Ageing-related changes in cognitive functions and their associations with event-related potentials and physical fitness in healthy older adults

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

Heikki Mannila

Juhani Vainio

Master's thesis

Department of Psychology

University of Jyväskylä

December 2019

Acknowledgements

Ageing manifests at least on the levels of structural degradation of the brain, as well as a slight cognitive decline even in healthy older adults. It is widely recognized that better physical fitness helps preserve cognitive abilities in older age and delay the inevitable cognitive decline. In this follow-up study, we investigated in what relation cognitive abilities, physical fitness and event-related potentials (ERPs) develop over approximately five years, and whether the development of cognitive abilities is possible to be predicted with the baseline of physical fitness or the ERPs.

Out of the 90 older women who participated in the baseline measurement, 35 participated in this follow-up. We measured ageing-sensitive ERPs with an electroencephalography (EEG) to either somatosensory tactile stimulation or auditory sinusoidal sound stimulation. To assess the participants' cognitive capacity, we conducted a battery of neuropsychological tests that assessed cognitive domains sensitive to ageing. In addition to these, the participants carried out a six-minute walk test to measure their aerobic physical fitness.

The somatosensory P3a and the auditory P2 showed significant change over time, which reflects their ageing-sensitive nature. Decreased six-minute walk test distance was associated with a decline in explicit memory. The enlargement of the auditory P2 standard response associated with improved explicit memory, indicating neural compensation. In addition, the enlargement of the deviant response of the auditory P2 associated with worsened six-minute walk test performance. Attenuation of the deviant response of the somatosensory P3a associated with a prolonged performance time in *Stroop 3* test. Higher baseline of six-minute walk test performance predicted better preservation of explicit memory. Additionally, higher (more negative) baseline of the auditory N1 predicted improvement in explicit memory performance and an increase in error susceptibility.

The results of this longitudinal study indicate a connection between aerobic physical fitness and explicit memory. Additionally, we got further support indicating that the auditory P2 seems to be linked to neural compensation, as activity in fronto-central topography increased as well from the baseline measurement to the follow-up. The somatosensory P3a seems to be associated to executive functions and attentional modulation. Better preservation of explicit memory was able to be predicted with the baseline of physical fitness. In addition to P2, the predicting ability of N1 on cognitive preservation may reflect neural compensation in older adults. The results of the present study are of great importance to older adults who wish to maintain their cognitive abilities for as long as possible, and to other researchers, as valuable evidence of complicated ageing-related processes.

Keywords: event-related potential (ERP), ageing, cognitive performance, physical fitness, neuropsychological assessment, follow-up study

JYVÄSKYLÄN YLIOPISTO

Psykologian laitos

MANNILA, HEIKKI & VAINIO, JUHANI: Terveeseen ikääntymiseen liittyvät kognitiiviset muutokset ja niiden yhteydet aivojen herätevasteiden ja fyysisen kunnon muutoksiin

Pro gradu -tutkielma, 51 s., 2 liitettä.

Ohjaaja: Juho Strömmer

Psykologia

Joulukuu 2019

Ikääntyminen näkyy aivojen rakenteellisena rappeutumisena sekä lievänä kognitiivisten kykyjen heikkenemisenä jopa terveillä ikääntyneillä. On myös laajalti todettu, että hyvä fyysinen kunto auttaa säilyttämään hyvän kognitiivisen toimintakyvyn pidempään. Tämän seurantatutkimuksen tarkoitus oli selvittää, kuinka aivojen herätevasteiden, fyysisen kunnon ja kognitiivisten kykyjen muutokset ovat yhteydessä toisiinsa viiden vuoden seurannassa. Lisäksi tarkastelimme, kuinka herätevasteiden ja fyysisen kunnon lähtötasolla voidaan ennustaa kognitiivisten kykyjen muutosta.

Ensimmäisen mittauksen 90 ikääntyneestä naisesta 35 osallistui tähän seurantaan. Mittasimme elektroenkefalografialla (EEG) sekä somatosensorisia että auditorisia aivovasteita. Koehenkilöiden kognitiivisia kykyjä mitattiin neuropsykologisella testipatterilla, joihin oli valittu ikääntymiselle herkkiä kognition osa-alueita mittaavia testejä. Koehenkilöt suorittivat myös aerobista fyysistä kuntoa mittaavan kuuden minuutin kävelytestin.

Herätevasteista somatosensorisessa P3a-komponentissa sekä auditorisessa P2-komponentissa havaittiin tilastollisesti merkitsevä muutos seurannan aikana, mikä on merkki niiden herkkyydestä ikääntymiselle. Kuuden minuutin kävelytestin tuloksen heikkeneminen oli yhteydessä eksplisiittisen muistisuoriutumisen heikkenemiseen. Auditorisen P2-komponentin standardiärsykkeen vasteen voimakkuuden kasvu oli yhteydessä parantuneeseen eksplisiittiseen muistiin, mikä voidaan nähdä merkkinä hermostollisesta kompensaatiosta. Lisäksi poikkeavan ärsykkeen vasteen voimakkuuden kasvu oli yhteydessä huonontuneeseen kuuden minuutin kävelytestin tulokseen. Somatosensorisen P3a-komponentin vasteen voimakkuuden lasku poikkeavaan ärsykkeeseen oli yhteydessä pidentyneeseen suoritusaikaan *Stroop 3* -testissä. Kuuden minuutin kävelytestin korkeampi lähtötaso ennusti puolestaan parempaa eksplisiittisen muistin säilymistä, minkä lisäksi auditorisen N1-vasteen amplitudiltaan korkeampi lähtötaso ennusti suurempaa eksplisiittisen muistisuoriutumisen paranemista ja kohonnutta virhealttiutta.

Tämän pitkittäistutkimuksen tulokset viittaavat eksplisiittisen muistin ja aerobisen kunnon väliseen yhteyteen. Saimme tutkimuksen myötä lisää todisteita siitä, että auditorinen P2-komponentti heijastaisi hermostollista kompensaatiota, mitä tukee myös aktiivisuuden lisääntyminen otsa- ja päälaenlohkojen alueella mittausten välillä. Somatosensorinen P3a-komponentti puolestaan vaikuttaa olevan yhteydessä toiminnanohjaukseen ja tarkkaavuuden säätelyyn. Eksplisiittisen muistin säilymistä voitiin ennustaa aerobisen kunnon lähtötasolla. P2-vasteen tavoin auditorisen N1-vasteen kyky ennustaa eksplisiittisen muistin kasvua voi olla merkki hermostollisesta kompensaatiosta ikääntyneillä. Tämän tutkimuksen tulokset ovat tärkeää tietoa niin ikääntyneille, jotka haluavat hidastaa kognitiivista rappeutumista, kuin myös tutkijoille lisäinformaationa monimutkaisista ikääntymiseen liittyvistä kognitiivisista prosesseista.

Avainsanat: herätevaste, ikääntyminen, kognitiiviset kyvyt, fyysinen kunto, neuropsykologinen arviointi, pitkittäistutkimus

Table of Contents

INTRODUCTION

Increased life expectancy and declined fertility rate have led to a rapid increase in the ageing population (United Nations, 2017). From 1970 to 2015, life expectancy has risen by 10.5 years to reach an average of 80.6 years in OECD countries (OECD, 2017). In the European Union the number of over 65 year-olds from total population is projected to rise from 19 % to 29 % between the years 2016 and 2060 (European Commission, 2018), while in Finland the share is supposed to reach even 31 % from the current 22 % by 2060 (SVT, 2018). This rapid ageing of the population worldwide has led to neurological disorders being the leading source of disability globally (Feigin et al., 2017). For instance, Prince et al. (2013) estimate that there were 35.6 million people suffering from dementia in 2010, with expectation for the number to nearly double every 20 years, to around 66 million in 2030 and 115 million in 2050.

This rapid ageing creates new issues for societies, which must respond to new requirements of the ageing population. The increase in ageing population creates broader social and economic challenges also for younger generations, who are responsible to secure the pension and care for the aged. This increase in both the ageing population and age-related neurological disorders has brought up urgency for a better understanding of cognitive ageing and the prevention of clinical age-related cognitive deterioration. To secure the quality of life of aged people and to facilitate the economic and social burden, it is necessary to investigate ways to maintain and support their physical and psychosocial well-being. By understanding how healthy older adults have been able to maintain and preserve their cognitive functioning, we are able to develop tools and guidelines to contribute to the prevention of cognitive diseases. Therefore, especially research for normal ageing is required, to extend the target group outside of the pathological conditions and get valuable information about normal cognitive ageing.

Ageing and cognition

Many neurobiological and cognitive changes occur in normal ageing. On both cortical and subcortical levels, ageing is accompanied with a decrease in gray matter volume and an increased shrinking of white matter (Fjell et al., 2009; Raz et al., 1997, 2005; Resnick, Pham, Kraut, Zonderman & Davatzikos, 2003). As the degradation of brain structures occurs increasingly in ageing, many cognitive and sensory processes, such as memory, attention and executive functions are impaired,

while for instance vocabulary and decision making can even be improved (Grady, 2012; Kaup, Mirzakhanian, Jeste & Eyler, 2011; Park et al., 2002; Verhaeghen, 2003). In addition to slight loss of brain structure, many other factors also contribute to the inevitable cognitive degradation in ageing. Lifestyle factors, such as intellectual hobbies, physical exercise, education and sleep all have an effect on cognition and can help slow down the cognitive decline (Bastien et al., 2003; Colcombe & Kramer, 2003; Davidson & Winocur, 2010; Dzierzewski, Dautovich, & Ravyts, 2018; Strömmer et al., 2017). As the spectrum of age-related cognitive changes is very broad, it is necessary to consider multiple interacting processes while trying to understand the nature of cognitive decline.

Typically, older adults show less brain activity and perform worse in cognitive tasks compared to young adults (see e.g. Grady, 2012). However, in areas such as the aforementioned vocabulary, semantic priming, reading fluency, or emotional processing, they can even perform as well as, or better than young adults (Carstensen et al., 2011; Carstensen, Fung, & Charles, 2003; Christensen et al., 1999; Laver, 2009). From a neurophysiological perspective, this is suggested to be due to neural compensation. The compensation hypothesis suggests that older adults' brains compensate for the ongoing cognitive decline by overactivation of certain brain areas, especially the prefrontal cortex, and implementing new neural networks (Cabeza et al., 2004; Reuter-Lorenz & Park, 2010). This compensation occurs mostly in tasks requiring higher-level cognitive processing, such as working memory, executive functions and inhibitory control (Martins, Joanette, & Monchi, 2015). Typically, according to the neural compensation hypothesis, the amplitudes of event-related potential (ERP) components related to cognitive control decrease in the posterior areas, while the components in the prefrontal cortical areas increase with age (Kropotov, Ponomarev, Tereshchenko, Müller, & Jäncke, 2016; Lubitz, Niedeggen, & Feser, 2017).

Another theory trying to account for some of the age-related cognitive changes is the processing speed theory (Salthouse, 1996). It suggests, that as age increases, the processing time of information increases, as well. This leads to impairment of behavioral cognitive functioning because of the increased time it takes to execute cognitive processes. The processing speed theory has been proven to be valid in several different experimental settings. It appears to be an important mediator of skill acquisition (Brigman & Cherry, 2002). Processing speed, as many other cognitive factors as well, is affected by aerobic endurance (Zettel-Watson, Suen, Wehbe, Rutledge, & Cherry, 2017). Zettel-Watson et al. (2017) found that better dynamic balance and aerobic endurance predicted better performance in cognitive tasks measuring processing speed, working memory and inhibition in older adults.

The inhibitory deficit hypothesis (Hasher & Zacks, 1988) is another noteworthy proposal for the underlying reasons for cognitive-behavioral changes associated with ageing. The

hypothesis suggests that under some circumstances, as in ageing, the efficiency to inhibit irrelevant sensory information is reduced. In other words, as information enters working memory during a task that requires selective attention, the brain's ability to select which information is goal-relevant is impaired. The weakened inhibitory control affects cortical early sensory processing and is also associated with impaired performance on behavioral level (Bolton & Staines, 2012). On both somatosensory and auditory modalities, the ability to inhibit task-irrelevant sensory information has been found to be declined in older adults in ERP studies (Bolton & Staines, 2012; Chao & Knight, 1997).

As elaborated above, normal ageing affects cognitive processing widely, with declines seen at least in working and episodic memory, information processing speed and inhibiting irrelevant information, respectively (Harada, Natelson, Love, & Triebel, 2013; Rönnlund, Nyberg, Bäckman, & Nilsson, 2005; Strömmer et al., 2017). In addition, interindividual variation in age-related change in cognitive performance is broad and can be seen especially in cognitive domains vulnerable to ageing, such as working memory and processing speed (Christensen et al., 1999; Wilson et al., 2002). The interconnectedness of these processes alongside with the aforementioned deterioration of brain structure makes cognitive ageing particularly challenging to study, and many underlying factors still remain unclear. Moreover, most changes in the brain's anatomy, physiology and in neuropsychological performance in ageing occur simultaneously, which makes their separate analysis rather challenging.

Event-related potentials (ERPs) have proven useful in deepening understanding of cognition-related automatic sensory processes. Many ERPs reflecting specific ageing-related cognitive processes have been discovered (see e.g. Kisley, Davalos, Engleman, Guinther, & Davis, 2005; Stothart & Kazanina, 2016; Strömmer et al., 2017; Van Dinteren, Huster, Jongsma, Kessels, & Arns, 2018).

ERPs

Event-related potentials (ERPs) are electric potentials created by electromagnetic activity in the brain and measured with an electroencephalography (EEG) from the scalp. They occur after sensory stimulation, related to sensory, motor or cognitive events (Sanei & Chambers, 2007). ERPs represent time-specific activation in the brain, as a result of voltage fluctuations caused by a large number of simultaneous action potentials. This simultaneous activation of thousands of neurons creates a

challenge for accurate location of the source of stimulation-related electromagnetic activity. However, ERPs obtained with EEG are a direct measure of the brain's activity and temporally very accurate. Therefore, they are an extremely useful tool for studying the brain's pre-attentive sensory processing and complex conscious processes that are related to higher, cognitive or affective brain functions (Partanen, Lang, Valkonen-Korhonen & Cheour, 2006).

Ageing and somatosensory change detection ERP components

Sensory gating is a pre-attentive mechanism of the brain's inhibitory function, which helps the brain differentiate between relevant and irrelevant stimuli. It occurs during early cortical processing of sensory stimuli and appears as an attenuated response to recurring identical stimuli (Freedman et al., 1987). The attenuation is suggested to reflect inhibition of irrelevant sensory information and thus facilitate higher-order cognitive processing (Braff & Geyer, 1990; Wan et al., 2008). The sensory gating ability is what is thought to deteriorate according to the inhibitory deficit hypothesis (Hasher & Zacks, 1988). Sensory gating has been widely studied with ERPs, and the component P50 is mostly linked to this inhibitory function mechanism (see e.g. Freedman et al., 1987; Lijffijt et al., 2009). P50 has been studied broadly at least in the somatosensory and auditory modalities and shown to be independent of attention, which means that its measured amplitude and latency stay the same whether the stimulus is attended or not (Bolton & Staines, 2011, 2012; Jerger, Biggins, & Fein, 1992; Lijffijt et al., 2009; Strömmer et al., 2014).

Ageing has been found to influence the P50 amplitude, in a way that older adults tend to elicit responses of higher amplitude compared to young adults in auditory and somatosensory modalities (Friedman, 2012; Strömmer et al., 2017), indicating weaker sensory inhibition. Additionally, ageing affects the latency of the P50 component, as it was found to be delayed in responses to deviant stimuli of older adults in a passive experimental procedure (Strömmer et al. 2017) and for both attended and unattended stimuli in an active somatosensory oddball experiment (Bolton and Staines, 2012). However, there is also evidence that neither P50 amplitude nor latency differed between groups of older and young adults in the auditory modality (Gmehlin, Kreisel, Bachmann, Weisbrod & Thomas, 2011).

Previous research about age-related changes in the somatosensory N80 component is scarce. Strömmer et al. (2014) found that older adults showed prolonged N80 responses to standard recurring stimuli compared to young adults. In their later study, Strömmer et al. (2017) found that older adults elicited responses of higher amplitude to both standard and deviant stimuli.

The mismatch negativity (MMN), or in the somatosensory modality, the mismatch response (MMR) is an automatic cortical event-related response, which appears when the brain detects a stimulus that differs from a flow of recurring stimuli (Näätänen, Gaillard, & Mäntysalo, 1978). Somatosensory mismatch response (sMMR) is not as broadly studied as its auditory counterpart. There is ambiguous evidence of the polarity of sMMR, due to differences in stimulus onset asynchronies (SOA) and experimental procedures. A few studies have found a negative polarity somatosensory mismatch response (Kekoni et al., 1997; Shen, Smyk, Meltzoff, & Marshall, 2018), but there are studies having found a positive polarity mismatch response as well (Akatsuka et al., 2005; Strömmer et al., 2014). Strömmer et al. (2014) measured the sMMR at a latency range of 250−290 ms after stimulus onset and found that the sMMR was attenuated in older adults compared to young adults. Similarly, Strömmer et al. (2017) found that in comparison to young adults, older adults had attenuated sMMRs. In addition, they found that within the group of older adults, higher sMMR amplitude was related to better performance in cognitive tasks related to executive functions.

The P3 component occurs around 300 ms post-stimulus (Friedman, Cycowicz & Gaeta, 2001; Squires, Squires & Hillyard, 1975), depending on factors such as stimulus modality, task conditions and subject age (Polich, 2007). P3 represents automatic shift of attention, while its subcomponent, the P3a is suggested to reflect the evaluation of proper behavioral action to a deviant stimulus that the brain perceives as significant (Friedman, Cycowicz & Gaeta, 2001; Polich, 2007). In oddball conditions, P3a peaks in latency ranges of approximately 220–280 ms (Squires et al., 1975), occuring in fronto-central topography (Nieuweinhuis, Aston-Jones & Cohen, 2005; Polich 2007; Yamaguchi & Knight, 1991). Investigating the somatosensory P3a, Strömmer et al. (2017) reported, that the group of older adults had prominent activation to deviant stimuli in somatosensory P3a only in central electrode sites, while the group of young adults had prominent activation in fronto-central topography. Bolton and Staines (2012) found, that in response to attended stimuli, the somatosensory P3 response was attenuated in amplitude and delayed in latency with older adults. Similar results have been found in auditory and visual modalities (Pontifex, Hillman, & Polich, 2009; Walhovd & Fjell, 2003). Ito, Takamatsu and Kimura (1996) replicated the results with delayed latency of somatosensory P3, but they did not find significant effect of ageing in amplitude. In general, short latency and high amplitude of the P3 are associated with good cognitive performance (Jausovec & Jausovec, 2000; Polich, 1996, 2007).

Studies on the effects of physical activity on the P3a component are scarce, mostly because of the specificity of the component. Strömmer et al. (2017) found a robust positive correlation between the somatosensory P3a amplitude and walk test performance, suggesting that better automatic stimulus evaluation is connected with better aerobic endurance. However, in active visual

oddball task Pontifex et al. (2009) did not find effects on P3a amplitude or latency changes relative to cardiorespiratory fitness.

Ageing and auditory change detection ERP components

N1, sometimes referred to as N100, is an ageing-affected ERP component. N1 reflects auditory stimulus detection and discrimination (Tomé, Barbosa, Nowak, & Marques-Teixeira 2015), and it occurs in the auditory cortex approximately 100 ms after tone onset (Näätänen et al., 1978). Näätänen et al. (1978) showed that N1 is independent of attentional focus. Typically, older adults elicit N1 responses of higher amplitude, when compared to young adults (Amenedo & Diaz, 1998; Anderer, Semlitsch, & Saletu, 1996; Chao & Knight, 1997; Kisley et al., 2005, Talsma, Kok, & Ridderinkhof, 2006). Especially in cases where recurring similar stimuli are presented, N1 is less suppressed in older adults than in young adults, which may be a sign of poorer inhibitory control (Fabiani et al., 2006). Strömmer et al. (2017) got similar results, wherein the auditory N1 response to standard stimuli had larger amplitude in older than in young adults. However, Stothart and Kazanina (2016) got controversial results in their study, where age had no effect on the amplitude of the response. Tomé et al. (2015) summarized in a review, that N1 amplitude stays approximately the same after adolescence when measured from the central electrode (Cz) (see Appendices for all electrode locations and electrode pools used in the present study). However, N1 amplitude measured from the frontal electrode (Fz) seems to decrease until the age of 60 years, and then increase rapidly in over 60-year-olds. There is no clear evidence of the N1 amplitude relating to cognitive performance in older adults. Some studies have measured N1 and cognitive performance between older and young adults but have not reported effects of N1 on cognition (Strömmer et al., 2017; Talsma et al., 2006).

The mismatch negativity (MMN) has mostly been recognized and studied in the auditory modality and is considered a valuable tool in studying stimulus change discrimination, central auditory information processing and pre-attentive auditory memory (Schröger, 1997). The MMN is elicited in paradigms where attentional focus or behavioral responses are not required, and thus it is a very optimal tool for studying age-related changes in task-irrelevant sensory processing (Alain & Woods, 1999). The range of the auditory MMN (aMMN) latency is typically 150–250 ms post-stimulus (Näätänen, Astikainen, Ruusuvirta & Huotilainen, 2010). Numerous studies and meta-analyses show that the effects of ageing to the aMMN amplitude are controversial: some studies report the amplitude to decline with age (see for example Alain & Woods, 1999; Cheng, Hsu & Lin, 2013; Kisley et al., 2005), whereas other studies report young and older adults to show similar amplitudes

(Müller, Brehmer, Von Örtzen, Li, & Lindenberger, 2008; Ruzzoli et al., 2012; Strömmer et al., 2017). Of the importance of this paper, cross-sectional research between older and young adults have found decreased aMMN amplitude to be related with poorer cognitive performance in older adults, especially in verbal memory and executive functions (Foster et al., 2013; Kisley et al., 2005).

P2 is another component reflecting age-related cognitive changes that occurs approximately 150–250 ms after stimulus onset (Crowley & Colrain, 2004). The auditory P2 component reflects stimulus classification, as well as processing its physical and contextual features (Garcia-Larrea, Lukaszewicz & Mauguiére, 1992; Yingling & Nethercut, 1983), reaching the peak amplitude in vertex-centered topography (Čeponienė, Westerfield, Torki & Townsend, 2008; Picton, Stuss, Champagne & Nelson, 1984; Yingling & Nethercut, 1983). It is not clear how ageing affects the P2 component. In active oddball condition ageing related increase in amplitude of the auditory P2 component was reported in frontal scalp location and in non-target tones (Amenedo & Díaz, 1998; Anderer et al., 1996; Ford & Pfefferbaum, 1991). However, there are also studies with results of no significant change in amplitude (Barrett et al., 1987; Brown, Marsh & LaRue, 1983; Iragui et al. 1993). The rarely used passive oddball condition was used in a cross-sectional study, that indicates a decrease of the P2 amplitude in ageing (Czigler, Csibra & Csontos, 1992). Strömmer et al. (2017) found that similarly to the auditory N1 component, the auditory P2 response to standard stimuli had larger amplitude in older than in young adults. The P2 amplitude was found to be lower in cognitively normal older adults when compared to older adults with probable mild cognitive impairment (MCI) measured with the Montreal Cognitive Assessment (Lister et al., 2015). However, a 4-month cognitive training with older adults decreased the latency of P2 but did not have significant effect on amplitude (Küper, Gajewski, Frieg, & Falkenstein, 2017).

The connection between physical activity and the P2 component is not broadly studied. There is some evidence that higher physical activity decreases the latency of the P2 component (Gajewski & Falkenstein, 2015; Polich & Lardon, 1997; Özkaya et al., 2005). In addition, findings of Özkaya et al. (2005) indicate, that compared to endurance training, physical strength training increases the amplitude of the P2 component. However, other studies have not found statistical significance between the P2 amplitude and physical activity (Gajewski & Falkenstein, 2015; Polich & Lardon, 1997). Thus, the consensus about how physical fitness and ERPs together affect cognition is still missing and the area requires more research.

Physical fitness and cognition

Overall, physical activity has been found to improve the quality of life in older adults (Morris, Vaccaro and Harris, 1983). Additionally, even low levels of physical activity help avoid cognitive decline especially in women (Laurin, Verreault, Lindsay, MacPherson & Rockwood, 2001; Sofi et al., 2011). The earlier physical activity is started in life, the better it protects from cognitive impairment (Middleton, Barns, Lui & Yaffe, 2010). However, even if started later in life, physical activity was proven to lower the risk of cognitive impairment in women.

Various physical fitness measures have been used when studying the connections between physical and cognitive performance. Zettel-Watson et al. (2017) and Strömmer et al. (2017) found that aerobic endurance, measured with a six-minute walk test, was connected to better cognitive performance. Ji et al. (2018) showed in a study with older adults that a six-week dance training improved the participants' memory performance and gait speed in a walk test. Another study conducted by Ponce-Bravo, Ponce, Feriche and Padial (2015) concluded that functional physical training program of 20 sessions, but not recreational training, showed improvements in cognitive performance measured by reaction times in older adults.

Colcombe and Kramer (2003) presented in their meta-analysis, that any kind of fitness training can enhance cognitive abilities in older adults. The effect was especially robust on executive functioning, but other cognitive domains were also improved. The best result was achieved with a combination of strength and aerobic training. However, participants only participating in aerobic training still showed improvements in cognitive performance. Furthermore, Smith et al. (2010) reported in their meta-analysis of randomized controlled trials, that aerobic exercise mediated improvements in performance in various cognitive domains in healthy older adults, including executive functions, processing speed and attention, but not working memory. They found, similarly to Colcombe and Kramer (2003), that aerobic exercise combined with strength training increased neurocognitive performance more than aerobic exercise alone. However, Young, Angevaren, Rusted and Tabet (2015) reported in a review, that no significant effects were found between aerobic physical activity and cognitive performance. This contradictory result may be due to other moderating factors not considered in the review.

These findings speak for the importance of physical activity especially in the ageing population. Whether physically active lifestyle actually mediates better cognitive performance in later life, is still somewhat inconclusive. The present longitudinal study sheds more light on this issue.

Aims of this study

In the present follow-up study, we investigated changes in neural sensory processing, as well as in cognitive and physical performance in healthy older adults in a five-year follow-up. We measured how cognitive performance, physical fitness and automatic event-related brain potentials have changed over time. Furthermore, we investigated how changes in physical fitness and automatic brain potentials and change in cognitive performance were connected, and how the baseline in physical fitness and brain responses were related to changes in cognitive performance.

Research on the subject suggests that physical fitness affects cognitive abilities, either helping preserve or enhance them. Additionally, research on brain functions, more specifically ERPs, shows that they are a valid tool in studying the brains automatic pre-attentive sensory-cognitive processing and thus provide a valuable help in investigating comprehensively how ageing affects cognition. In this field, the consensus is still missing, and thus, the present study elucidates the relationship between ageing-related ERPs and cognition.

Based on the inhibitory deficit hypothesis (Hasher & Zacks, 1988), we hypothesize that the participants' sensory gating ability is impaired, and thus show larger activation in the follow-up measurements in the P50 component, compared to the baseline measurement circa five years earlier (1). Additionally, as stated in the processing speed theory (Salthouse, 1996), we hypothesize that in the follow-up measurement, pre-attentive sensory processing is slowed down i.e. the latencies of the ERP components are longer (2). Based on the previous study of Strömmer et al. (2017), we hypothesize that declined physical fitness is associated with declined cognitive performance, and vice versa (3). Additionally, due to the missing consensus about how ERPs and cognition are related, we hypothesize that any change observed in the ERPs is associated with a change in cognitive performance (4). Finally, based on the statements of Colcombe and Kramer (2003), we hypothesize that better baseline in physical fitness predicts better preservation of cognitive abilities (5).

METHODS

Participants

We conducted a follow-up measurement for participants that had previously been measured in 2013–2014. The results of that study are reported by Strömmer et al. (2017). For the present study, we were able to recruit 35 of the original 90 older women who participated in 2013–2014. All measurements were carried out between October 2018 and January 2019. All participants were women, right-handed and lacked any history of neurological or psychiatric illness or brain operations. The age of the participants ranged from 67 to 85 (mean ± SD, 72.7 ± 4.7). The educational background of the participants was assessed in the background questionnaire as follows: elementary level (n = 8), upper secondary level (n = 18), undergraduate level (n = 3), and graduate level (n = 9). The participants were originally recruited from the University of the Third Age in Jyväskylä and the Society of the Retired in Jyväskylä as well as through an announcement in the local newspaper. Ethical approval for the study was obtained from the ethical committee of the Central Finland Health Care District. The experiments were undertaken in accordance with the Declaration of Helsinki. Written informed consent was acquired from all participants.

Neuropsychological assessment

The neuropsychological assessment was carried out in the Department of Psychology at the University of Jyväskylä. We used a test battery consisting of a variety of neuropsychological tests to measure the participants' cognitive performance. We used all the same tests as Strömmer et al. (2017) did in 2013–2014 measurements, with a few additional tests, in order to substantiate assessment of ageing-affected cognitive domains. Tests that measure ageing-sensitive cognitive functions such as processing speed, cognitive inhibition and working memory were selected into the battery. The testing procedure was fixed, with the same order of tests and same instructions for every participant. A psychologist or a trained psychology student carried out the testing right before the EEG measurement, except for the delay-sections of some of the tests. The testing lasted for about one hour. The following tests were included in the test battery:

The Stroop Color-Word test (Alvarez & Emory, 2006): The test engages especially the prefrontal cortex and measures executive functions and cognitive inhibition of information. The test consists of

three sections. The first section measures reading fluency and consists of an A4-sheet that has color-words written on it. The objective is to read the words out loud. The second section consists of an A4-sheet that has letter 'Xs' written on it with colors of either yellow, blue, green, or red. The objective is to name the colors. The second section measures color-naming fluency. The final section consists of an A4-sheet that has color-words written in an incongruent color. The objective is to name the colors that the words are printed in, thus prompting inhibition to read the word out loud. The section measures the ability to inhibit irrelevant stimuli. In all sections, the participant was instructed to perform as fast as possible, while avoiding mistakes.

The logical memory task (WMS-R): The test was used to measure immediate and delayed declarative and auditory memory. The participant was told a short story and then asked to immediately repeat the story as precisely as possible. The participant was then told a second story and again asked to repeat it as precisely as possible. Thereafter, the participant was told to memorize the two stories and told that they would be asked to repeat the stories again in about one hour, in between the EEG experiments.

Visual reproduction task (WMS-R): The test was used to measure immediate and delayed memory of visual stimuli. The test consists of five images, each of which are shown for seconds, which after the participant was asked to copy the image from memory as accurately as possible. The delay-part was after about one hour and the participant was asked to draw the images from memory in no specific order.

The digit span task (Ramsay & Reynolds, 1995): The test was used to measure auditory working memory. In the test, the participant was told sequences of random numbers and then asked to repeat them in the same order. If the sequence was recalled correctly, the sequences eventually grew in length. The same test was also performed backwards.

The letter-number sequencing task (Crowe, 2000): The test measures complex auditory working memory and executive functions. The participant was told sequences containing both letters and numbers. Thereafter, the participant was told to repeat the sequences, numbers first in numerical order and then letters in alphabetical order.

Trail Making Test A & B (TMT-A, TMT-B) (Bowie & Harvey, 2006): Trail Making Test A was used to measure basic attention of the participants. In the test the participant was instructed to connect 25

number circles with a pencil in numerical ascending order. TMT-B, respectively, was used to measure fast attention-shifting capacity. The test consists of circles of both letters and numbers on an A4-sheet and the participant was instructed to connect the circles of letters and numbers in turn, the numbers in numerical ascending order and the letters in alphabetical order (1-A-2-B-3-C etc.). In both tests, the participant was instructed to complete the test as fast as possible, while avoiding mistakes.

Rey Auditory Verbal Learning Test (AVLT) (Rey, 1964): The test was used to measure participants' verbal auditory short-term memory. The participant was told a list of 15 words and instructed to repeat as many of the words as possible, in no particular order. This process was then repeated four times. TMT-A was done after the five repetitions of the 15-word lists, which after the participant was told to once more repeat the word list without the instructor reading the list first. The participant was not told about the delayed recall section, which was done after about one hour.

The finger tapping task (Ruff & Parker, 1993): The test was used to assess the participants' motor control and psychomotor speed. The participant was handed a mechanical tally counter and instructed to hold it in their hand, placing the metal ring of the tally counter through their middle finger, so the device is steadily in their palm. They were then instructed to press the button in the counter with their thumb, making the counter of the device advance one number per press. The participant was instructed to press the button as many times as possible in ten seconds. Measurements were performed three times from both hands.

The handgrip task (Saehan, SH5001, Belgium): We used a hand dynamometer to measure the participants' isometric handgrip strength. Handgrip is known to predict declines in cognition, functional status and mobility in older adults (for a review, see Rijk, Roos, Deckx, van den Akker, & Buntinx, 2015). The handle of the dynamometer was adjusted individually for each participant, in order to make it fit their hand. The participant was instructed to hold the handle in their hand, with the hand in the air (not resting on the table), the elbow at about 90 degrees angle by the side of the body. The participant was then told to squeeze the handle with maximum effort for a couple of seconds. The maximum handgrip strength was recorded in kilograms. The procedure was completed three times for both hands.

Physical fitness measures

Participants' physical fitness was assessed with four separate measures: total body fat percentage, body mass index (BMI), a six-minute walk test (Crapo, 2002) and self-reported physical activity. Total body fat percentage was measured with dual-energy X-ray absorptiometry (DXA) (Delphi QDR series, Hologic, Bedford, MA, USA). Participants were instructed to avoid eating right before the measurement. During the DXA-scan, participants lay still in the machine for about 10 minutes. After completing the DXA measurement, participants took part in a six-minute walk test. The instructions for the walk test were to walk as long a distance as possible in six minutes. The test was done on a 200-meter indoor track. Participants' heart rates were monitored after every minute. BMI was calculated using the formula $BMI = \frac{mass(kg)}{height^2(m)}$. Self-reported physical activity was assessed in the background questionnaire with a five-scale question of medium-intensity exercise hours per week. The scale was <1, 1–2, 2–3, 3–4, and >5 hours.

Stimuli and procedure

The measurements were carried out in the EEG laboratory of the Department of Psychology at the University of Jyväskylä. During the whole experiment, the participant was sitting on a chair in an electrically shielded room, about 1.5 meters away from a TV screen. They were instructed to avoid any excessive movement, talking, muscle tension and blinking during the measurement. The participant was monitored the whole time via a video camera, and the instructors were able to talk to the participant from the control room via a microphone, and vice versa.

The actual measurement included two separate experiments: a somatosensory experiment and an auditory experiment. In both experiments, a run of 1000 stimuli of two types varying in either location (somatosensory) or frequency (auditory) were presented. Stimuli were delivered with a randomly varying stimulus onset asynchrony (SOA) of 400, 450, or 500 ms. The relatively short SOA was selected based on earlier findings by Strömmer et al. (2014), showing age-related changes in the amplitude of the sMMR with SOA of 500 ms providing thus a solid basis for the cross-modal investigation. During the experiments, the participant was watching a movie. Both experiments were carried out in a passive oddball condition: 'standard' stimuli were frequently presented at a probability of 86 %, and rare 'deviant' stimuli were presented at a probability of 14 %. The somatosensory experiment was always followed by the auditory one.

We used a constant current stimulator to generate the somatosensory stimuli (Digitimer Ltd, model DS7A, Welwyn Garden City, UK). The electrical pulses were delivered via metal ring electrodes attached to the participant's index and little fingers in a way, that one electrode was attached to the cathode above the proximal phalanx and one to the anode above the distal phalanx of both fingers. The electrodes were first moistened with conductive gel (Technomed Europe Ltd, Maastrich, Netherlands), and a piece of gauze was tied between the two electrodes of each finger to prevent conductivity between the electrodes in the same finger. The stimuli delivered were 200 µs long. Stimulus intensities were set individually for each participant, the intensity used in the experiment being twice the subjective sensory threshold. The sensory thresholds were determined by giving electrical pulses starting from very low intensities and going up in steps of 1 mA at a time. The participant was asked to report as soon as they would sense any stimulation at all. This procedure was done separately for index and little fingers of each participant. 10 participants reported the experiment intensity to be too high after setting the intensity to twice the somatosensory threshold. For these 10 participants the stimulus intensities used during the experiment were set lower, in order to make sure they are able to sit through the measurement. Stimuli to both fingers were applied as standard and deviant stimulus in a counterbalanced order to all participants.

The auditory stimulation was presented from a loudspeaker placed 90 cm above the participant. The stimuli were sinusoidal sounds, 50 ms in duration with a 10 ms onset and offset time. The sounds were 75 dB in intensity and either 1000 Hz or 750 Hz in frequency. Both frequencies were applied as standard and deviant stimuli and the conditions were counterbalanced across all participants. Individual hearing thresholds of 500 Hz and 1000 Hz for both ears were measured prior to the EEG experiment, using an audiometer (Mediroll SA-51, Mediroll Ltd, Debrecen, Hungary). While measuring the hearing thresholds, we started from very low intensities going slowly up in steps of 5 dB, until the participant reported hearing the sound. The somatosensory thresholds and the hearing thresholds are reported in Table 1.

Table 1. *Auditory and somatosensory thresholds and stimulus intensities. The differences between measurement times were tested with paired samples t-tests (two-tailed). SD, standard deviation; CI, confidence interval; p, statistical significance; p(fdr), FDR-corrected p-value; * p ≤ .05; ** p ≤ .01; *** p ≤ .001; d, Cohen's d.*

Sensory threshold and intensity	T1 mean ± SD	T2 mean ± SD	Mean difference (95 % CI)	p	p(fdr)	d
Somatosensory						
Forefinger threshold (mA)	24.0 ± 6.7	24.4 ± 6.1	0.4 (-1.8 to 2.3)	.727	.824	.07
Little finger threshold (mA)	21.2 ± 5.5	22.4 ± 7.1	1.3 (-1.1 to 4.0)	.389	.778	.19
Forefinger intensity (mA)	48.0 ± 13.4	45.4 ± 10.3	-2.6 (-6.8 to 0.9)	.204	.778	-.22
Little finger intensity (mA)	42.5 ± 10.8	42.0 ± 11.6	-0.5 (-4.7 to 3.9)	.824	.824	-.04
Auditory						
Hearing threshold right ear 1000 Hz (dB)	15.3 ± 9.6	19.4 ± 13.7	4.1 (1.6 to 7.1)	.023*	.092	.35
Hearing threshold right ear 500 Hz (dB)	22.6 ± 11.5	24.7 ± 12.2	2.1 (-1.7 to 5.7)	.307	.409	.18
Hearing threshold left ear 1000 Hz (dB)	13.0 ± 8.0	15.4 ± 11.1	2.4 (0.1 to 4.9)	.053	.106	.25
Hearing threshold left ear 500 Hz (dB)	21.4 ± 11.3	22.6 ± 12.7	1.1 (-2.3 to 5.0)	.564	.564	.10

EEG-recording

The EEG data were recorded with a high-impedance amplifier and a 128-channel EGI Sensor Net (Electrical Geodesics Inc., Hydrogel GSN 128, 1.0). The impedances were kept below 80 kΩ throughout the whole experiment. The sampling rate was 1000 Hz and an online bandpass filter from 0.1 Hz to 400 Hz was applied during the measurements.

EEG data analysis

We used Brain Vision Analyzer 2.1 software to preprocess and analyze the EEG data (Brain Products Gmbh). Noise patterns originating from eye blinks were removed with the Gratton & Coles method (Gratton, Coles, & Donchin, 1983). Channels that had poor skin contact or excessive noise were interpolated with spherical spline model. The average response of each channel was set as a new reference. An offline filter was applied with a low cut-off point of 0.1 Hz and a high cut-off point of 20 Hz (24 dB/octave roll-off). We extracted segments of 700 ms in total around each deviant stimulus

and each standard stimulus preceding a deviant stimulus. The time-window extracted started 200 ms before stimulus onset and ended 500 ms after stimulus onset. The 200 ms period before stimulus onset served as a baseline. After the data segmentation, segments with excessive variance in amplitudes (peak to peak amplitudes ranging from -100 to 100 µV) were rejected from final analyses. In addition, periods with activity ranging less than 0.5 µV in a 100-ms range were rejected from further analyses. From the total of 140 trials of deviant (70) and standard (70) stimuli, 133 trials on average were accepted for the grand averages in the somatosensory modality, and 115 in the auditory modality, respectively. Then, the trials were averaged over all participants. Two blocks of data from the somatosensory experiment and one block from the auditory experiment had to be entirely rejected from further analyses because of excessive noise on most channels.

The same ERP components were included in this follow-up as in the original measurement. Of the early components (somatosensory P50 and N80), peak amplitude was extracted from a specific time window (30–70 ms for P50, 60–100 ms for N80, respectively), and was exported along with the latency at peak amplitude for statistical analyses. After visual inspection of the data, we chose to export the mean peak amplitudes and mean latencies from only one channel for each of the two components (channel 93 for P50 and channel 104 for N80, respectively). The same export method was used for both the first and the follow-up measurement in order to conduct reliable comparisons over time.

For the latter ERP components, the same time windows and electrode sites were used as in the original measurement. Originally, permutation tests with BESA Statistics 1.0 (BESA Gmbh) software were used to determine the regions and time windows. The procedure is described by Strömmer et al. (2017). We decided to use the same time windows and electrode locations for better reliability for comparisons. The electrode locations and time windows are listed in Appendix 1 for the somatosensory ERPs, and Appendix 2 for the auditory ERPs, respectively.

Statistical analysis

All statistical analyses were performed with IBM SPSS Statistics 24. We compared the mean amplitude differences of all the ERP components between the baseline and follow-up measurements, as well as differences between stimulus types (standard and deviant) with repeated measures multivariate analysis of variance (MANOVA). For the somatosensory P50 and N80, we used mean peak amplitudes in the analyses, because of the sharply peaking nature of both components. For all

the latter ERP components, we used mean amplitude extracted from the time range individually determined for each component. Additionally, for the early somatosensory ERP components P50 and N80, we compared the differences in mean latency at peak between time points and stimulus types with MANOVA. Whenever a main effect of either time point or stimulus type was statistically significant, paired samples t-test was used to compare the mean differences. False discovery rate (FDR) (Benjamini & Hochberg, 1995) was used to correct the p-values of the t-tests, as we often conducted multiple comparisons within the same test. We used partial eta squared (η_p^2) scores to estimate effect sizes for MANOVA and Cohen's d for t-tests.

Strömmer et al. (2017) conducted a principal component analysis (PCA) in the 2013–2014 measurements to reduce the dimensionality of the 14 cognitive test scores. They ended up finding four principal components (eigenvalue > 1) that were named *Executive function*, *Error susceptibility*, *Working memory* and *Explicit memory*. To be able to conduct reliable comparisons between the original measurement and the follow-up, we conducted a reliability analysis for the individual variables that had the highest loadings to each of the four original principal components. The highest loading variables to each component were TMT-A, TMT-B, Stroop 1, Stroop 2, and Stroop 3 for *Executive function*; Stroop 2 Errors and Stroop 3 Errors for *Error susceptibility*; Digit-span, Digit-span backwards and Digit-Letter for *Working memory*; and Visual reproduction, Visual reproduction delayed, Logical memory and Logical memory delayed for *Explicit memory*. In addition, we decided to add the Rey Auditory Verbal Learning Test (AVLT) scores to the reliability analyses even though the test was not included in the original measurement with all participants (17 of 35 participants completed AVLT in the first measurements). AVLT measures critical age-affected memory processes, so it was profitable to include those variables in the analysis. The computation of the four mean sum scale variables reduced the effect of the missing AVLT scores. We added AVLT 1 (first repetition) and AVLT 5 (fifth repetition) to *Working memory* sum scale, and AVLT immediate (immediate, delayed repetition) and AVLT delay (later, delayed repetition) to *Explicit memory* sum scale. After comparing the reliabilities of the sum scales of the same principal component between the time points and getting assurance that the sum scale variables would be intercomparable, we standardized each of the cognitive test score variables and computed them into new variables calculating the means of each of the variables that had loaded highest for the original principal components. With this procedure, we managed to create four variables for both time points that mimicked the original four principal components, and were reliably comparable over the time points. This was done in order to follow the same methodological approach as Strömmer et al. (2017) did. We used Cronbach's α for the estimate of reliability. The reliabilities within the same variable were very similar in both time points with each variable, indicating a good basis for comparisons. Results

of the reliability analysis for each of the four estimated sum scales in both time points are presented in Table 7.

We calculated the change from the baseline measurement to the follow-up for each of the measured ERPs, cognitive test scores and the four mean sum scale variables of cognitive domains, as well as the physical fitness measures by subtracting the measure of the latter time point from the earlier (T1-T2). These subtractions we computed into new variables.

Two-tailed Pearson's correlation coefficients and partial correlations with age and education as covariates were computed to inspect the relationships between changes in the ERPs, the computed mean sum scale variables of the cognitive domains, as well as the individual cognitive test scores, and the physical fitness measures. Apart from the mean sum scale variables, we only included variables that showed statistical significance over time in the t-tests. In addition, relationships between ERPs and physical fitness measures at T1, and the change in mean variables of cognitive domains were examined with Pearson's correlation coefficients. Statistical significance threshold was $p < .05$.

RESULTS

Somatosensory ERP components

We found no statistically significant differences in the mean latencies of early somatosensory ERP components P50 and N80 between the baseline and follow-up measurements. There was a statistically significant difference between stimulus types in the N80 latencies ($p = .044$), in a way that the latency to deviant stimuli were longer (Table 2). For N80 and P50 peak amplitudes, a statistically significant effect was found between stimulus types, but not between time points. The deviant stimulus had larger peak amplitudes for both P50 and N80 (more positive for P50 and more negative for N80, respectively). There was no statistically significant change in peak amplitudes over time in the two ERP components.

Both later somatosensory ERP components (sMMR and sP3a) showed a statistically significant difference between stimulus types (Table 3), in a way that the response to deviant stimuli were larger in amplitude than to standard stimuli (Table 5). The average of deviant and standard stimuli of the somatosensory P3a component, as well as the deviant response alone showed a statistically significant decrease in amplitude over time (Table 4). The results lost statistical

significance after FDR correction. Nonetheless, the average response of P3a still showed a tendency of a statistically significant decrease ($p = .070$).

Table 2. *Results of the two-way repeated measures MANOVA of the latencies of early somatosensory ERP components in response to deviant and standard stimuli in the time points T1 and T2. Df, degrees of freedom; p, statistical significance; * $p \le .05$; ** $p \le .01$; *** $p \le .001$; η_p^2, partial eta squared.*

	Time point			Stimulus type			Stimulus type × Time point		
	Main effect			Main effect			Interaction		
	F (df, error df)	*p*	η_p^2	*F (df, error df)*	*p*	η_p^2	*F (df, error df)*	*p*	η_p^2
sP50	2.41 (1,34)	.130	.066	1.46 (1,34)	.235	.041	.00 (1,34)	.978	.000
sN80	2.11 (1,34)	.156	.058	4.39 (1,34)	.044*	.114	.04 (1,34)	.840	.001

Table 3. *Results of the two-way repeated measures MANOVA of somatosensory and auditory ERP components in response to deviant and standard stimuli in the time points T1 and T2. Df, degrees of freedom; p, statistical significance; * $p \le .05$; ** $p \le .01$; *** $p \le .001$; η_p^2, partial eta squared.*

	Time point			Stimulus type			Stimulus type × Time point		
	Main effect			Main effect			Interaction		
	F (df, error df)	*p*	η_p^2	*F (df, error df)*	*p*	η_p^2	*F (df, error df)*	*p*	η_p^2
sP50	1.03 (1,34)	.318	.029	105.21 (1,34)	<.001***	.756	.31 (1,34)	.584	.009
sN80	.17 (1,34)	.681	.005	11.37 (1,34)	.002**	.251	.09 (1,34)	.771	.003
sMMR	2.61 (1,34)	.116	.071	12.14 (1,34)	.001***	.263	.49 (1,34)	.491	.014
sP3a	6.00 (1,34)	.020*	.150	30.90 (1,34)	< .001***	.476	1.87 (1,34)	.181	.052
aN1	.28 (1,31)	.598	.009	199.94 (1,31)	< .001***	.866	.00 (1,31)	.996	.000
aMMN	.89 (1,31)	.353	.028	68.70 (1,31)	< .001***	.689	.05 (1,31)	.825	.002
aP2	8.58 (1,31)	.006**	.217	16.10 (1,31)	< .001***	.342	.01 (1,31)	.927	.000

Auditory ERP components

All auditory ERP components (aMMN, aN1, and aP2), similarly to the somatosensory ones, showed a statistically significant difference between stimulus types, in a way that the response to deviant stimuli were larger in amplitude than to standard stimuli in all components (Table 5). The average response amplitude of deviant and standard stimuli of the auditory P2 component, as well as the deviant and standard responses alone, enlarged statistically significantly from the first measurement to the follow-up (Table 4). However, after FDR correction, only the average response remained significant (p = .042). The waveforms and scalp topographies of the ERP components that showed a statistically significant change over time (sP3a and aP2) are presented in Figure 1.

Table 4. *Mean amplitude values, standard deviations and results of the paired samples t-tests (two-tailed) comparing the response amplitudes between the time points T1 and T2 in the ERP components in response to standard and deviant stimuli, and their average. Only components that showed a statistically significant change over time are reported. SD, standard deviation; SEM, standard error of mean; CI, confidence interval; df, degrees of freedom; p, statistical significance; p(fdr), FDR-corrected p-value; * p ≤ .05; ** p ≤ .01; *** p ≤ .001; d, Cohen's d.*

	Mean amplitude (µV) ± SD		Difference between time points					
	T1	T2	Mean [SEM]	95 % CI	t (df)	p	p(fdr)	d
sP3a								
std	.14 ± .29	.04 ± .29	.10 [.07]	-04. to .24	1.48 (34)	.149	.347	-.34
dev	.64 ± .61	.40 ± .45	.24 [.10]	.04 to .44	2.39 (34)	.022*	.147	-.45
avg	.39 ± .38	.22 ± .27	.17 [.07]	.03 to .31	2.45 (34)	.020*	.070	-.52
aP2								
std	.16 ± .38	.36 ± .35	-.20 [.09]	-.38 to -.01	-2.16 (31)	.039*	.273	.55
dev	.55 ± .60	.73 ± .61	-.18 [.09]	-.36 to -.01	-2.12 (31)	.042*	.147	.30
avg	.35 ± .40	.55 ± .37	-.19 [.07]	-.32 to -.06	-2.93 (31)	.006**	.042*	.52
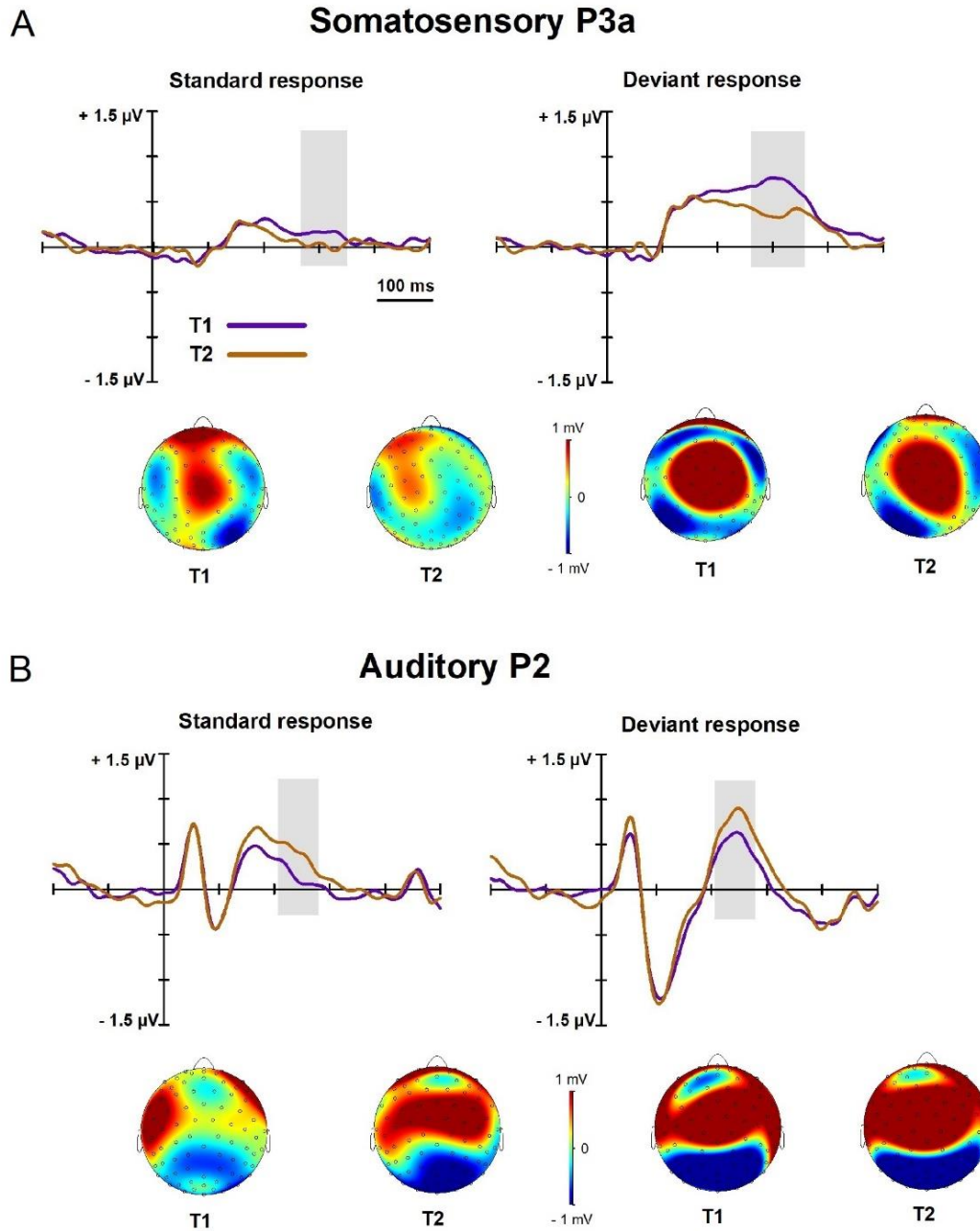
Figure 1. *Grand-averaged waveforms and scalp topographies of sP3a (**A**) and aP2 (**B**). Waveforms to standard and deviant stimuli represent averaged responses from response-specific electrode pools (see Appendices). The grey area indicates the latency ranges of 258-358 ms for sP3a, and 208-280 ms for aP2, from which the response amplitudes were averaged. Different time points are marked with purple for the first measurement (T1) and light brown for the follow-up measurement (T2). The scalp topographies indicate the average voltage distributions at the response-specific time points (see above), to standard and deviant stimuli.*

Mean sum scale variables of the four cognitive domains

We computed the original cognitive test scores to four new mean sum scale variables that were named, similarly to the principal components in the first measurement, *Executive function, Error susceptibility, Explicit memory* and *Working memory*. The reliabilities (Cronbach's α) are marked with an index of either 1 or 2, indicating time point of measurement. The calculated Cronbach's α for each variable are $\alpha_1 = .643$, $\alpha_2 = .699$ for *Executive function;* $\alpha_1 = .549$, $\alpha_2 = .490$ for *Error susceptibility;* $\alpha_1 = .840$, $\alpha_2 = .835$ for *Explicit memory*; $\alpha_1 = .748$, $\alpha_2 = .751$ for *Working memory*. The reliabilities of *Explicit memory* and *Working memory* are in both time points above $\alpha = .700$, which is considered acceptable. In other words, the sum scales measure the latent factors that they are supposed to measure. The reliability of *Executive function* is nearly acceptable in both time points. Only for *Error susceptibility* the reliabilities are quite low, which is because only two variables are included in the sum scale variable. Comparing T1 and T2, the reliabilities of the sum scales are very similar in each variable, meaning the internal consistency of the variables is not time dependent. Thus, further comparisons can be made using these variables as representatives for the cognitive domains.

We compared the differences between the variables in the two time points using paired samples *t*-test for *Working memory* and *Explicit memory*. Visual inspection of the normal distribution plots, as well as Shapiro-Wilk test of normality showed that *Executive functions* and *Error susceptibility* were not normally distributed, and therefore we used Wilcoxon signed-rank test for them. None of the cognitive domain variables showed statistically significant change over time.

Table 5. *Mean amplitude values, standard deviations and results of the paired samples t-tests (two-tailed, averaged over time) comparing the response amplitudes between deviant and standard stimulus types in the somatosensory and auditory ERP components. SD, standard deviation; SEM, standard error of mean; CI, confidence interval; df, degrees of freedom; p, statistical significance; p(fdr), FDR-corrected p-value; * p ≤ .05; ** p ≤ .01; *** p ≤ .001; d, Cohen's d.*

	Mean amplitude (µV) ± SD		Difference between stimulus types					
	Dev	Std	Mean [SEM]	95 % CI	t (df)	p	p(fdr)	d
sP50								
avg	1.40 ± .58	.91 ± .42	.48 [.05]	.39 to .58	10.26 (34)	<.001***	<.001***	.97
sN80								
avg	-.72 ± .47	-.48 ± .34	-.24 [.07]	-.38 to -.09	-3.37 (34)	.002**	.002**	-.59
sMMR								
avg	.41 ± .42	.16 ± .26	.25 [.07]	.10 to .39	3.49 (34)	.001***	.002**	.72
sP3a								
avg	.52 ± .44	.09 ± .20	.43 [.08]	.27 to .58	5.60 (34)	<.001***	<.001***	1.26
aN1								
avg	-1.19 ± .59	-.02 ± .58	-1.17 [.08]	-1.33 to -1.00	-14.35 (34)	<.001***	<.001***	-2.00
aMMN								
avg	-.54 ± .81	.63 ± .57	-1.17 [.13]	-1.43 to -.09	-9.24 (34)	<.001***	<.001***	-1.67
aP2								
avg	.68 ± .58	.30 ± .29	.38 [.10]	.18 to .57	3.97 (34)	<.001***	.001***	.83

Physical fitness measures and cognitive test scores

The change from the first measurement to the follow-up reached statistically significant *p*-values in distance of the six-minute walk test, BMI and total body fat (Table 6). Additionally, the increase in self-reported physical activity (*hrs/week*) had a tendency of statistical significance. The participants walked less meters in the six-minute walk test in the follow-up (mean ± SD, 570 ± 115) than in the first measurement (616 ± 80). BMI of the participants rose from the first measurement (26.4 ± 4.5) to the follow-up (26.6 ± 4.5), and total body fat rose as well (from 38.3 ± 6.7 to 39.3 ± 6.9) between the baseline and the follow-up measurements. After performing the FDR correction, the changes in the six-minute walk test distance and BMI remained statistically significant (*p* = .026 for six-minute walk test distance, and *p* = .001 for BMI, respectively).

The changes in the test scores of *Stroop 3*, *Handgrip left,* and *Handgrip right* were statistically significant. All three variables showed a decrease in performance (kilograms in *Handgrip right* and *Handgrip left* decreased, and the time used in *Stroop 3* increased, respectively). After performing FDR correction, only the change in *Handgrip left* remained significant (*p* = .043).

Table 6. *Physical fitness measures and individual cognitive test scores. Differences between the time points were tested using paired samples t-tests (two-tailed). SD, standard deviation; CI, confidence interval; p, statistical significance; p(fdr), FDR-corrected p-value; * p ≤ .05; ** p ≤ .01; *** p ≤ .001; d, Cohen's d.*

Test	T1 Mean ± SD	T2 Mean ± SD	Mean Difference (95% CI)	*p*	*p(fdr)*	*d*
Physical activity and fitness						
Six-minute walk test distance (*meters, more = better*)	617 ± 80	570 ± 115	-46 (-74 to -19)	.002**	.026*	-.46
Total body fat	38.3 ± 6.7	39.3 ± 6.9	1.0 (.1 to 1.9)	.032*	.164	.15
BMI	26.4 ± 4.5	26.6 ± 4.5	.2 (.1 to .3)	<.001***	.001***	.04
Self-reported physical activity (*hrs/week*)	3.0 ± 1.2	3.5 ± 1.3	.5 (.0 to 1.1)	.051	.189	.40
Cognitive test scores						
Tapping right (*clicks/10 s, more = better*)	42 ± 5	42 ± 5	0 (-1 to 2)	.977	.977	.00
Tapping left (*clicks/10 s, more = better*)	38 ± 5	37 ± 5	-1 (-1 to 0)	.326	.579	-.20
Handgrip right (*more = better*)	26 ± 4	24 ± 3	-2 (-4 to -1)	.010**	.065	-.57
Handgrip left (*more = better*)	24 ± 4	21 ± 3	-3 (-5 to -1)	.005**	.043*	-.85
TMT-A (*seconds, less = better*)	40 ± 15	40 ± 15	0 (-4 to 6)	.769	.833	.00
TMT-B (*seconds, less = better*)	97 ± 53	109 ± 75	12 (-15 to 39)	.370	.580	.18
Logical memory (*points, more = better*)	22 ± 6	22 ± 6	0 (-2 to 1)	.328	.569	.00
Logical memory delayed (*points, more = better*)	18 ± 7	16 ± 7	-2 (-3 to 1)	.251	.544	-.29
Stroop 1 – reading (*seconds, faster = better*)	55 ± 11	57 ± 11	2 (-3 to 5)	.501	.628	.18
Stroop 2 – colour labelling (*seconds, less = better*)	76 ± 19	79 ± 19	3 (-3 to 10)	.229	.541	.16
Stroop 3 – inhibition (*seconds, less = better*)	131 ± 37	145 ± 57	14 (0 to 28)	.047*	.189	.29
Stroop 2 errors (*points, less = better*)	1 ± 1	1 ± 1	.1 (-.4 to .7)	.626	.740	.00
Stroop 3 errors (*points, less = better*)	3 ± 4	4 ± 4	1.0 (-.4 to 2.4)	.146	.380	.25
Visual reproduction (*points, more = better*)	35 ± 4	35 ± 5	0 (-2 to 2)	.945	.977	.00
Visual reproduction delayed (*points, more = better*)	30 ± 9	31 ± 9	1 (-1 to 4)	.326	.569	.11
Digit span (*points, more = better*)	7 ± 2	7 ± 1	-.2 (-.7 to .3)	.379	.580	.00
Digit span backwards (*points, more = better*)	6 ± 2	6 ± 2	.2 (-.4 to .8)	.471	.628	.00
Digit-letter (*points, more = better*)	10 ± 3	9 ± 3	-.8 (-1.6 to .0)	.059	.192	-.33
AVLT 1 (*points, more = better*)	6 ± 2	7 ± 2	.4 (-.9 to 1.7)	.507	.628	.50
AVLT 5 (*points, more = better*)	12 ± 2	12 ± 3	-.4 (-1.4 to .7)	.490	.628	.00
AVLT immediate delayed (*points, more = better*)	11 ± 3	10 ± 3	-1.0 (-2.1 to .1)	.068	.196	-.33
AVLT delayed (*points, more = better*)	10 ± 3	10 ± 4	-.2 (-1.2 to .9)	.722	.816	.00

Associations between changes in the cognitive domains, individual cognitive test scores, ERPs, and physical fitness measures

As seen in Table 7, without controlling any variables the only statistically significant correlation including any of the cognitive mean sum scale variables, was the change in six-minute walk test distance and the change in explicit memory performance (r = .360, *p* = .047). After controlling for education, positive correlation between the changes in explicit memory and aP2 standard response amplitude was found, meaning that the enlarged amplitude of the response was associated with improved explicit memory performance (and vice versa) (r = .624, *p* = .040). In addition, controlling for education derived robust negative correlation with the change in working memory sum scale variable and the change of *Handgrip right* (r = -.704, *p* = .016), meaning that weakened handgrip was associated with improved working memory performance.

Apart from the cognitive domains, the change in six-minute walk test distance had an additional statistically significant negative correlation with the change in aP2 deviant stimulus response amplitude (r = -.392, *p* = .035), which means that worsened performance in the six-minute walk test was associated with an enlarged aP2 response. Statistically significant correlation was also found between the changes of *Handgrip right* and aP2 deviant response (r = .596, *p* = .025), meaning that strengthened handgrip was associated with an enlarged aP2 response. The change of performance time in *Stroop 3* test had a statistically significant negative correlation with the change of sP3a deviant response amplitude (r = -.358, *p* = .035), meaning that prolonged performance time in the test was associated with an attenuated response amplitude. After controlling for age, also the change between the amplitude of sP3a standard response and change of performance time in *Stroop 3* showed a significant negative correlation, meaning that attenuated sP3a response was associated with a prolonged performance time in *Stroop 3* (r = -.633, *p* = .037). After controlling for education, there were statistically significant partial correlations for the change of total body fat, as well. Significant positive correlation was found between the changes of total body fat and sP3a deviant response amplitude (r = .651, *p* = .030). This means, that increased total body fat was associated with an enlarged sP3a deviant response amplitude. Furthermore, the change in amplitude of aP2 standard response had significant negative correlation with the change of total body fat (r = -.773, *p* = .005), meaning that attenuated aP2 standard response is related to increased total body fat. All correlations with at least a tendency to statistical significance (*p* < .1) are listed in Table 7.

Table 7. *Reliabilities of the cognitive domain variables and correlations between changes of the cognitive and physical measures and the ERPs. Age or education (edu) refers to significant partial correlations after controlling for the variable in parentheses. α, Cronbach's α; r, Pearson's correlation coefficient; p, significance (two-tailed); All correlations showing at least a tendency to significance (p < .1) are listed. * p ≤ .05; ** p ≤ .01; *** p ≤ .001.*

Test	Reliability analysis		Variable	Correlations	
	α_{T1}	α_{T2}		*r*	*p*
Executive function	.643	.643	sP3a deviant (edu)	-.528	.095
Error susceptibility	.549	.490	-		
Explicit memory	.840	.835	Six-minute walk distance	.360	.047*
			aP2 standard (edu)	.624	.040*
Working memory	.748	.751	Handgrip right (edu)	-.704	.016*
			Handgrip left (edu)	-.536	.090
			aP2 standard (age)	-.597	.052
Six-minute walk distance			aP2 deviant	-.392	.035*
			sP3a standard	-.330	.070
Total body fat			sP3a deviant (edu)	.651	.030*
			sP2 standard (edu)	-.773	.005**
Handgrip right			sP3a standard	.475	.054
			aP2 deviant	.596	.025*
Stroop 3			sP3a deviant	-.358	.035*
			sP3a standard (age)	-.633	.037*

Associations of changes in the cognitive domains with the baseline measurement of the ERPs and physical fitness measures

When comparing the changes of the four mean sum scale variables of cognitive domains with the first measurement variables of physical fitness measures and ERPs, we found the most robust negative correlation between explicit memory change and the baseline of six-minute walk test distance (r = -.563, *p* < .001), meaning that the more meters a person originally walked, the more their explicit memory had improved (or the less it had worsened). aN1 deviant response amplitude at first

measurement (T1) had a statistically significant positive correlation with the changes of both explicit memory (r = .359, *p* = .043) and error susceptibility (r = .370, p = .037). This means, that more negative deviant response amplitude of auditory N1 component is related to more improved explicit memory and more increased error susceptibility in the next five years. In addition, aN1 standard amplitude at the first measurement had also a positive correlation with error susceptibility (r = .395, p = .025). Table 8 displays also correlations that showed a tendency to significance (p < .1).

Table 8. *Correlations between the changes in the cognitive domains and the baseline (T1) measurements of the ERPs and the physical fitness measures. r, Pearson's correlation coefficient; p, significance (two-tailed); All correlations showing at least a tendency to significance (p < .1) are listed. * p ≤ .05; ** p ≤ .01; *** p ≤ .001.*

Cognitive domain	Variable	*r*	*p*
Executive functions	sMMR deviant	.304	.075
Error susceptibility	aN1 standard	.395	.025*
	aN1 deviant	.370	.037*
Explicit memory	aN1 deviant	.359	.043*
	Six-minute walk test	-.563	< .001***
	BMI	.304	.075
Working memory	sP3a deviant	.308	.072

Loss analysis

We conducted a loss analysis for our data, because of the loss we had on participants. We were able to recruit 35 of the original 90 older women that participated in the first measurements. Therefore, it is relevant to inspect how our 35 participants might differ from the 55 participants who did not take part in this follow-up study, even though in this study we only used the data of the 35 participants who participated in both measurements. We performed independent samples *t*-tests to inspect differences between the follow-up participants (FU) and non-follow-up participants (NFU) in all cognitive test score variables, age, BMI, six-minute walk test distance, and ERPs. To inspect

differences in the categorical variables from the background questionnaire of *Income, Smoking, Exercise, Education, Alcohol* and *Sleep,* we performed a Pearson's Chi-Squared test.

In the physical fitness measures and cognitive test scores, the only statistically significant difference between FU participants and NFU participants was in the six-minute walk test distance ($p < .001$; Cohen's $d = .76$). The FU participants walked a longer distance than the NFU participants. Of the categorical variables, only *Exercise* showed statistically significant difference between the participant groups, in a way that FU participants reported exercising more hours in a week on average than the NFU participants ($\chi2= 11.441$; df = 4; $p = .022$). In the ERPs, the FU participants had nearly statistically significantly higher amplitudes in the standard response of the somatosensory P3a components when compared to the NFU participants ($p = .052$). Additionally, the difference in the deviant response of the aP2 component showed tendency of a statistical significance ($p = .063$), in a way that FU participants had higher amplitudes on average than did the NFU participants.

DISCUSSION

The aim of this study was to find out how older adults' somatosensory and auditory ERPs, physical fitness, and cognitive performance have changed over an approximately five-year period, and how the change in the ERPs and physical fitness affected the change in cognitive performance. Additionally, we investigated how the ERPs and physical fitness measured circa five years ago could predict change in cognitive performance.

Our main results were the findings of enlarged standard and deviant responses of the aP2 component, as well as the attenuated deviant response of the sP3a component. We found improved aerobic physical fitness being related to improved explicit memory, which supports our hypothesis (3) of the connection between physical fitness and cognitive performance. Enlarged aP2 response was associated with improved explicit memory (indicating neural compensation) and worsened aerobic fitness, and attenuated sP3a response associated with declined executive functions. This is in line with our hypothesis (4), that the change in the ERPs would associate with a change in cognitive performance. Additionally, better baseline of aerobic fitness predicted improved explicit memory (which supports hypothesis (5), that change in cognitive performance could be predicted with physical fitness), and higher baseline amplitude of aN1 predicted improved executive functions

and increased error susceptibility. Additionally, we hypothesized that the P50 amplitude would enlarge over time (1), and that the latencies of the ERP components would be delayed in the follow-up measurements (2). These two hypotheses did not get support, which may be due to a lack of statistical power, or the time gap between the measurements being too short to show differences in inhibitory control and sensory gating or in processing speed.

Changes in ERPs

As longitudinal research settings of ERPs are scarce, we needed to base our findings mainly on previous cross-sectional research. ERPs elicited by somatosensory finger stimulation and auditory sinusoidal sound stimulation showed statistically significant change over time in average response amplitudes of the somatosensory P3a and the auditory P2 components. More specifically, the deviant response amplitude showed a statistically significant attenuation in the somatosensory P3a component, whereas in the auditory P2 both standard and deviant response amplitudes showed a statistically significant enlargement over the measurements. Based on this finding, the analysis is extended to focus on different stimulus types of somatosensory P3a and auditory P2 response amplitudes, as we found statistically significant changes in them between the measurement times and stimulus types, even though results of two-way repeated measures MANOVA (Table 3) does not show statistically significant interaction between them. Furthermore, the nature of the sP3a and aP2 components is dependent on stimulus types, as P3a is suggested to reflect the evaluation of proper behavioral action to a deviant stimulus that the brain perceives as significant (Friedman et al., 2001; Polich, 2007), and the auditory P2 component reflects stimulus classification, as well as processing its physical and contextual features (Garcia-Larrea et al., 1992; Yingling & Nethercut, 1983). Thus, it is crucial for this study to analyze the stimulus types separately, to offer detailed and specific information about how the stimulus-related response amplitudes of the ERP components are sensitive for ageing as a biomarker.

In previous studies related to the aP2 component, there have been inconsistent results about age-related change in its amplitude. Regarding the standard response of the aP2 component, our finding is consistent with the active, cross-sectional oddball studies of Amenedo and Diaz (1998) and Ford and Pfefferbaum (1991), which both focused to compare age groups of older and young adults. In addition, Strömmer et al. (2017) had similar findings for standard response in a cross-sectional passive oddball procedure. Exceptionally, we found statistically significant age-related increase in the amplitude of the deviant response as well, unlike the previous studies. This divergent

finding can be partly explained by the rarely used passive oddball condition. Topographic maps for the aP2 responses show increasing activity from the baseline measurement to the follow-up in fronto-central location, as expected according to previous studies (Amenedo & Díaz, 1998; Anderer et al., 1996; Čeponienė et al., 2008). This supports the neural compensation hypothesis, according to which older adults' brains compensate for the declination of cognition by overactivation of especially the prefrontal cortex (Cabeza et al., 2004; Kropotov et al., 2016; Reuter-Lorenz & Park, 2010).

Figure 1 shows, that the attenuation in the somatosensory P3a component deviant response showed statistical significance. This is a trend towards less accurate acknowledging of possible significant features of a deviant stimulus, and a declining ability to modulate the attentional shift. This attenuation of the response was similar to the one found by Bolton and Staines (2012), who measured responses to attended somatosensory stimuli. Even though Strömmer et al. (2017) did not find statistically significant differences between older and young adults in the sP3a amplitude, it does not exclude the possibility, that by the time of the follow-up measurements participants had reached an age high enough to be more vulnerable to cognitive decline, and this is beginning to show on the level of pre-attentive sensory processing. Thus, it seems that ageing has a negative effect on the process of evaluation of deviant stimuli and the shift of attention in both attended and unattended situations. The voltage distribution of the deviant sP3a response was in the central topography in both measurement time points. This is coherent with the findings of Strömmer et al. (2017), who found that the group of older adults had prominent activation to deviant stimuli of the sP3a only in central electrode sites. In the same study, the group of young adults had prominent activation in fronto-central topography. Thus, judging by the longitudinal data of the scalp topographies, it seems that neural compensation is not visible in the sP3a response.

The short period of time between the two measurements was probably one major reason for why we did not observe differences in average amplitude in any other ERP components. Additionally, it could be the case that from the 90 original participants, the 35 that participated in the follow-up were less susceptible to the slowing down or impairment of pre-attentive sensory processing. The fact that we did not find statistically significant differences in P50 amplitude over time may indicate, that the sensory gating ability of the participants was still preserved. We did not observe differences in the latencies either, which in turn indicates that the automatic sensory processing speed may not change in such a short period of time.

Changes in physical fitness and cognition

As expected, the participants' physical fitness showed a slight degeneration between the measurements. Their total body fat rose and they walked less meters in the six-minute walk test than in the first measurement. This may be explained simply by the time difference between the measurements. Additionally, as shown by the loss analysis of the present study, the follow-up participants had a higher baseline of physical fitness than the non-follow-up participants, measured with the six-minute walk test. Thus, during the first measurement, the follow-up participants may not have yet reached an age high enough to show signs of reduced physical fitness, but during the follow-up they have passed that certain point. In addition to body fat and the walk test, their body mass index showed a slight increase. The nearly significant increase in self-reported physical activity was an interesting and relieving finding and may be due to the participants' motivation to exercise more after the first measurement. Nonetheless, the decreased distance in the six-minute walk test speaks for a decline in physical fitness. This indicates clearly that even a five-year period at such an advanced age can lead to physical reduction.

The change between the measurements was not significant in any of the mean sum scale variables of the cognitive domains. This is due to the variables being computed of individual standardized variables, which in turn means that the means of the variables are all close to zero. While changes in cognitive performance were not significant in the mean sum scale variables, it does not exclude the possibility to compare the correlations with the changes in the ERPs and the physical fitness measures, since all cognitive domain variables showed some change regardless whether or not visible in the average level. Nonetheless, it seems that the changes in cognition are subtle in a five-year time period, as only performance in *Stroop 3* and *Handgrip* -tests showed statistically significant change between the measurements (Table 6).

Associations between changes in cognitive performance, ERPs and physical fitness

In this study, we were especially interested to see what kind of association the change in cognitive performance over time had with the changes in physical fitness and the ERPs. Without any controlled variables, there was only one statistically significant correlation to be found between one of the mean sum scale variables of cognitive performance and any of the physical fitness or ERP change variables. The change in explicit memory performance and the change in six-minute walk test distance had a positive correlation, meaning that as explicit memory performance improved, the walk test

32

performance improved as well (and vice versa). This confirms the conclusion of the meta-analysis by Colcombe and Kramer (2003), who stated that any physical exercise improves cognitive performance in older adults. Thus, this finding further indicates the importance of physical exercise in preserving cognitive abilities in old age. As importantly, it shows that even in a five-year period the lack of physical exercise can start to deteriorate explicit memory in older adults.

We found a nearly significant correlation ($p = .052$) between changes in working memory and auditory P2 standard response amplitude, after controlling for age. This association indicates that worsened performance in working memory tasks is related to increased amplitude of the aP2 component, and vice versa. Lister et al. (2015) found, that older adults with probability to mild cognitive impairment (MCI) had higher amplitude of the P2 response. Even though our focus is on normal ageing, the results of Lister et al. (2015) supports our finding about how the enlarged amplitude of the P2 response may indicate also ageing-related decrease in some cognitive abilities, such as in working memory. In addition, McEvoy, Pellouchoud, Smith and Gevins (2000) found advanced age being associated with increased amplitude in frontal P2 component in spatial working memory task. Higher P2 amplitude has also been found to associate with a more challenging working memory task compared to a less straining task (Wolach & Pratt, 2001).

The positive correlation between the changes of the aP2 standard response and explicit memory suggests that improved performance in explicit memory tasks is related to an enlarged aP2 standard response. Even though this does not speak for higher response amplitude being related to cognitive decline (Lister et al., 2005), it might indicate neural compensation in explicit memory - related tasks. The topographic maps support this assumption of neural compensation as well, as fronto-central activity of the aP2 component was found to increase from the baseline measurement to the follow-up (see Figure 1). According to the findings of associations between cognitive performance and the auditory P2 component, it is possible that working memory is a more sensitive ageing-related cognitive ability, but explicit memory might be preserved better by means of neural compensation.

The negative correlation we found between the change in the six-minute walk test distance and the change in the deviant response amplitude of the auditory P2 component states, that the participants whose result worsened in the six-minute walk test, had an increased response amplitude to the P2 deviant stimulus. Previous studies have found mainly associations between physical activity and the latency of P2 component (Gajewski & Falkenstein, 2015; Polich & Lardon, 1997; Özkaya et al., 2005). Therefore, this study extends the view that also amplitude of the P2 component may be related to physical fitness, possibly through a shared association with cognitive functions. In this study, the negative correlation between total body fat and the aP2 standard response

was found as well. This means, that participants with less body fat after the baseline measurement had an increased aP2 response amplitude to standard stimuli, which supports also the association between auditory P2 component and physical fitness.

Ageing-related cognitive decline is observable also in the correlations between the changes of cognitive measurements and the sP3a deviant response change from the baseline measurement to the follow-up. Prolonged *Stroop 3* performance time was associated with an attenuated sP3a deviant response amplitude. The finding indicates that a decrease in the sP3a amplitude is associated with weakened executive functions, and more specifically selective attention. In addition, after controlling for education the sP3a deviant response amplitude showed tendency to a significant negative correlation with the sum scale variable of executive function. The findings are coherent with the results of Strömmer et al. (2017), as they found that higher response amplitude of the sP3a associated with better performance in executive functions in the cross-sectional study. In addition, the study of Fjell and Walhovd (2003) support our results, as they found tendency to a significant negative correlation between auditory P3a peak amplitude and *Stroop 3* performance time. Our findings support the hypothesis that the P3a amplitude is associated with age and attentional shift.

Some research suggests that the latency of P3 were more susceptible to ageing than the amplitude (Fjell & Walhovd, 2003; Ito et al., 1996; Walhovd & Fjell, 2003). Unfortunately, the present study was unable to address the latencies of the later ERP components, because the components were extracted from previously defined latency ranges (see Appendices). This procedure was performed in order to follow the same methodological approach as Strömmer et al. (2017), and because the nature of the later ERP components is more round-shaped, with no clear observable peak (see Figure 1).

The predictability of cognitive development with ERPs and physical fitness

Another significant point of interest was to see how the ERPs and physical fitness measured circa five years ago were associated with the change in cognitive performance. Correlational analysis showed us that the more meters the participants walked in the first measurements, the more their explicit memory had changed for the better. Respectively, the fewer meters they walked in the first measurements, the more their explicit memory had changed for the worse (or remained unchanged). As previous longitudinal experimental settings about the subject are scarce, this study is a unique indication of how physical aerobic fitness in advanced age can predict either preservation or decline of explicit memory.

Another statistically significant effect was seen between aN1 deviant response amplitude (T1) and the change in explicit memory, as well as both standard and deviant response amplitudes (T1) and the change in error susceptibility. The correlations were all positive, indicating that the larger (more negative) the amplitude, the more explicit memory had improved in the next five years, and the more error susceptibility had increased. The predicting effect of higher N1 amplitude on improving explicit memory in older age is rather unique, as studies showing such effects are scarce. One possible explanation is that higher N1 predicted improvement of explicit memory because of compensatory processes. The results of Strömmer et al. (2017), as well as Talsma et al. (2006) showing a larger amplitude in N1 elicited in older adults compared to young adults, and yet not showing a decrease in cognitive performance may be a sign of the same compensatory phenomenon. The predicting effect of higher N1 amplitude on increased error susceptibility may be an indication of worsened inhibition and a lack of attentional focus.

Limitations and advantages

Our study had some limitations. It would have been optimal to get as many participants from the original measurement to participate in this follow-up as possible, but we were able to recruit about one third of them. Even though the 35 participants that we were able to recruit is not little, it creates a situation where the participants may differ statistically to the ones that chose not to participate. In this case, they seemed to differ in at least physical performance, as the participants in the follow-up measurement had better aerobic fitness and they reported exercising more than the people not participated in the follow-up measurement. There may be other differing factors as well that we simply could not take into account. Naturally, the results of this study apply to women only, since all our participants were females. One limitation with the experimental procedure is that we adjusted the somatosensory stimulus intensities individually, but the auditory stimulus intensities were the same for all participants. This procedure was identical with the one of Strömmer et al. (2017). The individual adjustment of the somatosensory stimuli was important to perform, as it would be nearly impossible to find an intensity that is strong enough for every participant and yet not painful to any of them. In addition, we had measured hearing thresholds for all participants and that way confirmed that they all could hear the auditory sinusoidal sound stimuli. As was the case for Strömmer et al. (2017), individual MRI data were not measured, so source localization of the event-related potentials was not possible.

A definite advantage of the present study was the longitudinal research setting. This enabled us to ensure that the same methodological procedures were performed in both measurements. Additionally, changes observed in the group are likely to occur because of differences in the performance of the participants, and not because of differences between a confounding variable related to the differing groups, if this was a cross-sectional study. One disadvantage of this setting is that as the participants were familiar with the experimental procedure, they may have been less anxious during the follow-up measurements, and remembered parts of the neuropsychological test battery, both of which are reasons that may have contributed to them performing better. This is related to such a short follow-up period as four to five years. Probably because of this, any of the changes of cognitive test scores used in the cognitive sum scales, except used time in *Stroop 3*, were not able to reach significance between the baseline measurement and the follow-up. This leads to more uncertainty for using and interpreting the cognitive sum scale variables. In addition, with such a large number of variables, some of the correlations may have falsely reached statistical significance, as we did not use *p*-value correction because of probable loss of statistically significant correlations. This, in turn may have led to a narrowed view of ageing-related effects on cognition and physical fitness.

Conclusions and future research propositions

Based on findings of this study, we suggest future research to focus more on the connection between a decline in physical fitness and a decline in cognitive performance. The results of the present study indicate a connection between explicit memory and aerobic physical fitness, and this should be addressed in future longitudinal research in more detail and larger sample sizes. We found that higher baseline in physical fitness can even predict better preservation of explicit memory. This is information of great importance to all older adults who wish to slow down the inevitable cognitive decline. This cognition-preserving effect of aerobic fitness should be researched further, to acquire a better understanding of the complex processes behind it. Another proposition for future research is the association between changes in pre-attentive sensory information-processing and cognitive decline, as the present study was somewhat unable to comprehensively enlighten the relationship. The connection found in the present study between the amplified aP2 response and decreased six-minute walk test distance suggests that there may be a neural compensatory process occurring in relation to worsened aerobic fitness. As those changes are both individually implying of declined cognition, further research of this association is required. Finally, in order to get a sophisticated

understanding of the nature of cognitive ageing, studying older men in addition to women would provide valuable information.

To conclude, we found that certain changes over the four to five-year period in physical fitness and pre-attentive sensory processing in older women affected the way their cognitive performance had changed. Especially, we were able to further corroborate the connection between aerobic fitness and explicit memory, as well as suggest neural compensation to be linked with certain auditory event-related potentials. A remarkable finding was the preserving ability of aerobic fitness on explicit memory. Among older adults, it is essential therefore to adopt the habit of being physically active, in order to stay cognitively in shape and promote one's physical and psychological health, as well. The findings of the present study, along with any that follow it, must be brought to the attention of the people it touches, or otherwise we may be facing a societal issue that will concern everybody. The increasing ageing population brings a number of difficulties with it but educating older adults (of the present and the future) about the connections between physical fitness and cognitive abilities helps them maintain a better quality of life and helps the rest of us maintain a society that is not over-encumbered by neurocognitive disorders of older adults.

REFERENCES

Akatsuka, K., Wasaka, T., Nakata, H., Inui, K., Hoshiyama, M., & Kakigi, R. (2005). Mismatch responses related to temporal discrimination of somatosensory stimulation. *Clinical Neurophysiology*, *116*(8), 1930–1937.

Alain, C., & Woods, D. L. (1999). Age-related changes in processing auditory stimuli during visual attention: evidence for deficits in inhibitory control and sensory memory. *Psychology and aging*, *14*(3), 507.

Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: a meta-analytic review. *Neuropsychology review*, *16*(1), 17-42.

Amenedo, E., & Díaz, F. (1998). Aging-related changes in processing of non-target and target stimuli during an auditory oddball task. *Biological Psychology*, *48*(3), 235–267.

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.

Bastien, C. H., Fortier-Brochu, É., Rioux, I., LeBlanc, Ḿelanie, Daley, M., & Morin, C. M. (2003). Cognitive performance and sleep quality in the elderly suffering from chronic insomnia: Relationship between objective and subjective measures. *Journal of Psychosomatic Research*, *54*(1), 39–49.

Barrett, G., Neshige, R., & Shibasaki, H. (1987). Human auditory and somatosensory event-related potentials: effects of response condition and age. *Electroencephalography and Clinical Neurophysiology*, *66*(4), 409-419.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289-300.

Bolton, D. A. E., & Staines, W. R. (2011). Transient inhibition of the dorsolateral prefrontal cortex disrupts attention-based modulation of tactile stimuli at early stages of somatosensory processing. *Neuropsychologia*, *49*(7), 1928–1937.

Bolton, D. A. E., & Staines, W. R. (2012). Age-related loss in attention-based modulation of tactile stimuli at early stages of somatosensory processing. *Neuropsychologia*, *50*(7), 1502–1513.

Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the Trail Making Test. *Nature protocols*, *1*(5), 2277.

Braff, D. L., & Geyer, M. A. (1990). Sensorimotor gating and schizophrenia: human and animal model studies. *Archives of general psychiatry*, *47*(2), 181-188.

Brigman, S., & Cherry, K. E. (2002). Age and skilled performance: Contributions of working memory and processing speed. *Brain and Cognition*, *50*(2), 242–256.

Brown, W. S., Marsh, J. T., & LaRue, A. (1983). Exponential electrophysiological aging: P3 latency. *Electroencephalography and clinical Neurophysiology*, *55*(3), 277-285.

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*(4), 364–375.

Carstensen, L. L., Fung, H. H., & Charles, S. T. (2003). Socioemotional selectivity theory and the regulation of emotion in the second half of life. *Motivation and emotion*, *27*(2), 103-123.

Carstensen, L. L., Turan, B., Scheibe, S., Ram, N., Ersner-Hershfield, H., Samanez-Larkin, G. R., ... & Nesselroade, J. R. (2011). Emotional experience improves with age: evidence based on over 10 years of experience sampling. *Psychology and aging*, *26*(1), 21.

Čeponiene, R., Westerfield, M., Torki, M., & Townsend, J. (2008). Modality-specificity of sensory aging in vision and audition: Evidence from event-related potentials. *Brain Research*, *1215*, 53–68.

Chao, L. L., & Knight, R. T. (1997). Prefrontal deficits in attention and inhibitory control with aging. *Cerebral cortex (New York, NY: 1991)*, *7*(1), 63-69.

Cheng, C. H., Hsu, W. Y., & Lin, Y. Y. (2013). Effects of physiological aging on mismatch negativity: A meta-analysis. *International Journal of Psychophysiology*, *90*(2), 165–171.

Christensen, H., Mackinnon, A. J., Korten, A. E., Jorm, A. F., Henderson, A. S., Jacomb, P., & Rodgers, B. (1999). An analysis of diversity in the cognitive performance of elderly community dwellers: individual differences in change scores as a function of age. *Psychology and aging*, *14*(3), 365.

Colcombe, S., & Kramer, A. F. (2003). Fitness Effects on the Cognitive Function of Older Adults: A Meta-Analytic Study. *Psychological Science*, *14*, 125.

Crapo, R. O. et al. ATS statement: Guidelines for the six-minute walk test. (2002). *Am. J. Respir. Crit. Care Med. 166*, 111–117.

Crowe, S. F. (2000). Does the letter number sequencing task measure anything more than digit span?. *Assessment*, *7*(2), 113–117.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology*, *115*(4), 732–744.

Czigler, I., Csibra, G., & Csontos, A. (1992). Age and inter-stimulus interval effects on event-related potentials to frequent and infrequent auditory stimuli. *Biological Psychology*, *33*(2–3), 195–206.

Davidson, P. S. R., & Winocur, G. (2010). Aging and Cognition. *Encyclopedia of Behavioral Neuroscience,* 20–26.

Dzierzewski, J. M., Dautovich, N., & Ravyts, S. (2018). Sleep and Cognition in Older Adults. *Sleep Medicine Clinics*, *13*(1), 93–106.

Elwood, R. W. (1991). The Wechsler Memory Scale – Revised: psychometric characteristics and clinical application. *Neuropsychology review*, *2*(2), 179–201.

European Commission (2018). *The 2018 Ageing Report: Economic and Budgetary Projections for the EU Member States (2016-2070).* Luxembourg: Publications Office of the European Union.

Fabiani, M., Low, K. A., Wee, E., Sable, J. J., & Gratton, G. (2006). Reduced suppression or labile memory? Mechanisms of inefficient filtering of irrelevant information in older adults. *Journal of cognitive neuroscience*, *18*(4), 637–650.

Feigin, V. L., Abajobir, A. A., Abate, K. H., Abd-Allah, F., Abdulle, A. M., Abera, S. F., ... & Aichour, M. T. E. (2017). Global, regional, and national burden of neurological disorders during 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet Neurology, 16*(11), 877–897.

Fjell, A. M., Westlye, L. T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., ... & Dale, A. M. (2009). High consistency of regional cortical thinning in aging across multiple samples. *Cerebral cortex*, *19*(9), 2001–2012.

Fjell, A. M., & Walhovd, K. B. (2003). P3a and neuropsychological 'Frontal'tests in aging. *Aging, Neuropsychology, and Cognition*, *10*(3), 169–181.

Foster, S. M., Kisley, M. A., Davis, H. P., Diede, N. T., Campbell, A. M., & Davalos, D. B. (2013). Cognitive function predicts neural activity associated with pre-attentive temporal processing. *Neuropsychologia*, *51*(2), 211–219.

Ford, J. M., & Pfefferbaum, A. (1991). Event-related potentials and eyeblink responses in automatic and controlled processing: effects of age. *Electroencephalography and clinical neurophysiology*, *78*(5), 361-377.

Freedman, R., Adler, L. E., Gerhardt, G. A., Waldo, M., Baker, N., Rose, G. M., ... & Franks, R. (1987). Neurobiological studies of sensory gating in schizophrenia. *Schizophrenia bulletin*, *13*(4), 669-678.

Friedman, D. (2012). *The Components of Aging*. *The Oxford Handbook of Event-Related Potential Components*.

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience & Biobehavioral Reviews*, *25*(4), 355-373.

Gajewski, P. D., & Falkenstein, M. (2015). Long-term habitual physical activity is associated with lower distractibility in a Stroop interference task in aging: Behavioral and ERP evidence. *Brain and cognition*, *98*, 87-101.

García-Larrea, L., Lukaszewicz, A. C., & Mauguiére, F. (1992). Revisiting the oddball paradigm. Non-target vs neutral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia*, *30*(8), 723-741.

Gmehlin, D., Kreisel, S. H., Bachmann, S., Weisbrod, M., & Thomas, C. (2011). Age effects on preattentive and early attentive auditory processing of redundant stimuli: Is sensory gating affected by physiological aging? *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, *66 A*(10), 1043–1053.

Grady, C. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience*, *13*(7), 491–505.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and clinical neurophysiology*, *55*(4), 468-484.

Harada, C. N., Natelson Love, M. C., & Triebel, K. L. (2013). Normal cognitive aging. *Clinics in Geriatric Medicine*, *29*(4), 737–752.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view BT - The psychology of learning and motivation. *The Psychology of Learning and Motivation*, *22*, 193–225.

Iragui, V. J., Kutas, M., Mitchiner, M. R., & Hillyard, S. A. (1993). Effects of aging on event-related brain potentials and reaction times in an auditory oddball task. *Psychophysiology*, *30*(1), 10-22.

Ito, J., Takamatsu, N., & Kimura, J. (1996). Effects of aging on P3 components of somatosensory event-related potentials elicited by different stimulus conditions. Conventional averages and single-trial analysis. *International Journal of Neuroscience, 85(3–4),* 165–172.

Jaušovec, N., & Jaušovec, K. (2000). Correlations between ERP parameters and intelligence: a reconsideration. *Biological Psychology*, *55*(2), 137-154.

Jerger, K., Biggins, C., & Fein, G. (1992). P50 suppression is not affected by attentional manipulations. *Biological psychiatry*, *31*(4), 365-377.

Ji, L., Pearlson, G. D., Zhang, X., Steffens, D. C., Ji, X., Guo, H., & Wang, L. (2018). Physical exercise increases involvement of motor networks as a compensatory mechanism during a cognitively challenging task. *International Journal of Geriatric Psychiatry*, *33*(8), 1153–1159.

Kaup, A. R., Mirzakhanian, H., Jeste, D. V., & Eyler, L. T. (2011). A Review of the Brain Structure Correlates of Successful Cognitive Aging. *Journal of Neuropsychiatry*, *23*(1), 6–15.

Kekoni, J., Hämäläinen, H., Saarinen, M., Gröhn, J., Reinikainen, K., Lehtokoski, A., & Näätänen, R. (1997). Rate effect and mismatch responses in the somatosensory system: ERP-recordings in humans. *Biological Psychology*, *46*(2), 125–142.

Kisley, M. A., Davalos, D. B., Engleman, L. L., Guinther, P. M., & Davis, H. P. (2005). Age-related change in neural processing of time-dependent stimulus features. *Cognitive Brain Research*, *25*(3), 913–925.

Kropotov, J., Ponomarev, V., Tereshchenko, E. P., Müller, A., & Jäncke, L. (2016). Effect of aging on ERP components of cognitive control. *Frontiers in Aging Neuroscience*, *8*(APR), 1–15.

Küper, K., Gajewski, P. D., Frieg, C., & Falkenstein, M. (2017). A randomized controlled erp study on the effects of multi-domain cognitive training and task difficulty on task switching performance in older adults. *Frontiers in human neuroscience*, *11*, 184.

Laurin, D., Verreault, R., Lindsay, J., MacPherson, K., & Rockwood, K. (2001). Physical activity and risk of cognitive impairment and dementia in elderly persons. *Archives of neurology*, *58*(3), 498-504.

Laver, G. D. (2009). Adult aging effects on semantic and episodic priming in word recognition. *Psychology and Aging*, *24*(1), 28.

Lijffijt, M., Lane, S. D., Meier, S. L., Boutros, N. N., Burroughs, S., Steinberg, J. L., ... & Swann, A. C. (2009). P50, N100, and P200 sensory gating: relationships with behavioral inhibition, attention, and working memory. *Psychophysiology*, *46*(5), 1059-1068.

Lister, J. J., Harrison Bush, A. L., Andel, R., Matthews, C., Morgan, D., & Edwards, J. D. (2016). Cortical auditory evoked responses of older adults with and without probable mild cognitive impairment. *Clinical Neurophysiology*, *127*(2), 1279–1287.

Lubitz, A. F., Niedeggen, M., & Feser, M. (2017). Aging and working memory performance: Electrophysiological correlates of high and low performing elderly. *Neuropsychologia*, *106*(April), 42–51.

Martins, R., Joanette, Y., & Monchi, O. (2015). The implications of age-related neurofunctional compensatory mechanisms in executive function and language processing including the new Temporal Hypothesis for Compensation. *Frontiers in Human Neuroscience*, *9*(April), 1–17.

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

Middleton, L. E., Barnes, D. E., Lui, L. Y., & Yaffe, K. (2010). Physical activity over the life course and its association with cognitive performance and impairment in old age. *Journal of the American Geriatrics Society*, *58*(7), 1322–1326.

Morris, A.F., Vaccaro, P., & Harris, R. (1983). Life quality changes of older adults during various physical-activity programs. *Medicine and Science in Sports and Exercise*, *15(2),* 117.

Müller, V., Brehmer, Y., von Örtzen, T., Li, S. C., & Lindenberger, U. (2008). Electrophysiological correlates of selective attention: A lifespan comparison. *BMC Neuroscience*, *9*, 1–21.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus - norepinephrine system. *Psychological bulletin*, *131*(4), 510.

Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta psychologica*, *42*(4), 313-329.

Näätänen, R., Astikainen, P., Ruusuvirta, T., & Huotilainen, M. (2010). Automatic auditory intelligence: An expression of the sensory-cognitive core of cognitive processes. *Brain Research Reviews*, *64*(1), 123–136.

OECD (2017). Life expectancy at birth. In *Health at a Glance 2017: OECD Indicators*, OECD Publishing, Paris.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, *17*(2), 299–320.

Partanen, J., Lang, H., Valkonen-Korhonen, M. & Cheour, M. (2006). Tapahtumapotentiaalit (ERP). In Partanen, J., Björn, F., Hasan, J., Jäntti, V., Salmi, T. & Tolonen, U. (Eds.) In *Kliininen neurofysiologia (pp. 330–347).* Helsinki: Duodecim.

Picton, T. W., Stuss, D. T., Champagne, S. C., & Nelson, R. F. (1984). The effects of age on human event-related potentials. *Psychophysiology*, *21*(3), 312-326.

Polich, J. (1996). Meta-analysis of P300 normative aging studies. *Psychophysiology*.

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

Polich, J., & Lardon, M. T. (1997). P300 and long-term physical exercise. *Electroencephalography and clinical neurophysiology*, *103*(4), 493-498.

Ponce-Bravo, H., Ponce, C., Feriche, B., & Padial, P. (2015). Influence of two different exercise programs on physical fitness and cognitive performance in active older adults: Functional resistance-band exercises vs. recreational oriented exercises. *Journal of Sports Science and Medicine*, *14*(4), 716–722.

Pontifex, M. B., Hillman, C. H., & Polich, J. (2009). Age, physical fitness, and attention: P3a and P3b. *Psychophysiology*, *46*(2), 379–387.

Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., & Ferri, C. P. (2013). The global prevalence of dementia: A systematic review and meta-analysis. *Alzheimer's and Dementia*, 9(1), 63–75.

Ramsay, M. C., & Reynolds, C. R. (1995). Separate digits tests: A brief history, a literature review, and a reexamination of the factor structure of the Test of Memory and Learning (TOMAL). *Neuropsychology review*, *5*(3), 151-171.

Raz, N., Gunning, F. M., Head, D., Dupuis, J. H., McQuain, J., Briggs, S. D., ... & Acker, J. D. (1997). Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. *Cerebral cortex (New York, NY: 1991)*, *7*(3), 268-282.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., … Acker, J. D. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex*, *15*(11), 1676–1689.

Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., & Davatzikos, C. (2003). Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. *Journal of Neuroscience*, *23*(8), 3295–301.

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*, *65*(4), 405-415.

Rey, A. (1964). L'examen clinique en psychologie. *Paris: Presses Universitaries de France.*

Rijk, J. M., Roos, P. R. K. M., Deckx, L., Van den Akker, M., & Buntinx, F. (2016). Prognostic value of handgrip strength in people aged 60 years and older: A systematic review and meta-analysis. *Geriatrics and Gerontology International*, *16*(1), 5–20.

Ruff, R. M., & Parker, S. B. (1993). Gender-and age-specific changes in motor speed and eye-hand coordination in adults: normative values for the Finger Tapping and Grooved Pegboard Tests. *Perceptual and motor skills*, *76*(3_suppl), 1219-1230.

Ruzzoli, M., Pirulli, C., Brignani, D., Maioli, C., & Miniussi, C. (2012). Sensory memory during physiological aging indexed by mismatch negativity (MMN). *Neurobiology of Aging*, *33*(3), 625.e21-625.e30.

Rönnlund, M., Nyberg, L., Bäckman, L., & Nilsson, L. G. (2005). Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population-based study. *Psychology and Aging*, *20*(1), 3–18.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological review*, *103*(3), 403.

Sanei, S., & Chambers, J. A. (2007). EEG signal processing. John Wiley & Sons. Chichester.

Schröger, E. (1997). On the detection of auditory deviations: A pre-attentive activation model. *Psychophysiology*, *34*(3), 245-257.

Shen, G., Smyk, N. J., Meltzoff, A. N., & Marshall, P. J. (2018). Neuropsychology of human body parts: exploring categorical boundaries of tactile perception using somatosensory mismatch responses. *Journal of cognitive neuroscience*, *30*(12), 1858–1869.

Sofi, F., Valecchi, D., Bacci, D., Abbate, R., Gensini, G. F., Casini, A., & Macchi, C. (2011). Physical activity and risk of cognitive decline: A meta-analysis of prospective studies. *Journal of Internal Medicine*, *269*(1), 107–117.

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and clinical neurophysiology*, *38*(4), 387–401.

Stothart, G., & Kazanina, N. (2016). Auditory perception in the aging brain: the role of inhibition and facilitation in early processing. *Neurobiology of Aging*, *47*, 23–34.

Strömmer, J. M., Tarkka, I. M., & Astikainen, P. (2014). Somatosensory mismatch response in young and elderly adults. *Frontiers in Aging Neuroscience*, *6*(October), 1–9.

Strömmer, J. M., Põldver, N., Waselius, T., Kirjavainen, V., Järveläinen, S., Björksten, S., … Astikainen, P. (2017). Automatic auditory and somatosensory brain responses in relation to cognitive abilities and physical fitness in older adults. *Scientific Reports*, *7*(1), 1–12.

Suomen virallinen tilasto (SVT) (2018). Väestöennuste [e-publication]. ISSN=1798-5137. Helsinki: Tilastokeskus [referred: 29.10.2019]. Access method: https://www.tilastokeskus.fi/tup/suoluk/suoluk_vaesto.html?fbclid=IwAR3VmLcCpGLDJ4AE 0-gvWL05g4T_BgNfMamecdTUydKoQP7SjExtUGLTB4E#väestöennuste

Talsma, D., Kok, A., & Ridderinkhof, K. R. (2006). Selective attention to spatial and non-spatial visual stimuli is affected differentially by age: Effects on event-related brain potentials and performance data. *International Journal of Psychophysiology*, *62*(2), 249–261.

Tomé, D., Barbosa, F., Nowak, K., & Marques-Teixeira, J. (2015). The development of the N1 and N2 components in auditory oddball paradigms: a systematic review with narrative analysis and suggested normative values. *Journal of Neural Transmission*, *122*(3), 375–391.

Tremblay, K. L., Billings, C., & Rohila, N. (2004). Speech evoked cortical potentials: Effects of age and stimulus presentation rate. *Journal of the American Academy of Audiology, 15(3)*, 226–237.

United Nations, Department of Economic and Social Affairs, Population Division (2017). World Population Ageing 2017 (ST/ESA/SER.A/408).

Van Dinteren, R., Huster, R. J., Jongsma, M. L. A., Kessels, R. P. C., & Arns, M. (2018). Differences in Cortical Sources of the Event-Related P3 Potential Between Young and Old Participants Indicate Frontal Compensation. *Brain Topography*, *31*(1), 35–46.

Verhaeghen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology and Aging*, *18*(2), 332–339.

Walhovd, K. B., & Fjell, A. M. (2003). The relationship between P3 and neuropsychological function in an adult life span sample. *Biological Psychology*, *62*(1), 65-87.

Wan, L., Friedman, B. H., Boutros, N. N., & Crawford, H. J. (2008). P50 sensory gating and attentional performance. *International Journal of Psychophysiology*, *67*(2), 91–100.

Wilson, R. S., Beckett, L. A., Barnes, L. L., Schneider, J. A., Bach, J., Evans, D. A., & Bennett, D. A. (2002). Individual differences in rates of change in cognitive abilities of older persons. *Psychology and Aging*, *17*(2), 179–193.

Wolach, I., & Pratt, H. (2001). The mode of short-term memory encoding as indicated by event-related potentials in a memory scanning task with distractions. *Clinical neurophysiology*, *112*(1), 186-197.

Yamaguchi, S., & Knight, R. T. (1991). Anterior and posterior association cortex contributions to the somatosensory P300. *Journal of Neuroscience*, *11*(7), 2039-2054.

Yingling, C. D., & Nethercut, G. E. (1983). Evoked responses to frequency shifted tones: Tonotopic and contextual determinants. *International Journal of Neuroscience*, *22*(1–2), 107–118.

Young, J., Angevaren, M., Rusted, J., & Tabet, N. (2015). Aerobic exercise to improve cognitive function in older people without known cognitive impairment. *Cochrane Database of Systematic Reviews*, (4).

Zettel-Watson, L., Suen, M., Wehbe, L., Rutledge, D. N., & Cherry, B. J. (2017). Aging well: processing speed inhibition and working memory related to balance and aerobic endurance. *Geriatrics & gerontology international*, *17*(1), 108-115.

Özkaya, G. Y., Aydin, H., Toraman, F. N., Kizilay, F., Özdemir, Ö., & Cetinkaya, V. (2005). Effect of strength and endurance training on cognition in older people. *Journal of sports science & medicine*, *4*(3), 300.

APPENDICES
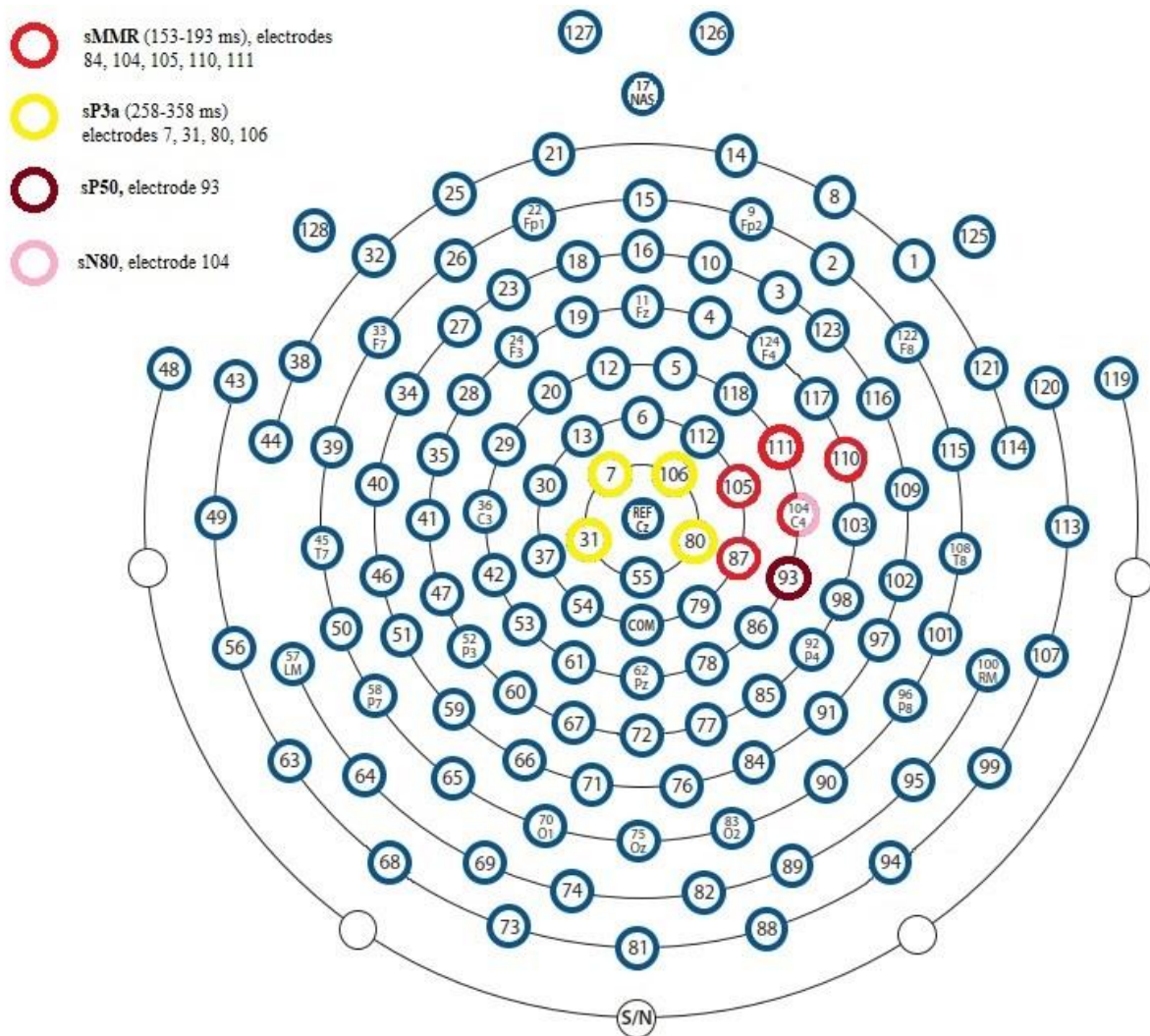
Appendix 1. *Electrode locations and latency ranges of the somatosensory ERP components sMMR (red), sP3a (yellow), sP50 (brown) and sN80 (pink). The latency range for sMMR was 153-193 ms, and for sP3a 258-358 ms, respectively.*
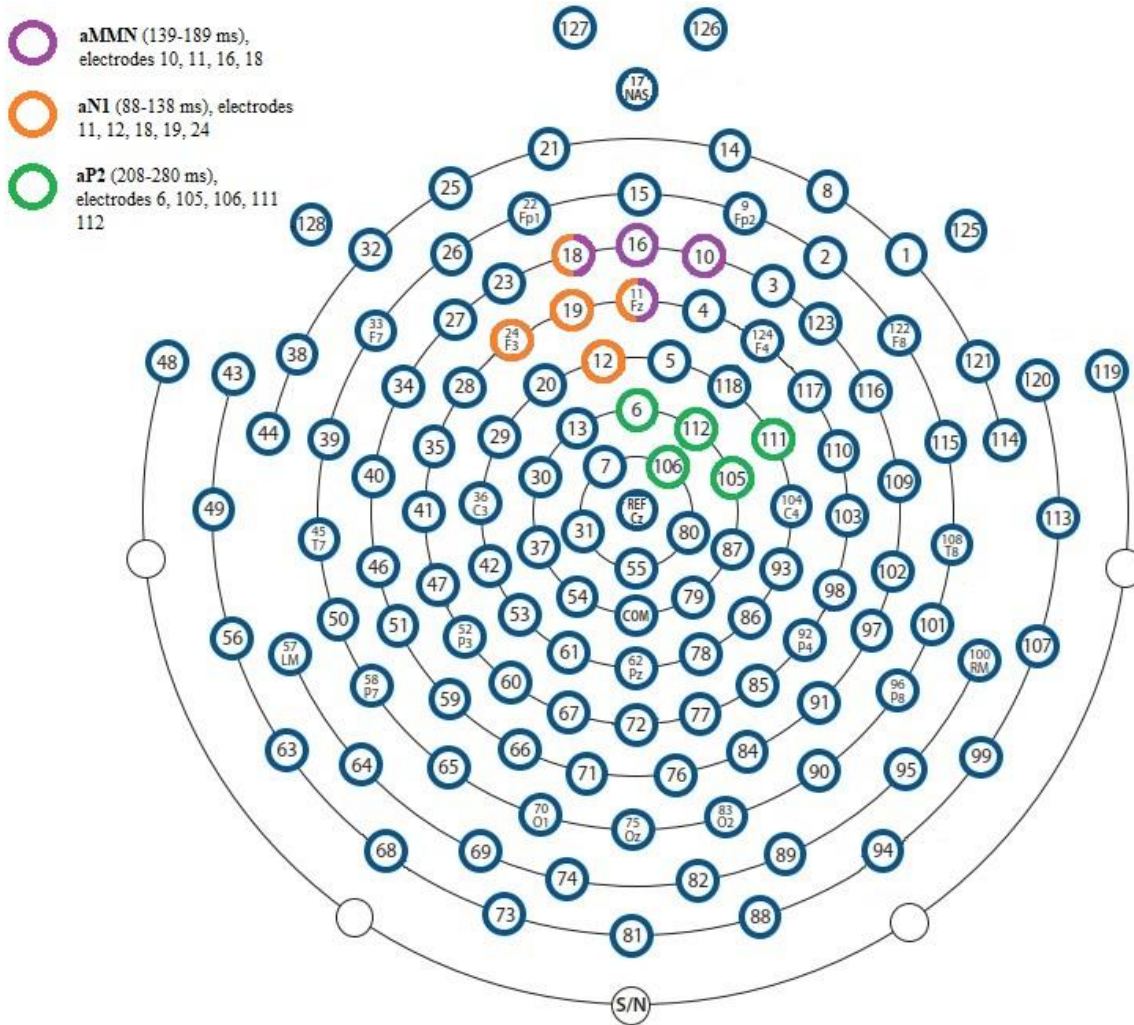
Appendix 2. *Electrode locations and latency ranges for the auditory ERP components aMNN (purple), aN1 (orange) and aP2 (green). The latency ranges for the components were 139-189 ms for aMMN, 88-138 ms for aN1, and 208-280 ms for aP2, respectively.*