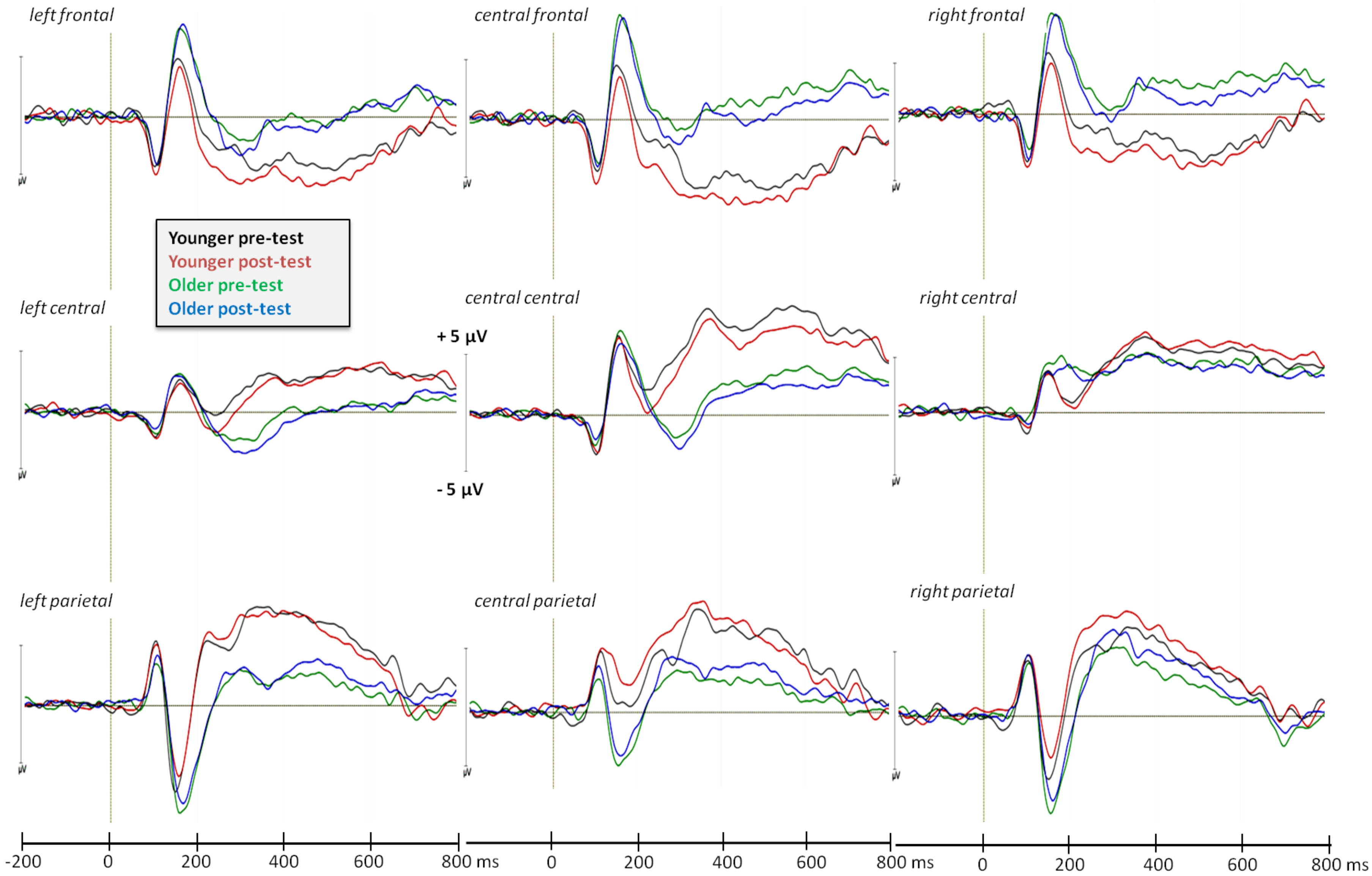

Figure 16. **No-contact control group**: ERPs for the verbal 2-back task at the nine ROIs for younger and older adults at pre-test and post-test.

Spatial N-back Task

The N2 component

The analyses of the amplitude and latency of the N2 component did not show significant effects for the spatial 2-back task gain scores (see Figures 17, 18, and 19 for ERP waveforms). The same was true for the repeated ANOVA on amplitude measures for the WM load cost gain scores.

The one-way ANOVA on Cz latency for these WM load gain scores revealed a marginally significant Age x Training Group interaction ($F(2, 66) = 3$, $p = .057$, $\eta^2_p = .08$; see Appendix F, Figures 4, 5, and 6 for ERP waveforms of the 0-back task). Pairwise comparisons revealed significant differences between the training groups in older adults only: the WM training group showed significantly smaller (negative) gain scores compared to both control groups whereas younger adults did not differ between training groups. This indicates that the older WM training group showed the N2 peak significantly later than the older control groups.

The P3 component

The repeated ANOVA on mean amplitude gain scores for the spatial 2-back task revealed a significant interaction X-axis x Age ($F(2, 132) = 3.3$, $p = .039$, $\eta^2_p = .05$) and Y-axis x Training Group ($F(4, 132) = 3.1$, $p = .019$, $\eta^2_p = .09$; see Appendix F, Table F3 and see Figures 17, 18, and 19 for ERP waveforms). Pairwise comparisons revealed that younger adults showed in general more change in central ROIs whereas older showed more change in the left ROIs. Further, the pairwise comparisons indicated that the frontal ROIs showed more change than the central and parietal ROIs in the WM training group only. That is, the frontal ROIs showed a positive change whereas the central and parietal ROIs demonstrated a negative change. The changes in both control groups were similar in all ROIs. The WM training group and the no-contact control group differed in the frontal ROIs significantly.

As regards the latency gain score measures taken at the local Pz electrode peak, no significant effects were observed.

The WM load costs, that is, the difference between 0-back and 2-back conditions did not reveal significant training related results.

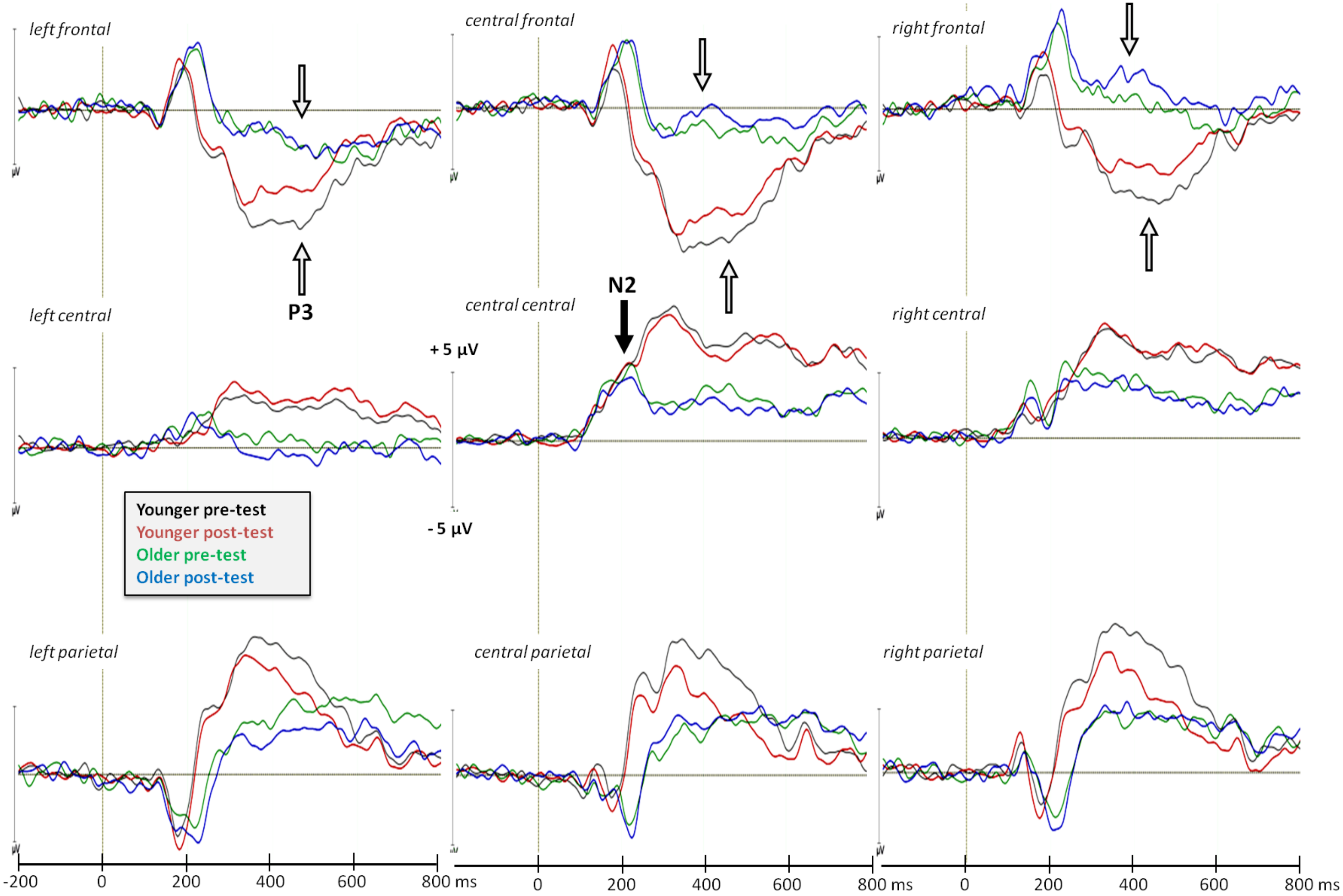

Figure 17. **WM training group**: ERPs for the spatial 2-back task at the nine ROIs for younger and older adults at pre-test and post-test; significant changes from pre- to post-test compared to the control groups are indicated with black arrows for the N2 component and with white arrows for the P3 component.

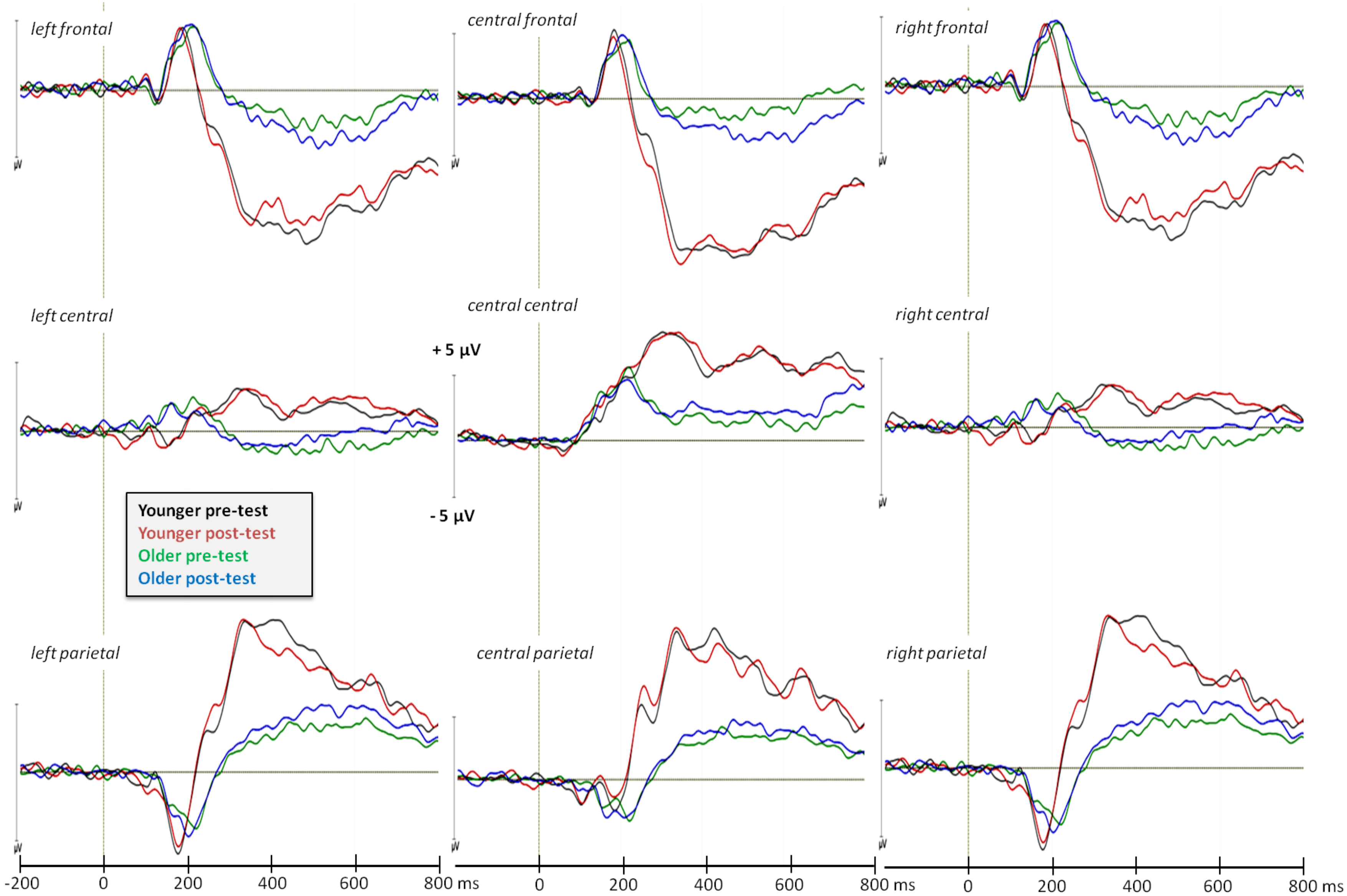

Figure 18. **Implicit training group**: ERPs for the spatial 2-back task at the nine ROIs for younger and older adults at pre-test and post-test.

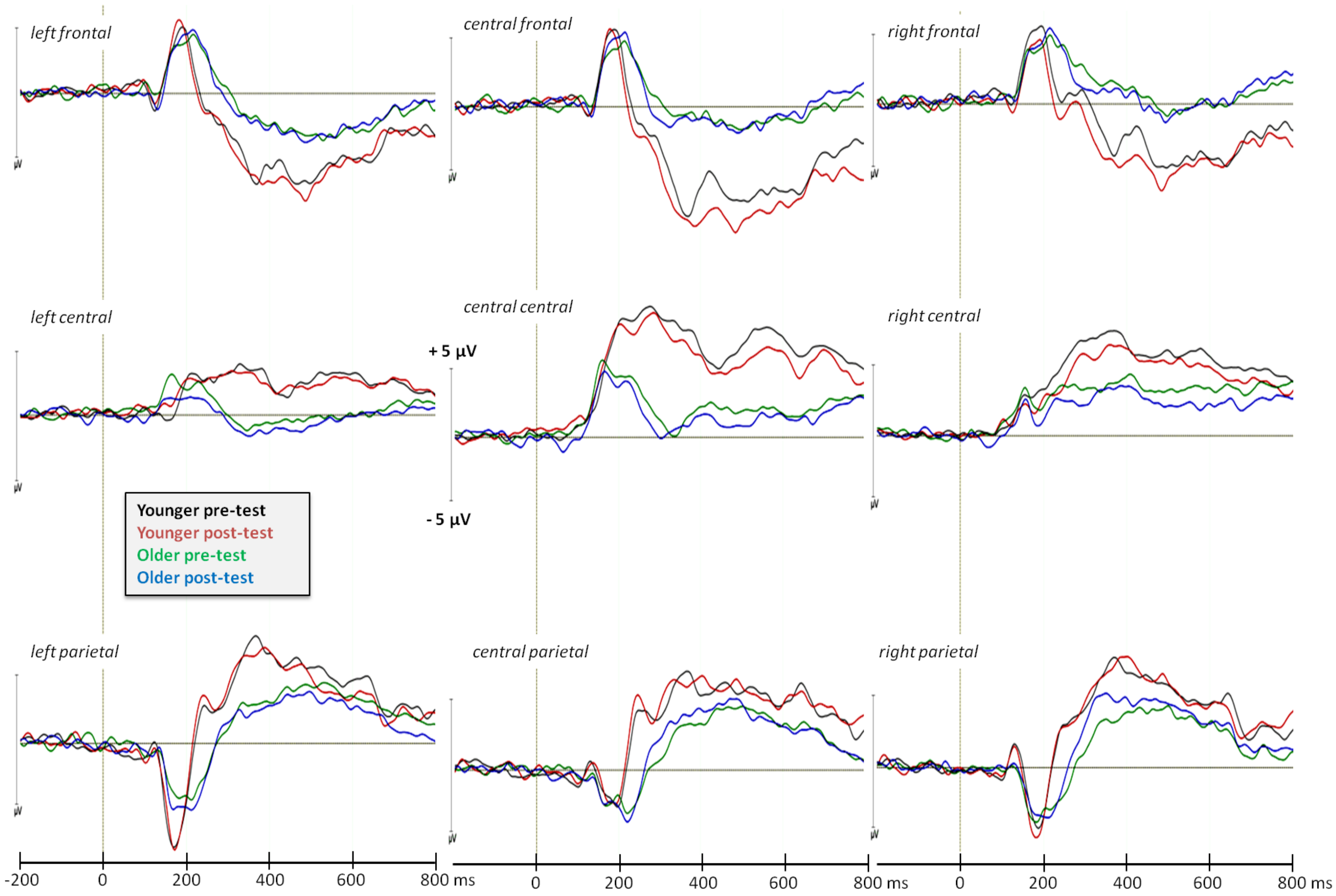

Figure 19. **No-contact control group**: ERPs for the spatial 2-back task at the nine ROIs for younger and older adults at pre-test and post-test.

Summary

With regards to age differences at pre-test in amplitude and latency measures, we observed the classic age-related pattern: Younger adults showed in general shorter latencies and larger amplitudes for both the N2 and the P3 component in all task conditions as compared to older adults.

The classic ERP analyses on gain scores revealed a different pattern for the verbal task as compared to the spatial task, for a summary see Figure 20.

As regards the verbal 2-back task, there were on the one hand WM training specific changes which were similar for both age groups. These effects were found for the N2 amplitude measures in the left-central and the central-central ROI. The amplitude was increased in negativity in both age groups after WM training. This indicates that the N2 negativity became more pronounced and distributed whereas the positive activity became less distributed over the scalp and more selectively activated in posterior regions.

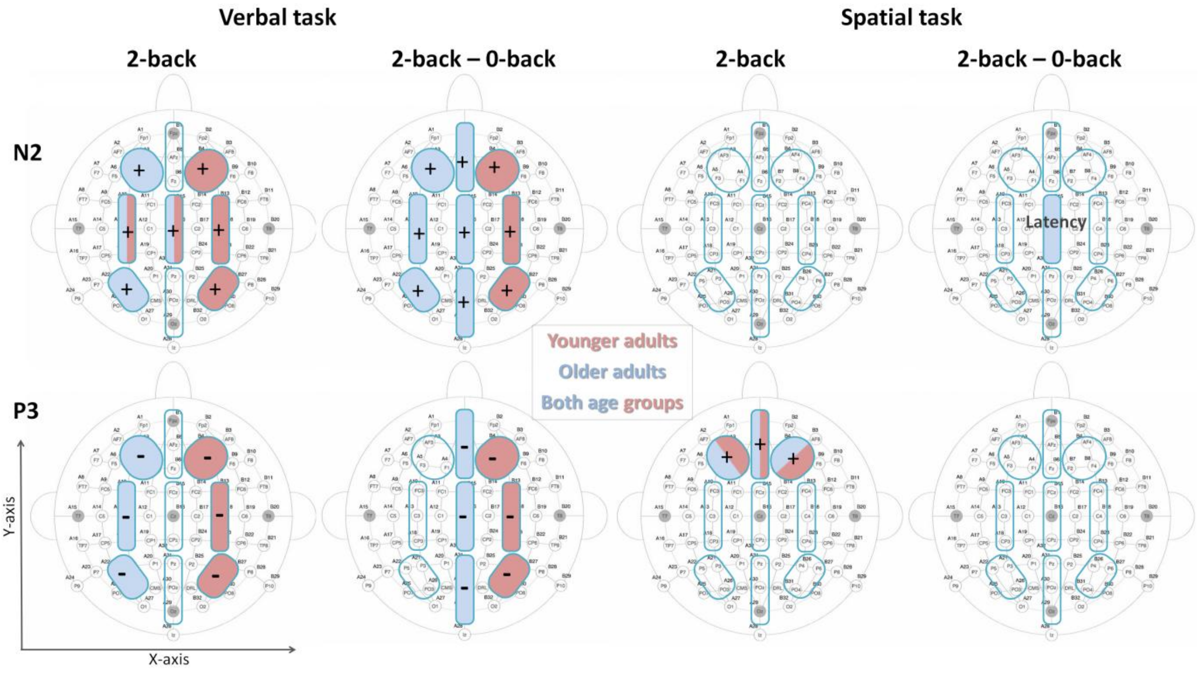

Figure 20. Summary of the ERP results: Significant ROI amplitude changes in the WM training group from pre-test to post-test compared to both control groups are marked in color, ‘+’ indicates that the amplitude increased, that is, more negative for the N2 component and more positive for the P3 component, ‘-‘ indicates that the amplitude decreased, respectively; left panels for the verbal task, right panels for the spatial task, upper panels for the N2 component and lower panels for the P3 component; the bicolored ROI indicate that both age groups changed similarly.

On the other hand, there were age-specific ERP changes after WM training. Younger adults showed an activation change at the right ROIs, an increased negativity for N2 and a decreased positivity for P3. The older adults in contrast showed these activation changes at the left ROIs for both ERP components.

The results for the verbal 2-back revealed in general, that the activation became less distributed over the scalp towards a more selective recruitment in central and posterior regions.

As concerns the untrained spatial 2-back task, less changes were observed. There was a training specific effect in the P3 amplitude: Frontal ROIs showed an increased amplitude after training in both age groups.

This result contradicts the findings from the trained verbal 2-back task where less frontal activation was observed after training. However, there seems to be a general training induced change which is not age-specific. Generally, the WM training groups showed less distributed, more central N2 activity and more posterior P3 activation in the trained verbal task. In the untrained spatial task in contrast, they showed more distributed and more anterior activation. These results have further to be confirmed by spatio-temporal microstate segmentation.

The lack of latency changes with training reflects the behavioral findings where no changes in response times were found.

Spatio-Temporal Microstate Segmentation

Age Differences in Microstates

First, we examined age differences in microstate map occurrence. The analysis revealed that there were age specific microstate maps. Maps 3 and 4 appeared more frequently in younger adults whereas the maps 5 and 6 were more frequently found in older adults as compared to younger adults for all task conditions (see Table 13). But for the verbal 0-back task, the map 6 was equally present in both age groups.

The map 3 reflected a typical youth-like N2 component and so did the map 4 which represented a youth-like P3 component (see Figure 21). Descriptively speaking, the map 3, or the N2 component, seemed more frontally oriented for the spatial 2-back task as compared to the verbal N2 map (see Figure 22). The map 3 was not present in the spatial 0-back condition. As regards the maps in older adults, they were all more anteriorly oriented as compared to the youth-like maps. These maps for the N2 (map 5) as well as the P3 (map 6) component showed more frontally distributed positivity than the maps in younger adults. Our population exhibited therefore the classic age-related ERP pattern at pre-test (Anderer et al., 1996; Daffner et al., 2010; Fabiani et al., 1998; Friedman, 2008).

Table 13. *T*-tests for the difference between younger and older adults on the duration of presence for microstate maps at pre-test.

	Younger *M (SD)*	Older *M (SD)*	*t*	*p*
0-back tasks				
Verbal map 3	18.56 (14.14)	4.75 (8.60)	5.00	< **.001**
map 4	11.50(13.80)	1.28 (3.65)	4.30	< **.001**
map 5	2.33 (6.98)	9.14 (11.69)	-3.00	**.004**
map 6	3.17 (7.36)	4.53 (7.81)	-0.76	.449
Spatial map 4	64.86 (29.47)	18.28 (26.35)	7.07	< **.001**
map 5	3.25 (8.18)	26.58 (32.60)	-4.17	< **.001**
map 6	20.89 (29.21)	44.14 (32.96)	-3.17	**.002**
2-back tasks				
Verbal map 3	15.94 (14.50)	3.89 (8.65)	4.28	< **.001**
map 4	10.17 (15.01)	0.92 (4.23)	3.56	**.001**
map 5	3.36 (6.88)	13.06 (12.79)	-4.01	< **.001**
map 6	5.81 (10.68)	12.03 (12.75)	-2.45	**.028**
Spatial map 3	21.86 (26.38)	11.69 (20.24)	1.83	.071
map 4	56.81 (32.13)	18.22 (24.63)	5.71	< **.001**
map 5	4.33 (9.62)	23.28 (26.27)	-4.06	< **.001**
map 6	6.00 (12.27)	35.81 (29.22)	-5.64	< **.001**

Note. Duration in time frames (TF); 1 TF = 4 ms. df = 70. For the corresponding voltage maps see Figure 21 and Figure 22.

Verbal N-back Task

Verbal 2-back task

For the verbal 2-back task, the microstate analysis yielded a solution with 7 stable microstate maps[1]. We analyzed the maps 1 - 6 for the period 175 - 360 ms after stimulus onset (see Figure 21 for microstate results). We observed a significant effect in the duration of presence and the GEV (global explained variance) for map 1 (see Appendix F, Table F4 for descriptive results). The Age x Training Group ANOVA for the duration gains scores revealed a significant effect of training group ($F(2, 66) = 4.07$, $p = .022$, $\eta^2_p = .11$). Pairwise comparisons showed that the WM training group ($M = 9.63$, $SD = 12.26$) exhibited a higher gain score for map 1 duration than both control groups ($M = 1.21$, $SD = 10.14$ for the implicit

[1] The microstate analysis provides a solution with an optimal number of maps. But only the maps of interest are subsequently analyzed.

group; *M* = 3.58, *SD* = 8.95 for the no-contact group). However, the difference to the no-contact control groups was only marginally significant. No further effects were found as regards map 1 duration.

As regards the GEV gain scores of map 1, similar effects were found: There was a significant effect of training group ($F(2, 66) = 4.61$, $p = .013$, $\eta^2_p = .12$), but no significant effect of age or an interaction. Pairwise comparison revealed that the WM training groups (*M* = 0.17, *SD* = 0.22) had significantly larger gains in GEV scores than both control groups (*M* = 0.03, *SD* = 0.16 for the implicit group; *M* = 0.04, *SD* = 0.14 for the no-contact group).

The analyses of the maps 2 - 6 did not yield another significant training related gain score or an interaction with the map 1. However, we conducted two post-hoc analyses as we expected a decrease in duration time for maps 3 and 4 in younger adults and for map 5 and 6 in older adults. The Session x Training Group ANOVA conducted separately for each age group on the mentioned map revealed no significant main effects or interactions. However, the pairwise comparisons yielded a trend according to our expectations.

In younger adults, the decrease from pre-test to post-test in map 4 duration was significant for the WM training group (pre-test: *M* = 11.33, *SD* = 17.64; post-test: *M* = 3.5, *SD* = 7.1) in contrast to the pre-post difference in both training groups (pre-test: *M* = 8.58, *SD* = 15.73; post-test: *M* = 5.17, *SD* = 9.44 for the implicit group; pre-test: *M* = 10.58, *SD* = 12.41 post-test: *M* = 9.17, *SD* = 9.40 for the no-contact group).

Similarly in older adults, map 6 duration decreased significantly from pre-test to post-test for the WM training group (pre-test: *M* = 15.17, *SD* = 13.11; post-test: *M* = 9.0, *SD* = 13.64) whereas the sessions did not differ for the implicit training group (pre-test: *M* = 12.83, *SD* = 13.75; post-test: *M* = 10.92, *SD* = 13.61). However, a significant decrease in map 6 was also found for the no-contact control group (pre-test: *M* = 8.08, *SD* = 11.31; post -test: *M* =0.67, *SD* = 2.31).

These analyses revealed in sum that younger and older adults showed similar changes with training by adding a new map after training. Both age groups showed increased duration in the same posterior positivity map and both decreased the duration of age-specific frontally oriented maps. These effects were significantly more present in the WM training group compared to both control groups. On the grand average microstate analyses it seems as if the map 6 was replaced by map 1 in older adults. But the interaction map 1 X map 6 was not significant; this effect was thus not confirmed.

Verbal 0-back task

For the verbal 0-back task, the microstate analysis yielded a solution with 10 stable microstate maps. We analyzed the maps 1 - 5, 7 and 8 for the period 175 - 360 ms after stimulus onset (see Figure 21 for microstate results; for younger adults, the maps were similar for the 0-back and the 2-back conditions). We obtained similar results on map 1 for the 0-back condition as those obtained for the 2-back condition: There was a significant effect in the duration of presence and the GEV for map 1. The Age x Training Group ANOVA for the duration gains scores revealed a significant effect of training group (F(2, 66) = 3.07, p = .053, η^2_p = .09). Pairwise comparisons revealed that the WM training group (M = 5.88, SD = 11.40) exhibited a higher gain score for map 1 duration than the implicit training group (M = -0.25, SD = 5.66) but not than the no-contact control group (M = 4.96, SD = 9.71). The results were similar for the GEV gain score: The effect of training group was significant (F(2, 66) = 3.32, p = .042, η^2_p = .09), but again, the mean GEV gain score was only significantly different between the WM training group (M = 0.12, SD = 0.18) and the placebo control group (M = 0.01, SD = 0.09) but did not differ from the no-contact control group (M = 0.07, SD = 0.14). No further significant training related gain score difference was observed for the remaining maps.

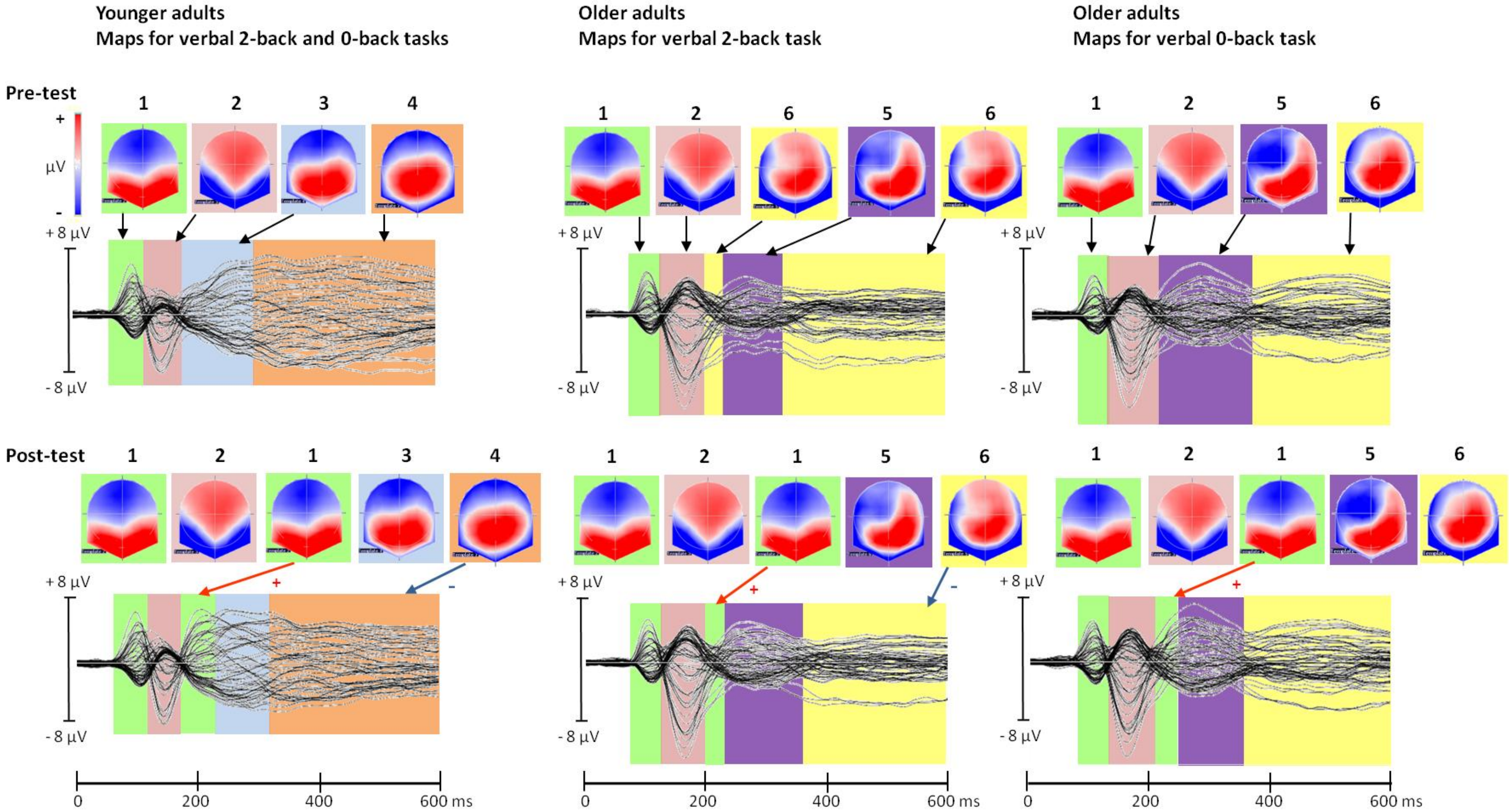

Figure 21. WM training group: Results of spatio-temporal microstate analyses of the verbal 2-back and 0-back task in younger (left panels, similar maps for 2-back and 0-back task) and older (middle panels for the 2-back and right panels for the 0-back task) adults for pre-test (upper panels) and post-test (lower panels) sessions; the presence of microstates is marked in color on the superimposed grand average ERP waveforms for all 64 electrodes; scalp voltage maps of the microstates are framed in the corresponding colors and are numbered; the maps represent a 2-D view from above, nasion on the top; blue arrows in the post-test panels indicate significant decreases of map duration of presence from pre- to post-test compared to both control groups whereas red arrows indicate increases of duration of presence.

Spatial N-back Task

Spatial 2-back task

For the spatial 2-back task, the microstate analysis offered a solution with 8 stable microstate clusters. We analyzed the maps 1 - 6 for the period 260 - 600 ms after stimulus onset (see Figure 22 for microstates results and Appendix F, Table F5 for descriptive results).

The analysis of map 3 revealed (marginally) significant effects of training for its gain in duration ($F(2, 66) = 2.78$, $p = .069$, $\eta^2_p = .08$) as well as in GEV ($F(2, 66) = 3.48$ $p = .036$, $\eta^2_p = .1$), but no effects of age or an interaction were observed. The WM training group decreased significantly in map 3 duration as well as in GEV as compared to the implicit training group who increased duration and GEV. However, both training groups did not differ significantly from the no-contact control group.

As regards map 5, also a significant main effect of training for gains in duration ($F(2, 66) = 3.13$, $p = .05$, $\eta^2_p = .09$) as well as in GEV ($F(2, 66) = 4.56$, $p = .014$, $\eta^2_p = .12$) were observed. However, the opposite pattern compared to map 3 was found: The WM training group increased significantly in duration as well as in GEV as compared to the implicit training group who decreased duration and GEV in map 5. But still no significant differences to the no-contact control group were found for both training groups.

With respect to map 4, a significant interaction Training Group x Age ($F(2, 66) = 3.21$, $p = .047$, $\eta^2_p = .05$) for the duration revealed that only the older training groups differed from each other. The older WM training group decreased in duration whereas the implicit training group in turn increased, again, no differences to the no-contact control group were observed.

Finally, for map 6, a significant interaction Training Group x Age was found as well, but in addition to duration ($F(2, 66) = 3.4$, $p = .040$, $\eta^2_p = .09$) also for GEV ($F(2, 66) = 2.83$, $p = .067$, $\eta^2_p = .08$). The pairwise comparisons also revealed that only within the older group training groups differed from each other. The inverse pattern to map 4 was found: The older WM training group increased in duration as well as GEV whereas the implicit training group in turn decreased, again, no differences to the no-contact control group were observed.

These analyses revealed that mainly the WM training groups and the implicit training group showed differential change due to training. The WM training groups' ERPs were better explained by the maps 5 and 6 with training whereas the implicit training groups' ERPs were better explained by the maps 3 and 4. This effect was on the one hand similar for older and younger adults (maps 3 and 5) and on the other hand found in additional maps for older adults only (maps 4 and 6). However, these changes were not significantly different from the no-contact control group.

In conclusion, the presence of frontally distributed maps increased in the WM training group, whereas the duration for the more posteriorly oriented maps increased in the implicit group.

Spatial 0-back task

For the spatial 0-back task, the microstate analysis provided a solution with 14 stable microstate clusters. We analyzed the maps 1, 2, 4, 5 and 6 for the period 260 - 600 ms after stimulus onset (see Figure 22 for microstates results). No significant gain score differences from pre-test to post-test were observed.

Summary

In terms of age differences at pre-test, results revealed the expected pattern and confirmed the findings from the classic ERP analyses: Older adults exhibited more frontally oriented maps as compared to the maps present in younger adults.

The results on training-related changes from the microstates analyses of the verbal and the spatial *N*-back tasks confirmed the results of the classic ERP analyses.

On the one hand, they revealed a similar pattern with regard to age differences, neither of the tasks did show a differential change between younger and older adults. Both age groups appeared to change similarly with WM training.

On the other hand, the verbal and spatial tasks yielded different patterns of change. The verbal task revealed that younger as well as older adults exhibited a new map with more posteriorly oriented negativity. For the verbal 2-back task, there was additionally a trend towards a reduction of the frontally distributed positivity pattern. This reflected a reorganization pattern for both age groups. The additional voltage map appearing after training indicates that less attentional processes were involved after training.

As regards the spatial task, no new maps occurred with training but significant changes in the already present maps were found. That reflected a redistribution pattern for both age groups which was slightly more pronounced in older adults. The changes showed an increase in frontally distributed maps while the more posteriorly distributed maps decreased. These frontally distributed maps are usually found in older adults and in higher load condition. We conclude that the changes went towards an increased recruitment of attentional processes.

In sum, we found a reorganization pattern towards a less attentional functioning in the trained task and a redistribution pattern toward a more attentional functioning in the near transfer task.

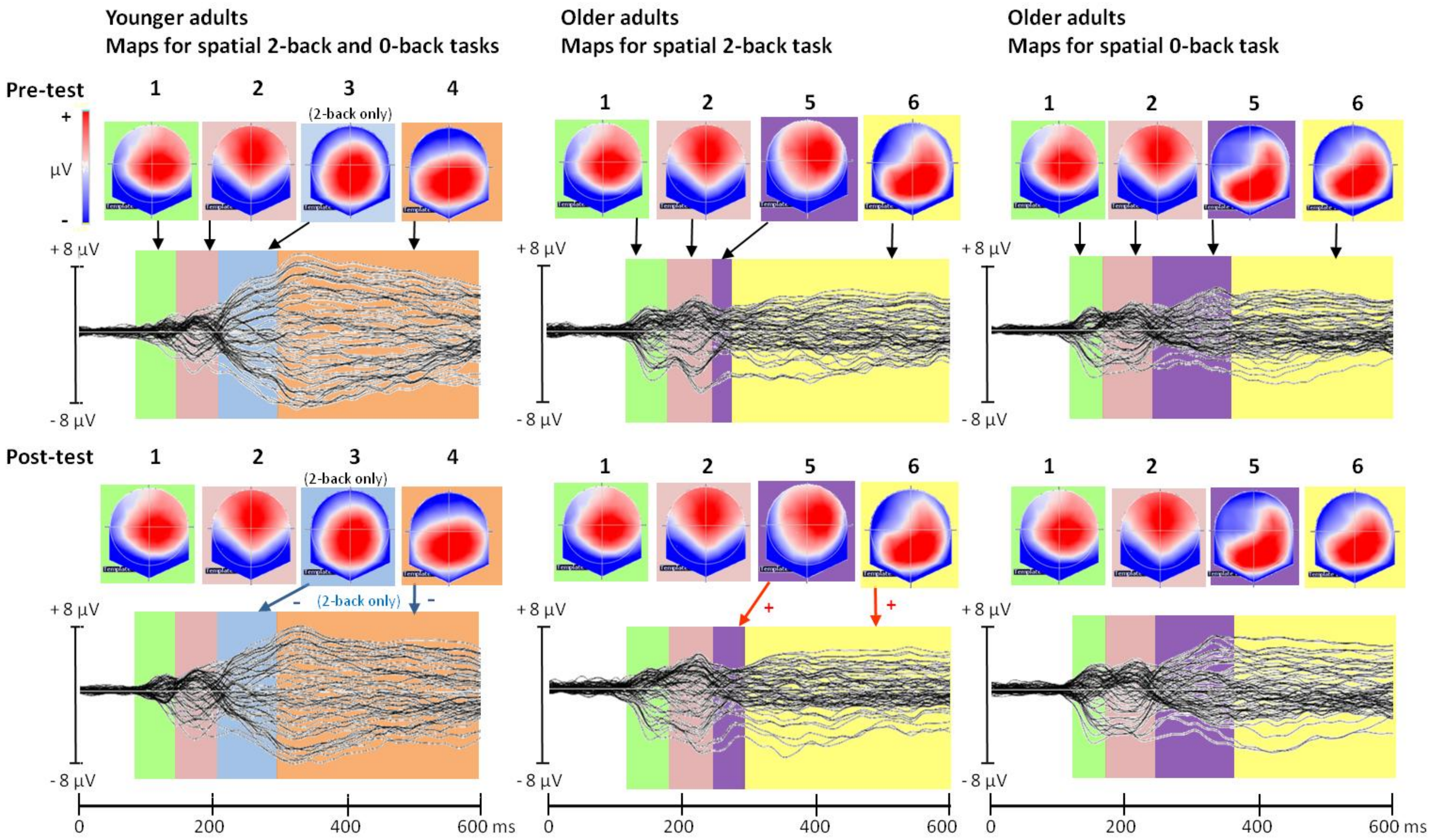

Figure 22. WM training group: Results of spatio-temporal microstate analyses of the spatial 2-back and 0-back task in younger (left panels, similar maps for 2-back and 0-back task with the exception of map 3) and older (middle panels for the 2-back and right panels for the 0-back task) adults for pre-test (upper panels) and post-test (lower panels) sessions; the presence of microstates is marked in color on the superimposed grand average ERP waveforms for all 64 electrodes; scalp voltage maps of the microstates are framed in the corresponding colors and are numbered; the maps represent a 2-D view from above, nasion on the top; blue arrows in the post-test panels indicate significant decreases of map duration of presence from pre- to post-test compared to both control groups whereas red arrows indicate increases of duration of presence.

DISCUSSION

For the discussion section, we first summarize the main results with reference to the hypotheses formulated before. Then, we discuss these results and integrate them with respect to the main research questions at the behavioral and the cerebral levels of analysis. Next, we conclude with an outlook on theoretical, methodological and practice-related implications of our results. Finally, we delineate limitations of the present work and suggestions for future research.

SUMMARY AND DISCUSSION OF THE FINDINGS

The aim of the study was to investigate the training-induced cognitive and brain plasticity in younger and older adults by means of a 10-day-WM-training procedure. We considered to implement an improved methodological training approach compared to previous studies. First, we used a promising process-based and adaptive training procedure adapted from Jaeggi and colleagues (2008; Chicherio, 2006; Ludwig et al., 2008), which has never been implemented with older adults before. Second, we included a younger and an older group of participants in order to study age-related differences in plasticity. Third, we assessed changes both in behavior and brain by the means of ERP measures. We included cerebral measures also in a near transfer task in addition to the ones in the trained task, which has, to our knowledge, only been done in one fMRI study but not with EEG measures to date. Fourth, we included an active placebo training group as well as a no-contact control group.

By this improved training study design, we intended to investigate age differences in and predictors of learning of a WM task. Moreover, we aimed to investigate age differences in the impact of a WM training on the trained WM task, near WM transfer tasks, and far transfer tasks.

A further purpose was to identify age-related differences in ERP correlates underlying the training-related changes in a trained and in an untrained structurally similar near transfer task. This aim, finally, enabled further insights into the interpretation of the cerebral reorganization which occurs at older age.

Behavioral Findings

Recall of the Hypotheses

With regard to the WM training, we hypothesized that there would be cognitive plasticity in both younger and older adults, but more pronounced in younger adults. This would lead to magnified age differences by the end of the WM training. However, both age groups would reach the limits of their performance and therefore the asymptote during ten sessions of WM training. Age in turn would be identified as a unique predictor beyond the cognitive variables assessed at pre-test of individual differences in training performance by the end of training. Further, we expected larger gains in the trained task for the WM training group than for both control groups; they would also be larger in younger adults. Next, we hypothesized near transfer effects to other, structurally similar as well as structurally different WM tasks in both age groups, but more pronounced or even exclusively present in younger adults. Finally, according to previous studies, we expected a far transfer effect in a fluid intelligence measure at least in younger adults (Jaeggi et al., 2008; Jaeggi, Studer-Luethi et al., 2010).

Summary of the Behavioral Findings

With respect to pre-test differences on cognitive tasks, the classic pattern of age-related differences was found: Younger adults showed more accurate and faster responses in all cognitive tasks from the gf domain (de Ribaupierre, 2001; Salthouse, 1991). The inverse pattern was observed for the gc measures (Baltes et al., 1998; Horn & Cattell, 1967). This pattern confirms that our sample of younger and older adults shows similar age differences in cognitive abilities as generally found in the literature. However, unexpectedly, an age effect was not found in the Reading Span task, in which younger and older adults did not differ significantly in performance (Delaloye et al., 2008).

As concerns the training phase, the results of the training procedures revealed that the performance of both age groups improved over the ten WM training sessions. Younger adults exhibited a generally higher training level than older adults. Moreover, younger adults improved faster during training and reached the asymptote later than older adults, which resulted in magnified age differences at the end of the training.

Further, age explained an important part of the variance in training data at the end of training, beyond the variance explained by cognitive abilities. However, at the beginning of the WM training, age did not explain more of the individual differences in training performance beyond the cognitive variables. These findings were confirmed by the LGCM analysis. It revealed that fluid intelligence was a significant predictor of the initial training performance, but not of the training improvement. Age group, by contrast, was a significant predictor of the improvement during training, but not of the initial training performance. Therefore, variability in fluid intelligence explained individual differences in initial training performance independent of age group. Age group, however, accounted for individual differences in the growth curve beyond individual differences in fluid intelligence.

A clear effect of training in younger and older WM training groups, compared to both control groups, was observed for the trained verbal 2-back task (in the following, whenever we indicate a significant results for the WM training groups, we mean as compared to both control groups in both age groups that showed no significant gain). The WM load cost, i.e., the difference between the low load 0-back condition and the higher load 2-back condition, also decreased significantly for both WM training groups. Further analyses comparing low versus high pre-test performers revealed that mainly the low performers benefited from training and improved performance due to training. In general, no interaction with age was observed. Therefore, no significant difference of training gain was found between younger and in older adults. However, the effect sizes were larger for older adults than for younger adults.

Moreover, a near transfer effect was observed in the spatial *N*-back task, which was present for the 2-back score as well as for the WM load costs. Similar to the verbal task, the effects were comparable for both age groups even though older adults exhibited larger effect sizes than younger adults. When further investigating high versus low performers in the spatial 2-back task, results revealed a similar pattern as in the verbal task, i.e., low performers benefited more from training, which was in the spatial task more pronounced for older adults. This indicated that mainly the older low performers benefited from the training and showed a near transfer effect.

As regards the structurally different WM tasks, that is, the Updating task and the Reading Span task, no further near transfer effects were found.

In terms of far transfer effects, there was an effect in the Stroop task, which was however also present in the placebo control group. No far transfer effects were found for the fluid intelligence measure or for the speed tasks.

Finally, we linked the training performance of the WM training group to their gains from pre-test to post-test in all cognitive variables. Analyses revealed that there was a positive relationship between training improvement and gain scores for spatial WM load costs, the Updating task, the Raven task and the SRT task, after controlling for age effects and fluid intelligence differences. This emerging relationship was found in tasks where a transfer effect has been hypothesized (except for the speed tasks), indicating that at least within the training group a link between training improvement and gains in transfer tasks existed.

Discussion of the Behavioral Research Questions

Reliable age-differences at pre-test

In line with the literature we found age-related differences in performance on all cognitive measures in favor of the younger adults on gf related measures and in favor of the older adults on the gc measure (e.g., Baltes & Mayer, 1999; Horn & Cattell, 1967). However, the unexpectedly similar performance between age groups on the Reading Span task was not in line with the literature, as a better performance would be expected for younger adults compared to older adults (Delaloye et al., 2008; Robert et al., 2009). According to the norms based on a French sample of younger and older adults, the performance for younger (*z* score = -0.50) and for older adults (*z* score = 0.12) were within the normal range (+/- 1.65) (Delaloye et al., 2008). Both age groups could therefore be considered as representative samples.

Increasing relevance of age in training performance

The results of the performance during WM training were in line with our predictions and with the cognitive plasticity literature, which we will discuss now in further detail. We found preserved cognitive plasticity in older adults, but less pronounced than in younger adults, which is in line with the literature (e.g., Baltes & Kliegl, 1992; Brehmer et al., 2012; Shing, Brehmer, & Li, 2008). Further, both age groups reached the asymptote performance in the course of training, the participants were therefore pushed to their performance limits (Kliegl et al., 1989). However, older adults reached the asymptote earlier than younger adults, which resulted in a magnification of age differences by the end of the ten WM training sessions. This magnification effect finding was in line with the results from strategy training research (Kliegl et al., 1990) as well as from process-based training (Brehmer et al., 2012), where age effects were also found to be magnified with training. However, a magnification of age differences with training has not consistently been reported for all process-based training studies (Dahlin, Stigsdotter Neely et al., 2008; S.-C. Li et al., 2008). In these studies, for example, older adults showed the same or even more training performance improvement than younger adults. However, these studies possibly suffered from ceiling effects in the training task and from an insufficient adaptive training procedure.

Further, the relevance of age as a limiting factor of cognitive plasticity was shown to increase in the course of training. That is, a significant amount of variance uniquely associated with age group was found in the second half of the training, which explained training performance over and beyond the cognitive variables assessed at pre-test. In a further analysis, age group was found to predict the learning curve significantly beyond fluid intelligence, whereas the initial training performance was predicted by individual differences in fluid intelligence but not by age group. This is in line with previous cognitive plasticity research, where age has been reported to explain unique variance at the end of the method of loci training for serial word recall (Kliegl et al., 1990). To our knowledge, this has never been replicated with a process-based cognitive training approach. In sum, with this study we showed that within an individually adapted training procedure, the effect of age became increasingly important and that the limits of performance can be reached in a WM training in only ten sessions of training.

But what mechanism could underlie the variable age, when the mediating performance between age and cognition from tasks assessing working memory, speed, inhibition, as well as fluid and crystallized intelligence has been controlled? First, we have to consider the possibility that not all cognitive measures managed to perfectly control for the intended cognitive ability. However, we covered a considerable amount of the cognitive tasks in the gf functioning and even the gc domain. The variable age, however, may reflect the increasingly negative influence of biological factors and the decreasing efficacy of cultural factors with advancing age, which go beyond the measured cognitive mediators (Baltes, 1997; de Ribaupierre, Poget, & Pons, 2005; de Ribaupierre, Pons, & Poget, 2003). This implies that even when controlling for increased processing speed and susceptibility to interference as well as for decreased WM capacity and fluid intelligence performance in older age (e.g., Hasher, Stoltzfus, Zacks, & Rypma, 1991; Salthouse, 1991), the negative influence of biology in older age becomes more and more apparent when the limits of performance are reached. This biological factor might be related to the general neurobiological mechanism described in the common-cause hypothesis (Baltes & Lindenberger, 1997; K. Z. H. Li & Lindenberger, 2002). The hypothesis suggests that in older age one common factor influences the integrity of the brain across a wide range of regions and so the performance in many different tasks is affected and thus maybe also cognitive plasticity.

As concerns the implicit memory training, we expected no magnification of age difference by the end of training. Implicit memory is known to be largely preserved with age (Parkin, 1993). The implicit training performance revealed that the response times of both age groups decreased in the course of training and that younger adults reached the asymptote earlier than the older adults. However, age differences remained stable when comparing the initial training session with the final session, in which the participants performed at their limits. Response times were generally higher in older adults, but the amount of learning, that is the plasticity, was similar between groups. This confirms the findings that the implicit learning rate is attenuated with age but that the total amount of plasticity in implicit learning is preserved in older adults (Daselaar et al., 2003). With further controlling for the cognitive variables, age explained the similarly unique amount of the variance in training performance throughout the ten training sessions. This indicates that the relevance of the variable age did not change in the course of training, in contrast to the WM training. This result, however, has to be interpreted with caution as we did not control for individual differences in implicit memory performance or in motor function. So we do not know whether age group is confounded with individual differences in implicit memory or motor performance.

However, we can conclude that cognitive plasticity is affected by age in the WM task, but that implicit memory plasticity seems to be preserved and can be found to the same amount in older adults as in younger adults.

Potentials and limits in training gain and transfer effects

Similar training gain in younger and older adults. We found a similar amount of training gain in younger and older WM trainings groups, which partially confirmed our hypothesis. The results showing that the performance of the WM training group increased on the trained verbal 2-back task and the WM load decreased to a larger extent compared to both control groups were in line with our predictions. However, we expected the training gains to be more pronounced in younger adults. These findings are therefore only to a certain extent in line with the training literature, for example with the study of S.-C. Li and colleagues (2008), where younger and older adults also exhibited similar training gain. However, other WM training studies revealed a larger training gain in younger adults compared to the older group (Brehmer et al., 2012; Dahlin, Nyberg et al., 2008).

The further comparison of training gain in low and high performers was in line with a recent study (Zinke, Zeintl, Eschen, Herzog, & Kliegel, 2011). The authors showed that the training gain was driven by the low performers only and no reliable training gain was found in high performers. We were able to confirm this differential finding in older as well as in younger adults.

However, these findings in low and high performers are in contrast to subsequent analyses we drew by linking training improvements with training gain. They revealed a positive association between the 2-back pre-test performance and the performance improvement in the course of the training. This indicates that participants with high pre-test performance also exhibited a steeper training improvement and therefore reached higher *N*-back levels than the low performers. These results are in line with the findings of Bissig and

Lustig (2007), for instance, who also reported a positive correlation between cognitive capacity and training performance.

The contradiction in our study between the results of the comparison of low versus high performers and the observed positive association of the training performance and the gain scores could, on the one hand, be interpreted in terms of a differential effect of training in low versus high performers. That is, low pre-test performers increased less during training, which suggests that they trained all through the ten training sessions at a lower *N*-back level than the high performers. This implies that the low performers trained in total on more blocks of the 2-back level than the high performers who trained only the first block of training on the 2-back level since they increased the *N*-back level immediately after the first block. That could explain why low performers improved more on the 2-back task from pre- to post-test and therefore showed a reliable training gain on the 2-back task. This connection would also be supported by the negative correlation between the training gain and initial training performance on the one hand and the training improvements on the other. So the ones who showed more training gains started at a lower initial training level and improved less during training. On the other hand, this evidence might also indicate that the verbal 2-back task exhibited a potential ceiling effect for high performers. That is, high performers were not able to further increase their performance on the 2-back task, whereas low performers had more space to improve.

Reliable near transfer effect in younger and older adults. The near transfer effect in a structurally similar spatial *N*-back task was in keeping with our predictions. This finding is in line with previous studies which reported near transfer effects to structurally similar tasks to the one trained in younger and in older adults (e.g., Brehmer et al., 2012; Buschkuehl et al., 2008; Karbach & Kray, 2009; S.-C. Li et al., 2008; Richmond et al., 2011; Schmiedek et al., 2010). The fact that younger and older adults exhibited similar transfer effects was reported, albeit less frequently, in the study of Li and colleagues (2008) and of Karbach and Kray (2009). Our findings, however, contradict the results of Dahlin and colleagues (2008), who reported transfer effects in younger adults but not in older adults. This is rather surprising as the training regimen was larger in the Dahlin and colleagues study than in the present one. Their training program involved three times 45 minutes of training a week during 5 weeks, which is much more than the ten sessions of 30 minutes of training provided in the present study. However, it might be that the more demanding 3-back task used by Dahlin and his colleagues exhibited a larger test-retest effect in older adults than in younger adults, as they may have needed more time for familiarization with the task (and this was very probably the case, as we found when consulting their raw data of the 3-back task: the older no-contact control group improved even more than the older training group in terms of numerical comparisons). In our study, we could frequently observe that the instructions for the 2-back tasks were not correctly understood at the first attempt by the older adults. In our view, this could be a reason for the lack of near transfer effects in Dahlin and co's training program.

We further found more near transfer effects in low performers compared to high performers; this was even more pronounced in older adults. Similar arguments as for the verbal 2-back task can be provided for the spatial one. First, low performers and especially older adults trained during more blocks on the 2-back level, which would foster near transfer to the structurally similar 2-back task. Second, ceiling effects also have to be considered for the spatial task. However, Jaeggi and colleagues (2008) also reported more transfer effects for low performers form the *N*-back training to a fluid intelligence task in a study with younger adults. They used fluid intelligence measures like the Raven's Advanced Progressive Matrices, which should, however, not suffer from ceiling effects (J. C. Raven, 1962). This phenomenon therefore also occurs in clearly unbiased measures.

Limited near and far transfer effects. As regards further near and far transfer effects, we could not confirm our hypotheses. No additional transfer effects were found, which is on the one hand in line with two recent WM training studies (S.-C. Li et al., 2008; Zinke et al., 2011). These studies reported no far transfer effects in younger or in older adults. Moreover, Zinke and colleagues (2011) reported a marginal far transfer effect in the Stroop interference score comparable to the one revealed by our analysis. However, in our study, the active control group exhibited a similar far transfer effect, so that the effect was not reliable.

On the other hand, our findings are partially in contrast with two studies where far transfer effects in younger adults but not in older adults were found (Dahlin, Nyberg et al., 2008; Schmiedek et al., 2010). Furthermore, the findings are fully in contrast with several other recent WM training studies, which reported far transfer effects for younger and older adults in different cognitive tasks, such as episodic memory tasks or in cognitive function in daily life (Borella et al., 2010; Brehmer et al., 2012; Buschkuehl et al., 2008; Mahncke et al., 2006; Richmond et al., 2011).

According to our predictions, we expected a near transfer effect to the Updating task, at least for younger adults, which would confirm the results of Dahlin, Nyberg and colleagues (2008). However, when dissecting the training tasks and the WM transfer tasks used in their study, the difference with our results seems less pronounced than initially assumed. Dahlin and his colleagues trained four different updating tasks, including a number updating task, where the participants were asked to recall the last four presented items in the correct order. The near transfer effect was found in a structurally similar number 3-back updating task. For the remaining near transfer tasks, however, no effects were found. These near transfer tasks included a computation span which follows the same task principle as the Reading Span task but with arithmetic problem solving and holding the final digit from each problem in memory. Further, in the Digit span forward and backward, no transfer effects were observed either. These results are in fact very similar to ours as we also found a near transfer effect for a structurally similar task, but no further transfer to structurally different WM tasks (i.e., Reading Span task and a number Updating task not structurally similar to the N-back task). Moreover, material used in our tasks differed more than the one in their tasks. We used verbal and visuo-spatial material in the *N*-back tasks, whereas Dahlin and colleagues used the same material (numbers) but with slightly changed task instructions. One could even question if Dahlin and colleagues assessed near transfer or whether it was qualified as training gain. We may therefore conclude from our study that, even though the trained mechanisms (updating and maintenance) were required in a transfer task, no transfer effect can be found whenever the difference between the task structures is too large.

We further made the prediction for near transfer effects in the Reading Span task. This hypothesis was drawn from previous *N*-back training studies which reported transfer effects into WM tasks in younger and older adults (Schmiedek et al., 2010). The lack of transfer effects in these tasks could be due to the relatively small number of training sessions (10) in the present study. Schmiedek and colleagues trained their participants over 100 days.

Furthermore, we hypothesized a far transfer effect in the Raven's Progressive Matrices at least in younger adults. This hypothesis was drawn from previous *N*-back training studies from which we had adapted our training procedure (Jaeggi et al., 2008). The lack of transfer effects in this task could be again due to the relatively small number of training sessions. Jaeggi, Buschkuehl and colleagues reported a positive association between training session and transfer effects (Jaeggi et al., 2008). In our study, we further used a single *N*-back task training and not the dual *N*-back task training as used by Jaeggi and colleagues (2008). We preferred the single task training to the dual task training as we considered the latter as too demanding and therefore frustrating for older adults. More recently, Jaeggi and colleagues supported our findings from the Reading Span task as they reported no transfer effect in a WM span task after a single *N*-back task training (Jaeggi, Studer-Luethi et al., 2010). However, they were still able to replicate their findings for fluid intelligence as they found far transfer effects after both single and dual task training in several fluid intelligence measures.

The studies by Jaeggi, Buschkuehl and colleagues recently became the subject of a controversial debate which was even discussed in The New York Times newspaper (Hurley, 2012). Several researchers raised doubts about the evidence that WM training should lead to reliable fluid intelligence improvements (e.g., Colom et al., 2010; Moody, 2009; Shipstead, Redick, & Engle, 2012). Engle and colleagues had attempted replicating the findings of Jaeggi and colleagues' WM training studies (Redick et al., in press). They provided 20 sessions of adaptive dual *N*-back training to a younger sample. The WM training group was compared to an adaptive visual search training group and a no-contact control group. The authors included a large battery of cognitive tasks to investigate near and far transfer effects on composite scores. However, they failed to replicate Jaeggi and colleagues' findings and could not show any reliable transfer effect. The authors further picked holes in Jaeggi and colleagues (2008) results by separately analyzing the four studies which were merged in their article. According to Redick and his colleagues, only one out of the four studies exhibited reliable far transfer results to fluid intelligence measures. For the other three studies, no far transfer was observed. This debate of whether WM training can improve intelligence is far from closed. On the one hand, there is a growing body of evidence supporting this view (Buschkuehl & Jaeggi, 2010; Jaeggi, Buschkuehl, Jonides, & Shah, 2011; Klingberg, 2010; Perrig et al., 2009; R. J. Sternberg, 2008), on the other hand, there is just as much evidence which counters this connection (Colom et al., 2010; Conway & Getz, 2010; Moody, 2009; Morrison & Chein, 2011; Owen et al., 2010; Salminen, Strobach, & Schubert, 2012; Shipstead et al., 2010; Shipstead et al., 2012). With the present study we supported the latter view.

Cerebral Findings

Recall of the Hypotheses

In terms of the ERP measures, we formulated the following predictions. In analogy to the behavioral measures we expected to observe training-related plastic changes at the cerebral level for both age groups. Before training, older adults would exhibit greater anterior positivity than younger adults. Given the scarce and inconclusive literature, precise predictions with respect to the specific nature of the plastic changes were difficult to make. In terms of ERP amplitudes, we expected a decrease in amplitude in both component of interest. Within the present work we aimed to investigate whether this more diffused and anterior positivity would decrease or increase with training. In accordance with several pieces of evidence from practice effects and low-versus-high performer studies, we hypothesized that older adults would show decreased and less frontally distributed positivity than before training. Similar changes were expected in younger adults, as we expected them to be high performers after training, who in general show decreased frontal activation compared to low performers. In terms of the change patterns proposed by Kelly and Garavan (2005), we expected a reorganization pattern for younger adults and a redistribution pattern for older adults. This in accordance with the literature, we therefore expected different change at the brain level for younger adults as compared to older adults.

Summary of the ERP Findings

With regards to age differences at pre-test in amplitude and latency measures, we observed the classic age-related pattern: Younger adults showed in general shorter latencies and larger amplitudes for both the N2 and the P3 component in all task conditions as compared to older adults. Microstate maps occurring at pre-test showed that older adults exhibited a more frontally oriented activation as compared to younger adults.

The classic ERP analyses on the gain scores for the verbal 2-back task revealed that the N2 amplitude increased for central sites and became more negative. This indicates that the N2 negativity became more important and shifted toward central sites and the positivity became less distributed and more local in posterior sites. This was true for both age groups of the WM training participants. Similar findings were found for the P3 amplitude, the positivity of which was decreased after training. There were, however, other ways by which younger and older adults came to a more posterior activity: Younger adults reduced activity in right sites whereas older adults reduced activity on the left sites. A similar effect was found for the P3 amplitude, which decreased for right sites in younger and for left sites in older adults. This also resulted in a pattern reflecting a less distributed posterior positivity.

These results of the classic ERP analyses on gain scores were confirmed by the microstate segmentation analysis. It revealed that an additional microstate, a marked central negativity, was present after training for a longer duration than in the other training groups. This state was not present before training in both age groups and can therefore be qualified as a reorganization pattern. There was also a trend towards a decreased duration of a microstate exhibiting a frontally distributed positivity. This map was mainly present in older adults before training. The additional map was also found in the 0-back condition in both age groups. Since for older adults, and in contrast to younger adults, the additional map was very different in scalp distribution from the map present before training, we could consider this change as a youth-like change. There was a considerable change from a frontally distributed map towards a new and youth-like process reflected by a posteriorly distributed map.

In sum, the results for the verbal 2-back were only partially according to our predictions. Younger adults showed a shift from a frontally towards a centrally oriented N2 negativity which was reflected by an amplitude increase at central sites. For the P3 component, they showed a decrease in amplitude at frontal sites. This resulted in a more posteriorly focused and less frontal positivity overall. Older adults showed similar trends but with a more pronounced change from a frontal positivity towards a posterior positivity/central negativity. In older adults, the activation became less distributed over the scalp than in the pre-test, tending towards a more selective recruitment in central and posterior regions.

As concerns the untrained spatial *N*-back task, we observed the inverse pattern: The frontal sites showed more positivity for the P3 component after training for the WM training groups, but only for the 2-back condition. This result was confirmed by the microstate segmentation. It revealed that the younger and older WM training groups showed a longer duration for the frontally oriented microstates after training and decreased presence for the posteriorly oriented maps during the spatial 2-back task. That was more extended for older adults, as the similar pattern was found in two additional maps. This pattern was, however, only significantly different from the implicit training group. The WM training group and the implicit training group differed systematically in all microstates after training while none of the training groups differed significantly from the no-contact control group. The WM training group showed more frontally distributed positivity, whereas the implicit training group exhibited more posteriorly oriented maps. This change was qualified as a redistribution pattern towards a frontally oriented map for the WM training group.

Considered together, the results from the verbal and the spatial *N*-back tasks were contradictory: A change towards a more posterior positivity was observed during the verbal task processing, whereas a more frontal positivity was observed during the spatial task processing. The WM training participants showed a reorganization towards decreased presence of control processes in the trained task and a redistribution pattern towards an increased presence of control processes in the transfer task.

Discussion of the Cerebral Research Questions

HAROLD and PASA at pre-test

Age differences at pre-test in the amplitude and latency measures revealed the classic age-related pattern with increased amplitudes and decreased latencies in younger as compared to older adults (Fabiani et al., 1998; Fabiani & Gratton, 2005; Ford et al., 1979; Friedman, 2008; Polich, 1996; Rossini et al., 2007). Voltage maps revealed a more anteriorly oriented activation in older as compared to younger adults, which is usually observed as well (Anderer et al., 1996; Daffner et al., 2010; Fabiani et al., 1998; Friedman, 2008). This pattern can be linked to models of functional cerebral changes, to the HAROLD and the PASA patterns (Cabeza, 2002; Davis et al., 2008). Even though we cannot make a direct link to the asymmetry reduction in frontal activation in older adults as described by the HAROLD pattern, our findings revealed that older adults showed more a attentional processes related pattern than younger adults. Our voltage maps can more directly be linked to the PASA pattern, in which older adults are found to shift towards an increased frontal recruitment compared to younger adults. We were able to confirm this pattern, our results showed at pre-test a shift from a posterior positivity in younger adults to an anterior positivity in older adults.

Decreased control in the trained task

The cerebral changes in the verbal task were partially in keeping with our predictions. We expected that younger and older adults of the WM training group would show similar attention related changes due to the WM training. This was the case as both younger and older adults showed less frontal and more posterior positivity after training than before training. These changes in the N2 and P3 components are related to more efficient allocation of attentional resources, decision processes and updating processes (Folstein & Van Petten, 2008; Friedman, 2008; Gajewski et al., 2008; Polich, 2007; Ritter et al., 1979).

However, we expected that only the younger adults would show a reorganization pattern but for older adults less change and a redistribution pattern. We found a reorganization pattern in both age groups which pointed even more towards an age-common change due to training.

The findings correspond to some of the few existing WM training studies that examined training-related cerebral activity changes by fMRI scans. A decrease in frontal activation has been reported in younger adults after visual and auditory *N*-back task training (Schneiders et al., 2011) and visuo-spatial *N*-back task training (Hempel et al., 2004). However, Hempel and colleagues showed an increase in activation during the first two weeks of the four weeks of daily training. This contradicts our findings, since our total training duration was two weeks of daily training and not four weeks. As introduced we linked their findings to the scaffolding framework which explains this frontal activation increase during the first sessions of training to serve as scaffold for the new demand. After continued training, the scaffold was probably no longer used and decreased as a consequence.

A frontal activation decrease in fMRI scans has also been reported in older adults after verbal and visuo-spatial span training (Brehmer et al., 2011). Yet our findings contradict other fMRI WM training studies, which reported an increased frontal activation after training in younger adults (Olesen et al., 2004; Westerberg & Klingberg, 2007) as well as in older adults (Dahlin, Stigsdotter Neely et al., 2008).

As regards the similar training-related changes in younger and older adults and the increased similarity between younger and older adults after training, they are in line with almost all few age-comparative training studies we introduced in the conceptual background section: Similar activation or structural brain changes in younger and older adults have been reported in WM training and juggling training studies (Boyke et al., 2008; Dahlin, Stigsdotter Neely et al., 2008; Schmiedek et al., 2010). Our results are further in line with the findings of a dual-task training, which revealed an increased similarity between younger and older adults, that is, an age-common pattern of activity after the training (Erickson et al., 2007b).

When further comparing our results to low versus high performer studies, the younger as well as the older adults already exhibited a high-performer profile at pre-test according to the results of Daffner and colleagues (2011). These authors found less frontally oriented voltage maps in high performers compared to low performers during the 0-back low load condition and a more frontal positivity in the 2-back high load condition. For low performers, however, they observed the opposite. Low performers showed higher activation in low-load condition on the one hand, and decreased frontal positivity and more widespread activation in high load conditions on the other hand. This was true for both age groups, but with generally more frontal activity in older adults than in younger adults. We found a comparable pattern in our sample, but more pronounced for older adults since younger adults showed similar voltage maps for the verbal 0-back and 2-back condition. Older adults in turn showed an increase in frontal activation from the 0-back to the 2-back condition. Furthermore, older adults showed overall more frontal activation than younger adults. But with training, both age groups decreased in frontal activation and increased in central negativity and posterior positivity. Hence, for our WM training groups, we could claim that the high-load 2-back task was turned into a lower-load condition at post-test. The activation pattern of the trained group at post-test was therefore in keeping with that of the high performers at the low load condition in the Daffner and colleagues (2008) study.

Also in line with this study was the fact that younger and older adults showed a similar trend in change with training like the high performers with increasing load conditions. Daffner and his colleagues showed that younger and older high performers exhibited the same relative increase in frontal positivity with increasing load.

Two other ERP studies, where older low and high performers were compared to younger adults, were also in line with our findings (Duarte et al., 2006; Lorenzo-López et al., 2007). They reported a youth-like and less widespread brain topography in older high performers as compared to older low performers. Furthermore, an fMRI study confirmed our findings, as it found less frontal activity in older high performers, similar to the sample of young participants (Nagel et al., 2009).

Our findings were in turn contradictory to the HAROLD model which found higher performance in older adults correlated with additional recruitment in both PFC (Cabeza, 2002; Cabeza et al., 2004). The authors reported that only the older high performers showed this additional recruitment and differed in the activation pattern from the younger adults (Cabeza et al., 2002). The older low performers in contrast showed similar brain activation as the younger adults, but, as its label indicates, they exhibited a poorer performance than younger adults and older high performers. Our findings were not in line with these results, since we found higher performance in both age groups with training but combined with similar cerebral changes in younger and older adults. Our youth-like non-compensatory cerebral pattern in older adults was in contrast with Cabeza and colleagues' findings.

When further integrating our findings into the debate about the interpretation of age-related differences in cerebral activation, our findings support the non-selective recruitment approach. This approach interprets the over-activation pattern in older adults as the cause of cognitive decline and not as a compensational reaction to decline. Since we found similar training-induced changes in younger and older adults, we could show that no age-specific pattern of change towards an improved performance existed. Further, our results did not point towards a compensational recruitment of frontal areas in either of the two age groups but rather towards a more efficient allocation of attentional resources.

Our results showed that the initially anteriorly distributed and more diffused positivity in attentional-related components in older adults became less distributed and more specific with practice in the trained task. The same change was found for younger adults, who exhibited a posterior activity already before training. This became even more pronounced with training, indicating that they improved their allocation of attentional resource. This suggests that the less frontally distributed and more specific topographic activation pattern was, in both age groups, more beneficial for resolving the 2-back task than the initial posterior activation in younger adults and the widespread frontal activation pattern at pre-test in older adults. This conclusion can be drawn because task performance was increased in both WM training groups at post-test. The pre-test pattern in older adults might therefore reflect the difficulty in engaging and activating appropriate brain networks. These data thus reinforce the benefit of a youth-like pattern in older adults and do not provide support for the compensation hypothesis.

Increased control in the transfer task

The spatial *N*-back task revealed a different pattern compared to the trained verbal task. The increased frontal positivity in both age groups after training was partially in line with the three, and to our knowledge only, WM training studies where cerebral (fMRI) measures were assessed during a transfer task. On the one hand, the studies revealed an activation increase in younger adults during a near transfer task (Dahlin, Stigsdotter Neely et al., 2008; Westerberg & Klingberg, 2007), on the other hand, an activation decrease in older adults (Brehmer et al., 2011). The former findings are in line with our findings while the latter contradict our findings, since the increase in frontal positivity was even more pronounced in older adults than in younger adults.

In contrast to the trained verbal 2-back task, the spatial task revealed a pattern which may supports the compensation hypothesis when comparing the findings to low versus high performers studies. As seen in the previous section, increased frontal activation in older adults can in contrast either be associated to low performers (e.g., Lorenzo-López et al., 2007) or to high performers (Cabeza et al., 2002). A more frontally oriented P3 topography is commonly associated to older age, higher cognitive demand and novelty (Segalowitz et al., 2001). Novelty however can be excluded, since the task was not new at post-test. That would indicate that the spatial *N*-back task became more demanding for the older adults at post-test. This could indeed be what happened after training, older adults were able to achieve a superior performance level but only by recruiting additional attentional processes alike for a higher load task. So the initially anteriorly distributed and more diffused positivity in P3 was even more pronounced in order to achieve a higher performance level. These findings reinforce therefore the benefit of a frontally oriented P3 topography which is in line with the literature supporting the compensation approach (Cabeza, 2002; Cabeza et al., 2002; Dolcos et al., 2002; Reuter-Lorenz & Cappell, 2008).

However, the fact that also the younger adults did probably allocate more attentional ressources to the spatial task after training is not fully in keeping with the predictions of the

HAROLD model (Cabeza et al., 2002). According to their findings, older high performers compensate by over-activating frontal areas while younger adults do not. Nevertheless, since Cabeza and colleagues did not distinguish between low and high performers in younger adults, we do not know how they behave on the cerebral level. Would younger low and high performers differ the same way as older low versus high performers do? Or would rather younger low performers show a similar activation as the older high performer? Our findings show that also younger adults over-activate frontal areas to improve performance at post-test, which indicates that an increased frontal activation in younger and older adults is beneficial.

Finally, as regards the age-comparative training studies, our findings - similar increased activation redistribution in younger and older adults - was in line with two of them, which reported similar changes in both age groups (Dahlin, Stigsdotter Neely et al., 2008; Erickson et al., 2007b).

Integration of verbal and spatial task findings

The findings of the verbal and spatial task were in conflict with each other: The verbal task findings supported the non-selective recruitment approach and the benefit of a decreased fontal activation pattern in older adults whereas the spatial task findings supported the compensation view and the benefit of an increased frontal activation pattern. We were therefore able to show decrease in one task and increase in another, but structurally similar, task of frontal activation after one and the same training intervention. The training led to a reorganization and so to a qualitative change in underlying processes in the trained task. For the transfer task in turn, a quantitative redistribution of activation was observed, but no change in underlying processes. These findings of different patterns were in line with a recent literature review which discussed the neuronal changes after WM training interventions (Buschkuehl et al., 2012). The authors stated that there is currently no clear pattern of evidence in training-related cerebral changes. The findings of change are still manifold, indicating all four patterns of change: increase, decrease, redistribution, and reorganization of cerebral activity.

On the other hand, our finding fit well with the scaffolding framework theory (Park & Reuter-Lorenz, 2009; Petersen et al., 1998), which describes the phenomenon during learning when additional regions are recruited to serve as a scaffold. Once the processes are more efficient, the scaffolding network is no longer needed and decreases. In older adults, the scaffolds are the frontal regions which compensate for the non-selectively recruited task-specific regions (Goh & Park, 2009; Park & Reuter-Lorenz, 2009; Reuter-Lorenz & Park, 2010).

The scaffolds for WM tasks would be the frontal regions which provide support for coping with increased attentional demands. In line with the scaffolding framework, our findings revealed that with a training intervention the scaffolding networks became less nessecary or even fell away by the end of training for the trained task. The processes turned more efficient and even changed so that no more scaffolding was needed and therefore less frontal recruitment was required. For the near transfer task in contrast, the training has not been extended enough to allow the scaffolding networks to decrease for WM tasks addressing another modality. On the contrary, they were still present and very efficient as they allowed the participants already performing at a higher performance level.

More precisely, we propose to interpret the increased frontal P3 activation in the transfer task as a classic compensational scaffolding pattern during skill acquisition and a sign of increased attentional resource allocation (Daffner et al., 2011; Johnson, 2001; Mattay et al., 2006). For the trained task in turn, the scaffolding activation of the P3 component was no longer needed, since the processes became more efficient and even changed in quality due to the training. We may hypothesize that the activation in the trained verbal 2-back task changed from a frontally over-recruited P3 component towards a more efficient early processing manifested in the N2 component. This improved earlier N2 process would decrease the demands on the subsequent P3 component which attenuates in amplitude since less effort is required at this stage of processing (Daffner et al., 2011; Folstein & Van Petten, 2008).

In order to test this assumption, an intermediate EEG recording session should be added in a further training study. By this means, we could examine if there was initially an increased P3 component in the trained task as well. We could therefore confirm that there would be an evolution from an increased P3 amplitude towards an increased N2 amplitude combined with a decreased P3 amplitude in the trained task.

Integration of the Behavioral and Cerebral Research Questions

The global research questions of the present work were as follows: First, to what extent cognitive plasticity exists in older as compared to younger adults, second, what its underlying cerebral plastic change patterns in younger and older adults are, and third, whether training-related brain plasticity can provide insights into the interpretation of age-related cerebral changes, i.e., the frontal over-activation usually observed in older as compared to younger adults.

There was manifested plasticity during the training in younger and older adults, but to a larger extend in younger adults. In terms of transfer effects, we were able to show that with our training procedure not only task and domain specific skills were trained, but also that to a certain degree general attentional processes involved in *N*-back task performance were acquired. This was shown in the transfer effect to the spatial *N*-back task and in the partially confirmed transfer effect to the Stroop task. These effects were similar for younger and older adults, indicating that the training had comparable effects on both age groups, even though the basic age differences persisted over the training. The effects were similar at the cerebral level: Results revealed that younger and older adults showed the same cerebral plastic changes due to training in the trained as well as in the transfer task.

The findings from the Stroop task, which assessed the transfer effects to the capacity of inhibitory processes, can partially be linked to the cerebral findings in the trained verbal *N*-back task. The ERP components revealed an increased activation of the N2 component with training. The N2 component is not only related to general control processes but has often been linked to response inhibition tasks and was also assessed in Stroop tasks (Folstein & Van Petten, 2008; Lamm, Zelazo, & Lewis, 2005). There seems to be a trend towards a change in inhibitory processes associated with an *N*-back training. Similar behavioral results have been reported in a recent study in which children were trained on a single *N*-back task condition (Jaeggi, 2012).

Indeed, updating tasks do not only require updating processes of the information to be retained in WM, but also the processes of shifting and inhibition (Miyake et al., 2000; Schmiedek, Li, & Lindenberger, 2009). The former requires the shifting from one task, i.e., the updating and monitoring of the flow of information, to another task, i.e., the comparison of the present information with the information hold in WM and making a decision. The latter, inhibition, is similarly important to provide a correct answer: old information which is no longer relevant for task completion has to be inhibited and deleted from WM.

In term of the debate concerning the interpretation of cerebral aging patterns, we were able to contribute to the interpretation of age-related changes at the cerebral level. We provided evidence for the non-selective recruitment account, since younger and older adults changed similarly due to the training. Our results suggest on the one hand that both age groups were similarly able to change the quality of processes towards a less attentional recruitment for responding to the trained task. On the other hand, we were able to show that compensational processes, indexed by the increased recruitment of attentional processes, do not only occur in older but also in younger adults. This was found for the transfer task, in which both age groups increased the frontal attentional recruitment with training. It seems as if the compensational behavior was needed to respond to the transfer task but that the processing has already turned into a more efficient stage for the trained task. This pattern has been found in both age groups, since no age-specific changes with training were observed. By pushing younger and older individuals to their limits, we went beyond the assessment of the shallow baseline performance and were able to investigate cerebral processes which underlie the so-called developmental reserve (Baltes & Kliegl, 1992; Baltes & Lindenberger, 1988; Kliegl et al., 1989). We revealed that on the way from the baseline performance to the developmental reserve, similar and age-independent processes were operating.

We can therefore not confirm that a different cerebral activation pattern in older as compared to younger adults would be beneficial. We claim that a youth-like functioning is more beneficial in older adults than a senior-like functioning. This corroborates the non-selective recruitment approach (S.-C. Li et al., 2006; S.-C. Li & Sikström, 2002), which suggests that cerebral functional age-differences are a sign of the deterioration at the level of the brain occurring with age.

Finally, we have to state that there remains a basic qualitative difference in processing between younger and older adults. Even though the nature of change due to cognitive training seemed to be similar, age differences in behavioral performance as well as in ERP patterns were still present after training. This was further confirmed by regression analyses in which age was a significant predictor of the final training performance beyond the cognitive performance. This basic qualitative difference in cognitive processes between younger and older adults has to be due to biological factors which can't be changed by cultural means (Baltes, 1997; Lindenberger et al., 2007).

OUTLOOK

This last chapter concludes the present work with several considerations placing our research in the broader context of aging and training theory as well as research, and delineating connections to practice. Then, we discuss some methodological and theoretical limitations of the present research. Finally, we provide suggestions for future research.

Implications for Theory, Research, and Practice

In terms of implications for theory, we want to stress the following points: First, we confirmed age differences in cognitive and cerebral plasticity and its presence in both younger and older adults. We demonstrated that the limits of performance can not only be tested by a strategy-based training (Baltes & Kliegl, 1992; Baltes & Lindenberger, 1988; Kliegl et al., 1989), but also by the means of a 10 session process-based training. A similar age-related plastic pattern as the one found in method of loci studies by Baltes and Kliegl was observed. Second, we confirmed many studies in which limited transfer effects from training were reported (e.g., Dahlin, Nyberg et al., 2008; S.-C. Li et al., 2008; Zinke et al., 2011). We further claim that training research would profit from a more unified definition of transfer effects. The concept of transfer needs to be defined beyond near and far transfer in a more detailed way and appropriate measures should be found. This would allow to distinguish between changes in task-specific processes, in general attentional processes, or in strategy use. One way to disentangle transfer of strategy acquisition from transfer of basic processes is the investigation of a latent change (Lövdén, Bäckman et al., 2010; Schmiedek et al., 2010). This method would allow to find latent change factors which would exclude strategy use. Finally, our results provided a contribution to the debate on the interpretation of age-related cerebral changes. These kinds of studies do provide insight into the age-related cerebral organization beyond the basic performance level and allow uncovering the individual's limits and the study on its optimal level of performance.

Concerning methodological issues in aging and plasticity research, the present work yielded that reliable improvements in the trained task as well as in a transfer task were found after only10 sessions of training. We also confirmed that a process-based training needs to be adaptive to the individuals' performance in order to challenge the individual throughout the training. Furthermore, as we have demonstrated, the inclusion of a younger reference group was essential for the interpretation of the training-related changes, particularly for those at the cerebral level. Further, the use of EEG measures allowed us to gain insight into the difference in quality of change between the trained task and the transfer task. We stress the importance of investigating underlying cerebral changes in both training and transfer effects. Moreover, by the use of EEG, we were able to study the training-related changes on the process level. Here we were able to distinguish changes in earlier processes (N2) from changes in later processes (P3). This would not have been possible with fMRI measures for instance, because the bad temporal resolution would not allow distinguishing the different processes. Finally, we demonstrated the importance of including control groups in training studies. Test-retest effects and the placebo training effect (implicit training) were substantial; however, the latter did most often not exceed test-retest effects as regards gain scores. In terms of effects sizes, both control groups exhibited small effect sizes for the trained task. As regards the spatial 2-back task, the effect sizes were still small for the no-contact groups whereas they reached medium level for the placebo training group. There seems to be a trend towards a general training related effect in the spatial transfer task. In this task, for instance, we showed that the inclusion of a placebo training group besides a no-contact control group is advisable.

With respect to applied contexts, this study illustrated that an adaptive training procedure is useful and that process-based training approaches have a great potential to study plasticity. However, for applied contexts, we would insist more on a tailor-made approach in order to respond directly to the individuals' needs. As proposed by McDaniel and Bugg (2012), the individual's needs should be evaluated in everyday memory and an appropriate training regimen should then be proposed. For some individuals, it might be more useful to conduct a process-based training in order to improve basic cognitive processes. In this case, the use of strategies should be avoided in order to reach the largest effect (Kliegel & Bürki, 2012). For other individuals, a training regimen could just target one specific everyday problem, since compensational means have a great potential especially in real-world tasks (Kliegel, Martin, McDaniel, & Phillips, 2007). For instance, performance in prospective memory, a highly relevant aspect of cognition in everyday life, could be addressed (Einstein & McDaniel, 1996). Training regimens targeting prospective memory could foster the use of strategies, since strategy use has been found to be very efficient for young older adults (Schnitzspahn & Kliegel, 2009).

Further factors such as social aspects or metacognitive self-regulation have been recently proposed as powerful tools to foster successful training (Hertzog & Dunlosky, 2012; McDaniel & Bugg, 2012) and should thus be considered in applied training contexts.

However, as recently stated, "there is currently little evidence that memory interventions (…) are relevant to the ultimate goal of training, which is to support older adults to remain independent longer by reducing, remediating, or reversing functional impairments" (Zelinski, 2012, p. 57). This is why future research should focus more on ecological validity and target on effects that might transfer into older adults' everyday life.

Limitations of the Present Study

The present study has provided a number of interesting findings, which have important implications for aging and training research. Nevertheless we need to address some general limitations of the current work. At the same time we provide suggestions for future research.

The most important limitations concern the sample which was investigated in the present study. First, our population exhibited a rather high education. Close to 90% of the younger adults and approximately 50% of the older adults held a university-entrance diploma. In 2010, 20% of the 19 year olds obtained a university-entrance diploma in Switzerland, whereas the proportion was around 12% in 1985 (Bundesamt für Statistik, 2012; no information is provided about the rates before 1985). This indicates that our sample did not entirely represent the average (Swiss) population regarding education. To date, it is well known that education influences cognitive and brain plasticity over the adult lifespan. Education has a protective effect on age-related decline (e.g., Christensen, Anstey, Leach, & Mackinnon, 2008; Kramer, Bherer, Colcombe, Dong, & Greenough, 2004; Raz et al., 2005). Thus, we investigated a cognitively rather high functioning population. Second, the majority of our sample was female (i.e., about 70% women), which did not represent the average population either, in which men and women are equally present. Third, given the fact that women even less frequently obtained a university-entrance diploma four to five decades ago, the preceding two points are particularly pronounced in the older sample. Moreover, the older sample was rather young, with a mean age around 68 years. Hence, we can assume that our older adults were high functioning and active individuals which were more likely to show a youth-like functioning than the average older population. Finally, although the sample size per cell for the behavioral measures was adequate (n = 20-23), the number of participants was reduced for the EEG subsample (n = 12) as well as for the high versus low performers analyses (n = 10 - 12). This reduces the statistical power of the analyses and increases the probability that the statistical tests will reject the true alternative hypothesis (increased the type II error). However, this point is usually hard to realize in training studies since training procedures require lots of time and resources.

Another critical point for generalizing our findings is the ceiling effect in the accuracy scores of the *N*-back tasks, which we already addressed in previous sections (see also

Appendix D). This is a very common problem in developmental and aging research, since the challenge is to find a task which covers the performance level of younger and older adults at the same time. Ceiling effects are also very common in cognitive neuroimaging research, especially in fMRI studies. Similar performance levels are desired in fMRI studies, since BOLD signals cannot be separated for correct and false responses. The overall behavioral performance has to be comparable across groups in order to be able to compare the cerebral results. In age-comparative studies for instance, this can only be reached by ceiling effects. There is a number of training studies including fMRI measures, where ceiling effects were observed in pre- and post-test tasks. However, they found reliable training-related cerebral changes despite the ceiling effects (Brehmer et al., 2011; S.-C. Li et al., 2008; Olesen et al., 2004; Schmiedek et al., 2010; Westerberg & Klingberg, 2007). Moreover, in EEG studies, even though the epochs for false responses can be eliminated, it is preferable to implement a simple task which generates similar accuracy levels across groups. Errors can lead to activation of areas related to emotional processing of failure-related anxiety and to a lack of activation if the participant abandons the task when errors become frequent (Poldrack, 2000).

However, the ceiling effect may have masked potential age differences in training gain and transfer effects since we did not find age differences in any of the gain scores. Especially younger adults may have shown more transfer effects if there had been room for a larger performance improvement from pre-test to post-test. Ceiling effects may have also biased the high versus low performers analyses, since low performers had more space for improvements than high performers.

For future studies, a good solution would be to include a more challenging 3-back condition besides the 2-back condition. Another and maybe more fine grained way to adapt the task difficulty to the different age groups might be the change in presentation time by manipulating the ISI or the stimuli presentation time. This would allow using a shorter presentation time for younger adults and a longer one for older adults.

A potential shortcoming of our analyses might be the use of absolute gain scores. By the calculation of the simple difference between pre-test and post-test, we did not correct for age differences at pre-test. One possibility may be to divide the pre-post-difference by the performance at pre-test. By this bias the initial differences in performance level can be corrected. However, this measure also has some shortcoming. It does for example not fit for error rates, since a division by zero is not possible.

Another point of discussion in our analyses is the use of median split for the creation of the high and low performer groups. Median splits have several shortcomings: Among others, the power to detect effects and effect sizes are reduced (MacCallum, Zhang, Preacher, & Rucker, 2002) and individuals performing around the median of the group are assigned to one group or the other by chance. To address the latter problem, one could divide the sample into terciles and eliminate the middle tercile while using the first and the third tercile as high and low performers groups respectively (Ludwig, 2005). This has the advantage that the probability that a participant would be assigned to a group by chance due to the intra-individual variability in task performance is reduced. However, our samples were too small for this method. Another way to address the difference in gain scores of high versus low performers would be to conduct regression analyses of pre-test performance on the gain scores. This analysis however would not allow the investigation of higher way interactions between the independent variables such as age group and training group in one single analysis (e.g., ANOVA).

A final statistical shortcoming is the potential alpha error inflation due to the multiple statistical tests we conducted. However, since we had specific hypothesis, we did not correct the significance level for the ANOVAs. We did in turn adjust the alpha level using Bonferroni corrections in follow-up pairwise comparisons for significant interactions.

A potentially important point missing in our study is the assessment of long-term effects which are considered very important in training literature (e.g., La Rue, 2010; Vance et al., 2007; Willis & Schaie, 2009). It would be interesting to investigate whether the observed behavioral and even more the ERP changes would be maintained after several month or a years. However, since our training was rather short, we did not expect long-term maintenance and will not investigate this point.

A conceptual difficulty was that we were not able to conclude whether the initially more distributed cerebral activity was a compensatory activation or a non-selective recruitment. However, we were able to conclude that the activity after the training was more beneficial than the one before the training, since younger and older adults increased behavioral performance. In order to disentangle these patterns at pre-test we should examine low versus high performers in both age groups in order to qualify the pattern at pre-test.

Future Research

Based on our findings and the above discussed limitations we now provide suggestions for future research. We propose a study in which several limitations are addressed and which may provide further insights into the interpretation of our findings, in particular the occurrence of increased and decreased control processes.

First, we propose to examine the cerebral activation patterns in an intermediate session after a few training sessions (e.g., 4-5 sessions). This may provide further knowledge about the evolution from a redistribution pattern with increased attentional recruitment towards a reorganization pattern with a decreased attentional recruitment in the trained task, since we hypothesize a scaffolding pattern in the course of training. In order to investigate if a reorganization pattern would also occur in the transfer task, an extension of the training procedure has to be considered. By this means, the evolution from a more attentional towards a less attentional processing could be investigated in the trained as well as in the transfer task.

Second, we suggest including a group of oldest-old adults aged 80+. This would allow investigating whether oldest-old adults also show similar behavioral and cerebral changes due to training as younger adults or whether these training gains are limited to young older adults (aged 65 to 80).

These points would allow to study the vicariance (Reuchlin, 1978) of compensational and less controlled processes in a more fine grained way. The moment of occurrence of these processes does probably differ between younger, young older and oldest-old adults. Similar processes exist in younger and older adults, but they might have different weights in the course of adult development. It is possible that older adults aged 80+ would maybe no longer develop a less controlled processing due to training, but rely more on increased attentional processes.

Third, in order to classify the patterns at pre-test, we would analyze high and low performers in EEG measures too, as proposed above. This would allow better understanding the cerebral patterns at pre-test and revealing which pattern is associated with relatively higher and lower performances. Consequently, this would provide the possibility to investigate how these patterns change over the training sessions.

Finally, regarding cerebral analyses, we would additionally conduct a source analysis of ERP data using an inverse solution. This would provide insights into differential changes on potential generators of the ERP and allow the more precise investigation of HAROLD and PASA patterns.

FRENCH SUMMARY

Généralement, les fonctions cognitives restent malléables chez la personne âgée. Néanmoins, leurs capacités ont tendance à diminuer avec l'âge, même si cela ne s'applique pas à toutes les fonctions dans la même mesure ou au même moment. Dans ce travail, nous avons cherché à fournir un élément important pour mieux comprendre le vieillissement cognitif et ainsi développer des moyens pour favoriser l'autonomie des personnes âgées. Nous avons étudié les mécanismes comportementaux et cérébraux qui sous-tendent un entraînement cognitif.

Le premier but de ce travail était de mieux comprendre le potentiel de la plasticité comportementale chez les adultes jeunes et âgés en mettant en œuvre un entraînement de la mémoire de travail (MDT), soit un entraînement de la tâche *N*-back verbal. Deuxièmement, nous avons argumenté le débat sur les effets de transfert, à savoir le bénéfice d'un entraînement sur une tâche non entraînée. En effet, l'existence, le degré et l'étendue du transfert ne sont pas clairs à l'heure actuelle. Troisièmement, nous avons cherché à obtenir un aperçu des mécanismes de la plasticité cérébrale sous-jacents aux effets d'un entraînement ainsi qu'aux effets de transfert. À cette fin, pendant les mesures de pré-test et post-test sur une tâche entraînée et une tâche de transfert proche, nous avons enregistré l'électroencéphalogramme (EEG). Par ce moyen, nous avons réussi à contribuer à la problématique de la signification fonctionnelle des modifications cérébrales pendant le vieillissement cognitif qui n'est pas encore résolu aujourd'hui, à savoir la sur-activation frontale, qui se produit naturellement avec l'âge. Nous avons également pour but d'améliorer plusieurs problèmes méthodologiques dans les études d'entraînement empêchant la bonne compréhension des effets d'un entraînement. Premièrement, il est important d'inclure un groupe de personnes jeunes comme groupe de référence pour mieux comprendre les effets chez la personne âgée. Deuxièmement, afin de contrôler les effets test-retest et les variables parasites, pouvant se produire dans un programme d'intervention, nous avons inclus à la fois un groupe contrôle sans contact ainsi qu'un groupe contrôle actif qui complète un entraînement dit placebo.

Or, nos hypothèses sont les suivantes : Premièrement, nous nous attendons à observer des changements plastiques dans les deux groupes d'âge au cours de l'entraînement, mais plus prononcés chez les adultes jeunes. Deuxièmement, nous émettons l'hypothèse de trouver des effets de l'entraînement dans les tâches non entraînées qui mesurent la capacité la MDT (transfert proximal). Troisièmement, nous avons testé l'hypothèse selon laquelle l'entraînement de la MDT pourrait générer du transfert distal qui se manifesterait dans une tâche d'intelligence fluide. Ces effets de transfert devraient apparaître dans les deux groupes d'entraînement de MDT, plus fortement chez les adultes jeunes. Au niveau cérébral, nous avons émis l'hypothèse de changements dans les composantes des potentiels évoqués liés à la prise de décision et au recrutement des ressources attentionnelles, à savoir les composantes N2 et P3. Nous nous attendions à une amplitude décroissante avec l'entraînement. En effet, nous avions émis l'hypothèse qu'après l'entraînement, moins de ressources attentionnelles sont nécessaires pour répondre à la tâche entraîné et de transfert proximal. Nous nous attendions également à trouver des cartes topographiques de voltage moins orientées vers le lobe frontal pour les composantes N2 et P3. L'amplitude de la N2 et P3 est généralement diminuée et moins frontalement orientée dans des tâches cognitivement moins exigeantes.

Afin de tester ces hypothèses, nous avons conduit un entraînement de la MDT en utilisant une procédure de *N*-back verbal pendant dix séances quotidiennes de trente minutes. Avant et après l'entraînement, nous avons administré une batterie de tests afin d'évaluer les changements entre le pré-test et le post-test. La batterie de tests cognitifs comprenait des mesures de MDT, d'inhibition, de vitesse de traitement et d'intelligence fluide. Nous avons inclus un groupe d'adultes jeunes (entre 18 et 38 ans, $n = 63$) et un groupe d'adultes âgés (entre 60 et 84 ans, $n = 65$). Toutes les séances se sont passées dans un laboratoire de l'Université de Genève. Afin de tester les effets d'un entraînement, un entraînement placebo a été mis sur pied. Ce dernier était composé d'une tâche implicite dans une procédure similaire de dix séances d'environ trente minutes par jour. En outre, nous avons inclus un groupe contrôle, qui n'a effectué aucun entraînement entre les 2 passations de la batterie de tests cognitifs (groupe sans contact). Des mesures EEG ont été enregistrées durant la passation de la tâche *N*-back verbale (tâche d'entraînement) et spatiale (transfert proximal).

Au niveau comportemental, les résultats ont montré que la performance des deux groupes d'âge s'est améliorée au cours des dix séances d'entraînement de la MDT. Les adultes jeunes présentaient un niveau d'entraînement généralement plus élevé que les adultes âgés. En plus, les adultes jeunes montrent une amélioration plus rapide lors de l'entraînement en atteignant l'asymptote plus tard que les adultes âgés. Cela se traduit par une différence entre les deux groupes d'âge qui a augmenté à la dernière session de l'entraînement. D'autres analyses ont montré, qu'indépendamment du groupe d'âge, la performance dans la tâche d'intelligence fluide explique les différences individuelles dans le niveau de performance initial dans l'entraînement. En revanche, en ce qui concerne la courbe de croissance, nous avons montré que l'âge explique d'avantage les différences individuelles que l'intelligence fluide.

Concernant la deuxième hypothèse, un effet d'entraînement a été observé seulement dans le groupe d'entraînement (adultes jeunes et âgés) de la MDT pour la tâche 2-back verbale. Les coûts de la charge en MDT, c'est à dire la différence entre la condition avec charge faible 0-back et la condition avec charge élevée 2-back, ont également diminué de façon significative dans les groupes d'entraînement MDT comparé aux groupes contrôle. Les mêmes effets d'entraînement ont été trouvés chez les adultes jeunes et les adultes âgés. Un effet de transfert proximal a été trouvé dans la tâche *N*-back spatial, à savoir dans la tâche 2-back ainsi que pour les coûts de charge en MDT. Comme pour la tâche verbale, les effets étaient comparables pour les deux groupes d'âge dans la tâche spatiale. Aucun effet de transfert dans les autres mesures de la MDT, de l'intelligence fluide ou pour les tâches de vitesse de traitement ont été trouvées.

Les analyses des potentiels évoqués pour la tâche 2-back verbal ont révélé une augmentation de l'activation pour la composant N2 et une diminution de l'activation pour la composante P3 dans les deux groupes d'âge. Cela s'est traduit par un changement de positivité frontale vers une négativité centrale et une positivité postérieure. Dans l'ensemble, l'activation est devenue moins répartie sur le cerveau, tendant vers un recrutement plus sélectif dans les régions centrales et postérieures. Ces changements ont été trouvés pour décrire un modèle de réorganisation. Nous avons observé la tendance inverse pour la tâche non-entraînée 2-back spatial: Les sites frontaux ont montré plus de positivité de la composante P3 après l'entraînement pour les groupes d'entraînement de la MDT alors que les cartes orientées postérieurement ont diminué en présence. Il en résulte un pattern de redistribution, ce qui indique que des processus similaires ont été engagés dans la tâche 2-back spatial avant et après l'entraînement, néanmoins plus efficaces après l'entraînement. En revanche, les résultats de la tâche verbale ont révélé un pattern de réorganisation indiquant un changement de processus mises en œuvre du pré-test au post-test.

Les résultats de la performance au cours de l'entraînement sont en accord avec nos hypothèses, étant donné que les différences de performance dans la tâche *N*-back verbale entre les groupes d'âge ont augmentées après l'entraînement. En ce qui concerne les résultats de la tâche entrainée et de transfert, nous avons trouvé les gains attendus dans la tâche entraînée et dans une tâche de transfert proximal (*N*-back spatial). Par contre, aucun effet de transfert n'a été observé dans d'autres tâches MDT et dans les mesures d'intelligence fluide. Ces résultats sont en accord avec un nombre croissant de publications qui ont montré que la plasticité comportementale chez les personnes âgées reste préservée, même si elle est atténuée. En outre, les effets de transfert d'un entraînement MDT à une mesure de l'intelligence fluide sont actuellement en cours de débat, car les résultats sont encore contradictoires.

Les résultats des potentiels évoqués ont révélé que la plasticité cérébrale a également été préservée chez les personnes âgées. Nous étions en mesure de relier nos résultats avec la scaffolding theory, qui stipule que les régions frontales sont plus activées à cause d’un recrutement accru des ressources attentionnelles afin de répondre aux exigences croissantes au cours de l'apprentissage. Lorsque ces demandes persistent, les processus deviennent plus efficaces ce qui se reflète dans la diminution du recrutement frontale. La première étape, c’est-à-dire une augmentation de l’activité frontale était trouvée dans la tâche de transfert proximale. Pour la tâche entraînée, en revanche, une diminution de l’activité frontale dans la composante P3 et un processus précoce plus efficace indiqué par la composante N2 ont été observée.

Les résultats alimentent le débat sur l'interprétation des changements liés au vieillissement cérébral. Ils ont révélé qu'un changement similaire chez les personnes âgées par rapport aux adultes jeunes dans les processus cérébraux semble être bénéfique. Autrement dit, nous n'étions pas en mesure de montrer un changement différentiel dû à l’entraînement chez les adultes jeunes et les adultes âgés. Ainsi, nous avons conclu que la sur-activation frontale chez les adultes âgés par rapport aux adultes jeunes ne décrit pas nécessairement un fonctionnement bénéfique.

Dans l'ensemble, cette thèse fournit des preuves que les études d’entraînement ont un grand potentiel pour contribuer à la compréhension des mécanismes du vieillissement cognitif. Par ailleurs, nos résultats appellent à l'intégration des études d’entraînement dans des contextes appliquées en ciblant sur les effets qui pourrait être généralisés dans la vie quotidienne des personnes âgées.

REFERENCES

Ackerman, P. L., Kanfer, R., & Calderwood, C. (2010). Use it or lose it? Wii brain exercise practice and reading for domain knowledge. *Psychology and Aging, 25*(4), 753-766.

Anderer, P., Semlitsch, H. V., & Saletu, B. (1996). Multichannel auditory event-related brain potentials: effects of normal aging on the scalp distribution of N1, P2, N2 and P300 latencies and amplitudes. *Electroencephalography and Clinical Neurophysiology, 99*(5), 458-472.

Angel, L., Fay, S., Bouazzaoui, B., Baudouin, A., & Isingrini, M. (2010). Protective role of educational level on episodic memory aging: An event-related potential study. *Brain and Cognition, 74*(3), 312-323.

Bach, M. (1996). The "Freiburg Visual Acuity Test" - Automatic measurement of visual acuity. *Optometry and Vision Science, 73*, 49-53.

Bäckman, L., & Nyberg, L. (2010). Dopamine, cognition, and human aging: New evidence and ideas. In L. Bäckman & L. Nyberg (Eds.), *Memory, Aging and the Brain* (pp. 190-208). Hove: Psychology Press.

Bäckman, L., Nyberg, L., Lindenberger, U., Li, S.-C., & Farde, L. (2006). The correlative triad among aging, dopamine, and cognition: Current status and future prospects. *Neuroscience and Biobehavioral Reviews, 30*, 791–807.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). New York: Academic Press.

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., et al. (2002). Effects of cognitive training interventions with older adults: A randomized controlled trial. *Journal of American Medical Association, 288*, 2271-2281.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology on the dynamics between growth and decline. *Developmental Psychology, 23*, 611-626.

Baltes, P. B. (1997). On the incomplete architecture of human ontogeny: Selection, optimization, and compensation as foundation of developmental theory. *American Psychologist, 52*(4), 366-380.

Baltes, P. B., & Kliegl, R. (1992). Further testing of limits of cognitive plasticity: Negative age differences in a mnemonic skill are robust. *Developmental Psychology, 28*.

Baltes, P. B., Kliegl, R., & Dittmann-Kohli, F. (1988). On the Locus of Training Gains in Research on the Plasticity of Fluid Intelligence in Old Age. *Journal of Educational Psychology, 80*, 392–400.

Baltes, P. B., & Lindenberger, U. (1988). On the range of cognitive plasticity in old age as a function of experience: 15 years of intervention research. *Behavior therapy, 19*(3), 283-300.

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology and Aging, 12*, 12-21.

Baltes, P. B., Lindenberger, U., & Staudinger, U. M. (1998). Life-span theory in developmental psychology. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development* (Vol. 1, pp. 1029–1041). New York: Wiley.

Baltes, P. B., Lindenberger, U., & Staudinger, U. M. (2006). Lifespan theory in developmental psychology. In R. M. L. W. Damon (Ed.), *Handbook of child psychology: Theoretical models of human development* (Vol. 6, pp. 569-664). New York: Wiley.

Baltes, P. B., & Mayer, K. U. (1999). *The Berlin aging study: aging form 70 to 100*. New York: Cambridge University Press.

Bartzokis, G., Cummings, J. L., Sultzer, D., Henderson, V. W., Nuechterlein, K. H., & Mintz, J. (2003). White matter structural integrity in healthy aging adults and patients with Alzheimer disease: a magnetic resonance imaging study. *Archives of Neurology, 60*(3), 393.

Bellgrove, M. A., Hester, R., & Garavan, H. (2004). The functional neuroanatomical correlates of response variability:evidence from a response inhibition task. *Neuropsychologia*, *42*, 1910-1916.

Bissig, D., & Lustig, C. (2007). Training Memory: Who Benefits From Memory Training? *Psychological Science*, *8*, 720-726.

Borella, E., Carretti, B., Riboldi, F., & Beni, R. D. (2010). Working memory training in older adults: evidence of transfer and maintenance effects. *Psychology and Aging, 25*, 767-778.

Borella, E., Delaloye, C., Lecerf, T., Renaud, O., & de Ribaupierre, A. (2009). Do age differences between young and older adults in inhibitory tasks depend on the degree of activation of information? *The European Journal of Cognitive Psychology, 21*, 445-472.

Borella, E., Ghisletta, P., & de Ribaupierre, A. (2011). Age differences in text processing: The role of working memory, inhibition, and processing speed. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *66*(3), 311.

Boyke, J., Driemeyer, J., Gaser, C., Büchel, C., & May, A. (2008). Training-induced brain structure changes in the elderly. *The Journal of Neuroscience, 28*(28), 7031.

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A Parametric Study of Prefrontal Cortex Involvement in Human Working Memory. *NeuroImage, 5*(1), 49-62.

Brehmer, Y., Li, S.-C., Müller, V., Oertzen, T. v., & Lindenberger, U. (2007). Memory plasticity across the life span: Uncovering children’s latent potential. *Developmental Psychology, 43*, 465-478.

Brehmer, Y., Li, S.-C., Straube, B., Stoll, G., von-Oertzen, T., Muller, V., et al. (2008). Comparing memory skill maintenance across the life span: Preservation in adults, increase in children. *Psychology and Aging, 23*, 227-238.

Brehmer, Y., Rieckmann, A., Bellander, M., Westerberg, H., Fischer, H., & Bäckman, L. (2011). Neural correlates of training-related working-memory gains in old age. *NeuroImage, 58*(4), 1110-1120.

Brehmer, Y., Westerberg, H., & Bäckman, L. (2012). Working-memory training in younger and older adults: Training gains, transfer, and maintenance. *Frontiers in human neuroscience*.

Brooks, J. O., Friedman, L., Pearman, A. M., Gray, C., & Yesavage, J. A. (1999). Mnemonic Training in Older Adults: Effects of Age, Length of Training, and Type of Cognitive Pretraining. *International-Psychogeriatrics*, *11*, 75-84.

Brunet, D., & Murray, M. (2011). Spatiotemporal Analysis of Multichannel EEG: CARTOOL. *Computational Intelligence and Neuroscience*, *2011*.

Bundesamt für Statistik. (2012). Gymnasiale Maturitätsquote. Neuenburg, Switzerland: Retrieved July 7, 2012 from http://www.bfs.admin.ch/bfs/portal/de/index/themen/15/06/dos/blank/05/01.html.

Buschkuehl, M. (2007). *Arbeitsgedächtnistraining: Untersuchungen mit jungen und älteren Erwachsenen.* Unpublished Doctoral Thesis, University of Bern, Bern, Switzerland.

Buschkuehl, M., & Jaeggi, S. M. (2010). Improving intelligence: a literature review. *Swiss medical weekly: official journal of the Swiss Society of Infectious Diseases, the Swiss Society of Internal Medicine, the Swiss Society of Pneumology*.

Buschkuehl, M., Jaeggi, S. M., Hutchison, S., Perrig-Chiello, P., Däpp, C., Müller, M., et al. (2008). Impact of working memory training on memory performance in old-old adults. *Psychology and Aging, 23,* 743-753.

Buschkuehl, M., Jaeggi, S. M., & Jonides, J. (2012). Neuronal effects following working memory training. *Developmental Cognitive Neuroscience, 2, Supplement 1*(0), S167-S179.

Cabeza, R. (2002). Hemispheric Asymmetry Reduction in Older Adults, The HAROLD Model. *Psychology and Aging, 17,* 85-100.

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging Gracefully: Compensatory Brain Activity in High-Performing Older Adults. *NeuroImage, 17,* 1394-1402.

Cabeza, R., Daselaar, S. M., Dolcos, F., Prince, S. E., Budde, M., & Nyberg, L. (2004). Task-independent and Task-specific Age Effects on Brain Activity during Working Memory, Visual Attention and Episodic Retrieval. *Cerebral Cortex, 14,* 364-375.

Cabeza, R., Grady, C. L., Nyberg, L., McIntosh, A. R., Tulving, E., Kapur, S., et al. (1997). Age-Related Differences in Neural Activity during Memory Encoding and Retrieval: A Positron Emission Tomography Study. *The Journal of Neuroscience, 17,* 391-400.

Cabeza, R., Nyberg, L., & Park, D. C. (2005). *Cognitive neuroscience of aging: Linking cognitive and cerebral aging*: Oxford University Press, USA.

Carp, J., Gmeindl, L., & Reuter-Lorenz, P. A. (2010). Age differences in the neural representation of working memory revealed by multi-voxel pattern analysis. *Frontiers in human neuroscience, 4*.

Carp, J., Park, J., Polk, T. A., & Park, D. C. (2011). Age differences in neural distinctiveness revealed by multi-voxel pattern analysis. *NeuroImage, 56*(2), 736-743.

Carretti, B., Cornoldi, C., & Pelegrina, S. L. (2007). Which factors influence number updating in working memory? The effects of size distance and suppression. *British Journal of Psychology, 98,* 45-60.

Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: training and transfer effects with a complex working memory span task. *Psychonomic Bulletin & Review, 17,* 193-199.

Chen, Y. N., & Mitra, S. (2009a). Distinctions between Spatial and Verbal Working Memory: A Study Using Event-related Potentials. *Chang Gung Med J, 32*(4), 380-389.

Chen, Y. N., & Mitra, S. (2009b). 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.

Chicherio, C. (2006). *Contrôle exécutif et réseaux neurofonctionnels au cours du vieillissement normal: Un test de l'hypothèse de dé-différenciation cognitive.* Unpuplished Doctoral Thesis, Université de Genève, Genève, Suisse.

Christensen, H., Anstey, K. J., Leach, L. S., & Mackinnon, A. J. (2008). Intelligence, education, and the brain reserve hypothesis. In F. I. M. Craik & T. A. Salthouse (Eds.), *The Handbook of Aging and Cognition* (3 ed., pp. 133-188). New York; Hove: Psychology Press.

Church, J. A., Coalson, R. S., Lugar, H. M., Petersen, S. E., & Schlaggar, B. L. (2008). A Developmental fMRI Study of Reading and Repetition Reveals Changes in Phonological and Visual Mechanisms Over Age. *Cerebral Cortex, 18*(9), 2054-2065.

Cohen, J. D. (1988). *Statistical power analysis for the behavioral sciences*: Lawrence Erlbaum.

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., et al. (1997). Temporal dynamics of brain activation during a working memory task. *Nature, 386*(6625), 604-608.

Colom, R., Quiroga, M. Ã., Shih, P. C., MartÃnez, K., Burgaleta, M., MartÃnez-Molina, A., et al. (2010). Improvement in working memory is not related to increased intelligence scores. *Intelligence, 38*(5), 497-505.

Content, A., Mousty, P., & Radeau, M. (1990). Brulex, une base de données lexicales informatisées pour le français parlé et écrit. *L'Année Psychologique, 90*, 551-566.

Conway, A. R. A., & Getz, S. J. (2010). Cognitive Ability: Does Working Memory Training Enhance Intelligence? *Current Biology, 20*(8), R362-R364.

D'Esposito, M., Zarahn, E., Aguirre, G. K., & Rypma, B. (1999). The effect of normal aging on the coupling of neural activity to the bold hemodynamic response. *NeuroImage, 10*(1), 6-14.

Daffner, K. R., Chong, H., Sun, X., Tarbi, E. C., Riis, J. L., McGinnis, S. M., et al. (2011). Mechanisms underlying age-and performance-related differences in working memory. *Journal of Cognitive Neuroscience, 23*(6), 1298-1314.

Daffner, K. R., Sun, X., Tarbi, E. C., Rentz, D. M., Holcomb, P. J., & Riis, J. L. (2010). Does compensatory neural activity survive old-old age? *NeuroImage, In Press, Corrected Proof*.

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*(5), 405-419.

Dahlin, E., Nyberg, L., Bäckman, L., & Neely, A. S. (2008). Plasticity of executive functioning in young and older adults: Immediate training gains, transfer, and long-term maintenance. *Psychology and Aging, 23*, 720–730.

Dahlin, E., Stigsdotter Neely, A., Larsson, A., Bäckman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science, 320*, 1210-1212.

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

Daselaar, S. M., Rombouts, S. A. R. B., Veltman, D. J., Raaijmakers, J. G. W., & Jonker, C. (2003). Similar network activated by young and old adults during the acquisition of a motor sequence. *Neurobiology of Aging, 24*(7), 1013-1019.

Davis, S. W., Dennis, N. A., Buchler, N. G., White, L. E., Madden, D. J., & Cabeza, R. (2009). Assessing the effects of age on long white matter tracts using diffusion tensor tractography. *Neuroimage, 46*(2), 530-541.

Davis, S. W., Dennis, N. A., Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2008). Qué PASA? The Posterior--Anterior Shift in Aging. *Cerebral Cortex, 18*, 1201-1209.

de Ribaupierre, A. (2001). Working memory and attentional processes across the lifespan. In P. Graf & N. Otha (Eds.), *Lifespan development of human memory* (pp. 59-80). Cambridge, MA: MIT Press.

de Ribaupierre, A. (2005). Développement et vieillissement cognitifs. In J. L. J.-F. Richard (Ed.), *L'Intelligence* (pp. 211-226). Paris: Lavoisier.

de Ribaupierre, A., & Bailleux, C. (1995). Development of attentional capacity in childhood: A longitudinal study. In F. E. Weinert & W. Schneider (Eds.), *Memory performance and competencies: Issues in growth and development* (pp. 45-70). Mahwah: Lauwrence Erlbaum.

de Ribaupierre, A., Fagot, D., & Dirk, J. (2009). Déclin et plasticité des fonctions cognitives avec l’âge: une question de différences individuelles? In M. Oris, E. Widmer, A. de

Ribaupierre, D. Joye, D. Spini, G. Labouvie-Vief & J.-M. Falter (Eds.), *Transitions dans le parcours de vie et construction des inégalités* (pp. 313-333). Lausanne, Switzerland: Presses Polytechniques et Universitaires Romandes.

de Ribaupierre, A., Fagot, D., & Lecerf, T. (2011). Working memory capacity and its role in cognitive development. In P. Barrouillet & V. Gaillard (Eds.), *Cognitive development and working memory* (pp. 105-133). East Sussex: Psychology Press.

de Ribaupierre, A., Ghisletta, P., & Lecerf, T. (2006). Etude de la variabilité inter- et intra-individuelle au cours du cycle de vie. In R. M. C. Houssemand, & P. Dickes (Ed.), *Perspectives de psychologie différentielle* (pp. 169-193). Rennes: Presse Universitaires de Rennes.

de Ribaupierre, A., Ghisletta, P., Lecerf, T., Bürki, C., Dirk, J., Fagot, D., et al. (2008). *Inter- and intraindividual variability across the lifespan. Final report, Swiss National Science Foundation, Grant N° 100011-107764*. Geneva: University of Geneva.

de Ribaupierre, A., Lecerf, T., Leutwyler, J., & Poget, L. (1997). *Mémoire de travail et attention sélective au cours du life-span. Final report, Swiss National Science Foundation, Grant N°1114-040465*. Geneva: University of Geneva.

de Ribaupierre, A., & Ludwig, C. (2003). Age differences and divided attention: Is there a general deficit? *Experimental Aging Research*, *29*, 79-105.

de Ribaupierre, A., Olivier, M., Leutwyler, J., & Poget, L. (1999). Attention sélective et mémoire de travail au cours du lifespan. In M. H. J. Lautrey (Ed.), *Approches différentielles en psychologie* (pp. 33-40). Rennes: Presses Universitaires de Rennes.

de Ribaupierre, A., Poget, L., & Pons, F. (2005). The Age Variable in Cognitive Developmental Psychology. In C. Sauvain-Dugerdil, H. Leridon & N. Mascie-Taylor (Eds.), *Human Clocks. The Bio-Cultural Meanings of Age* (pp. 101-123). Bern, Switzerland: Peter Lang.

de Ribaupierre, A., Pons, F., & Poget, L. (2003). L'âge en psychologie du développement cognitif : Une variable explicative ? In J.-P. F. S. Cavalli (Ed.), *L'avenir : Attentes, projets (dés)illusions, ouvertures. Hommages à Christian Lalive d'Epinay* Lausanne: Réalités sociales.

Delaloye, C., Ludwig, C., Borella, E., Chicherio, C., & de Ribaupierre, A. (2008). L'Empan de lecture comme épreuve mesurant la capacité de mémoire de travail: normes basées sur une population francophone de 775 adultes jeunes et âgés. *Revue européenne de psychologie appliquée*, *58*, 89–103.

Della-Maggiore, V., Grady, C. L., & McIntosh, A. R. (2002). Dissecting the effect of aging on the neural substrates of memory: Deterioration, preservation or functional reorganization? *Reviews in Neurosciences, 13*, 167-181.

Deltour, J. J. (1993). *Echelle de Vocabulaire Mill Hill de J. C. Raven*. Braine de Chateau, Belgium: Editions L'Application des Techniques Modern S.P.R.L.

Dennis, A. D., & Cabeza, R. (2008). Neuroimaging of healthy cognitive aging. In F. I. Craik & T. A. Salthouse (Eds.), *The Handbook of Aging and Cognition* (3 ed., pp. 1-54). New York, Hove: Psychology Press.

Dien, J., & Santuzzi, A. M. (2005). Application of repeated measures ANOVA to high-denstiy ERP datasets: A review and tutorial. In T. C. Handy (Ed.), *Event-related potentials: a methods handbook*. Cambridge: MIT Press.

Dixit, N. K., Gerton, B. K., Kohn, P., Meyer-Lindenberg, A., & Berman, K. F. (2000). Age-related changes in rCBF activation during an N-back working memory paradigm occur prior to age 50. *NeuroImage*, *11*(5, Supplement 1), S94-S94.

Dolcos, F., Rice, H. J., & Cabeza, R. (2002). Hemispheric asymmetry and aging: right hemisphere decline or asymmetry reduction. *Neuroscience and Biobehavioral Reviews, 26*, 819-825.

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating?

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Changes in grey matter induced by training. *Nature, 427*, 311-312.

Driemeyer, J., Boyke, J., Gaser, C., Büchel, C., & May, A. (2008). Changes in gray matter induced by learning—revisited. *PLoS One, 3*(7), e2669.

Duarte, A., Ranganath, C., Trujillo, C., & Knight, R. T. (2006). Intact Recollection Memory in High-performing Older Adults: ERP and Behavioral Evidence. *Journal of Cognitive Neuroscience, 18*(1), 33-47.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and application* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Edwards, J. D., Wadley, V. G., Myers, R. S., Roenker, D. L., Cissell, G. M., & Ball, K. K. (2002). Transfer of a Speed of Processing: Intervention to Near and Far Cognitive Functions. *Gerontology, 48*, 329-340.

Einstein, G. O., & McDaniel, M. A. (1996). Retrieval processes in prospective memory: Theoretical approaches and some new empirical findings. *Prospective memory: Theory and applications*, 115-141.

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.

Erickson, K. I., Colcombe, S. J., Wadhwa, R., Bherer, L., Peterson, M. S., Scalf, P. E., et al. (2007a). Training-Induced Functional Activation Changes in Dual-Task Processing: An fMRI Study. *Cerebral Cortex, 17*, 192-204.

Erickson, K. I., Colcombe, S. J., Wadhwa, R., Bherer, L., Peterson, M. S., Scalf, P. E., et al. (2007b). Training-induced plasticity in older adults: Effects of training on hemispheric asymmetry. *Neurobiology of Aging, 28*, 272-283.

Fabiani, M., Friedman, D., & Cheng, J. C. (1998). Individual differences in P3 scalp distribution in older adults, and their relationship to frontal lobe function. *Psychophysiology, 35*(6), 698-708.

Fabiani, M., & Gratton, G. (2005). Electrophysiological and Optical Measures of Cognitive Aging. In D. Park, R. Cabeza & L. Nyberg (Eds.), *Cognitive neuroscience of aging: Linking cognitive and cerebral aging* (pp. 85-106). New York: Oxford University Press.

Fingelkurts, A. A. (2006). Timing in cognition and EEG brain dynamics: discreteness versus continuity. *Cognitive processing, 7*(3), 135-162.

Fjell, A. M., Rosquist, H., & Walhovd, K. B. (2009). Instability in the latency of P3a/P3b brain potentials and cognitive function in aging. *Neurobiology of Aging, 30*(12), 2065-2079.

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology, 45*(1), 152-170.

Ford, J. M., Roth, W. T., Mohs, R. C., Hopkins, W. F., & Kopell, B. S. (1979). Event-related potentials recorded from young and old adults during a memory retrieval task. *Electroencephalography and Clinical Neurophysiology, 47*(4), 450-459.

Fotenos, A. F., Snyder, A. Z., Girton, L. E., Morris, J. C., & Buckner, R. L. (2005). Normative estimates of cross-sectional and longitudinal brain volume decline in aging and AD. *Neurology, 64*(6), 1032-1039.

Friedman, D. (2003). Cognition and Aging: A Highly Selective Overview of Event-Related Potential (ERP) Data. *Journal of Clinical and Experimental Neuropsychology, 25*, 702-720.

Friedman, D. (2008). The components of aging. In S. J. Luck & E. S. Kappenman (Eds.), *Oxford Handbook of Event-Related Potential Components*. New York: Oxford University Press.

Friedman, D., & Johnson, R. (2000). Event-Related Potential (ERP) Studies of Memory Encoding and Retrieval: A Selective Review. *Microscopy Research and Technique 51,* 6-28.

Friedman, D., & Simpson, G. V. (1994). ERP amplitude and scalp distribution to target and novel events: Effects of temporal order in young, middle-aged and older adults. *Cognitive Brain Research, 2*(1), 49-63.

Gaillard, V., Destrebecqz, A., Michiels, S., & Cleeremans, A. (2009). Effects of age and practice in sequence learning: A graded account of ageing, learning, and control. *European Journal of Cognitive Psychology, 21*(2), 255-282.

Gajewski, P. D., Stoerig, P., & Falkenstein, M. (2008). ERP-Correlates of response selection in a response conflict paradigm. *Brain Research, 1189*(0), 127-134.

Gazzaniga, M. S. (2004). *The cognitive neurosciences*: The MIT Press.

Gevins, A., Smith, M. E., Le, J., Leong, H., Bennett, J., Martin, N., et al. (1996). High resolution evoked potential imaging of the cortical dynamics of human working memory. *Electroencephalography and clinical Neurophysiology, 98,* 327-348.

Ghisletta, P., Lecerf, T., & de Ribaupierre, A. (2003). Etude de la variabilité inter- et intra-individuelle dans des épreuves d'inhibition, de mémoire de travail et de vitesse de traitement. In H. C. A. vom Hofe, J.L. Bernaud, & D. Guédon (Ed.), *Psychologie différentielle: recherches et réflexions* (pp. 217-221). Rennes: Presses Universitaires de Rennes.

Ghisletta, P., & Lindenberger, U. (2003). Age-Based Structural Dynamics Between Perceptual Speed and Knowledge in the Berlin Aging Study: Direct Evidence for Ability Dedifferentiation in Old Age. *Psychology and Aging, 18*, 696-713.

Goh, J. O. S. (2011). Functional dedifferentiation and altered connectivity in older adults: neural accounts of cognitive aging. *Aging and disease, 2*(1), 30.

Goh, J. O. S., & Park, D. C. (2009). Neuroplasticity and cognitive aging: The scaffolding theory of aging and cognition. *Restorative Neurology and Neuroscience, 27*, 391–403.

Goldberg, T. E., Keefe, R. S. E., Goldman, R. S., Robinson, D. G., & Harvey, P. D. (2010). Circumstances Under Which Practice Does Not Make Perfect: A Review of the Practice Effect Literature in Schizophrenia and Its Relevance to Clinical Treatment Studies. *Neuropsychopharmacology, 35*(5), 1053-1062.

Grady, C. L., Maisog, J. M., Horwitz, B., Ungerleider, L. G., Mentis, M. J., Salerno, J. A., et al. (1994). Age-related changes in cortical blood flow activation during visual processing of faces and location. *Journal of Neuroscience, 14*(3, Pt 2), 1450-1462.

Green, C. S., & Bavelier, D. (2008). Exercising your brain: A review of human brain plasticity and training-induced learning. *Psychology and Aging, 23*, 692-701.

Greenwood, P. M., & Parasuraman, R. (2010). Neuronal and cognitive plasticity: A neurocognitive framework for ameliorating cognitive aging. *Frontiers in Aging Neuroscience, 2*.

Hager, W., & Hasselhorn, M. (1998). The effectiveness of the cognitive training for children from a differential perspective: A Metaevaluation. *Learning and Instruction, 8*, 411–438.

Handy, T. C. (2005). Basic principles of ERP quantification. In T. C. Handy (Ed.), *Event-related potentials: a methods handbook* (pp. 33-55). Cambridge: MIT Press.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(1), 163-169.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193-225). San Diego, CA, US: Academic Press.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention and Performance Xvii - Cognitive Regulation of Performance: Interaction of Theor and Application* (Vol. 17, pp. 653-675).

Hauert, C.-A. (2004). Apprentissage moteur et approche neuro-comportementale. In J.-P. Didier (Ed.), *La plasticité de la fonction motrice: Bases et management de la restauration, de la substitution et de la compensation* (pp. 317-340). Berlin: Springer Verlag.

Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: A view from cognitive neuroscience. *Nature reviews neuroscience, 5*, 87-96.

Hempel, A., Giesel, F. L., Caraballo, N. M. G., Amann, M., Meyer, H., Wüstenberg, T., et al. (2004). Plasticity of Cortical Activation Related to Working Memory During Training. *American Journal of Psychiatry*, *161*, 745–747.

Hennighausen, E., Heil, M., & Rösler, F. (1993). A correction method for DC drift artifacts. *Electroencephalography and Clinical Neurophysiology*, *86*(3), 199-204.

Hertzog, C., & Dunlosky, J. (2012). Metacognitive approaches can promote transfer of training: Comment on McDaniel and Bugg. *Journal of Applied Research in Memory and Cognition*, *1*(1), 61-63.

Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2009). Enrichment Effects on Adult Cognitive Development: Can the functional capacity of older adults be preserved and enhanced? *Psychological Science in the Public Interest*, *9*, 1-65.

Heuninckx, S., Wenderoth, N., & Swinnen, S. P. (2008). Systems neuroplasticity in the aging brain: Recruiting additional neural resources for successful motor performance in elderly persons. *Journal of Neuroscience*, *28*, 91-99.

Hof, P. R., & Morrison, J. H. (2004). The aging brain: morphomolecular senescence of cortical circuits. *Trends in neurosciences*, *27*(10), 607-613.

Horn, J.-L., & Cattell, R.-B. (1967). Age differences in fluid and crystallized intelligence. *Acta-Psychologica*, *26*, 107-129.

Howard, D. V., & Howard, J. H., Jr. (1992). Adult age differences in the rate of learning serial patterns: Evidence from direct and indirect tests. *Psychology and Aging*, *7*(2), 232.

Howard, D. V., & Howard, J. H., Jr. (in press). Dissociable forms of implicit learning in aging. In M. N.-B. N. Ohta (Ed.), *Perspectives on human memory and aging*. London: Psychology Press.

Howard, D. V., Howard, J. H., Jr., Japikse, K., DiYanni, C., Thompson, A., & Somberg, R. (2004). Implicit Sequence Learning: Effects of Level of Structure, Adult Age, and Extended Practice. *Psychology and Aging*, *19*(1), 79-92.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55.

Hultsch, D. F., MacDonald, S. W. S., Hunter, M. A., Levy-Bencheton, J., & Strauss, E. (2000). Intraindividual variability in cognitive performance in older adults: Comparison of adults with mild dementia, adults with arthritis, and healthy adults. *Neuropsychology*, *14*, 588-598.

Hurley, D. (2012). Can You Make Yourself Smarter? *The New York Times, April 22*, p. 38.
Jaeggi, S. M. (2012). *Improving intelligence by means of cognitive training – current challenges and potential resolutions*. Paper presented at the Formation continue organisée par le service de Neurorééducation, HUG, Geneva. .
Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proc. Natl. Acad. Sci. USA, 105*, 6829–6833.
Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short-and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences, 108*, 10081-10086.
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.
Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., Su, Y.-F., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning - Implications for training and transfer. *Intelligence, 38*, 625-635.
James, C. E., Britz, J., Vuilleumier, P., Hauert, C. A., & Michel, C. M. (2008). Early neuronal responses in right limbic structures mediate harmony incongruity processing in musical experts. *NeuroImage, 42*, 1597-1608.
Jäncke, L. (2009). The plastic human brain. *Restorative Neurology and Neuroscience, 27*(5), 521-538.
Jennings, J. M., & Jacoby, L. L. (2003). Improving memory in older adults: Training recollection. *Neuropsychological Rehabilitation, 13*, 417-440.
Johnson, M. H. (2001). Functional brain development in humans. *Nature Reviews Neuroscience, 2*(7), 475-483.
Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., et al. (1997). Verbal working memory load affects regional brain activation as measured by PET. *Journal of Cognitive Neuroscience, 9*(4), 462-475.
Jonides, J., & Smith, E. E. (1997). The architecture of working memory. In M. D. Rugg (Ed.), *Cognitive neuroscience* (pp. 243–276). Cambridge, MA: MIT Press.
Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M. J., Iragui, V., et al. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology, 37*(2), 163-178.
Kaasinen, V., Vilkman, H., Hietala, J., Någren, K., Helenius, H., Olsson, H., et al. (2000). Age-related dopamine D2/D3 receptor loss in extrastriatal regions of the human brain. *Neurobiology of Aging, 21*(5), 683-688.
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*(6), 978-990.
Karni, A., Meyer, G., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature, 377*, 155-158.
Kelly, A. M. C., Foxe, J. J., & Garavan, H. (2006). Patterns of normal human brain plasticity after practice and their implications for neurorehabilitation. *Arch Phys Med Rehabil, 87*, 20-29.
Kelly, A. M. C., & Garavan, H. (2005). Human functional neuroimaging of brain changes associated with practice. *Cerebral Cortex, 15*, 1089-1102.
Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology, 55*(4), 352-358.
Kliegel, M., & Bürki, C. (2012). Memory training interventions require a tailor-made approach. *Journal of Applied Research in Memory and Cognition, 1*, 58–60.

Kliegel, M., Martin, M., McDaniel, M. A., & Phillips, L. H. (2007). Adult age differences in errand planning: The role of task familiarity and cognitive resources. *Experimental Aging Research, 33*, 145–161.

Kliegl, R., Smith, J., & Baltes, P. B. (1989). Testing-the-Limits and the Study of Adult Age Differences in Cognitive Plasticity of a Mnemonic Skill. *Developmental Psychology, 25*, 247-256.

Kliegl, R., Smith, J., & Baltes, P. B. (1990). On the locus and process of magnification of age differences during mnemonic training. *Developmental Psychology, 26*(6), 894-904.

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences, 14*(7), 317-324.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., et al. (2005). Computerized training of working memory in children with ADHD - A randomized, controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*, 177-186.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002a). Increased brain activity in frontal and parietal cortex underlies the development of visuospatial working memory capacity during childhood. *Journal of Cognitive Neuroscience, 14*(1), 1-10.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002b). Training of Working Memory in Children With ADHD. *Journal of Clinical and Experimental Neuropsychology, 24*, 781-791.

Kramer, A. F., Bherer, L., Colcombe, S. J., Dong, W., & Greenough, W. T. (2004). Environmental influences on cognitive and brain plasticity during aging. *The Journals of Gerontology: Biological Sciences and Medical Sciences, 59*, 940-957.

Kramer, A. F., & Erickson, K. I. (2007). Capitalizing on cortical plasticity: influence of physical activity on cognition and brain function. *Trends in Cognitive Sciences, 11*, 342-348.

La Rue, A. (2010). Healthy Brain Aging: Role of Cognitive Reserve, Cognitive Stimulation, and Cognitive Exercises. *Clinics in Geriatric Medicine, 26*(1), 99-+.

Lamm, C., Zelazo, P. D., & Lewis, M. D. (2005). Neural correlates of cognitive control in childhood and adolescence: Disentangling the contributions of age and executive function. *Neuropsychologia, 10*, 013.

Lecerf, T., de Ribaupierre, A., & Neimer, J. (2008). Variabilité intra-individuelle dans des épreuves de temps de réaction et de mémoire de travail chez des jeunes adultes. In P. V. E. Loarer, J.-L. Mogenet, F. Cuisinier, H. Gottesdiener, P. Mallet. (Ed.), *Perspectives différentielles en psychologie* (pp. 143-146). Rennes, France: Presse Universitaires de Rennes.

Lehmann, D. (1987). Principles of spatial analysis. In A. G. a. A. Remond (Ed.), *Handbook of Electroencephalography and Clinical Neurophysiology, Vol. 1: Methods of Analysis of Brain Electrical and Magnetic Signals* (pp. 309-354). Amsterdam: Elsevier.

Li, K. Z. H., & Lindenberger, U. (2002). Relations between aging sensory/sensorimotor and cognitive functions. *Neuroscience and Biobehavioral Reviews, 26*, 777-783.

Li, K. Z. H., Lindenberger, U., Freund, A. M., & Baltes, P. B. (2001). Walking while memorizing: Age-related differences in compensatory behavior. *Psychological Science, 12*, 230-237.

Li, S.-C. (2002). Connecting the many levels and facets of cognitive aging. *Current Directions in Psychological Science, 11*, 38-43.

Li, S.-C., Brehmer, Y., Shing, Y. L., Werkle-Bergner, M., & Lindenberger, U. (2006). Neuromodulation of associative and organizational plasticity across the life span: Empirical evidence and neurocomputational modeling. *Neuroscience & Biobehavioral Reviews, 30*, 775-790.

Li, S.-C., & Lindenberger, U. (1999). Cross-level unification: A computational exploration of the link between deterioration of neurotransmitter systems and dedifferentiation of cognitive abilities on old age. In L.-G. Nilsson & H. J. Markowitsch (Eds.), *Cognitive Neuroscience of Memory* (pp. 103-146). Toronto: Hogrefe & Huber.

Li, S.-C., Lindenberger, U., & Frensch, P. A. (2000). Unifying cognitive aging: From neuromodulation to representation to cognition. *Neurocomputing, 32*(1), 879-890.

Li, S.-C., Schmiedek, F., Huxhold, O., Röcke, C., Smith, J., & Lindenberger, U. (2008). Working memory plasticity in old age: Practice gain, transfer, and maintenance. *Psychology and Aging*, *23*, 731–742.

Li, S.-C., & Sikström, S. (2002). Integrative neurocomputational perspectives on cognitive aging, neuromodulation, and representation. *Neuroscience and Biobehavioral Reviews*, *26*, 795-808.

Lindenberger, U., & Baltes, P. B. (1995). Kognitive Leistungsfähigkeit im hohen Alter: Erste Ergebnisse aus der Berliner Altersstudie / Cognitive abilities in old age: First results from the Berlin Aging Study. *Zeitschrift für Psychologie*, *203*, 283-317.

Lindenberger, U., Li, S. C., Lövdén, M., & Schmiedek, F. (2007). The Center for Lifespan Psychology at the Max Planck Institute for Human Development: Overview of conceptual agenda and illustration of research activities. *International Journal of Psychology, 42*(4), 229-242.

Lindenberger, U., Marsiske, M., & Baltes, P. B. (2000). Memorizing while walking: Increase in dual-task costs from young adulthood to old age. *Psychology and Aging*, *15*, 417-436.

Logan, J. M., Sanders, A. L., Snyder, A. Z., Morris, J. C., & Buckner, R. L. (2002). Under-Recruitment and Nonselective Recruitment: Dissociable Neural Mechanisms Associated with Aging. *Neuron, 33*, 827-840.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150-157.

Lorenzo-López, L., Amenedo, E., Pazo-Álvarez, P., & Cadaveira, F. (2007). Visual target processing in high- and low-performing older subjects indexed by P3 component. *Clinical Neurophysiology*, *37*, 53-61.

Lövdén, M. (2010). Decline-induced plastic changes of brain and behaviour in aging. In L. Bäckman & L. Nyberg (Eds.), *Memory, Aging and the Brain* (pp. 167-189). Hove: Psychology Press.

Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin*, *136*(4), 659-676.

Lövdén, M., Bodammer, N. C., Kühn, S., Kaufmann, J., Schütze, H., Tempelmann, C., et al. (2010). Experience-dependent plasticity of white-matter microstructure extends into old age. *Neuropsychologia*, *48*, 3878–3883.

Lövdén, M., Ghisletta, P., & Lindenberger, U. (2004). Cognition in the Berlin Aging Study (BASE): The First 10 Years. *Aging, Neuropsychology, and Cognition*, *11*, 104–133.

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

Ludwig, C. (2005). *Age and individual differences in attentional control: A behavioral study.* Unpuplished Doctoral Thesis, Université de Genève, Genève, Suisse.

Ludwig, C., Borella, E., Tettamanti, M., & de Ribaupierre, A. (2010). Adult age differences in the Color Stroop Test: A comparison between an Item-by-item and a Blocked version. *Archives of Gerontology and Geriatrics, 51*, 135-142.

Ludwig, C., & Chicherio, C. (2007). Pertes et gains dans le développement adulte: La perspective des neurosciences cognitives du vieillissement. *Gérontologie et Société, 123*, 109-134.

Ludwig, C., Chicherio, C., Terraneo, L., Magistretti, P., Ribaupierre, A. d., & Slosman, D. (2008). Functional imaging studies of cognition using 99mTc-HMPAO SPECT: empirical validation using the n-back working memory paradigm. *Eur J Nucl Med Mol Imaging, 35*, 695-703.

Lustig, C., & Flegal, K. E. (2008). Targeting latent function: Encouraging effective encoding for successful memory training and transfer. *Psychology and Aging, 23*, 754–764.

Lustig, C., Shah, P., Seidler, R., & Reuter-Lorenz, P. A. (2009). Aging, Training, and the Brain: A Review and Future Directions. *Neuropsychology Review, 19*(4), 504-522.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological methods, 7*(1), 19.

Macmillan, N. A., & Creelman, C. D. (1990). Response Bias: Characteristics of Detection Theory, Threshold Theory, and. *Psychological Bulletin, 107*(3), 401-413.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., et al. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *PNAS, 97*, 4398–4403.

Maguire, E. A., Spiers, H. J., Good, C. D., Hartley, T., Frackowiak, R. S. J., & Burgess, N. (2003). Navigation Expertise and the Human Hippocampus: A Structural Brain Imaging Analysis. *Hippocampus, 13*(2), 250-259.

Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London Taxi Drivers and Bus Drivers: A Structural MRI and Neuropsychological Analysis. *Hippocampus, 16*(12), 1091-1101.

Mahncke, H. W., Connor, B. B., Appelman, J., Ahsanuddin, O. N., Hardy, J. L., Wood, R. A., et al. (2006). Memory enhancement in healthy older adults using a brain plasticity-based training program: A randomized, controlled study. *Proceedings of the National Academy of Sciences, 103*(33), 12523-12528.

Martin, M., Clare, L., Altgassen, A. M., Cameron, M. H., & Zehnder, F. (2011). Cognition-based interventions for healthy older people and people with mild cognitive impairment. *Cochrane Database of Systematic Reviews, 1*.

Mattay, V. S., Fera, F., Tessitore, A., Hariri, A. R., Berman, K. F., Das, S., et al. (2006). Neurophysiological correlates of age-related changes in working memory capacity. *Neuroscience Letters, 392*, 32-37.

McBee, M. (2010). Modeling outcomes with floor or ceiling effects: An introduction to the Tobit model. *Gifted Child Quarterly, 54*(4), 314.

McDaniel, M. A., & Bugg, J. M. (2012). Memory training interventions: What has been forgotten? *Journal of Applied Research in Memory and Cognition, 1*(1), 45-50.

McEvoy, L. K., Smith, M. E., & Gevins, A. (1998). Dynamic cortical networks of verbal and spatial working memory: Effects of memory load and task practice. *Cerebral Cortex, 8*, 563-574.

Michel, C. M., Koenig, T., Brandeis, D., Gianotti, L. R. R., & Wackermann, J. (2009). *Electrical Neuroimaging*: Cambridge University Press.

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Peralta, R. G. d. (2004). EEG source imaging. *Clinical Neurophysiology, 115* 2195-2222.

Michel, C. M., Seeck, M., & Landis, T. (1999). Spatiotemporal dynamics of human cognition. *News Physiol. Sci., 14*, 206-214.

Missonnier, P., Gold, G., Leonards, U., Costa-Fazio, L., Michel, J.-P., Ibanez, V., et al. (2004). Aging and working memory: early deficits in EEG activation of posterior cortical areas. *Journal of Neural Transmission, 111*, 1141-1154.

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*, 49-100.

Moody, D. E. (2009). Can intelligence be increased by training on a task of working memory? *Intelligence, 37*(4), 327-328.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81*(2), 111.

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*(1), 46-60.

Mueller, V., Brehmer, Y., von Oertzen, T., Li, S.-C., & Lindenberger, U. (2008). Electrophysiological correlates of selective attention: A lifespan comparison. *BMC Neuroscience, 9*(1), 1-21.

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP Analyses: A Step-by-Step Tutorial Review. *Brain Topogr., 20*, 249-264.

Muthén, L., & Muthén, B. (2004). *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén.

Nagel, I. E., Preuschhof, C., Li, S. C., Nyberg, L., Bäckman, L., Lindenberger, U., et al. (2009). Performance level modulates adult age differences in brain activation during spatial working memory. *Proceedings of the National Academy of Sciences, 106*(52), 22552.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*(1), 1-32.

Nyberg, L., Sandblom, J., Jones, S., Neely, A. S., Petersson, K. M., Ingvar, M., et al. (2003). Neural correlates of training-related memory improvement in adulthood and aging. *PNAS, 100*, 13728-13733.

Oechslin, M. S., Imfeld, A., Loenneker, T., Meyer, M., & Jäncke, L. (2010). The plasticity of the superior longitudinal fasciculus as a function of musical expertise: a diffusion tensor imaging study. *Frontiers in Human Neuroscience, 3*, 1-12.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience, 7*, 75-79.

Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A. S., et al. (2010). Putting brain training to the test. *Nature*(Online).

Owen, A. M., Herrod, N. J., Menon, D. K., Clark, J. C., Downey, S. P. M. J., Carpenter, T. A., et al. (1999). Redefining the functional organization of working memory processes within human lateral prefrontal cortex. *European Journal of Neuroscience, 11*(2), 567-574.

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.

Park, D. C., Polk, T. A., Park, R., Minear, M., Savage, A., & Smith, M. R. (2004). Aging reduces neural specialization in ventral visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 13091-13095.

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annu. Rev. Psychol., 60*, 173-196.

Parkin, A. J. (1993). Implicit memory across the lifespan. In P. Graf & M. E. J. Masson (Eds.), *mplicit memory: New directions in cognition, development, and neuropsychology* (pp. 191-206). Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Pascual-Leone, A., Freitas, C., Oberman, L., Horvath, J., Halko, M., Eldaief, M., et al. (2011). Characterizing Brain Cortical Plasticity and Network Dynamics Across the Age-Span in Health and Disease with TMS-EEG and TMS-fMRI. *Brain Topography, 24*(3), 302-315.

Pascual-Leone, J., & Baillargeon, R. (1994). Developmental measurement of mental attention. *International Journal of Behavioural Development, 17*, 161-200.

Paus, T. (2005). Mapping brain maturation and cognitive development during adolescence. *Trends in Cognitive Sciences, 9*(2), 60-68.

Paus, T. (2009). Growth of white matter in the adolescent brain: Myelin or axon? *Brain and Cognition*.

Payer, D., Marshuetz, C., Sutton, B., Hebrank, A., Welsh, R. C., & Park, D. C. (2006). Decreased neural specialization in old adults on a working memory task. *Neuroreport, 17*, 487-491.

Perkins, D. N., & Salomon, G. (1992). Transfer of learning. *International encyclopedia of education, 2*.

Perrig, W. J., Hollenstein, M., & Oelhafen, S. (2009). Can We Improve Fluid Intelligence With Training on Working Memory in Persons With Intellectual Disabilities? *Journal of Cognitive Education and Psychology, 8*(2), 148-164.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology, 72*(2), 184-187.

Persson, J., Lustig, C., Nelson, J. K., & Reuter-Lorenz, P. A. (2007). Age differences in deactivation: a link to cognitive control? *Journal of Cognitive Neuroscience, 19*(6), 1021-1032.

Persson, J., Nyberg, L., Lind, J., Larsson, A., Nilsson, L.-G., Ingvar, M., et al. (2006). Structure-function correlates of cognitive decline in aging. *Cerebral Cortex, 16*, 907-915.

Petersen, S. E., van Mier, H., Fiez, J. A., & Raichle, M. E. (1998). The effects of practice on the functional anatomy of task performance. *Proceedings of the National Academy of Sciences, 95*(3), 853-860.

Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society, 9*(4), 456.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson Jr, R., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology, 37*, 127-152.

Picton, T. W., & Stuss, D. T. (1980). The Component Structure of the Human Event-Related Potentials. In H. H. Kornhubek & L. Deecke (Eds.), *Progress in Brain Research* (Vol. Volume 54, pp. 17-49). Amsterdam: Elsevier.

Pizzagalli, D. A. (2007). Electroencephalography and high-density electrophysiological source localization. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of psychophysiology* (Vol. 3, pp. 56-84). Cambridge: Univ Pr.

Poldrack, R. A. (2000). Imaging Brain Plasticity: Conceptual and Methodological Issues - A Theoretical Review. *NeuroImage, 12*, 1-13.

Polich, J. (1996). Meta-analysis of P300 normative aging studies. *Psychophysiology, 33*(4), 334-353.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*, 2128-2148.

Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling*: Sage Publications, Inc.

Pruchno, R. (2012). Not Your Mother's Old Age: Baby Boomers at Age 65 (Editorial of the special issue: Baby Boomers). *Gerontologist, 52*(2), 149-152.

Rajah, M. N., & D'Esposito, M. (2005). Region-specific changes in prefrontal function with age: A review of PET and fMRI studies on working and episodic memory. *Brain: A Journal of Neurology, 128*, 1964-1983.

Ramsey, N. F., Jansma, J. M., Jager, G., Van Raalten, T., & Kahn, R. S. (2004). Neurophysiological factors in human information processing capacity. *Brain, 127*(3), 517-525.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 5: The Mill Hill Vocabulary Scale.* . San Antonio, TX: Harcourt Assessment.

Raven, J., Raven, J. C., & Court, J. H. (1998, updated 2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 1: General Overview.* San Antonio, TX: Harcourt Assessment.

Raven, J. C. (1958). *Standard Progressive Matrices. Sets A, B, C, D & E*. Oxford: Psychologist Press Ldt.

Raven, J. C. (1962). *Advanced Progressive Matrices. Set II.* Oxford: Psychologist Press Ldt.

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of aging and cognition* (Vol. 2, pp. 1–90). Mahwah: Erlbaum.

Raz, N. (2004). The aging brain observed in vivo: differential changes and their modifiers. In R. Cabeza, L. Nyberg & D. C. Park (Eds.), *Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging* (pp. 17-55). New York: Oxford University Press.

Raz, N., Gunning, F. M., Head, D., Dupuis, J. H., McQuain, J., Briggs, S. D., et al. (1997). Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. *Cerebral Cortex, 7*(3), 368-382.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., et al. (2005). Regional Brain Changes in Aging Healthy Adults: General Trends, Individual Differences and Modifiers. *Cerebral Cortex, 15*, 1676--1689.

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

Raz, N., Rodrigue, K. M., Head, D., Kennedy, K. M., & Acker, J. D. (2004). Differential aging of the medial temporal lobe. *Neurology, 62*(3), 433-439.

Rebok, G. W., Carlson, M. C., & Langbaum, J. B. S. (2007). Training and Maintaining Memory Abilities in Healthy Older Adults: Traditional and Novel Approaches. *Journals of Gerontology: Psycholocigal Sciences, 62B*, 53-61.

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., et al. (in press). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology: General*.

Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology-Learning Memory and Cognition, 20*(3), 585-594.

Rehman, H. U., & Masson, E. A. (2001). Neuroendocrinology of ageing. *Age and ageing, 30*(4), 279-287.

Reisberg, D., Baron, J., & Kemler, D. G. (1980). Overcoming stroop interference: The effects of practice on distractor potency. *Journal of Experimental Psychology: Human, Perception and Performance, 6*, 140-150.

Reuchlin, M. (1978). Processus vicariants et différences individuelles. *Journal de Psychologie, 2*, 133-145.

Reuter-Lorenz, P. A., & Cappell, K. A. (2008). Neurocognitive Aging and the Compensation Hypothesis. *Current Directions in Psychological Science*, *17*, 177-182.

Reuter-Lorenz, P. A., Jonides, J., Smith, E. E., Hartley, A., Miller, A., Marshuetz, C., et al. (2000). Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. *Journal of cognitive neuroscience*, *12*, 174-187.

Reuter-Lorenz, P. A., & Park, D. C. (2010). Human Neuroscience and the Aging Mind: A New Look at Old Problems. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *65*(4), 405.

Reuter-Lorenz, P. A., Stanczak, L., & Miller, A. C. (1999). Neural recruitment and cognitive aging: Two hemispheres are better than one, especially as you age. *Psychological Science*, *10*, 494-500.

Richmond, L. L., Morrison, A. B., Chein, J. M., & Olson, I. R. (2011). Working Memory Training and Transfer in Older Adults. *Psychology and Aging*.

Rieckmann, A., & Bäckman, L. (2009). Implicit Learning in Aging: Extant Patterns and New Directions. *Neuropsychology Review*, *19*(4), 490-503.

Riis, J. L., Chong, H., Ryan, K. K., Wolk, D. A., Rentz, D. M., Holcomb, P. J., et al. (2008). Compensatory neural activity distinguishes different patterns of normal cognitive aging. *NeuroImage*, *39*(1), 441-454.

Ritter, W., Simson, R., Vaughan, H. G., Jr., & Friedman, D. (1979). A Brain Event Related to the Making of a Sensory Discrimination. *Science*, *203*(4387), 1358-1361.

Robert, C., Borella, E., Fagot, D., Lecerf, T., & de Ribaupierre, A. (2009). Working memory and inhibitory control across the lifespan: Intrusion errors in the Reading Span Test. . *Memory & Cognition*, *37*, 336-345.

Rossini, P. M., Rossi, S., Babiloni, C., & Polich, J. (2007). Clinical neurophysiology of aging brain: From normal aging to neurodegeneration. *Progress in Neurobiology*, *83*, 375-400.

Rypma, B., & D'Esposito, M. (2000). Isolating the neural mechanisms of age-related changes in human working memory. *Nature Neuroscience*, *3*, 509-515.

Salat, D. H., Tuch, D. S., Greve, D. N., van der Kouwe, A. J. W., Hevelone, N. D., Zaleta, A. K., et al. (2005). Age-related alterations in white matter microstructure measured by diffusion tensor imaging. *Neurobiology of aging*, *26*(8), 1215-1227.

Salminen, T., Strobach, T., & Schubert, T. (2012). On the impacts of working memory training on executive functioning. *Frontiers in human neuroscience, 6*.

Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science, 2,* 179-183.

Salthouse, T. A. (1996). The Processing-Speed Theory of Adult Age Differences in Cognition. *Psychological Review, 103,* 403-428.

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General*, *132*(4), 566.

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology, 27,* 763-776.

Salthouse, T. A., & Meinz, E. J. (1995). Aging, inhibition, working memory, and speed. *Journals of Gerontology: Psycholocigal Sciences, 50,* 297-306.

Schäfer, S., Huxhold, O., & Lindenberger, U. (2006). Healthy mind in healthy body? A review of sensorimotor-cognitive interdependencies in old age. *European Review of Aging and Physical Activity, 3*(2), 45-54.

Schaie, K. W. (1983). The Seattle Longitudinal Study: A 21-year exploration of psychometric intelligence in adulthood. In K. W. Schaie (Ed.), *Longitudinal studies of adult psychological development* (pp. 64-135). New York: Guilford Press.

Schaie, K. W. (1996). Intellectual development in adulthood. In J. E. B. K. W. Schaie (Ed.), *Handbook of the psychology of aging* (Vol. 4, pp. 266-286). San Diego, CA: Academic Press.

Schaie, K. W. (2005). What can we learn from longitudinal studies of adult development? *Research in Human Development, 2*, 133-158.

Schaie, K. W., & Willis, S. L. (1986). Can decline in adult intellectual functioning be reversed? *Developmental Psychology, 22*(2), 223-232.

Schaie, K. W., Willis, S. L., Jay, G., & Chipuer, H. (1989). Structural invariance of cognitive abilities across the adult life span: A cross-sectional study. *Developmental Psychology, 25*(4), 652-662.

Schmiedek, F., Li, S.-C., & Lindenberger, U. (2009). Interference and facilitation in spatial working memory: age-associated differences in lure effects in the n-back paradigm. *Psychology and Aging, 24*, 203-210.

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: findings from the COGITO study. *Frontiers in Aging Neuroscience, 2*.

Schneider-Garces, N. J., Gordon, B. A., Brumback-Peltz, C. R., Shin, E., Lee, Y., Sutton, B. P., et al. (2010). Span, CRUNCH, and beyond: Working memory capacity and the aging brain. *Journal of Cognitive Neuroscience, 22*(4), 655-669.

Schneiders, J. A., Opitz, B., Krick, C. M., & Mecklinger, A. (2011). Separating Intra-Modal and Across-Modal Training Effects in Visual Working Memory: An fMRI Investigation. *Cerebral Cortex, 21*(11), 2555-2564.

Schnitzspahn, K. M., & Kliegel, M. (2009). Age effects in prospective memory performance within older adults: the paradoxical impact of implementation intentions. *European Journal of Ageing, 6*(2), 147-155.

Segalowitz, S. J., Wintink, A. J., & Cudmore, L. J. (2001). P3 topographical change with task familiarization and task complexity. *Cognitive Brain Research, 12*, 451-457.

Shing, Y. L., Brehmer, Y., & Li, S.-C. (2008). Cognitive plasticity and training across the lifespan. In O.-S. Tan & A. S.-H. Seng (Eds.), *Cognitive modifiability in learning and assessment: International perspectives.* (pp. 59-82). Singapore: Thomson Learning.

Shing, Y. L., Schmiedek, F., Lövdén, M., & Lindenberger, U. (2012). Memory updating practice across 100 days in the COGITO study. *Psychology and Aging, 27*(2), 451-461.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica, 50, 3*(4), 245-276.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is Working Memory Training Effective? *Psychological Bulletin*.

Shucard, D. W., Chichelli, T. L., Lee, W. H., Covey, T. J., Lee, J. T., & Shucard, J. L. (2009). *ERP P3 amplitude change over trials during the performance of a working memory n-back task in older and younger adults*.

Skullerud, K. (1985). Variations in the size of the human brain. Influence of age, sex, body length, body mass index, alcoholism, Alzheimer changes, and cerebral atherosclerosis. *Acta neurologica Scandinavica. Supplementum, 102*, 1-94.

Slosman, D. O., Chicherio, C., Ludwig, C., Genton, L., de Ribaupierre, S., Pichard, C. Mayer, E. Annoni, J.-M., de Ribaupierre, A. (2001). 133Xe SPECT cerebral blood study in an healthy population: T-score analysis as a primary approach for clinical use. *Journal of Nuclear Medicine, 42*, 864-870.

Smith, E. E., & Jonides, J. (1997). Working Memory: A View from Neuroimaging. *Cognitive Psychology, 33*(1), 5-42.

Smith, E. E., & Jonides, J. (1999). Storage and Executive Processes in the Frontal Lobes. *Science*, *283*(5408), 1657-1661.

Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *nature neuroscience*, *6*(3), 309-315.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology, 22*(2), 461-479.

Sternberg, R. J. (2008). Increasing fluid intelligence is possible after all. *Proceedings of the National Academy of Sciences, 105*(19), 6791-6792.

Sternberg, S. (1966). High-Speed Scanning in Human Memory. *Science, 153*(3736), 652-654.

Sterr, A., Muller, M. M., Elbert, T., Rockstroh, B., Pantev, C., & Taub, E. (1998). Changed perceptions in Braille readers. *Nature, 391*(6663), 134-135.

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

Suhara, T., Fukuda, H., Inoue, O., Itoh, T., Suzuki, K., Yamasaki, T., et al. (1991). Age-related changes in human D1 dopamine receptors measured by positron emission tomography. *Psychopharmacology, 103*(1), 41-45.

Sullivan, E. V., Rohlfing, T., & Pfefferbaum, A. (2008). Quantitative fiber tracking of lateral and interhemispheric white matter systems in normal aging: Relations to timed performance. *Neurobiology of aging, 31*(3), 464-481.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions: III. Functions involving attention, observation and discrimination. *Psychological Review, 8*(6), 553-564.

Tisak, J., & Meredith, W. (1990). Descriptive and associative developmental models. In A. v. Eye (Ed.), *Statistical methods in longitudinal research* (Vol. 2, pp. 387-406). Boston: Academic Press.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica: Journal of the Econometric Society*, 24-36.

Tulving, E., Kapur, S., Craik, F. I. M., Moscovitch, M., & Hoile, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory:Positron emission tomography findings. *Proceedings of the National Academy of Sciences USA, 91*, 2016-2020.

Ungerleider, L. G., Doyon, J., & Karni, A. (2002). Imaging brain plasticity during motor skill learning. *Neurobiology of Learning and Memory, 78*(3), 553-564.

van Veen, V., & Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience, 14*(4), 593-602.

Vance, D. E., Dawson, J., Wadley, V. G., Edwards, J., Roenker, D. L., Rizzo, M., et al. (2007). The Accelerate Study: The Longitudinal Effect of Speed of Processing Training on Cognitive Performance of Older Adults. *Rehabilitation Psychology, 52*, 89-96.

Verhaeghen, P., & Basak, C. (2005). Ageing and switching of the focus of attention in working memory: Results from a modified N-Back task. *The quarterly journal of experimental psychology, 58*, 134-154.

Verhaeghen, P., Marcoen, A., & Goossens, L. (1992). Improving memory performance in the aged through mnemonic training: A meta-analytic study. *Psychology-and-Aging, 7*, 242-251.

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.

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore: Williams & Wilkins.

Westerberg, H., & Klingberg, T. (2007). Changes in cortical activity after training of working memory--a single-subject analysis. *Physiology & Behavior*, *92*(1-2), 186-192.

WHO. (2012). *Dementia: a public health priority*. Geneva: WHO Press.

Williams, K. N., & Kemper, S. (2010). Interventions to Reduce Cognitive Decline in Aging. *Journal of Psychosocial Nursing and Mental Health Services*, *48*(5), 42-51.

Willis, S. L. (2001). Methodological issues in behavioral intervention research with the elderly. In K. W. Schaie & J. E. Birren (Eds.), *Handbook of the psychology of aging* (5 ed., pp. 78-108). San Diego, CA, US: Academic Press.

Willis, S. L., Blieszner, R., & Baltes, P. B. (1981). Intellectual training research in aging: Modification of performance on the fluid ability of figural relations. *Journal of Educational Psychology*, *73*(1), 41-50.

Willis, S. L., & Schaie, K. W. (2009). Cognitive training and plasticity: Theoretical perspective and methodological consequences. *Restorative Neurology and Neuroscience*, *27*(5), 375-389.

Willis, S. L., Schaie, K. W., & Martin, M. (2009). Cognitive plasticity. In V. L. Bengtson, N. P. D. Gans & M. Silverstein (Eds.), *Handbook of theories of aging* (pp. 295-322). New York: Springer.

Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., et al. (2006). Long-term Effects of Cognitive Training on Everyday Functional Outcomes in Older Adults. *Journal of American Medical Association, 296*, 2805-2814.

Wintink, A. J., Segalowitz, S. J., & Cudmore, L. J. (2001). Task complexity and habituation effects on frontal P300 topography. *Brain and Cognition*, *46*(1-2), 307-311.

Wozniak, J. R., & Lim, K. O. (2006). Advances in white matter imaging: A review of in vivo magnetic resonance methodologies and their applicability to the study of development and aging. *Neuroscience & Biobehavioral Reviews*, *30*(6), 762-774.

Zehnder, F., Martin, M., Altgassen, M., & Clare, L. (2009). Memory training effects in old age as markers of plasticity: A meta-analysis. *Restorative Neurology and Neuroscience, 27*(5), 507-520.

Zelinski, E. M. (2012). Are strategies necessary to improve memory? *Journal of Applied Research in Memory and Cognition, 1*(1), 56-57.

Zimprich, D., Martin, M., Kliegel, M., Dellenbach, M., Rast, P., & Zeintl, M. (2008). Cognitive abilities in old age: Results from the Zurich Longitudinal Study on Cognitive Aging. *Swiss Journal of Psychology, 67*(3), 177-195.

Zinke, K., Zeintl, M., Eschen, A., Herzog, C., & Kliegel, M. (2011). Potentials and Limits of Plasticity Induced by Working Memory Training in Old-Old Age. *Gerontology, 58*(1), 79-87.

Zöllig, J., & Eschen, A. (2009). Measuring compensation and its plasticity across the lifespan. *Restorative Neurology and Neuroscience, 27*, 421-433.

APPENDIX

APPENDIX A

BEHAVIORAL METHOD

Table A1. Task order at pre-test and post-test for the general behavioral testing and the EEG subsample.

Standard Order Behavioral Testing	Subsample with EEG recording
1. Informed consent (at pre-test only)	1. Informed consent (at pre-test only)
2. Health questionnaire (at pre-test only)	2. Health questionnaire (at pre-test only)
3. Visual acuity (at post-test only)	3. Visual acuity (at post-test only)
4. Mill Hill (at post-test only)	4. Mill Hill (at post-test only)
5. Simple Reaction Time Task	5. Simple Reaction Time Task
6. Reading Span Task	6. Reading Span Task
7. Color Stroop Task	7. Color Stroop Task
8. Updating Task	8. Updating Task
9. Letter Comparison Task	9. Letter Comparison Task
10. Pattern Comparison Task	10. Pattern Comparison Task
11. Verbal 0-back task	11. Raven’s Progressive Matrices
12. Verbal 2-back task	*In EEG (second session)*
13. Spatial 0-back task	1. Verbal 0-back task
14. Spatial 2-back task	2. Verbal 2-back task
15. Raven’s Progressive Matrices	3. Spatial 0-back task
	4. Spatial 2-back task

Questionnaire on the strategy use after the WM training

1. Donnez une description aussi précise que possible de la manière / de vos méthodes pour mettre en œuvre cet exercice.

1)____

2)____

3)____

4)____

2. Jugez l'efficacité / l'utilité de chaque méthode que vous avez utilisée:

Méthode 1):

○ très utile
○ parfois utile
○ inutile

Méthode 3):

○ très utile
○ parfois utile
○ inutile

Méthode 2):

○ très utile
○ parfois utile
○ inutile

Méthode 4):

○ très utile
○ parfois utile
○ inutile

3. Avez-vous répondu aux n-backs de manière

- plutôt consciente/contrôlée
- parfois contrôlée, parfois intuitive
- plutôt intuitive (peu sûr que c'était un n-back ou non)

APPENDIX B

TRAINING TASKS: PAIRWISE COMPARISONS

Table B1. WM training. Pairwise comparisons for the Age Group x Session interaction.

Age Group	Session X	Session Y	Difference (X-Y)	*SE*	*p*
Younger	1	2	-0.12	0.15	1.000
		3	-0.59	0.12	.001
		4	-1.24	0.17	<.001
		5	-1.36	0.16	<.001
		6	-1.55	0.19	<.001
		7	-1.85	0.17	<.001
		8	-2.02	0.20	<.001
		9	-2.27	0.23	<.001
		10	-2.31	0.25	<.001
	2	1	0.12	0.15	1.000
		3	-0.48	0.12	.008
		4	-1.12	0.12	<.001
		5	-1.24	0.14	<.001
		6	-1.44	0.13	<.001
		7	-1.73	0.17	<.001
		8	-1.90	0.21	<.001
		9	-2.16	0.19	<.001
		10	-2.19	0.21	<.001
	3	1	0.59	0.12	.001
		2	0.48	0.12	.008
		4	-0.65	0.13	<.001
		5	-0.77	0.13	<.001
		6	-0.96	0.17	<.001
		7	-1.26	0.15	<.001
		8	-1.43	0.19	<.001
		9	-1.68	0.19	<.001
		10	-1.71	0.22	<.001
	4	1	1.24	0.17	<.001
		2	1.12	0.12	<.001
		3	0.65	0.13	<.001
		5	-0.12	0.11	1.000
		6	-0.31	0.15	1.000
		7	-0.61	0.14	.005
		8	-0.78	0.18	.005
		9	-1.03	0.18	<.001
		10	-1.07	0.19	<.001
	5	1	1.36	0.16	<.001
		2	1.24	0.14	<.001
		3	0.77	0.13	<.001
		4	0.12	0.11	1.000

	6	-0.19	0.13	1.000
	7	-0.49	0.12	.011
	8	-0.66	0.18	.035
	9	-0.91	0.17	<.001
	10	-0.95	0.19	<.001
6	1	1.55	0.19	<.001
	2	1.44	0.13	<.001
	3	0.96	0.17	<.001
	4	0.31	0.15	1.000
	5	0.19	0.13	1.000
	7	-0.29	0.15	1.000
	8	-0.46	0.20	1.000
	9	-0.72	0.15	.001
	10	-0.75	0.17	.002
7	1	1.85	0.17	<.001
	2	1.73	0.17	<.001
	3	1.26	0.15	<.001
	4	0.61	0.14	.005
	5	0.49	0.12	.011
	6	0.29	0.15	1.000
	8	-0.17	0.15	1.000
	9	-0.43	0.16	.524
	10	-0.46	0.17	.390
8	1	2.02	0.20	<.001
	2	1.90	0.21	<.001
	3	1.43	0.19	<.001
	4	0.78	0.18	.005
	5	0.66	0.18	.035
	6	0.46	0.20	1.000
	7	0.17	0.15	1.000
	9	-0.25	0.15	1.000
	10	-0.29	0.16	1.000
9	1	2.27	0.23	<.001
	2	2.16	0.19	<.001
	3	1.68	0.19	<.001
	4	1.03	0.18	<.001
	5	0.91	0.17	<.001
	6	0.72	0.15	.001
	7	0.43	0.16	.524
	8	0.25	0.15	1.000
	10	-0.03	0.13	1.000
10	1	2.31	0.25	<.001
	2	2.19	0.21	<.001
	3	1.71	0.22	<.001
	4	1.07	0.19	<.001
	5	0.95	0.19	<.001
	6	0.75	0.17	.002

		7	0.46	0.17	.390
		8	0.29	0.16	1.000
		9	0.03	0.13	1.000
Older	1	2	-0.17	0.15	1.000
		3	-0.29	0.12	.989
		4	-0.57	0.17	.064
		5	-0.71	0.16	.003
		6	-0.75	0.19	.017
		7	-0.91	0.17	<.001
		8	-0.97	0.20	.001
		9	-1.16	0.23	<.001
		10	-1.15	0.25	.002
	2	1	0.17	0.15	1.000
		3	-0.12	0.12	1.000
		4	-0.40	0.12	.064
		5	-0.54	0.14	.012
		6	-0.57	0.13	.004
		7	-0.74	0.17	.005
		8	-0.79	0.21	.020
		9	-0.98	0.19	<.001
		10	-0.97	0.21	.002
	3	1	0.29	0.12	.989
		2	0.12	0.12	1.000
		4	-0.29	0.13	1.000
		5	-0.43	0.13	.116
		6	-0.46	0.17	.445
		7	-0.62	0.15	.008
		8	-0.68	0.19	.042
		9	-0.87	0.19	.002
		10	-0.86	0.22	.014
	4	1	0.57	0.17	.064
		2	0.40	0.12	.064
		3	0.29	0.13	1.000
		5	-0.14	0.11	1.000
		6	-0.17	0.15	1.000
		7	-0.34	0.14	.970
		8	-0.39	0.18	1.000
		9	-0.58	0.18	.091
		10	-0.57	0.19	.223
	5	1	0.71	0.16	.003
		2	0.54	0.14	.012
		3	0.43	0.13	.116
		4	0.14	0.11	1.000
		6	-0.03	0.13	1.000
		7	-0.20	0.12	1.000
		8	-0.25	0.18	1.000
		9	-0.44	0.17	.463

	10	-0.43	0.19	1.000
6	1	0.75	0.19	.017
	2	0.57	0.13	.004
	3	0.46	0.17	.445
	4	0.17	0.15	1.000
	5	0.03	0.13	1.000
	7	-0.17	0.15	1.000
	8	-0.22	0.20	1.000
	9	-0.41	0.15	.466
	10	-0.40	0.17	.893
7	1	0.91	0.17	<.001
	2	0.74	0.17	.005
	3	0.62	0.15	.008
	4	0.34	0.14	.970
	5	0.20	0.12	1.000
	6	0.17	0.15	1.000
	8	-0.05	0.15	1.000
	9	-0.24	0.16	1.000
	10	-0.23	0.17	1.000
8	1	0.97	0.20	.001
	2	0.79	0.21	.020
	3	0.68	0.19	.042
	4	0.39	0.18	1.000
	5	0.25	0.18	1.000
	6	0.22	0.20	1.000
	7	0.05	0.15	1.000
	9	-0.19	0.15	1.000
	10	-0.18	0.16	1.000
9	1	1.16	0.23	<.001
	2	0.98	0.19	<.001
	3	0.87	0.19	.002
	4	0.58	0.18	.091
	5	0.44	0.17	.463
	6	0.41	0.15	.466
	7	0.24	0.16	1.000
	8	0.19	0.15	1.000
	10	0.01	0.13	1.000
10	1	1.15	0.25	.002
	2	0.97	0.21	.002
	3	0.86	0.22	.014
	4	0.57	0.19	.223
	5	0.43	0.19	1.000
	6	0.40	0.17	.893
	7	0.23	0.17	1.000
	8	0.18	0.16	1.000
	9	-0.01	0.13	1.000

Table B2. Implicit sequence training. Pairwise comparisons for the Age Group x Session interaction.

Age Group	Session X	Session Y	Difference (X-Y)	*SE*	*p*
Younger	1	2	114	8	<.001
		3	186	10	<.001
		4	237	11	<.001
		5	268	11	<.001
		6	287	14	<.001
		7	301	14	<.001
		8	309	15	<.001
		9	311	15	<.001
		10	320	16	<.001
	2	1	-114	8	<.001
		3	73	6	<.001
		4	123	8	<.001
		5	154	9	<.001
		6	173	11	<.001
		7	187	12	<.001
		8	195	12	<.001
		9	197	13	<.001
		10	206	14	<.001
	3	1	-186	10	<.001
		2	-73	6	<.001
		4	50	6	<.001
		5	82	8	<.001
		6	101	10	<.001
		7	115	11	<.001
		8	122	12	<.001
		9	124	12	<.001
		10	133	14	<.001
	4	1	-237	11	<.001
		2	-123	8	<.001
		3	-50	6	<.001
		5	31	5	<.001
		6	50	7	<.001
		7	64	9	<.001
		8	72	9	<.001
		9	74	9	<.001
		10	83	11	<.001
	5	1	-268	11	<.001
		2	-154	9	<.001
		3	-82	8	<.001

	4	-31	5	<.001
	6	19	5	.016
	7	33	6	<.001
	8	41	6	<.001
	9	42	7	<.001
	10	52	8	<.001
6	1	-287	14	<.001
	2	-173	11	<.001
	3	-101	10	<.001
	4	-50	7	<.001
	5	-19	5	.016
	7	14	5	.208
	8	22	5	.005
	9	24	5	.001
	10	33	7	.001
7	1	-301	14	<.001
	2	-187	12	<.001
	3	-115	11	<.001
	4	-64	9	<.001
	5	-33	6	<.001
	6	-14	5	.208
	8	8	4	1.000
	9	9	4	1.000
	10	19	6	.136
8	1	-309	15	<.001
	2	-195	12	<.001
	3	-122	12	<.001
	4	-72	9	<.001
	5	-41	6	<.001
	6	-22	5	.005
	7	-8	4	1.000
	9	2	3	1.000
	10	11	4	.320
9	1	-311	15	<.001
	2	-197	13	<.001
	3	-124	12	<.001
	4	-74	9	<.001
	5	-42	7	<.001
	6	-24	5	.001
	7	-9	4	1.000
	8	-2	3	1.000
	10	9	5	1.000

	10	1	-320	16	<.001
		2	-206	14	<.001
		3	-133	14	<.001
		4	-83	11	<.001
		5	-52	8	<.001
		6	-33	7	.001
		7	-19	6	.136
		8	-11	4	.320
		9	-9	5	1.000
Older	1	2	97	8	<.001
		3	162	10	<.001
		4	210	11	<.001
		5	249	11	<.001
		6	270	14	<.001
		7	296	14	<.001
		8	318	15	<.001
		9	335	15	<.001
		10	340	16	<.001
	2	1	-97	8	<.001
		3	65	6	<.001
		4	113	8	<.001
		5	153	9	<.001
		6	173	11	<.001
		7	199	12	<.001
		8	221	12	<.001
		9	238	13	<.001
		10	243	14	<.001
	3	1	-162	10	<.001
		2	-65	6	<.001
		4	48	6	<.001
		5	88	8	<.001
		6	108	10	<.001
		7	134	11	<.001
		8	157	12	<.001
		9	173	12	<.001
		10	179	14	<.001
	4	1	-210	11	<.001
		2	-113	8	<.001
		3	-48	6	<.001
		5	40	5	<.001
		6	60	7	<.001
		7	86	9	<.001
		8	109	9	<.001
		9	125	9	<.001

	10	130	11	<.001
5	1	-249	11	<.001
	2	-153	9	<.001
	3	-88	8	<.001
	4	-40	5	<.001
	6	20	5	.006
	7	46	6	<.001
	8	69	6	<.001
	9	85	7	<.001
	10	91	8	<.001
6	1	-270	14	<.001
	2	-173	11	<.001
	3	-108	10	<.001
	4	-60	7	<.001
	5	-20	5	.006
	7	26	5	<.001
	8	49	5	<.001
	9	65	5	<.001
	10	70	7	<.001
7	1	-296	14	<.001
	2	-199	12	<.001
	3	-134	11	<.001
	4	-86	9	<.001
	5	-46	6	<.001
	6	-26	5	<.001
	8	22	4	<.001
	9	39	4	<.001
	10	44	6	<.001
8	1	-318	15	<.001
	2	-221	12	<.001
	3	-157	12	<.001
	4	-109	9	<.001
	5	-69	6	<.001
	6	-49	5	<.001
	7	-22	4	<.001
	9	16	3	<.001
	10	22	4	<.001
9	1	-335	15	<.001
	2	-238	13	<.001
	3	-173	12	<.001
	4	-125	9	<.001
	5	-85	7	<.001
	6	-65	5	<.001
	7	-39	4	<.001
	8	-16	3	<.001

	10	6	5	1.000
10	1	-340	16	<.001
	2	-243	14	<.001
	3	-179	14	<.001
	4	-130	11	<.001
	5	-91	8	<.001
	6	-70	7	<.001
	7	-44	6	<.001
	8	-22	4	<.001
	9	-6	5	1.000

APPENDIX C

BEHAVIORAL RESULTS *N*-BACK TASKS: DESCRIPTIVE STATISTICS

Table C1. Descriptive statistics for mean accuracy data of the *N*-back tasks presented as proportions.

	Younger			Older		
Task	WM	Implicit	No-contact	WM	Implicit	No-contact
Verbal 0-back						
Pre-test						
M	.96	.97	.98	.98	.97	.97
SD	.04	.02	.02	.02	.03	.03
Range	.88-1.0	.89-1.0	.92-.1.0	.92-1.0	.89-1.0	.84-1.0
Post-test						
M	.98	.97	.98	.99	.98	.97
SD	.05	.03	.02	.01	.02	.03
Range	.80-1.0	.90-1.0	.90-1.0	.97-1.0	.94-1.0	.86-1.0
Verbal 2-back						
Pre-test						
M	.91	.93	.91	.86	.90	.84
SD	.08	.06	.05	.11	.08	.12
Range	.73-.99	.75-.99	.78-.98	.58-1.0	.72-.99	.62-.98
Post-test						
M	.96	.94	.92	.96	.92	.87
SD	.05	.05	.08	.05	.08	.11
Range	.80-1.0	.76-1.0	.62-.99	.80-1.0	.74-.99	.58-1.0
Spatial 0-back						
Pre-test						
M	.94	.97	.97	.95	.96	.94
SD	.06	.02	.02	.05	.04	.04
Range	.75-.99	.90-.99	.93-.1.0	.81-1.0	.86-1.0	.87-.99
Post-test						
M	.97	.96	.97	.97	.97	.96
SD	.03	.03	.02	.04	.02	.05
Range	.90-1.0	.87-1.0	.90-1.0	.84-1.0	.92-1.0	.75-1.0
Spatial 2-back						
Pre-test						
M	.92	.94	.91	.85	.91	.84
SD	.06	.06	.08	.13	.08	.14
Range	.77-.99	.72-.99	.63-.97	.47-.98	.71-.98	.47-.98
Post-test						
M	.93	.93	.92	.93	.94	.86
SD	.08	.06	.10	.06	.05	.13
Range	.64-1.0	.77-1.0	.55-.98	.77-.99	.80-1.0	.60-.99

Table C2. Descriptive statistics for mean *d* prime data of the *N*-back tasks.

	Younger			Older		
Task	WM	Implicit	No-contact	WM	Implicit	No-contact
Verbal 0-back						
Pre-test						
M	4.00	4.15	4.23	4.34	4.16	4.17
SD	0.82	0.65	0.52	0.52	0.76	0.61
Range	2.47-5.16	2.41-5.16	2.88-4.92	2.68-4.92	2.52-5.16	2.63-5.16
Post-test						
M	4.53	4.14	4.15	4.87	4.20	4.28
SD	0.60	0.68	0.51	0.35	0.60	0.61
Range	3.00-5.16	2.52-5.16	3.26-5.16	4.05-5.16	3.04-5.16	3.19-5.16
Verbal 2-back						
Pre-test						
M	3.00	3.23	2.85	2.74	2.94	2.53
SD	0.88	0.88	0.66	1.18	1.10	1.10
Range	1.03-4.40	1.14-4.92	1.70-4.11	0.49-5.16	1.14-4.66	0.64-4.50
Post-test						
M	4.04	3.42	3.18	4.03	3.28	2.80
SD	0.89	0.79	0.97	0.85	1.08	1.10
Range	2.00-5.16	1.22-4.90	0.47-4.43	2.45-5.16	1.18-4.89	1.11-4.90
Spatial 0-back						
Pre-test						
M	3.47	3.96	3.92	3.83	3.80	3.53
SD	0.88	0.55	0.56	0.98	0.99	0.56
Range	1.19-4.66	2.50-4.66	2.98-4.90	1.65-4.92	2.06-5.16	2.50-4.37
Post-test						
M	4.21	3.76	3.82	4.48	4.06	3.86
SD	0.86	0.76	0.59	0.81	0.68	0.69
Range	2.68-5.16	2.25-5.16	2.81-5.16	2.62-5.16	3.00-5.16	2.50-5.16
Spatial 2-back						
Pre-test						
M	3.07	3.37	2.96	2.72	3.14	2.52
SD	0.81	0.75	0.85	0.98	0.84	1.02
Range	1.31-4.63	0.94-4.38	0.24-4.02	0.78-4.63	1.81-4.19	0.50-4.12
Post-test						
M	3.34	3.32	3.17	3.36	3.60	2.80
SD	0.92	0.85	0.95	0.78	0.89	1.12
Range	1.20-4.91	1.32-4.91	0.55-4.20	1.47-4.63	2.03-4.92	0.89-4.63

Table C3. Descriptive statistics for accuracy and *d* prime gain scores (post-test – pre-test score) of the 2-back tasks and its difference to the 0-back task (Δ 0-back = 2-back gain score - 0-back gain score).

	Younger			Older		
Task	WM	Implicit	No-contact	WM	Implicit	No-contact
Verbal 2-back						
Accuracy						
Gain score						
M	.05	.01	.01	.11	.02	.03
SD	.07	.04	.06	.10	.06	.07
Range	-.16-.19	-.08-.08	-.16-.07	-.02-.30	-.08-.19	-.12-.16
Δ 0-back						
M	.04	.01	.01	.09	.02	.03
SD	.08	.04	.06	.09	.06	.06
Range	-.16-.25	-.09-.08	-.13-.07	-.02-.30	-.10-.19	-.06-.16
d prime						
Gain score						
M	1.05	0.20	0.33	1.29	0.33	0.27
SD	0.70	0.62	0.83	1.06	0.80	0.53
Range	-0.73-2.16	-1.03-1.36	-2.04-1.36	-0.55-3.24	-0.87-1.90	-0.63-1.30
Δ 0-back						
M	0.51	0.20	0.41	0.76	0.29	0.15
SD	0.83	0.75	0.78	0.96	0.76	0.51
Range	-1.02-2.11	-1.28-1.89	-1.78-1.36	-1.04-3.11	-0.91-1.88	-0.61-1.54
Spatial 2-back						
Accuracy						
Gain score						
M	.02	-.01	.00	.08	.03	.02
SD	.06	.03	.08	.11	.04	.09
Range	-.16-.17	-.09-.05	-.21-.22	-.13-.33	-.03-.09	-.15-.21
Δ 0-back						
M	-.02	.01	.00	.06	.02	.01
SD	.06	.05	.08	.09	.05	.10
Range	-.15-.09	-.10-.15	-.21-.20	-.04-.30	-.08-.09	-.18-.21
d prime						
Gain score						
M	0.27	-0.06	0.21	0.64	0.46	0.28
SD	0.71	0.69	0.82	0.78	0.59	0.70
Range	-1.00-1.82	-1.68-1.27	-1.07-1.89	-1.11-1.73	-0.81-1.68	-0.81-1.54
Δ 0-back						
M	-0.47	0.14	0.30	-0.01	0.20	-0.05
SD	1.03	1.06	0.89	0.70	0.96	1.07
Range	-2.78-0.93	-1.79-3.69	-1.22-2.67	-1.19-1.46	-1.75-1.29	-1.84-2.07

Table C4. Descriptive statistics for the verbal and spatial 2-back task accuracy for high and low performers (median split groups).

	Younger			Older		
Task	WM	Implicit	No-contact	WM	Implicit	No-contact
Verbal 2-back						
High						
Pre-test						
M	.96	.97	.95	.95	.96	.93
SD	.01	.01	.02	.04	.02	.05
Range	.94-.99	.94-.99	.93-.98	.88-1	.93-.99	.85-.98
Post-test						
M	.97	.96	.96	.98	.97	.95
SD	.05	.03	.04	.02	.01	.04
Range	.83-1	.88-1	.85-.99	.94-1	.95-.99	.85-1
Low						
Pre-test						
M	.85	.88	.87	.76	.83	.74
SD	.07	.05	.04	.08	.07	.07
Range	.73-.93	.75-.94	.78-.92	.58-.88	.72-.92	.62-.82
Post-test						
M	.94	.91	.89	.94	.87	.79
SD	.06	.06	.09	.06	.08	.11
Range	.8-1	.76-.98	.62-.97	.8-.99	.74-.98	.58-.98
Spatial 2-back						
High						
Pre-test						
M	.96	.97	.96	.93	.97	.94
SD	.02	.01	.01	.03	.01	.03
Range	.94-.99	.96-.99	.94-.97	.89-.98	.94-.98	.91-.98
Post-test						
M	.97	.96	.95	.95	.98	.94
SD	.02	.03	.04	.06	.01	.05
Range	.91-1	.88-1	.88-.98	.77-.99	.94-1	.83-.99
Low						
Pre-test						
M	.88	.91	.87	.77	.85	.75
SD	.06	.07	.09	.14	.06	.15
Range	.77-.93	.72-.96	.63-.94	.47-.89	.71-.93	.47-.9
Post-test						
M	.90	.91	.89	.91	.90	.79
SD	.09	.07	.12	.06	.05	.13
Range	.64-.97	.77-.98	.55-.98	.77-.99	.8-.97	.6-.97

APPENDIX D

BEHAVIORAL RESULTS *N*-BACK TASKS: ANALYSES ADDRESSING CEILING EFFECTS

The lack of interactions with age however could be due to a ceiling effect. Whenever the task is too easy for one or both age groups at post-test or at both sessions, an interaction with age can potentially be suppressed. A ceiling effect is assumed when the sum of the mean and its standard deviation exceeds the maximum possible score (proportion of 1 in our case). In order to exclude this possible bias we reanalyzed accuracy data by excluding ceiling participants. First, we identified participants who showed a positive gain score and excluded those with a zero or negative gain score. The assumption was that subjects performing at ceiling may not be able to improve their performance from pre-test to post-test. Second, we identified the top third in accuracy performance at pre-test for each experimental group separately and excluded them. This method allows analyzing the same *n* per group without the potential subjects performing at ceiling who might no longer be able to improve their performance. However, these analyses are exploratory and have several serious shortcomings (e.g., the selection of relative low score subject). We further checked for robustness of the ANOVA by a Tobit model recommended for dealing with censored data (Bellgrove, Hester, & Garavan, 2004). The term *censoring* describes a process that results in partial information loss (McBee, 2010; Tobin, 1958). Such data is often found in economic surveys of households. In our case, the Tobit model takes into account that the participants cannot attain a score better than all correct, that is data are censored at proportion 1 (the model was also tested with a censored proportion score of .98, it did not change the results). We conducted the Tobit model regression analyses in EViews. The post-test score was regressed on pre-test score, training group and age group.

Verbal 2-back task

Here the possible bias by a ceiling effect on the verbal 2-back data was addressed. For the first method, positive gain sample only, the main effects of age ($F(1, 86) = 5.41$, $p = .022$, $\eta 2p = .06$) and training ($F(2, 86) = 6.21$, $p = .003$, $\eta 2p = .13$) were significant, interaction however did still not reach significance ($F(2, 86) = 1.2$, $p = .306$, $\eta 2p = .03$). Further, the number of subjects retained per group was larger for the WM training group than for the other groups for younger as well as for older with 20 participants out of 22 each (see Table D1

Table D1). Regarding the second method, the top third eliminated sample, there was also a significant main effect of age ($F(1, 79) = 5.69$, $p = .017$, $\eta 2p = .07$) and training (F(2,

79) = 12.85, p < .001, η2p = .25), but no significant interaction (F(2, 79) = 1.46, p = .238, η2p = .05). Both analyses indicated that the WM training groups showed a significantly larger gain score than the other groups and that older adults had generally larger gain scores (see Table D1). However, when considering the non-significant interactions in all three samples for accuracy data the pairwise comparisons indicated always the same trend: Gain scores were significantly larger for older but not for younger adults in the WM group compared to both control groups. This indicates that training gains were more pronounced for older adults than for younger adults.

Table D1. Verbal 2-back task gain scores for correct responses. Mean and standard deviation for all participants, for participants with positive gain score only and for the remaining participants after elimination of the top third for each group separately.

	All		Gain only		Without top third	
	n	*M (SD)*	*n*	*M (SD)*	*n*	*M (SD)*
Young						
WM	22	.05 (.07)	20	.06 (.05)	15	.07 (.05)
Implicit	20	.01 (.04)	9	.04 (.02)	13	.02 (.04)
No-contact	21	.01 (.06)	16	.04 (.02)	14	.02 (.06)
Older						
WM	22	.11 (.10)	20	.12 (.10)	15	.15 (.09)
Implicit	20	.02 (.06)	13	.05 (.06)	13	.04 (.07)
No-contact	23	.03 (.07)	15	.07 (.08)	16	.04 (.08)

By the third method, the robustness of the ANOVA was checked by a Tobit model. There were 13 left censored observations at post-test for the verbal 2-back task. The Tobit model regression analysis was conducted in EViews. Pre-test score, training group and age group was regressed on the post-test score. The results however did not change, that is, there was still a significant effect of WM training (z = 4.63, p < .001), but no significant effect of age or Age x WM training interaction. In addition there was a significant effect of pre-test score (z = 9.58, p < .001) and a significant interaction WM Training x Pre-test score (z = -4.25, p < .001). These results indicated that WM training and a low score at pre-test predicted a high score at post-test. The age group however did not influence the score at post-test significantly even after considering the censored nature of the data but they confirm the results reported in the results section.

Spatial 2-back task

The same analyses were conducted with the spatial task (see Table D2

Table D2). However, all analyses show the same effects as the whole sample ANOVA: There is still no significant Age x Training interaction which would indicate an age-difference in the WM training groups. The main effects of age were (marginally) significant for both subsamples ($F(1, 78) = 3.74$, $p = .057$, η2p = .05 for gain only; $F(1, 79) = 7.27$, $p = .009$, η2p = .08 for without top third) indicating that older show more gain than younger. The main effects of training were marginally significant for both subsamples as well ($F(2, 78) = 2.77$, $p = .069$, η2p = .07 for gain only; $F(2, 79) = 2.88$, $p = .062$, η2p = .07 for without top third) indicating that the WM group gained marginally more than both control groups. As for the verbal task an interaction with age is not significant.

Table D2. Spatial 2-back task gain scores for correct responses. Mean and standard deviation for all participants, for participants with positive gain score only and for the remaining participants after elimination of the top third for each group separately.

	All		Gain only		Without top third	
	n	*M (SD)*	*n*	*M (SD)*	*n*	*M (SD)*
Young						
WM	22	.02 (.06)	14	.04 (.04)	15	.02 (.07)
Implicit	20	-.01 (.03)	8	.02 (.01)	13	.00 (.03)
No-contact	21	.00 (.08)	12	.05 (.06)	14	.01 (.10)
Older						
WM	22	.08 (.11)	19	.10 (.10)	15	.11 (.11)
Implicit	20	.03 (.04)	17	.04 (.03)	13	.05 (.04)
No-contact	23	.02 (.09)	15	.07 (.07)	16	.03 (.11)

However, the pairwise comparisons for all the non significant interactions reveal the same trend as for the verbal task: Older adults with WM training differ at least marginally from both control groups whereas the younger WM training group doesn’t show any significant difference to both control groups. For the spatial task we also conducted a Tobit model regression for dealing with the left censored nature of the data. However for the spatial 2-back task there were only 3 left censored observations at post-test .Results revealed the same pattern as for the verbal task: There is still a significant effect of WM training ($z = 2.66$, $p = .008$), but no significant effect of age or interaction Age x Training. In addition there was a significant effect of pre-test score ($z = 8.79$, $p < .001$) and a significant interaction WM training x Pre-test score ($z = -2.54$, $p = .011$). These results indicate that WM training and a low score at pre-test predicted a high score at post-test. The age group however did not influence the score at post-test significantly even after considering the censored nature of the data.

APPENDIX E

BEHAVIORAL RESULTS PRE-TEST AND POST-TEST TASKS: DESCRIPTIVE STATISTICS

Table E1. Descriptive statistics for the proportion of correctly recalled numbers in the Updating Task; the average number of words correctly recalled per item for the Reading Span task; the total of correct responses in the Raven Progressive Matrices; the interference ratio of the Stroop task.

	Younger			Older		
Task	WM	Implicit	No-contact	WM	Implicit	No-contact
Updating						
Pre-test						
M	.59	.63	.46	.57	.44	.39
SD	.20	.21	.25	.22	.20	.21
Range	.23-.92	.31-.92	.08-1	.15-.85	.08-.77	.15-.77
Post-test						
M	.70	.66	.52	.49	.44	.43
SD	.21	.21	.21	.24	.21	.19
Range	.38-.92	.15-1	.15-.85	.23-.85	.08-.77	.15-.69
Reading Span						
Pre-test						
M	2.74	2.84	2.63	2.76	2.82	2.71
SD	.52	.38	.46	.35	.47	.45
Range	1.63-3.31	2.13-3.31	1.44-3.25	2.19-3.5	2-3.5	1.56-3.38
Post-test						
M	2.91	2.97	2.87	2.88	2.88	2.82
SD	.41	.41	.30	.34	.48	.32
Range	1.69-3.38	2-3.44	2.25-3.44	2.13-3.44	2.06-3.5	2.19-3.31
Raven						
Pre-test						
M	36.23	34.85	35.57	27.05	30.30	25.17
SD	5.82	6.12	5.48	7.55	6.87	6.10
Range	24-45	24-46	22-43	9-36	14-42	14-36
Post-test						
M	37.41	35.95	36.86	28.86	31.20	27.61
SD	6.43	7.35	6.55	7.10	6.67	6.82
Range	23-46	19-47	21-47	14-39	13-41	9-38
Stroop						
Pre-test						
M	0.29	0.25	0.31	0.34	0.32	0.34
SD	0.11	0.12	0.16	0.12	0.13	0.14
Range	0.13-0.57	0.09-0.49	0.05-0.58	0.1-0.53	0.14-0.57	0.04-0.64
Post-test						
M	0.22	0.20	0.30	0.29	0.29	0.33
SD	0.13	0.11	0.23	0.10	0.15	0.13
Range	-0.05-0.49	-0.01-0.39	0.05-0.98	0.08-0.48	0.11-0.73	0.1-0.63

Table E2. Descriptive statistics for the average response time in seconds for the Letter Comparison task and the Patter Comparison task; average response time in ms for the Simple Reaction Time (SRT) task.

	Younger			Older		
Task	WM	Implicit	No-contact	WM	Implicit	No-contact
Letter Comparison						
Pre-test						
M	65.59	59.23	58.49	81.51	83.73	77.70
SD	17.86	16.12	17.27	15.75	20.29	18.09
Range	34.5-111.63	35.6-115.56	33.68-108.66	57.69-115.59	57.6-120.42	30.25-117.54
Post-test						
M	63.79	60.22	60.00	80.70	78.16	75.65
SD	21.83	20.20	16.94	15.44	17.76	11.81
Range	32.23-121.29	27.42-111.55	30.34-88.23	58.9-116.11	53.67-115.38	59.89-102.46
Pattern Comparison						
Pre-test						
M	52.00	51.67	49.25	75.00	70.70	73.28
SD	13.14	11.29	14.05	18.79	18.85	19.31
Range	33-82.33	31.48-73.77	32.81-94.93	41.86-126.38	46.47-116.08	48.38-120.07
Post-test						
M	51.33	49.92	48.41	73.78	67.82	71.35
SD	16.34	11.55	15.48	10.76	11.63	16.19
Range	34.59-90.37	28.07-72.27	27.19-91.93	53.66-98.81	44.42-98.11	52.78-119.32
SRT						
Pre-test						
M	304	278	289	342	328	309
SD	53	44	39	41	49	49
Range	239-449	232-419	227-368	270-410	247-452	246-414
Post-test						
M	295	274	292	353	317	321
SD	48	32	56	55	54	73
Range	240-398	216-332	221-456	262-495	246-477	251-521

APPENDIX F

CLASSIC ERP RESULTS: DESCRIPTIVE STATISTICS

Table F1. Mean amplitude gain scores in µV for the N2 component in the verbal 2-back task.

	Younger			Older		
ROI	WM	Implicit	No-contact	WM	Implicit	No-contact
Frontal left						
M	0.29	0.75	1.04	0.55	-0.07	0.37
SD	1.09	1.94	1.78	1.85	1.48	1.04
Range	-1.36-1.85	-2.16-4.75	-0.86-5.36	-3.01-3.54	-2.9-2.78	-1.11-2.44
Central left						
M	0.99	0.42	0.61	1.37	0.05	0.27
SD	1.24	0.85	0.83	1.28	0.88	1.01
Range	-1.01-3.01	-0.89-2.13	-0.55-2.0	0.07-4.12	-1.26-1.68	-2.42-1.41
Parietal left						
M	-0.82	-0.25	-0.37	0.19	-0.04	-0.17
SD	1.34	1.74	1.51	1.18	1.03	0.97
Range	-4.49-0.34	-3.91-2.11	-3.47-1.45	-1.58-3.13	-2.62-1.32	-2.15-1.75
Frontal central						
M	0.09	0.67	1.20	0.41	-0.12	0.48
SD	1.50	2.11	2.01	2.45	1.74	0.74
Range	-2.68-2.99	-2.31-5.24	-1.24-5.76	-3.61-6.13	-2.59-3.71	-0.65-1.49
Central central						
M	1.93	0.56	0.79	1.35	0.53	0.33
SD	2.00	0.93	0.99	0.97	1.19	0.86
Range	-2.54-4.59	-1.49-2.26	-0.79-2.59	-0.6-2.8	-1.54-2.77	-0.89-2.09
Parietal central						
M	0.31	-0.42	-0.94	0.19	-0.01	-0.45
SD	1.27	1.60	1.68	2.06	1.07	1.02
Range	-1.68-2.45	-3.78-1.75	-3.77-2.13	-3.19-4.59	-2.21-2.02	-2.28-0.67
Frontal right						
M	0.51	0.56	0.98	0.44	0.07	0.42
SD	1.63	1.53	1.23	2.04	1.46	0.82
Range	-2.08-3.65	-1.41-2.75	-1.0-3.17	-3.4-3.01	-2.8-2.51	-1.09-1.6

Central right						
M	1.44	0.18	0.19	-0.04	0.23	0.12
SD	1.02	1.21	0.53	0.92	0.89	0.85
Range	-0.72-2.76	-2.17-1.96	-0.57-1.24	-1.69-1.33	-0.77-2.01	-1.03-1.63
Parietal right						
M	.64	-.82	-1.01	-.89	-.63	-.42
SD	1.76	1.75	1-34	2.22	1.67	1.04
Range	-1.65-4.22	-3.94-2.32	-3.30-.95	-4.25-3.77	-4.08-2.35	-2.31-.95

Table F2. Mean amplitude gain scores in µV for the P3 component in the verbal 2-back task.

	Younger			Older		
ROI	WM	Implicit	No-contact	WM	Implicit	No-contact
Frontal left						
M	0.19	-0.28	-0.66	-0.59	0.25	-0.19
SD	1.91	1.72	1.15	1.82	1.12	1.73
Range	-3.24-3.56	-2.95-2.73	-2.25-0.89	-3.7-1.48	-1.98-1.55	-3.29-1.64
Central left						
M	-0.66	0.28	-0.25	-0.65	0.57	-0.27
SD	1.66	0.72	1.20	1.27	1.00	1.17
Range	-2.57-2.53	-1.35-1.44	-1.87-1.44	-3.11-1.29	-2.16-2.01	-2.28-1.81
Parietal left						
M	0.09	0.00	-0.12	0.24	0.32	0.37
SD	1.33	1.90	1.32	1.42	1.05	1.37
Range	-1.51-1.85	-2.72-4.47	-2.33-2.36	-2.65-2.36	-0.98-1.96	-1.77-2.65
Frontal central						
M	0.11	-0.56	-0.79	-0.63	0.17	-0.56
SD	2.35	2.01	2.15	2.52	1.13	1.43
Range	-3.81-3.64	-3.51-2.53	-4.04-3.84	-7.18-2.36	-1.72-1.51	-2.86-2.16
Central central						
M	-1.99	-0.52	-0.53	-0.61	-0.01	-0.45
SD	2.49	1.29	1.41	1.61	1.26	0.98
Range	-4.44-3.84	-3.75-1.3	-2.89-1.16	-4.18-1.54	-1.73-1.93	-2.16-1.26
Parietal central						
M	-0.65	-0.05	0.44	0.18	-0.12	0.51
SD	2.08	1.60	1.46	1.70	0.96	0.91
Range	-3.24-4.16	-1.88-4.28	-1.85-3.46	-2.63-2.97	-1.82-0.94	-1.67-1.55

Frontal right						
M	-0.55	-0.68	-0.59	-0.70	-0.01	-0.55
SD	1.89	1.65	1.35	2.43	1.56	1.62
Range	-3.19-2.62	-2.96-2.32	-2.69-1.58	-4.88-3.69	-2-3.67	-3.07-2.21
Central right						
M	-1.42	-0.04	0.18	0.06	0.08	-0.10
SD	1.19	1.52	0.84	1.18	1.25	1.11
Range	-2.9--0.02	-2.43-3	-1.36-1.55	-2.52-2.19	-1.92-2.16	-2.93-1.08
Parietal right						
M	-1.26	0.55	0.62	0.54	0.53	0.53
SD	2.54	1.73	1.50	2.00	1.68	1.67
Range	-4.33-2.81	-3.24-2.54	-1.29-4.1	-1.81-4.55	-1.79-4.63	-1.91-3.38

Table F3. Mean amplitude gain scores in µV for the P3 component in the spatial 2-back task.

	Younger			Older		
ROI	WM	Implicit	No-contact	WM	Implicit	No-contact
Frontal left						
M	0.97	0.35	-0.36	0.15	-0.56	-0.07
SD	1.76	1.53	1.17	1.41	1.42	1.30
Range	-1.95-3.21	-1.88-3.3	-3.39-0.75	-1.37-3.2	-3.99-1.46	-1.34-3.0
Central left						
M	0.33	0.15	-0.15	-0.45	0.32	-0.26
SD	1.53	1.44	0.73	1.16	1.08	0.84
Range	-1.81-3.05	-2.28-2.22	-1.13-1.3	-3.56-1.02	-1.15-2.37	-2.14-0.77
Parietal left						
M	-0.82	-0.43	-0.23	-0.90	0.51	-0.34
SD	1.55	1.06	1.07	1.38	1.35	1.28
Range	-4.66-0.82	-2.35-0.99	-1.74-2.24	-4.24-0.36	-1.62-3.09	-2.09-1.79
Frontal central						
M	0.72	-0.04	-0.88	0.56	-0.82	-0.07
SD	1.85	2.06	1.26	2.07	1.56	1.83
Range	-2.26-3.6	-3.17-3.63	-3.19-0.89	-2.77-4.25	-3.86-0.92	-3.05-3.26
Central central						
M	-0.29	0.18	-0.54	-0.33	0.37	-0.40
SD	1.67	1.36	1.53	1.66	1.31	0.79
Range	-3.3-3.08	-1.62-3.2	-2.6-2.96	-4.32-1.69	-1.34-2.92	-1.67-1.05

Parietal central						
M	-1.37	-0.20	0.10	0.16	0.36	0.15
SD	1.82	1.62	1.27	1.57	1.31	1.50
Range	-4.14-1.68	-2.27-1.99	-2.21-2.47	-2.08-3.14	-1.71-2.94	-3.11-2.86
Frontal right						
M	0.89	-0.04	-0.69	0.78	-0.23	-0.17
SD	2.07	1.45	1.50	1.77	1.05	1.66
Range	-1.98-4.44	-2.47-2.42	-3.87-1.07	-1.54-3.38	-1.62-1.91	-2.63-2.93
Central right						
M	0.15	0.27	-0.37	-0.12	-0.12	-0.47
SD	0.59	0.98	1.38	0.82	0.95	0.80
Range	-0.85-1.53	-0.84-2.56	-2.25-1.99	-1.97-1.05	-1.49-1.97	-2.12-0.6
Parietal right						
M	-1.44	-0.47	0.10	0.13	0.38	0.36
SD	2.00	1.63	1.08	1.01	1.27	1.31
Range	-4.08-2.01	-3.27-3.31	-1.23-2.41	-1.31-1.68	-1.13-3.3	-3.02-2.13

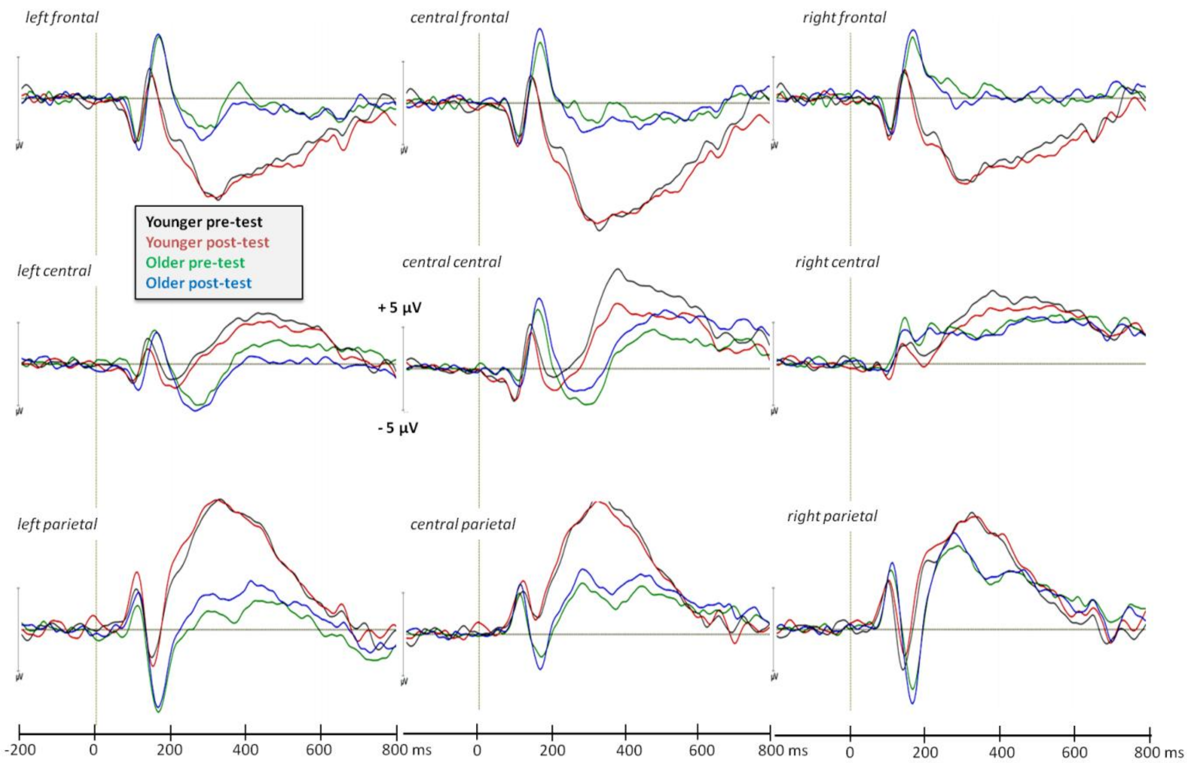

Figure F1. **WM training group**: ERPs for the verbal 0-back task.

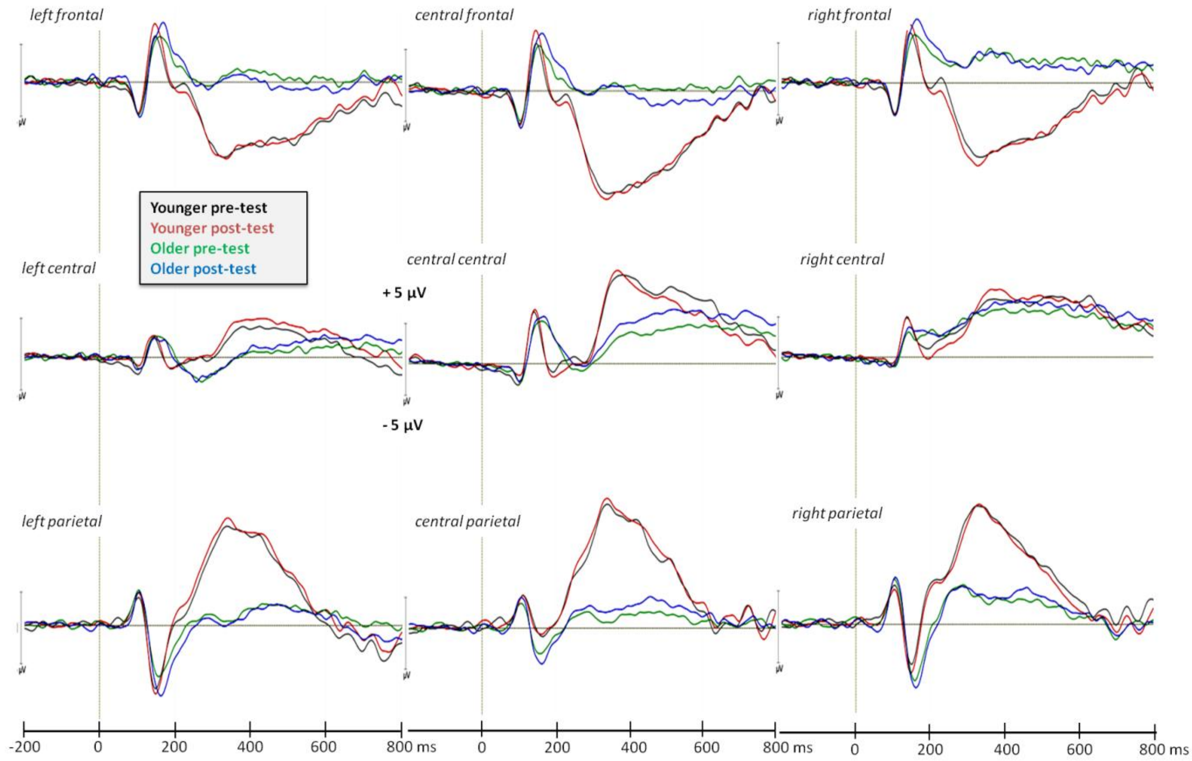

Figure F2. **Implicit training group**: ERPs for the verbal 0-back task.

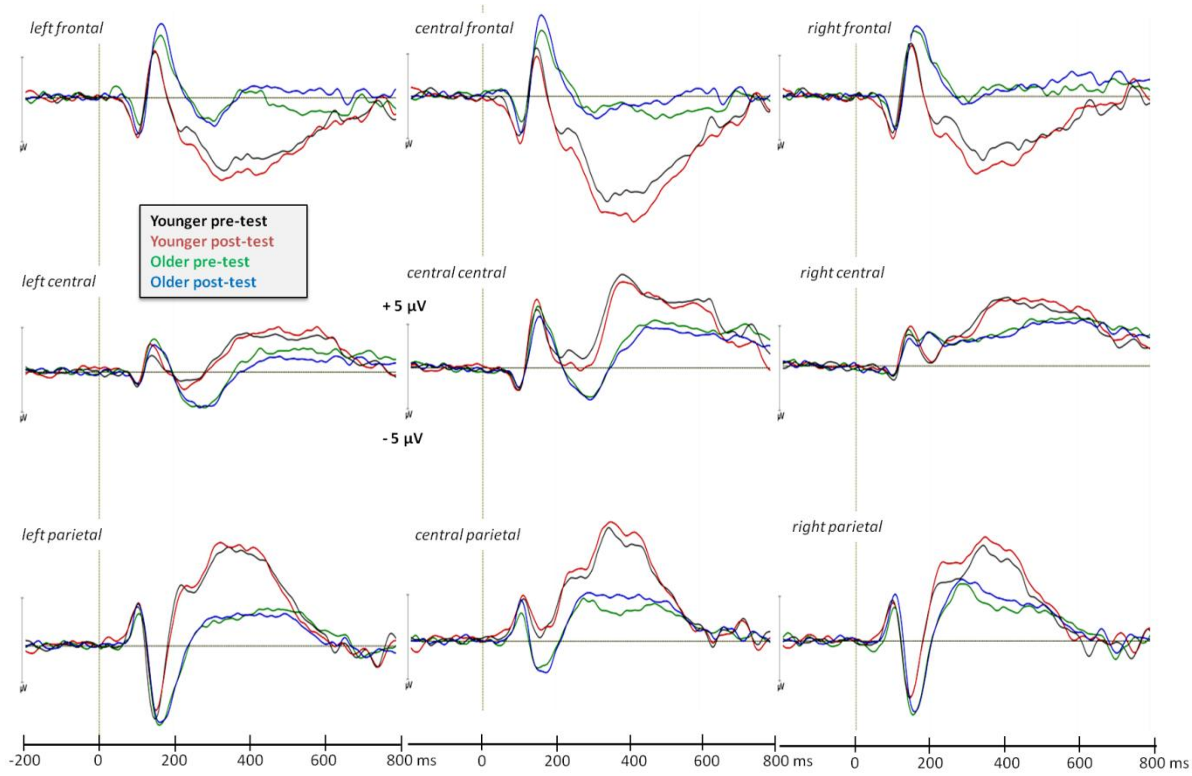

Figure F3. **No-contact control group**: ERPs for the verbal 0-back task.

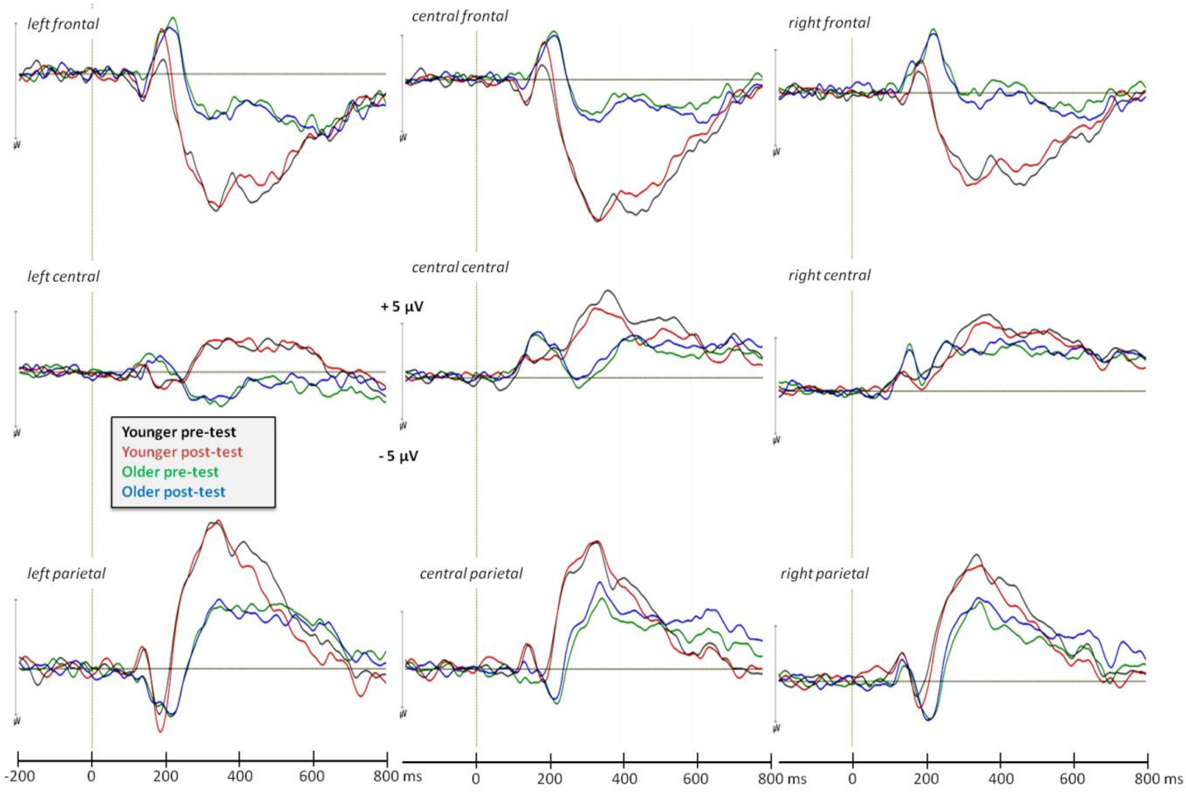

Figure F4. **WM training group**: ERPs for the spatial 0-back task.

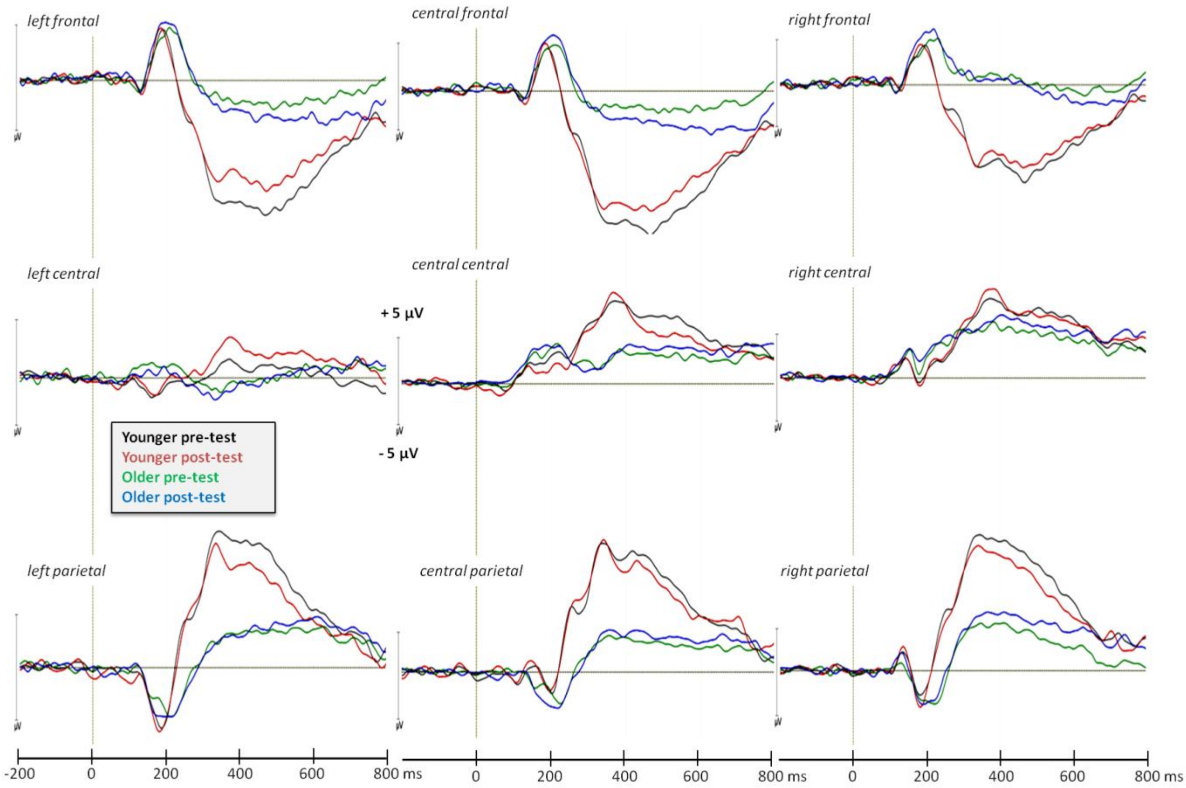

Figure F5. **Implicit training group**: ERPs for the spatial 0-back task.

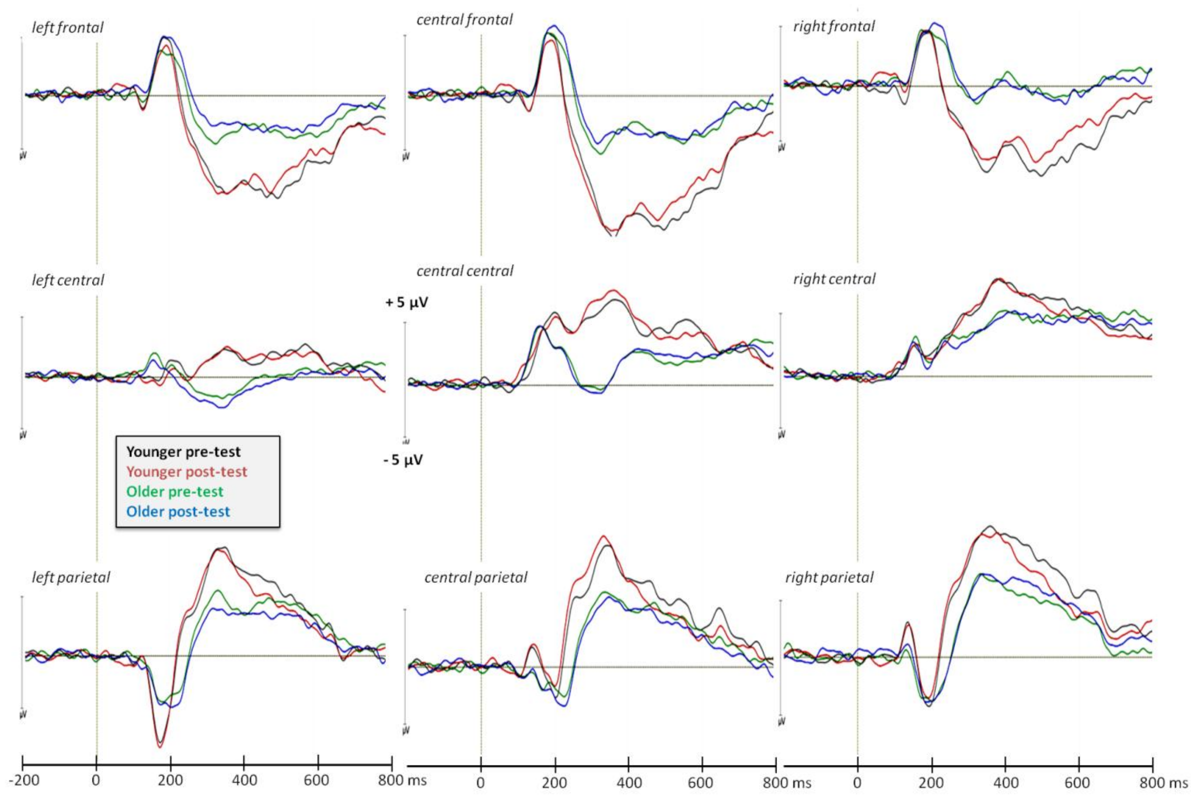

Figure F6. **No-contact control group**: ERPs for the spatial 0-back task.

MICROSTATE SEGMENTATION RESULTS: DESCRIPTIVE STATISTICS

Table F4. Duration of presence in time frames (TF; 1 TF = 4 ms) and global explained variance (GEV) for the map 1 in the verbal 2-back task.

	Younger			Older		
Variable	WM	Implicit	No-contact	WM	Implicit	No-contact
Duration						
Pre-test						
M	5.92	13.42	6.75	2.67	2.50	8.00
SD	9.49	14.11	10.48	6.46	8.66	15.37
Range	0-24	0-41	0-29	0-20	0-30	0-40
Post-test						
M	15.50	13.58	13.92	12.33	4.75	8.00
SD	10.88	14.22	14.80	18.65	11.26	15.26
Range	0-33	0-41	0-47	0-47	0-33	0-40
GEV						
Pre-test						
M	.06	.16	.07	.04	.01	.13
SD	.11	.29	.14	.09	.02	.26
Range	0-.29	0-.81	0-.37	0-.27	0-.07	0-.73
Post-test						
M	.26	.18	.16	.18	.05	.13
SD	.24	.24	.22	.29	.13	.26
Range	0-.67	0-.81	0-.68	0-.76	0-.38	0-.69

Table F5. Duration of presence in time frames (TF; 1 TF = 4 ms) and global explained variance (GEV) for the map 3 to 6 in the spatial 2-back task.

	Younger			Older		
Task	WM	Implicit	No-contact	WM	Implicit	No-contact
Duration Map 3						
Pre-test						
M	22.67	10.83	32.08	17.58	7.92	9.58
SD	31.33	19.58	24.55	23.04	22.85	13.88
Range	0-89	0-66	0-82	0-81	0-79	0-41
Post-test						
M	17.33	13.50	20.67	7.75	20.33	14.42
SD	25.35	20.41	27.19	14.91	35.21	14.66
Range	0-69	0-51	0-89	0-46	0-89	0-38
GEV Map 3						
Pre-test						
M	0.22	0.09	0.30	0.15	0.06	0.04
SD	0.30	0.18	0.22	0.22	0.17	0.08
Range	0-0.88	0-0.61	0-0.79	0-0.67	0-0.59	0-0.23
Post-test						
M	0.14	0.13	0.20	0.07	0.17	0.08
SD	0.21	0.22	0.24	0.14	0.30	0.08
Range	0-0.55	0-0.53	0-0.75	0-0.43	0-0.79	0-0.24
Duration Map 4						
Pre-test						
M	55.67	73.92	40.83	24.75	10.67	19.25
SD	33.98	24.96	30.14	21.57	21.06	30.21
Range	0-89	0-89	0-89	0-58	0-59	0-89
Post-test						
M	53.67	69.33	53.42	17.17	32.42	21.42
SD	35.41	30.46	26.30	22.01	32.66	25.61
Range	0-89	7-89	0-89	0-56	0-81	0-78
Duration Map 5						
Pre-test						
M	6.33	2.42	4.25	19.25	29.08	21.50
SD	13.59	4.70	8.91	20.50	31.68	26.73
Range	0-41	0-14	0-28	0-65	0-89	0-70
Post-test						
M	15.42	3.75	10.83	30.33	14.50	15.92
SD	29.48	7.74	17.02	34.23	28.38	29.59
Range	0-89	0-25	0-54	0-82	0-89	0-83

GEV Map 5						
Pre-test						
M	0.05	0.02	0.03	0.08	0.23	0.11
SD	0.11	0.04	0.06	0.09	0.27	0.17
Range	0-0.34	0-0.12	0-0.19	0-0.25	0-0.69	0-0.54
Post-test						
M	0.12	0.03	0.06	0.22	0.10	0.07
SD	0.25	0.05	0.12	0.29	0.21	0.18
Range	0-0.8	0-0.15	0-0.31	0-0.78	0-0.6	0-0.63
Duration Map 6						
Pre-test						
M	4.33	1.83	11.83	27.42	41.33	38.67
SD	6.75	4.47	18.82	16.21	33.52	34.82
Range	0-18	0-14	0-55	0-56	0-89	0-89
Post-test						
M	2.58	2.42	4.08	33.75	21.75	37.25
SD	7.27	6.71	7.50	27.95	27.73	33.66
Range	0-25	0-23	0-19	0-89	0-89	0-89
GEV Map 6						
Pre-test						
M	0.02	0.01	0.10	0.14	0.24	0.23
SD	0.03	0.03	0.16	0.10	0.21	0.27
Range	0-0.09	0-0.12	0-0.43	0-0.25	0-0.57	0-0.75
Post-test						
M	0.01	0.01	0.02	0.16	0.12	0.22
SD	0.04	0.04	0.05	0.15	0.21	0.24
Range	0-0.13	0-0.12	0-0.13	0-0.45	0-0.74	0-0.66