Original Article

QJEP

Quarterly Journal of Experimental Psychology
1–20

DOI: 10.1177/1747021818791994
qjep.sagepub.com
SAGE

Working memory in older adults declines with age, but is modulated by sex and education

Christos Pliatsikas[1], João Veríssimo[2], Laura Babcock[3], Mariel Y Pullman[4], Dana A Glei[5], Maxine Weinstein[5], Noreen Goldman[6] and Michael T Ullman[7]

Abstract

Working memory (WM), which underlies the temporary storage and manipulation of information, is critical for multiple aspects of cognition and everyday life. Nevertheless, research examining WM specifically in older adults remains limited, despite the global rapid increase in human life expectancy. We examined WM in a large sample ($N=754$) of healthy older adults (aged 58-89) in a non-Western population (Chinese speakers) in Taiwan, on a digit *n*-back task. We tested not only the influence of age itself and of load (1-back vs. 2-back) but also the effects of both sex and education, which have been shown to modulate WM abilities. Mixed-effects regression revealed that, within older adulthood, age negatively impacted WM abilities (with linear, not nonlinear, effects), as did load (worse performance at 2-back). In contrast, education level was positively associated with WM. Moreover, both age and education interacted with sex. With increasing age, males showed a steeper WM decline than females; with increasing education, females showed greater WM gains than males. Together with other findings, the evidence suggests that age, sex, and education all impact WM in older adults, but interact in particular ways. The results have both basic research and translational implications and are consistent with particular benefits from increased education for women.

Keywords

Ageing; sex differences; education; working memory; *n* back

Received: 26 February 2018; revised: 25 May 2018; accepted: 28 June 2018

Introduction

Working memory (WM) is generally considered to be the domain of human cognition that underlies the temporary storage and manipulation of information (Baddeley, 1992, 2003a, 2012; Cowan, 1998, 1999, 2010). As such, this capacity appears to play an important role mediating between the processing of stored or incoming information and its use for specific cognitive goals, as diverse as orientation, reasoning, language processing, planning, and spatial processing (Cansino et al., 2013; D'Esposito, 2007). WM is generally conceptualised as involving various components that work together. An executive or attentional component is often assumed to focus on the relevant information, which is thought to be maintained in either temporary or long-term storage (Baddeley, 1992, 2003a, 2012; Cowan, 1998, 1999, 2010). The executive/attentional component also seems to underlie various functions such as focusing attention, switching between information and tasks, and interfacing with long-term memory (Baddeley, 2012; Cowan, 1999). The capacity of WM is quite limited, in that the amount of information that it can maintain is finite and relatively small, although the size of this

[1]School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK
[2]Potsdam Research Institute for Multilingualism, University of Potsdam, Potsdam, Germany
[3]Department of Neuroscience, Karolinska Institutet, Solna, Sweden
[4]Icahn School of Medicine at Mount Sinai, New York, NY, USA
[5]Center for Population and Health, Georgetown University, Washington, DC, USA
[6]Office of Population Research, Princeton University, Princeton, NJ, USA
[7]Department of Neuroscience, Georgetown University, Washington, DC, USA

Corresponding author:
Michael T Ullman, Department of Neuroscience, Georgetown University, Box 571464, Washington, DC 20057-1464, USA.
Email: michael@georgetown.edu

capacity ("span") varies as a function of various factors (Cowan, 2010; Miller, 1956).

Although WM has been the focus of a very large literature (e.g., Baddeley, 2003a, 2007; Conway et al., 2005; D'Esposito, 2007; D'Esposito & Postle, 2015), there has somewhat been less research on the effects of ageing on this domain. Yet, given the importance of WM in various aspects of cognition (e.g., language, maths; Baddeley, 2003b; Raghubar, Barnes, & Hecht, 2010) and everyday life (e.g., reading, typing, orienting in space, planning what to do and when to do it; G. Cohen & Conway, 2007; Kane, Brown, et al., 2007), and considering the rapidly ageing population globally (Phillips, 2002; Rechel et al., 2013), a thorough understanding of ageing and WM is warranted. In addition, because healthy ageing typically constitutes the baseline comparison for disorders that are associated with ageing as well as WM deficits, such as Alzheimer's disease, Parkinson's disease, and aphasia (Pfeiffer, Løkkegaard, Zoetmulder, Friberg, & Werdelin, 2014; Whitwell et al., 2015), elucidating WM in healthy ageing may have important translational impacts.

Of particular interest here is the fact that although quite a number of studies have examined how WM abilities may change between younger and older adulthood, less research has investigated WM trajectories *within* old age. Yet, the nature of potential WM changes during old age, including the rate of any changes and the factors and mechanisms involved, could be different from WM changes *between* younger and older adults; indeed, evidence suggests that some cognitive abilities show nonlinear declines over the adult lifespan (Nyberg, Lövdén, Riklund, Lindenberger, & Bäckman, 2012). Importantly, the average duration of "old age" (reasonably defined from about 60 till the inevitable demise; Scullin & Bliwise, 2015) and individuals' activity during old age are both increasing (Cassel, 2001). Thus, a thorough understanding of WM trajectories across older years seems valuable.

In the remainder of the Introduction, we, first, briefly review the literature on WM and ageing as examined in comparisons between younger and older adults. We then more comprehensively review the small number of studies focusing on the issue examined in this study, that is, research that probes whether ageing within older adults is associated with differential WM effects.

Brief review of effects of age, sex, and education across younger and older adults

Most studies examining ageing effects on WM have compared younger and older adults. These have generally found that ageing detrimentally affects various aspects of WM, for both verbal and non-verbal information (Grady & Craik, 2000; Orsini et al., 1986; Park et al., 2002; Reuter-Lorenz & Sylvester, 2005), including in tasks probing verbal span, visual object manipulation, updating and switching, and the temporary storage of information (Atkinson, Baddeley, & Allen, 2018; Bopp & Verhaeghen, 2005; Federico, Delogu, & Raffone, 2014; I. E. Nagel et al., 2011; Pertzov, Heider, Liang, & Husain, 2015; Peterson & Naveh-Benjamin, 2016; van Gerven, Meijer, Prickaerts, & Van der Veen, 2008); for a recent comprehensive review, see Bopp and Verhaeghen (2018). For example, Johnson, Logie, and Brockmole (2010), who tested a large sample of participants from early adulthood to old age, grouped in 5-year age cohorts, reported significant age-related declines in tasks tapping several aspects of working and short-term memory. Some evidence also suggests that verbal WM might be less severely affected by ageing across the adult lifespan than visuospatial WM (Hale et al., 2011). Of particular relevance here, Cansino and colleagues (2013) tested a large cohort of healthy participants aged between 21 and 80 on "verbal" (visual presentation of letters) and visuospatial versions of the *n*-back task, each with both 1-back and 2-back subtasks to probe different WM loads (which seems to tap WM span; see below for details about the *n*-back task, which is also employed in this study). The authors reported that ageing negatively impacted WM performance, across both versions of the task and in both subtasks (note that here and below we use the term "subtask" only to refer to subtasks of a given WM task with different loads; e.g., 1-back vs. 2-back subtasks.) In addition, they observed both a main effect of subtask (worse performance at the 2-back than 1-back subtask) and an interaction between age and subtask, with declines in performance observed during older ages for both subtasks, but during younger ages mainly for the 2-back subtask (across both the verbal and the visuospatial domains).

The mechanisms underlying WM changes between younger and older adults are not yet clear, and various explanatory accounts have been proposed. Cognitive accounts have attributed the observed age-related WM changes to a general slowing of cognitive processing (Salthouse, 1996), to declines in attentional resources (Craik & Byrd, 1982), to reduced efficiency of inhibitory processes (Hasher & Zacks, 1988; Rypma & D'Esposito, 2000), or to slower retrieval speed (Dehn, 2011). From a neural perspective, WM changes have been linked to age-related changes in the prefrontal cortex, a region important for WM (Braver et al., 1997). Consistent with this view, in functional neuroimaging studies of WM, age-related reductions in activation in the left prefrontal cortex have been observed, although these can be accompanied by increased activity in the right prefrontal cortex, which may play a compensatory role (Esposito, Kirkby, Van Horn, Ellmore, & Berman, 1999; Reuter-Lorenz et al., 2000).

However, the age-related changes in WM also seem to be modulated by factors that may not be well captured by the proposed explanatory accounts, and thus warrant further examination. Importantly, for this study, these include the key demographic factors of sex and education.

First, some evidence suggests that WM may be differentially impacted by ageing in males and females, although findings have been inconsistent. In younger adults, a number of studies suggest that females show better performance than males at verbal WM tasks, whereas males outperform females at visuospatial WM tasks (Duff & Hampson, 2001; Kaufman, 2007; Lejbak, Crossley, & Vrbancic, 2011; Loring-meier & Halpern, 1999; Lynn & Irwing, 2008; Postma, Jager, Kessels, Koppeschaar, & van Honk, 2004; Voyer, Voyer, & Saint-Aubin, 2017). However, other evidence suggests that males can outperform females in verbal as well as visuospatial WM (Zilles et al., 2016). Moreover, a fair number of studies examining younger adults report no sex differences at all in a variety of WM tasks (Brockmole & Logie, 2013; K. L. Evans & Hampson, 2015; Goldstein et al., 2005; T. Li, Luo, & Gong, 2010; Robert & Savoie, 2006; Schmidt et al., 2009).

The picture is also somewhat mixed in older adults. While some studies of older adults have reported no sex differences in verbal WM (Doppelt & Wallace, 1955) or visuospatial WM (Ruggiero, Sergi, & Iachini, 2008), others have found male advantages in aspects of verbal, visual, or visuospatial WM (Cansino et al., 2013; Fournet et al., 2012). Of interest here, in their *n*-back study Cansino and colleagues (2013) reported better performance in males than females on visuospatial WM (across the 1-back and 2-back subtasks) between ages 41 and 70, on verbal WM between ages 41 and 50, and on the 2-back subtasks (across the verbal and visuospatial versions) between ages 21 and 30 and again between 41 and 60, with no differences on the 1-back subtasks. No male advantages were observed at the highest age range, 71-80. In addition, in no case was superior WM performance observed in females compared with males. Overall, the evidence seems to suggest that sex differences in WM are often (though not always) observed, but that this pattern is at least somewhat modulated by age, with the possibility of male advantages across both verbal and non-verbal WM in mid-to-older ages. However, the available evidence is still relatively scarse, and further elucidation of the potential effects of sex on WM and ageing seems warranted.

Evidence also suggests that education may play a role in WM, although this has been less well studied than the role of sex in WM. Indeed, we are aware of only a handful of studies that have examined the relation between education and WM across younger and older adults. van Gerven, Meijer, and Jolles (2007) found that across both younger adults and (somewhat) older adults (aged 50-60), participants with higher education outperformed those of lower education on a numerical *n*-back task; however, education did not have differential effects in the younger and older groups. Dorbath, Hasselhorn, and Titz (2013), who examined aspects of verbal WM, also found better performance on high-educated compared with low-educated participants, but only in older adults (59-80 years), not in younger adults (19-35 years). In Cansino and colleagues (2013), education was not included as a factor in their regression models, but was examined separately, though without testing for interactions with age, sex, load, or verbal/visuospatial WM. The analyses revealed that higher education predicted better performance at the *n*-back task. Similarly, Brockmole and Logie (2013), who examined a wide age range between childhood and old age, found that education correlated positively with visual WM performance across the lifespan, although that study did not report interactions with age or examine education separately at different ages. In sum, education appears to show a positive relation with WM, perhaps especially at older ages, although there is still little research on the role of education on WM, let alone on how it may interact with ageing and WM.

In sum, the literature examining WM effects across adulthood thus far suggests that WM shows declines between younger and older adulthood, and that sex and perhaps education might modulate WM declines. However, these findings do not in themselves shed light on how WM is affected by age or other factors *over the course of old age.* Indeed, most studies examining younger and older adults have grouped older adults together across a range of ages, which moreover can be quite large (e.g., 55-81 years in Atkinson et al., 2018), thus precluding the examination of age effects *within* old age. Such coarse-grained categorisation of age seems to implicitly assume that few if any changes in WM abilities take place within older ages, although such patterns are still unclear.

Review of effects of age (and sex and education) within older adults

We are aware of three studies that have examined effects of ageing on WM within older adults (Cansino et al., 2013; Fournet et al., 2012; Kumar, Priyadarshi, & Sah, 2017). All three of these studies treated age as a categorical rather than continuous variable. Although treating age as a categorical variable can provide advantages, such as reducing the effect of extreme age outliers, examining age as a continuous variable can reveal more fine-grained patterns of ageing, including more easily revealing the exact (linear or nonlinear) shape of declines. Moreover, extreme age outlier effects can be addressed through other means in studies using age as a continuous variable (see the "Methods" section). Note that there is also a rich literature investigating how training impacts WM in older age (S.-C. Li et al., 2008; Schmiedek, Lövdén, & Lindenberger, 2009; see also Karbach & Verhaeghen, 2014, for a meta-analysis). However, as these studies examine effects of training, rather than unveiling the trajectory of WM decline in older years, and how this is predicted by education and sex, we do not review this literature here.

Here, we summarise the results and gaps of the three studies. First, in the earliest of these studies, which most clearly focused on old age, Fournet and colleagues (2012) tested a large group of older participants (55-85 years) on a set of tasks tapping verbal, visual, and visuospatial WM. They found that age (grouped in age decades: 55-65, 66-75, and 76-85 years) predicted WM declines in all domains, with a steeper decline for visuospatial than verbal WM (visual WM was not included in this comparison), consistent with Hale et al. (2011). The analyses did not reveal whether the declines were linear or nonlinear. Fournet and colleagues (2012) also found male advantages and positive education effects, across verbal, visual, and visuospatial WM, but did not examine interactions between either sex or education and age. Second, as described above, Cansino and colleagues (2013) examined WM in adults by age decades from 21-30 to 71-80. With respect to older individuals (above 60 years), they found WM declines between the 61-70 decade and the 71-80 decade, but solely for men in visuospatial WM (main effect, across 1-back and 2-back subtasks), with no declines for verbal WM, nor for either the 1-back or 2-back subtasks (main effects, across verbal and visuospatial WM) (see Figure 2 in Cansino et al., 2013). As with Fournet and colleagues (2012), the analyses did not provide any indication of the shape of the decline. As indicated above, within the older age range, sex differences were observed only in the 61-70 decade for visuospatial WM, with a male advantage. Education was not examined separately in older adults. Third and most recently, Kumar and colleagues (2017) tested adults between 40 and ~85 years of age in tasks designed to tap aspects of spatial, visual, and visuospatial WM. Again, they analysed their participants in groups by age decade. They found that all WM abilities declined between the ages of 40 and 60, with this decline continuing across older ages only for the spatial and visuospatial WM tasks. The analyses did not reveal whether the decline was linear or nonlinear. Effects of sex and education were not examined.

In sum, the small number of studies published thus far suggests that WM may decline within old age, although the fine-grained pattern is not yet clear, including whether any declines are linear or nonlinear. There are also gaps regarding whether and how sex and/or education, as well as load and interactions among these variables, may modulate these declines.

The present study

Thus, although there is by now a reasonable body of literature examining how and why WM changes between younger and older adults, there are gaps regarding whether, how, and why there may be WM changes during ageing *within* older ages. This study attempts to address these lacunae.

The study examines effects of ageing on WM abilities in a relatively large sample ($N=754$) of older Chinese-speaking adults from Taiwan. Thus, unlike most research on WM (and (neuro)cognition more generally), this study investigates a non-Western population. The study should therefore elucidate the nature of WM in ageing beyond Western populations, who in fact constitute only a portion of the global population. The adults ranged in age from 58 to 89 (with further analyses including extreme-aged individuals up to 98). Unlike previous studies of ageing in older adults, we used mixed-effects regression modelling, with age examined as a continuous variable for both linear and nonlinear effects. We also controlled and tested for potential roles of both sex and years of education (across a wide range, from 0 to 17 years of education), as well as load. Moreover, all main effects as well as all interactions among these variables were examined, to fully reveal their influence on WM in old age.

Importantly, as we have seen, both sex and education are associated with WM, including in old age, and thus warrant careful examination. Moreover, both sex and education are potentially confounding variables with respect to age. Women tend to live longer than men (Austad, 2006; Ginter & Simko, 2013), and older individuals may be less well educated, especially in recently developing countries such as Taiwan (Thornton, Chang, & Sun, 1984; Tsai, Gates, & Chiu, 1994). In addition (or alternatively), lower educated individuals often have lower socioeconomic status (SES), and thus may have a shorter life expectancy (Marmot, 2005). Thus, if these variables are not taken into account, apparent age effects could in fact be due in part to sex and/or education.

To examine WM abilities, we gave participants an *n*-back task (Owen, McMillan, Laird, & Bullmore, 2005; Schmiedek, Li, & Lindenberger, 2009). In its most typical form, which probes what are generally considered verbal aspects of WM, participants view a series of letters or digits, presented one at a time for a brief duration in the middle of a computer screen, and have to indicate whether each item is the same as that presented *n* items earlier on the list. For example, in a 1-back task participants are asked to indicate whether each item is the same or not as the item just presented, whereas in a 2-back task they must indicate whether each item is the same or not as the item presented penultimately (i.e., two items previously). Such parametric differences (e.g., 1-back vs. 2-back) are often referred to as "load" or "difficulty," and seem to probe aspects of WM capacity, or span. Indeed, it has been shown that performance on *n*-back tasks correlates well with performance on tasks designed to measure WM span (e.g., counting, reading, or rotation span tasks), suggesting that both types of tasks measure (at least in part) the same construct (Schmiedek, Hildebrandt, Lövdén, Wilhelm, & Lindenberger, 2009; Schmiedek, Lövdén, & Lindenberger, 2014; Shamosh et al., 2008, but see also Redick & Lindsey, 2013, for a discussion). In *n*-back tasks, participants are generally assessed regarding the accuracy of their responses, which are often computed as d' measures to avoid bias (see the "Methods" section), although in some studies response times are also collected.

The *n*-back task is one of the most widely used tasks in the study of WM (Braver et al., 1997; J. D. Cohen et al., 1997; Jaeggi, Buschkuehl, Perrig, & Meier, 2010; Kane, Conway, Miura, & Colflesh, 2007; B. Nagel, Ohannessian, & Cummins, 2007; Owen et al., 2005). The task seems to involve several aspects of WM, including the temporary storage of items, binding items to their temporal order, item retrieval, updating both items and their order, and monitoring and control over non-target items (Cansino et al., 2013). Therefore, the task can capture broad WM functioning and so can indicate whether such broad functioning is indeed affected by ageing or other factors. The task also has a number of other desirable characteristics. In particular, it is not only a relatively conceptually simple task, but in addition it seems to be much less dependent on (and so less influenced by) extraneous information and processes (e.g., in language or maths) than various other tasks commonly employed to examine WM (e.g., listening, reading, or operation span tasks; Alptekin & Erçetin, 2009; Janusik, 2007; Unsworth, Heitz, Schrock, & Engle, 2005). Thus, it appears to be a relatively pure probe of WM-related processes.

In this study, participants were given a digit *n*-back test (rather than a version with letters; Owen et al., 2005), to ensure that all the participants would be familiar with the items, given that Chinese-speaking people in Taiwanese commonly use Arabic numerals. Participants were given both 1-back and 2-back subtasks, allowing the examination of load effects. A 3-back subtask was not included as it was deemed to be too taxing for most older participants (Grigorova, Sherwin, & Tulandi, 2006).

Based on previous findings from the WM literature examining age effects within old age, we expected that WM abilities would decline with increasing age. A main effect of load was also predicted, with worse performance at 2-back than 1-back, as is generally found in *n*-back studies (Cansino et al., 2013; van Gerven et al., 2007, 2008). In addition, we expected that the examination of main effects and interactions involving sex and education, and their interactions with age, might reveal patterns found in previous studies. In particular, we expected that males might show better WM performance than females, especially at the lower age range examined here (see discussion above), and that, perhaps across the age range and across both sexes, participants with higher education would perform better than those with lower education.

Methods

Participants

This study was part of the Social Environment and Biomarkers of Aging Study (SEBAS), which, together with its parent study (the Taiwan Longitudinal Study of Aging), has collected a wide range of social, demographic, and health-related data, as well as performance and biomarker measures, on elderly and near elderly in Taiwan (Cornman et al., 2016; Goldman et al., 2004; Weinstein et al., 2014). During the 2011 SEBAS data collection, three computer-based cognitive tasks were also included: the Attention Network Test (ANT) (Fan, McCandliss, Fossella, Flombaum, & Posner, 2005), a recognition memory task to examine learning in declarative memory (Hedenius, Ullman, Alm, Jennische, & Persson, 2013; Lukács, Kemény, Lum, & Ullman, 2017), and the *n*-back task of WM that is reported in this article.

In this collection wave, a variety of demographic and related information was also acquired. This included sex, date of birth, total years of education (0-17, where 17 also included any additional years of education), handedness as measured by four questions modified from the Edinburgh Handedness Inventory (Oldfield, 1971) for the population being tested (targeting writing and the use of chopsticks, scissors, and brushing teeth), and information on any history of neurological, psychiatric, learning, cognitive, or other brain-related problems. This research was approved by the Georgetown University Institutional Review Board and the University of Kent Research Ethics Committee (the first author was previously at the University of Kent). Data requests for this study should be sent to: Health Promotion Administration, Ministry of Health and Welfare, 6th Floor, No 95 Mincyuan Road, West District, Taichung City, Taiwan 40341, ROC.

A cohort of 1,031 individuals participated in the 2011 wave of SEBAS, of whom 963 were given the *n*-back task. All were native speakers of Chinese, in particular Hakka, Mandarin, or Taiwanese (Taiwanese Hokkien). Of these, 39 participants were excluded because they did not perform the entire task to completion without interruptions; five because of coding errors, which made it impossible to match their *n*-back performance data with their demographic measures; 71 because of a diagnosis of a neurological, psychiatric, or other brain-related disorder, including stroke, brain embolism, intracranial haemorrhage, cerebral vascular sclerosis, brain atrophy/degeneration, concussion, hypoxia, recurrent headaches and dizziness, Parkinson's disease, epilepsy, meningitis, brain tumour, schizophrenia, depression, and bipolar disorder; and another 50 because their date of birth could not be obtained. The ages of the remaining 798 participants were calculated by subtracting their date of birth from the date of testing. Finally, to avoid extreme age outlier effects, we excluded the small number of participants ($n=12$) in their ninth decade (aged 90 or above; range 90-98).

Performance on the task was assessed by computing d' scores for each of the 786 remaining participants, separately for the 1-back and 2-back subtasks (see below for details regarding the calculation of d'). Some participants

Table 1. Demographic information.

	N	Age	Years of education
Male	398	69.05 (8.79)	8.61 (4.33)
Female	356	67.82 (8.25)	6.11 (4.56)
Total	754	68.47 (8.56)	7.43 (4.61)

Mean age (in years) and years of education, with standard deviations in parentheses. Males and females differed both in age ($t(752) = 1.97$, $p = .049$) and in years of education ($t(752) = 7.69$, $p < .001$). Also see section "Data Analysis."

produced only a small number of valid responses ("same" or "different" responses within the allotted time) and/or showed reverse discrimination (negative d' scores, indicating that the participants may have been performing the task incorrectly) in one or both of the subtasks. We excluded from analyses any participant's subtask with 10 or fewer valid trials (i.e., about one quarter of the trials) and/or with negative d' scores. Thus, participants for whom *both* subtasks met one or both of these exclusion criteria were fully excluded ($n = 32$). Statistical analyses were performed on the data of the resulting 754 participants, as reported below. Mean age and years of education for these participants are presented in Table 1. Also see Table in Supplemental Material for a breakdown of these demographics for participants grouped into 5-year age brackets.

Materials and design

The *n*-back task was adapted from a similar task developed by Benjamin Robinson and Rebecca Fuller at the University of Maryland, School of Medicine (http://step.talkbank.org/scripts-plus/). Participants viewed a series of digits (0-9) presented one at a time on a computer screen. Each digit was presented in Palatino Linotype font (72 point). For each digit, participants were asked to judge whether that digit was or was not the same as the digit that appeared immediately prior (1-back) or that appeared two items previously (2-back).

The 1-back subtask always preceded the 2-back subtask. Each of the two subtasks consisted of a single experimental block of 45 items, preceded by 12 practice items. In each of the two subtasks, one third of the items were selected to be targets (i.e., identical to the item that appeared one or two items prior, respectively, in the two subtasks); the first three items in each experimental block were not selected as target items. The remaining two thirds of the items were randomly selected as digits between 0 and 9. Thus, the appearance of "lure" items (trials that match an earlier item in the sequence, but not the item *n* items back; e.g., not 2 back in the 2-back subtask) was random. Lures are therefore likely to occur with similar probability across the variables of interest (age, sex, education). Lure effects are not examined here; for discussion of lures and their effects on ageing, see Schmiedek, Li, and colleagues (2009). When a randomly selected item was identical to the item that was presented one or two items before, it was appropriately treated as a 1-back or 2-back item in the analysis.

Procedure

Participants were given written instructions with Chinese characters, which were read orally in their native Chinese dialect (Hakka, Mandarin, or Taiwanese). They were asked to judge whether each digit was the same as the digit presented one or two items previously (in the respective subtasks). To perform this judgement, they were asked to press one of two buttons (left or right) on a Psychology Software Tools Serial Response Box (SRBox). These indicated yes or no answers, with the left/right order counterbalanced across participants. A reminder indicating which button to press (left or right) was displayed at the bottom of the screen during every trial (a green circle for "yes," a red X for "no"). Participants received training to ensure that they understood the task. This training included a running display of previous items to help the participant. This running display of numbers did not appear in the practice or experimental sessions. After training, the participants proceeded to the practice session, where 12 items were presented in the same manner as the subsequent experimental items. In cases where the participant had clear difficulties or requested a repetition, the practice session was repeated.

In the experimental (and practice) blocks, each trial involved the presentation of a digit for 500 ms, followed by a blank screen for a maximum of 2,500 ms, or until a response was given, at which point the next trial began. The task was presented in black on a white background on a laptop with Windows XP, using E-Prime Version 2.0 (Schneider, Eschman, & Zuccolotto, 2002a, 2002b).

Data analysis

As stated above, d' was computed for both the 1-back and 2-back subtasks (experimental items only) for each participant (consistent with most previous studies of *n* back, we focus on accuracy rather than response times). According to signal detection theory (Stanislaw & Todorov, 1999), d' scores measure discrimination independent of response bias, that is, independent of any tendencies for participants to give one or the other type of response (in this case, yes or no). d' is calculated from hits, misses, false alarms, and correct rejections. In the context of this task, when the item is the same as the one presented *n* back, a correct response (yes) is a hit, whereas an incorrect response (no) is a miss. When the item is not the same as the one presented *n* back, an incorrect response (yes—that is incorrectly indicating that the item is the same as the item *n* back) is a false alarm, whereas a correct response (no) is a correct rejection.

To compute d', we first calculated the Hit Rate (HR) and the False-Alarm Rate (FAR) over valid trials, that is, trials for which a yes or no response was given within the time limit. The HR is the proportion of correct hits over hits plus misses. The HR was adjusted by the loglinear method, to avoid infinite or indeterminate d' scores (Hautus, 1995; Stanislaw & Todorov, 1999). That is, 0.5 was added both to hits and to misses in the computation of HR. Thus, HR = (hits + 0.5) / ((hits + 0.5) + (misses + 0.5)). The FAR is the proportion of false alarms over false alarms plus correct rejections. Thus, FAR = (false alarms + 0.5) / ((false alarms + 0.5) + (correct rejections + 0.5)). To compute d', z-scores were first computed from these raw probabilities, separately for HR and FAR for each subtask for each participant. Finally, d' for each subtask for each participant was computed by subtracting the FAR z-score from the HR z-score (Macmillan, 1993; Stanislaw & Todorov, 1999). Higher d' values reflect better discrimination. A value of zero corresponds to chance performance, whereas negative values reflect reverse discrimination (Stanislaw & Todorov, 1999).

The participant d' scores were analysed with mixed-effects linear regression, with participant as a random effect. The following fixed predictors were included, as well as all of their interactions: load (two levels: 1-back, 2-back), age in years (as a continuous variable), years of education (also as a continuous variable), and sex (two levels: males, females). To obtain estimates of "main effects" for all predictors (analogous to those obtained for main effects in AN(C)OVAs), continuous predictors (i.e., age and education) were mean-centred, whereas categorical predictors (i.e., load and sex) were assigned sum-coded contrasts (i.e., −0.5 and 0.5; e.g., Barr, Levy, Scheepers, & Tily, 2013; Levy, 2014). Note that an alternative type of "main effect" coding for categorical predictors is to convert them to numeric variables and then mean centring them (e.g., Fraundorf & Jaeger, 2016; Montero-Melis, Jaeger, & Bylund, 2016). We also ran the regression model using this coding approach. The exact same pattern of significance (i.e., $ps < .05$, $ps < .10$) for main effects and interactions (as shown in Table 3) was obtained when using the alternative approach as our primary approach in coding the two categorical variables.

Because all predictors were simultaneously included in the regression analyses, this allowed us to control for any correlations between them. Specifically, estimates in multiple regression, including with mixed-effects models, reflect the unique variance of each predictor (i.e., the part of each variable that cannot be predicted by all others in the regression model). Effects should therefore be interpreted as the "pure" contribution of each variable, beyond any correlations with the others (e.g., Wurm & Fisicaro, 2014). For example, any differences between males and females in age or education (Table 1) do not explain the observed sex differences.

Table 2. Mean d' values (and *SDs*) for the 1-back and 2-back subtasks.

	1-back	2-back
Male	2.36 (1.11)	1.52 (1.09)
Female	2.05 (1.29)	1.27 (1.01)
Total	2.21 (1.21)	1.40 (1.06)

SD: standard deviation.

Finally, we computed standardised effect sizes for all critical significant effects. The computation of standardised effect sizes for mixed-effects regression is not straightforward, and indeed, these are often not reported for mixed-effects models. Here, we follow Westfall, Kenny, and Judd (2014) and compute a mixed-effects model analogue of Cohen's d by dividing regression estimates (b) by the expected variation of individual data points. For the mixed-effects model reported here, this expected variation is defined as the square root of the total variance (i.e., the sum of the subject intercept variance and the residual variance). Note that, in the case of continuous predictors, regression estimates (including in mixed-effects regression) do not reflect a comparison between two groups (the typical use case for Cohen's d), but instead correspond to changes in the dependent variable for each unit in the predictor. Thus, to calculate an interpretable effect size, we calculated Cohen's d for the two continuous predictors (age, education, and their interactions with sex) by first fitting a regression model in which these predictors were standardised and then dividing the regression coefficients obtained in this model by the square root of the total variance, as above. Thus, Cohen's d value for each continuous predictor captures the effect size for each standard deviation in the predictor (i.e., for age, education), allowing for comparability across predictors. Interpretation of the magnitude of Cohen's d values follows Cohen's (1988) recommendation of 0.2 as a small effect size, 0.5 as a medium effect size, and 0.8 as a large effect size. For terminological precision, here we interpret 0.10-0.30 as small, 0.31-0.39 as small-to-medium, 0.40-0.60 as medium, 0.61-0.69 as medium-to-large, and 0.70 or above as large.

Results

Table 2 presents mean (by participant) d' scores in each of the n-back subtasks (1-back and 2-back), both across all participants and separately for males and females. The results of the mixed-effects regression model are shown in Table 3, which presents regression estimates (b), standard errors (*SE*), t-values, and p-values for every main effect and interaction.

Significant main effects were obtained for load, age, and years of education, whereas a borderline significant main effect was observed for sex. The main effect of load,

Table 3. Main effects and interactions from the mixed-effects linear regression model on the *n*-back *d′* scores.

	b	*SE*	*t*	*p*
Intercept (estimated grand mean)	1.9747	0.0319	61.87	<.001*
Load (1-back vs. 2-back)	0.8441	0.0465	18.16	<.001*
Age	−0.0345	0.0039	−8.76	<.001*
Education	0.0702	0.0068	10.30	<.001*
Sex (males vs. females)	0.1238	0.0638	1.94	.053#
Load × Age	0.0055	0.0058	0.95	.344
Load × Education	−0.0074	0.0100	−0.74	.461
Load × Sex	0.0578	0.0929	0.62	.534
Age × Education	0.0009	0.0008	1.07	.286
Age × Sex	−0.0171	0.0079	−2.18	.030*
Education × Sex	−0.0332	0.0136	−2.44	.015*
Load × Age × Education	0.0016	0.0012	1.32	.188
Load × Age × Sex	0.0110	0.0115	0.96	.339
Load × Education × Sex	−0.0294	0.0200	−1.47	.142
Age × Education × Sex	0.0009	0.0016	0.59	.557
Load × Age × Education × Sex	−0.0026	0.0024	−1.10	.274

*$p<.05$; #$p<.10$. The *p*-values were obtained from *t*-tests with 1373 degrees of freedom, calculated as number of data points (i.e., 1389) minus the number of fixed effect estimates (i.e., 16) (Baayen, Davidson, & Bates, 2008). All continuous predictors were mean-centred; all categorical predictors were assigned sum-coded contrasts (see the "Methods" section).

which showed a large effect size (Cohen's $d=0.95$; see the "Methods" section), was due to better *n*-back performance (i.e., higher *d′* scores) in the 1-back subtask than in the 2-back subtask (see Table 2). The borderline significant main effect of sex reflected the overall better performance by males than females (see Table 2). The main effects of the continuous variables of age and years of education on *d′* scores are displayed, respectively, in Figures 1 and 2. Whereas increasing age was associated with worse *n*-back performance, with a small-to-medium effect size (Cohen's $d=0.33$; Figure 1), a higher number of years of education was associated with better *n*-back performance, also with a small-to-medium effect size (Cohen's $d=0.36$; Figure 2).

The main effect of age was qualified by a significant interaction between age and sex, with a small effect size (Cohen's $d=0.16$). As shown in Figure 3, although the negative effect of age on *n*-back performance was present for both males ($b=-0.0430$, $SE=0.0046$, $t=-9.36$, $p<.001$; Cohen's $d=0.41$) and females ($b=-0.0259$, $SE=0.0064$, $t=-4.06$, $p<.001$; Cohen's $d=0.25$), the effect was more pronounced for males (medium effect size) than females (small effect size). (These follow-up analyses were computed by refitting the model with sex as a dummy-coded variable.) Moreover, at the minimum age in our sample (58 years), the predicted *d′* score (i.e., from the regression model) for males was significantly higher than for females ($b=0.2977$, $SE=0.0931$, $t=3.20$, $p=.001$; Cohen's $d=0.34$), with a small-to-medium effect size. (This "endpoint" comparison was computed by refitting the model after subtracting the minimum age value, that is, 58, from each participant's age; the endpoint analyses below were computed in an analogous manner.) In contrast, at the maximum age of 89, the predicted *n*-back performance of males and females did not significantly differ ($b=-0.2332$, $SE=0.1861$, $t=-1.25$, $p=.211$).

Moreover, the main effects of years of education were qualified by a significant interaction between education and sex, with a small effect size (Cohen's $d=0.17$); see Figure 4. Although increasing education had a positive effect on *n*-back performance in both males ($b=0.0536$, $SE=0.0093$, $t=5.78$, $p<.001$; Cohen's $d=0.27$) and females ($b=0.0868$, $SE=0.0100$, $t=8.69$, $p<.001$; Cohen's $d=0.45$), this effect was larger for females (medium effect size) than males (small effect size), as can also be seen by the steeper slope of the dashed line in Figure 4. In addition, at the minimum level of education in our sample (0 years), the predicted *d′* score for males was significantly larger than for females ($b=0.3777$, $SE=0.1173$, $t=3.22$, $p=.001$; Cohen's $d=0.41$), with a medium effect size. In contrast, at the maximum level of education (17 years), this difference between the two sexes disappeared ($b=-0.1875$, $SE=0.1474$, $t=-1.27$, $p=.204$).

We emphasise that the sex difference endpoint tests at low and high age and education are *not* simple comparisons of mean *d′* scores of males and females at those points, but rather comparisons of *predicted d′* scores from the regression model. Note also that for all male versus female comparisons at minimum and maximum age and education, both sexes were represented; that is, there were both males and females with 58 and 89 years of age and with 0 and 17 years of education; see the "Discussion" section regarding participants with 0 years of education.

These patterns were robust, with the exact same pattern of significance (i.e., $ps<.05$, $ps<.10$) for main effects

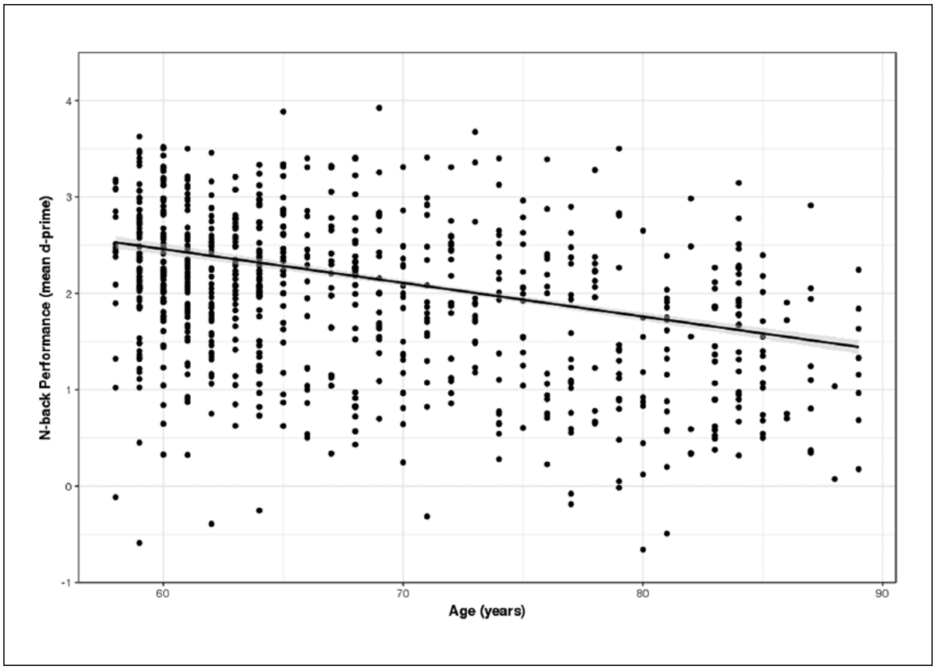

Figure 1. Performance on the *n*-back task as a function of age. In all figures, regression line(s) represent the effect(s) of interest while holding all other predictors constant at their means. Also in all figures, the value of each plotted data point is an individual participant's mean *d′* score in the *n*-back task, averaged over the 1-back and 2-back subtasks. Before computing this average, *d′* scores were adjusted by subtracting the summed effect of all predictors, except the plotted predictors of interest (see Prado & Ullman, 2009, p. 859, footnote 3). Shaded bands represent pointwise standard errors (95% confidence intervals are approximately twice the width of standard error bands).

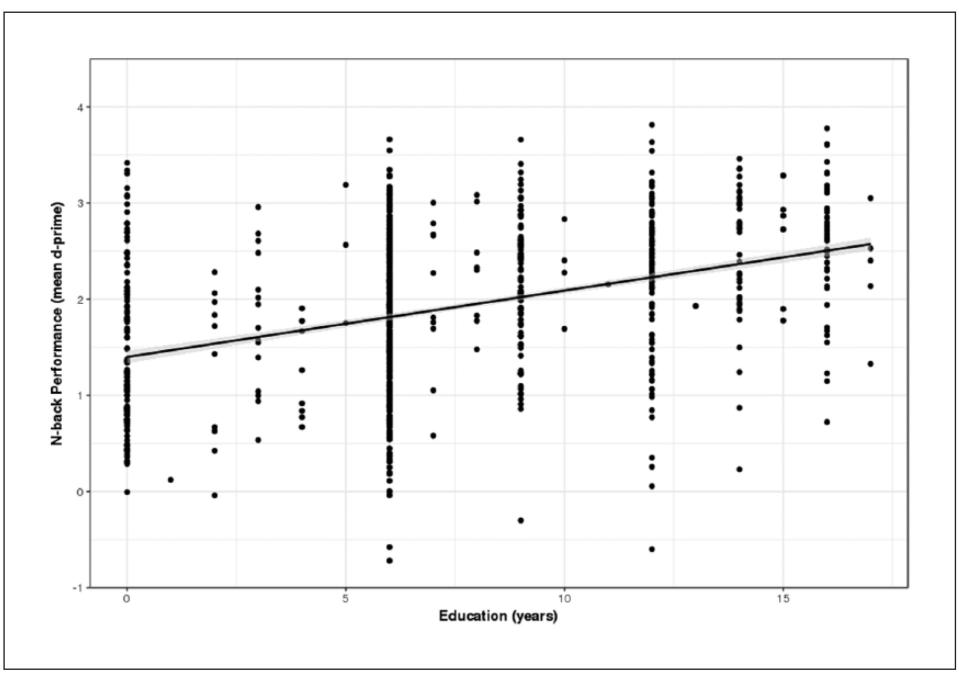

Figure 2. Performance on the *n*-back task as a function of years of education. Each data point represents an individual participant's *n*-back performance, which has been adjusted by subtracting the effect of all other predictors; see Note for Figure 1. Shaded bands represent pointwise standard errors (95% confidence intervals are approximately twice the width of standard error bands).

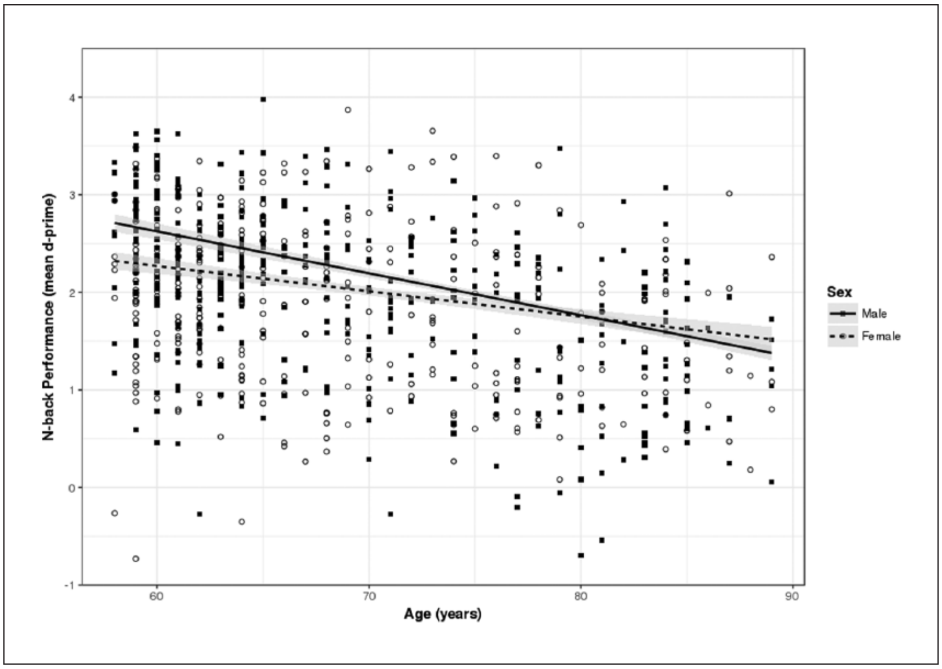

Figure 3. Performance on the *n*-back task as a function of age, separately for males (solid line, squares) and females (dashed line, circles). Each data point represents an individual participant's *n*-back performance, which has been adjusted by subtracting the effect of all other predictors; see Note for Figure 1. Shaded bands represent pointwise standard errors (95% confidence intervals are approximately twice the width of standard error bands).

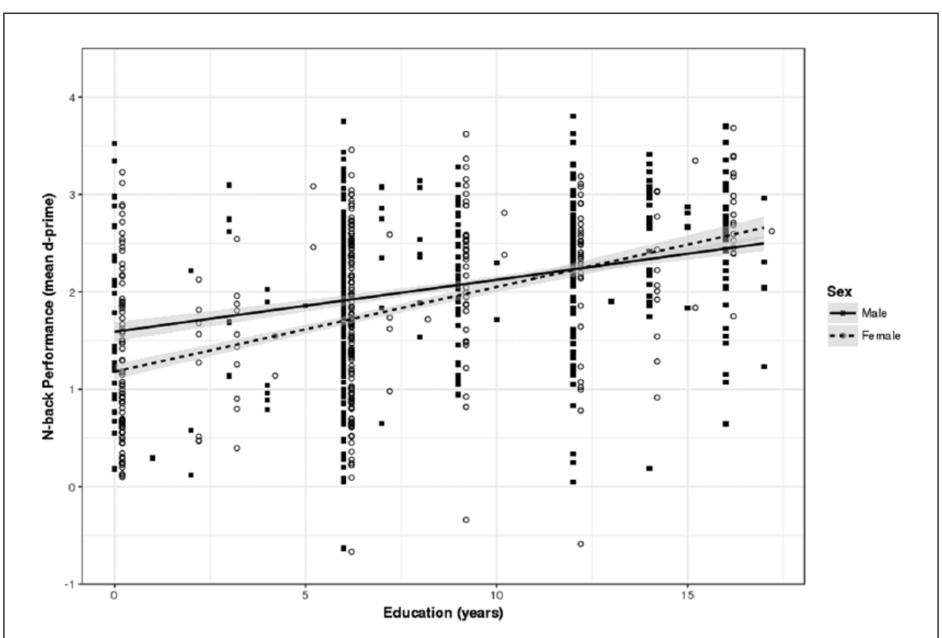

Figure 4. Performance on the *n*-back task as a function of years of education, separately for males (solid line, squares) and females (dashed line, circles). In this plot, an offset of 0.2 years was added to the education level of females, to avoid overlapping data points between males and females; thus, data points which are in close horizontal proximity correspond to the same number of years of education. Each data point represents an individual participant's *n*-back performance, which has been adjusted by subtracting the effect of all other predictors; see Note for Figure 1. Shaded bands represent pointwise standard errors (95% confidence intervals are approximately twice the width of standard error bands).

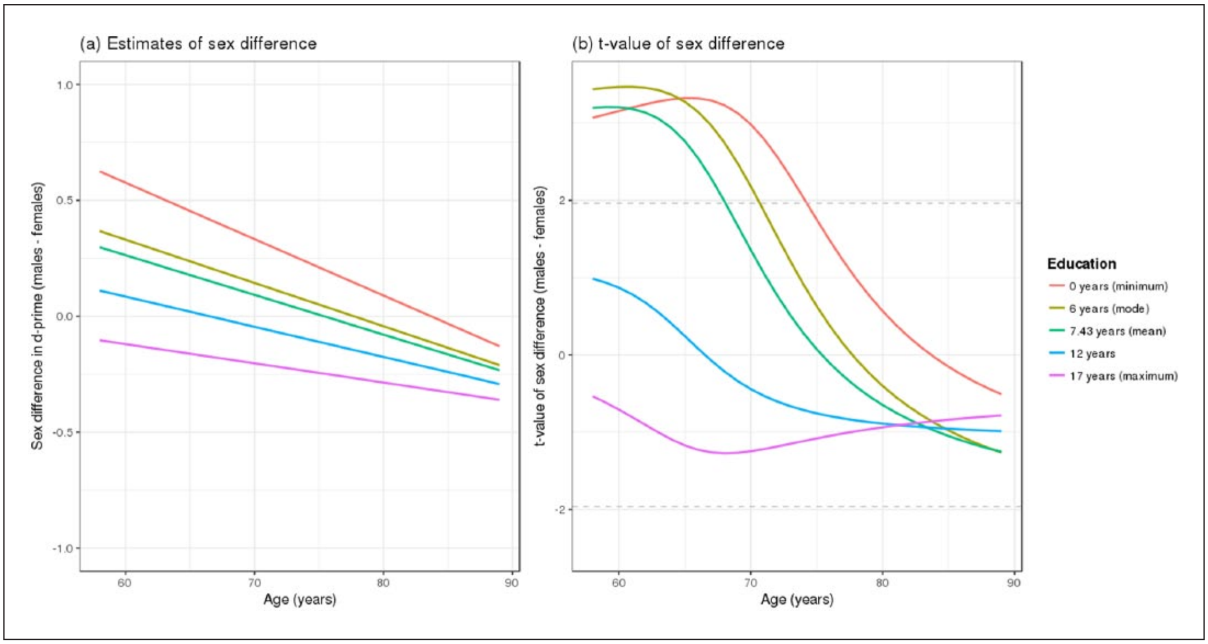

Figure 5. Estimated sex difference between males and females in d' scores (panel a) and t-value of that difference (panel b) as a function of age, for five different levels of education (coloured lines), obtained from the mixed-effects regression model. Dashed lines on panel b indicate statistical significance. See main text.

and interactions (as shown in Table 3) being obtained in a range of different alternate analyses. First, as indicated in the "Methods" section, the same patterns were obtained with two types of "main effect" coding, that is, regardless of whether categorical predictors were assigned sum-coded contrasts or were converted to numerical variables and then mean-centred. Second, it might be argued that the higher order interactions should not be retained in the regression model, as they were not significant, and might reduce the statistical power for lower order effects. However, when the four-way interaction and all four three-way interactions were removed from the model, again the same pattern of significance was obtained for the remaining effects. Third, it is possible that handedness, which was not included as a covariate, might bias the results. However, when handedness (coded as −100 to 100; Oldfield, 1971) was included as a covariate, for those 742 participants for whom handedness values were available, again the same pattern was obtained (though in fact in this case the main effect of sex reached statistical significance). Finally, the same pattern of significant results was also observed when the 10 participants who were 90 or older (and did not meet any other exclusion criterion) were added in to the analyses. Thus, even the inclusion of these extreme-aged individuals did not affect the results.

Given that both age and education had linear effects on d' scores, and given that both predictors interacted with sex, we performed exploratory analyses asking how these two interactions additively combined to determine sex differences. As described above, males outperformed females at the lower age endpoint of our sample, but increasing age was associated with a reduction and eventual elimination of sex differences. Similarly, at 0 years of education, males outperformed females, but this difference disappeared with increasing education. We can thus examine whether at younger ages (i.e., among older adults) higher levels of education eliminate sex differences and, conversely, whether at lower education higher ages eliminate sex differences.

Figure 5 displays the estimated sex difference (panel a) and the t-value (allowing for the computation of statistical significance) of this difference (panel b) as a function of age, for five different years of education (in the sample of 754 participants): (a) the minimum education in our cohort, 0 years ($n=109$); (b) the most common education level in our cohort, 6 years (i.e., the mode; $n=299$); (c) the mean level of education, 7.43 years; (d) the second most common education level in our cohort, 12 years ($n=109$); and (e) the maximum education level, 17 years ($n=5$). This provides estimations not only for common statistics in our sample (minimum, maximum, mean, mode), but more generally for levels of education that may be pertinent in the real world: no education (pertinent in developing or formerly developing countries such as Taiwan; see the "Discussion" section), some primary/middle school, secondary/high school education, and higher/university education. Estimates and t-values were computed by refitting the model presented in Table 3 with both education and age centred around specific values (every year of age between the minimum and maximum points, combined with the five levels of education described above).

As can be seen in Figure 5, at the lowest level of education (0 years), our regression model predicts that males have numerically higher d' scores than females throughout much of the age range, a difference that is significant from 58 until 74 years of age, inclusive (significance corresponds to an absolute t-value of 1.962 or higher, given the size of our sample; 1373 degrees of freedom). However, as expected, with increasing education, the age at which the male advantage disappears is progressively reduced. At 6 years of education, a significantly higher d' for males than females is predicted until 70 years of age. At the mean level of education in our sample (7.43 years), a significant male advantage is present until 68 years of age. Finally, at higher levels of education (12 and 17 years), *females* show a numerical advantage throughout much (at 12 years of education) or all (17 years) of the age range, although this advantage never reaches statistical significance. The finding that similar patterns are observed at both of these higher levels of education argues against spurious results from the small sample size at maximum education.

Thus, at the lowest education levels males show clear advantages, at least up to fairly old age, whereas at the highest education levels no sex differences are found, and indeed females generally show a quantitative advantage. From the perspective of age, at lower ages males tend to show advantages, except at the highest education levels, whereas at higher ages there are no significant sex differences at any education level, although females show a consistent quantitative advantage.

Finally, we tested for potential non-linearities in the relation between age and d' scores, by including an additional quadratic term for age in the mixed-effects regression model (for the original data set of 754 participants). To eliminate the correlation between the quadratic and linear terms of age, the quadratic term was included in the model as an orthogonal polynomial. A likelihood ratio test revealed that this model did not have a significantly higher goodness-of-fit than the linear model presented above; that is, the quadratic term for age failed to reach significance ($\chi^2(1) = 0.50$, $p = .477$). In addition, we ran a more complex model in which the quadratic term of age was allowed to interact with all other predictors, that is, with the (linear) predictors of load, education, and sex (thus, in this model, both age as a linear term and age as a quadratic term interacted with these other predictors). Again, a likelihood ratio test revealed that this more complex model did not differ in goodness-of-fit from the linear model ($\chi^2(8) = 9.28$, $p = .319$) or from the simpler model with only a quadratic term of age ($\chi^2(7) = 8.78$, $p = .269$). Because the inclusion of a quadratic term of age failed to improve model fit, cubic and other higher order polynomials were not tested for inclusion.

Discussion

This study investigated WM *within* older adults. Specifically, we examined WM in 754 healthy older adults in Taiwan (aged 58-89), on a "verbal" (digit) version of the n-back task, with both 1-back and 2-back subtasks. With mixed-effects linear regression, we investigated not only the influence of age and load (1-back vs. 2-back) but also the effects of sex and education, and all interactions among these variables.

Interpretation of results

The results suggest the following. First, the striking main effect of load is consistent with the pattern more generally observed in n-back studies (see Introduction). Indeed, the result is consistent with the finding from Cansino and colleagues (2013) of worse performance at 2-back than 1-back across younger and older adults. Thus, this study suggests that higher load also leads to greater difficulties specifically within older adults, even when accounting for age, sex, education, and their interactions. In addition, the results suggest that the effect of load is not particularly modulated by ageing (within old age), as load did not interact with age or indeed with any of the other factors. This in turn suggests that the effects of ageing on WM might primarily impact aspects of WM other than load (span) (Hasher & Zacks, 1988; Rypma & D'Esposito, 2000).

Second, and more importantly, the findings suggest that ageing has a detrimental effect on WM not only between younger and older adults (Introduction) but also *within* old age. The analyses revealed that this is a linear rather than a nonlinear effect. Given that the age range examined in this study is quite large (between 58 and 90, and even to about 100 years in the analyses including the extreme-aged participants), the findings suggest that age has a negative linear effect on WM across much of old age. Note that the age by sex interaction does not obviate the general negative effect of age on WM, as both males and females showed this pattern. Unlike the three previous studies that examined WM within old age (Cansino et al., 2013; Fournet et al., 2012; Kumar et al., 2017), our study included age as a continuous variable, and probed for nonlinear as well as linear effects, revealing only linear age-related declines. Verbal WM was examined in two of these studies (Cansino et al., 2013; Fournet et al., 2012), one of which reported declines (Fournet et al., 2012). This study suggests that, consistent with Fournet and colleagues, aspects of verbal WM indeed show declines within old age. This conclusion is strengthened by the fact that our study examined a large sample of older adults, that the findings held across both the 1-back and 2-back subtasks, that our analyses held constant certain potentially moderating factors, and the results were robust. Together with other studies (Cansino et al., 2013; Fournet et al., 2012; Kumar et al., 2017), the evidence suggests that verbal as well as visuospatial WM shows declines within old age, and thus, WM may weaken during old age quite generally. Finally, note that the absence of nonlinear effects within old age does not preclude nonlinear declines across the full adult lifespan (as have been found in other cognitive domains; Nyberg et al., 2012), because declines may be quite shallow during early adulthood, and only later show steep declines, which may be captured here.

Third, the finding that education has a positive linear association with WM abilities is consistent with previous studies of younger and older adults that have examined this issue. As we saw in the Introduction, previous studies have reported positive main effects of education across younger and older adults (Brockmole & Logie, 2013; Cansino et al., 2013; van Gerven et al., 2007), with one study finding positive effects in older but not younger adults (Dorbath et al., 2013). In addition, Fournet and colleagues (2012) reported that education was positively associated with WM performance in their sample of older adults (aged 55-85). Neither of the other two studies of WM within older adults (Cansino et al., 2013; Kumar et al., 2017) examined effects of education within old age. Together with this study, which reveals positive effects of education on verbal WM in older adults while accounting for the influence of age, the available research seems to suggest that higher education is indeed associated with improved WM quite generally, across both verbal and visual/spatial WM tasks, but perhaps in particular in older adults. Note, however, that observations of greater effects of education on WM in older than younger adults (Dorbath et al., 2013) could be partly due to decreased WM abilities at older ages, which may be accompanied by increased variability (thus increasing the likelihood of observing effects of education or other factors). Importantly, the education by sex interaction does not obviate the general positive association between education with WM, because both males and females showed this pattern. Note also that this study examined a very large range of education, including participants with 0 years of education, and thus constitutes an important extension of the investigation of the relation between education and WM.

Interpretation of the positive association between education and WM is not straightforward. One possibility is that the observed association is explained by a positive effect of education on WM. For example, education may lead to strengthened long-term memory representations (Ritchie, Bates, & Deary, 2015), which themselves are associated with better WM performance (Engle, Nations, & Cantor, 1990; Gregg, Freedman, & Smith, 1989; see also Cowan, 1999, for the relationship between long-term memory and WM). Such strengthened representations could come about from greater input and/or cognitive stimulation, from schooling itself and/or from resulting social or professional outcomes of increased education (Adey, Csapó, Demetriou, Hautamäki, & Shayer, 2007). Strengthened long-term memory representations could also be explained by improved learning and memory (declarative memory) as a result of greater education, because such improvements have been linked to increased studying (Draganski et al., 2006; Ullman & Pullman, 2015). Education may also have more direct benefits on WM. Indeed, some evidence suggests that WM training may improve WM performance (Morrison & Chein, 2011), although this remains controversial (Melby-Lervåg & Hulme, 2013; Ritchie et al., 2015), and the equivalence between WM training and greater education is not clear. More generally, although a causal effect of education on WM (whether indirect or direct) is difficult to specifically test for, some evidence supports such a causal view, at least for certain cognitive functions, though not WM (Ritchie et al., 2015). This causal perspective jibes with the view that education-related WM advantages within old age may be explained by education-related improvements in "cognitive reserve," which are posited to lead to decelerated rates of cognitive decline in more highly educated individuals (Anderson, Saleemi, & Bialystok, 2017; Dorbath et al., 2013; Haut et al., 2005; Stern, 2002). This notion of cognitive reserve is consistent with recent findings that higher levels of education are related to less age-related loss of volume in frontal regions of the brain, as well as higher activation of these regions in older participants (>67 years old) in the *n*-back task (Boller, Mellah, Ducharme-Laliberté, & Belleville, 2017). Thus overall, it is quite plausible that education leads to improved WM.

However, we emphasise that other accounts of the positive association between education and WM are also possible. For example, perhaps better WM leads to higher levels of education, that is, to more years of schooling. Note that this is a different and perhaps somewhat less likely possibility than the suggestion that higher WM leads to better educational outcomes, such as improved scores in reading and mathematics assessments (Pickering, 2006). It is also possible that one or more other factors (e.g., motivation, or perhaps SES) could lead to improvements in both WM and education. Indeed, SES generally correlates both with educational level (White, 1982) and chronic stress (G. W. Evans & Schamberg, 2009), which in turn can negatively impact WM (Lupien, Maheu, Tu, Fiocco, & Schramek, 2007). Thus overall, the positive association between education and WM must be interpreted with caution.

Fourth, the significant interaction between age and sex reveals that, holding education constant, age negatively impacts WM more in males than females within old age. Moreover, whereas males showed WM advantages at about 60, no sex differences were observed at about 90. These findings are consistent with, as well as extend, the results reported by Cansino and colleagues (2013). As discussed above, in that study male advantages were observed for verbal and visuospatial WM (with no female advantages), mainly at middle age to earlier stages of old age (i.e., in the 41-50, 51-60, and 61-70 age decades), with no sex differences observed at the oldest decade tested (71-80). Thus, both this study and that by Cansino and colleagues suggest that at earlier stages of older adulthood, males have WM advantages, but that these gradually disappear during old age.

This pattern may be at least partly explained as follows. First, females may show a particular decrease in WM abilities during menopause, likely due to oestrogen loss (Almela, van der Meij, Hidalgo, Villada, & Salvador, 2012; Weber & Mapstone, 2009; Weber, Mapstone, Staskiewicz, & Maki, 2012). Indeed, research has shown a positive association between oestrogen and WM performance (Grigorova et al., 2006; Keenan, Ezzat, Ginsburg, & Moore, 2001). A menopause-related decline in WM in females is also consistent with the suggestion from the broader literature that male advantages in WM are less reliably observed in younger adults, in particular for verbal WM (see Introduction). Menopause typically occurs between about 49 and 52 years of age (Palacios, Henderson, Siseles, Tan, & Villaseca, 2010; Takahashi & Johnson, 2015), suggesting that most, if not all, of the women in our sample had completed menopause, and thus, menopause-related declines in WM would likely have already occurred. Therefore, the male advantage observed at the lower end of our age range (58 years of age) seems likely to be at least partly explained by a decrease in WM abilities in women during menopause.

The observed "recovery" of females in older age, both in this study and in Cansino and colleagues (2013), is potentially an even more interesting finding. Rather than suggesting an improvement in females' WM, the observed effect seems instead to reflect a steeper WM decline in males than females over the course of old age, eventually resulting in similar WM abilities between the sexes. This in turn could be due in part to the gradual decrease of testosterone in men in old age, as in males oestrogen is derived from testosterone (Mooradian & Korenman, 2006). Thus, a gradual decrease in oestrogen in males could help account for the pattern, although it remains unclear whether or to what extent oestrogen in fact declines in men during old age (Mooradian & Korenman, 2006).

Fifth, the study shows for the first time an interaction between education and sex in older individuals, with greater WM gains related to education in females than males. Moreover, whereas we found a male advantage at 0 years of education, no sex differences were observed at a high level of education, that is, at 17 or more years of schooling, more or less corresponding to a university education.

The mechanisms underlying this pattern remain to be clarified. One possibility is that the female WM disadvantage at low education simply reflects a more general female WM disadvantage in old age, especially at earlier stages of older adulthood, perhaps due to the effects of menopause (see above). On this view, overlaying this effect there is a stronger positive association between education and WM in women than men, at least in older adults. Such an association could be due to various factors. For example, perhaps the female disadvantage from menopause allows for greater gains from education, leading to greater education benefits in women than men. Alternatively, as evidence suggests female advantages at declarative memory (Ullman, Miranda, & Travers, 2008), more education in females might lead to correspondingly stronger memory representations (see above) in females than males, thus providing greater female benefits for WM. Conversely, it is plausible that better WM is more likely to lead to higher levels of education in women than men, as in the sample examined here women were less likely to be educated (see Table 1, and Tsai et al., 1994), and thus, WM (or other) advantages might be more likely to lead to more schooling for girls or women.

Another possibility is that at lower levels of education men in this sample may have tended to have substantially more cognitive stimulation than women, because men may have been more likely to be employed, whereas women tended to stay at home and raise children (Thornton et al., 1984; Tsai et al., 1994). In contrast, at higher education levels the amount of cognitive stimulation might have been more similar between the sexes. On this view, it is not the case that higher education is more beneficial to women than men, but rather that low education does not adequately capture individuals' cognitive stimulation. Indeed, Ardila and colleagues (2010) found a similar pattern of greater sex differences at lower than higher education with respect to other cognitive abilities in adults in Latin America and posited a similar account. Note that such an explanation might be expected to hold not just in older adults but also at younger ages, as indeed was found by Ardila and colleagues. Overall, although the mechanisms of the observed education by sex interaction are unclear, further studies examining the pattern seem desirable, given the potential importance of the finding.

It is worth noting that a significant portion of our participants had 0 years of education. This is not surprising for older adults in Taiwan. Thornton and colleagues (1984) reported that a considerable portion of the population born in the 1930s and 1940s in Taiwan (i.e., people between the ages of 60-80 in our sample) received no formal education, with percentages ranging from 35.4% of women and 15.6% of men born in the early 1930s, to 16% of women and 1% of men born in the late 1940s. Thornton and colleagues (1984) attributed these percentages to cultural factors (e.g., the expectation that women will be less educated than the men they will marry; see Tsai et al., 1994) and socioeconomic reasons (e.g., father's education), as well as to the gradual establishment of a formal educational system during and after the Japanese colonial period

(1895-1945), which might also explain the drop in these percentages in the 1940s.

Sixth, we found that the combination of lower ages (within old age) and lower education was associated with a clear male advantage at WM, whereas the combination of higher ages and higher education was associated with no sex difference in WM, and in fact quantitative *advantage* for females. Thus, whatever the mechanisms for age, education, and their interactions with sex regarding WM abilities, it appears that together, the factors of age and education are associated with substantial sex differences that moreover differ across both age and education.

Implications

The findings have a number of implications and suggest various lines of future research. The study suggests that WM and other aspects of cognition can and should be examined not just across the lifespan but also specifically within old age. It supports the view that age can be profitably examined in studies of WM (and other cognitive functions) in older adults as a continuous predictor, which can reveal fine-grained linear or nonlinear patterns. The study also underscores the feasibility of examining non-Western populations, who have been under-represented in the populations examined in experimental psychology. The results emphasise the importance of including other, potentially moderating variables, in particular sex and education, as well as corresponding interactions, in the examination of WM and other effects in ageing. Although the mechanisms underlying the interplay between sex, age, and education cannot be unpacked with our study, the results clearly reveal that both age and education are associated with WM abilities within old age, and that both of these interact with sex.

The sex differences, including the male–female differences in WM found at lower old age and at low education, were not negligible. The results suggest that sex differences in WM may be real, despite the controversy over whether neurocognitive sex differences actually exist (Denworth, 2017; Ingalhalikar et al., 2014). However, as sex appears to interact with other factors, sex differences may not be apparent without examining such interactions.

The findings also suggest that greater education could have positive effects on WM in older adults. Indeed, based on the (unstandardised) regression estimates (b), *each additional year of education* has about as much positive impact as the negative impact of *two years of age* (b is more than twice as large for education as for age; see Table 3). Moreover, as we have seen, such effects may benefit women more than men. This clearly warrants further study, given that globally, females often still have lower levels of education (Bradley, 2014), and that at least at earlier points of old age, women appear to have lower WM abilities than men, especially at lower levels of education.

Finally, the findings have potentially important translational implications. Given the importance of WM in cognition and everyday life, the age-related declines in WM during old age suggest the possible value of prevention or remediation. For example, education, or a more targeted approach focusing on whichever mechanisms may underlie possible positive WM effects of education, could potentially either delay WM declines (from education early in life) or ameliorate them in old age (from further education in old age). This may provide an argument for further efforts to increase the educational level of women, in particular in non-Western societies such as Taiwan, where the educational level of females has only started to approach that of males in the past 50 years (Thornton et al., 1984; Tsai et al., 1994). It is also possible that pharmacological analogues of sex hormones, or perhaps other pharmacological agents that improve memory, could be employed in older adults to ameliorate WM (Grigorova et al., 2006; Keenan et al., 2001; Ullman & Pullman, 2015).

Conclusion

In conclusion, this study showed that WM abilities are affected by a number of factors in older adults, at least as tested in a verbal n-back task. These factors include not only load and age itself (i.e., increasing age within older adults) but also education. Crucially however, whereas age has a negative impact on WM, education has a positive association. Moreover, both age and education interact with sex, with greater declines during old age in males than females, and greater gains associated with more education in females than in males. The findings reveal important aspects of the nature of WM within old age and have a number of basic research and potential translational implications.

Acknowledgements

C.P. and J.V. contributed equally to this work.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This work was supported by NIH R01 AG016790 (Princeton University; subcontract to Michael Ullman), NIH R01 AG016661 (Georgetown University), NIH R01 AG016661 (Georgetown University), The Graduate School of Arts and Sciences, Georgetown University, and a Franklin Grant awarded by the American Philosophical Society to Christos Pliatsikas.

Supplementary material

Supplementary material is available at journals.sagepub.com/doi/suppl/10.1177/1747021818791994.

ORCID iD

Christos Pliatsikas https://orcid.org/0000-0001-7093-1773

References

Adey, P., Csapó, B., Demetriou, A., Hautamäki, J., & Shayer, M. (2007). Can we be intelligent about intelligence? Why education needs the concept of plastic general ability. *Educational Research Review*, *2*, 75–97. doi:10.1016/j.edurev.2007.05.001

Almela, M., van der Meij, L., Hidalgo, V., Villada, C., & Salvador, A. (2012). The cortisol awakening response and memory performance in older men and women. *Psychoneuroendocrinology*, *37*, 1929–1940. doi:10.1016/j.psyneuen.2012.04.009

Alptekin, C., & Erçetin, G. (2009). Assessing the relationship of working memory to L2 reading: Does the nature of comprehension process and reading span task make a difference? *System*, *37*, 627–639.

Anderson, J. A. E., Saleemi, S., & Bialystok, E. (2017). Neuropsychological assessments of cognitive aging in monolingual and bilingual older adults. *Journal of Neurolinguistics*, *43*, 17–27. doi:10.1016/j.jneuroling.2016.08.001

Ardila, A., Bertolucci, P. H., Braga, L. W., Castro-Caldas, A., Judd, T., Kosmidis, M. H., ... Rosselli, M. (2010). Illiteracy: The neuropsychology of cognition without reading. *Archives of Clinical Neuropsychology*, *25*, 689–712. doi:10.1093/arclin/acq079

Atkinson, A. L., Baddeley, A. D., & Allen, R. J. (2018). Remember some or remember all? Ageing and strategy effects in visual working memory. *The Quarterly Journal of Experimental Psychology*, *71*, 1561–1573. doi:10.1080/17470218.2017.1341537

Austad, S. N. (2006). Why women live longer than men: Sex differences in longevity. *Gender Medicine*, *3*, 79–92. doi:10.1016/S1550-8579(06)80198-1

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, 59, 390–412.

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

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

Baddeley, A. D. (2003b). Working memory and language: An overview. *Journal of Communication Disorders*, *36*, 189–208.

Baddeley, A. D. (2007). *Working memory, thought, and action* (Vol. 45). Oxford, UK: Oxford University Press.

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

Barr, D., Levy, R., Scheepers, C., & Tily, H. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. Retrieved from http://www.sciencedirect.com/science/article/pii/S0749596X12001180

Boller, B., Mellah, S., Ducharme-Laliberté, G., & Belleville, S. (2017). Relationships between years of education, regional grey matter volumes, and working memory-related brain activity in healthy older adults. *Brain Imaging and Behavior*, *11*, 304–317. doi:10.1007/s11682-016-9621-7

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *The Journals of Gerontology, Series B: Psychological Sciences & Social Sciences*, *60*, P223–P233. doi:10.1093/geronb/60.5.P223

Bopp, K. L., & Verhaeghen, P. (2018). Aging and n-back performance: A meta-analysis. *The Journals of Gerontology, Series B: Psychological Sciences & Social Sciences*. Advance online publication. doi:10.1093/geronb/gby024

Bradley, K. (2014). Gender inequality in education. In D. Levinson, P. Cookson & A. Sadovnik (Eds.), *Education and sociology: An encyclopedia* (p. 295), London: Routledge.

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage*, *5*, 49–62. doi:10.1006/nimg.1996.0247

Brockmole, J. R., & Logie, R. H. (2013). Age-related change in visual working memory: A study of 55,753 participants aged 8–75. *Frontiers in Psychology*, *4*, Article 12. doi:10.3389/fpsyg.2013.00012

Cansino, S., Hernández-Ramos, E., Estrada-Manilla, C., Torres-Trejo, F., Martínez-Galindo, J. G., Ayala-Hernández, M., ... Rodríguez-Ortiz, M. D. (2013). The decline of verbal and visuospatial working memory across the adult life span. *Age*, *35*, 2283–2302. doi:10.1007/s11357-013-9531-1

Cassel, C. K. (2001). Successful aging. How increased life expectancy and medical advances are changing geriatric care. *Geriatrics*, *56*, 35–39; quiz 40. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11196337

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum

Cohen, G., & Conway, M. A. (2007). *Memory in the Real World. London: Psychology Press.*

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, *386*, 604–608. doi:10.1038/386604a0

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

Cornman, J. C., Glei, D. A., Goldman, N., Chang, M.-C., Lin, H.-S., Chuang, Y.-L., & Weinstein, M. (2016). Cohort profile: The Social Environment and Biomarkers of Aging Study (SEBAS) in Taiwan. *International Journal of Epidemiology*, *45*, 54–63. doi:10.1093/ije/dyu179

Cowan, N. (1998). Visual and auditory working memory capacity. *Trends in Cognitive Sciences*, *2*(3), 77. doi:10.1016/S1364-6613(98)01144-9

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (*Vol. 20*, pp. 62–101). New York, NY: Cambridge University Press.

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

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik (Ed.), *Aging and cognitive processes* (pp. 191–211). New York: Plenum Press.

Dehn, M. J. (2011). *Working memory and academic learning: Assessment and intervention*. Hoboken, NJ: John Wiley.

Denworth, L. (2017). Is there a "female" brain? *Scientific American*, *317*, 38–43. doi:10.1038/scientificamerican 0917-38

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

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, *66*, 115–142. doi:10.1146/annurev-psych-010814-015031

Doppelt, J. E., & Wallace, W. L. (1955). Standardization of the Wechsler Adult Intelligence Scale for older persons. *The Journal of Abnormal and Social Psychology*, *51*, 312–330. doi:10.1037/h0044391

Dorbath, L., Hasselhorn, M., & Titz, C. (2013). Effects of education on executive functioning and its trainability. *Educational Gerontology*, *39*, 314–325. doi:10.1080/0360 1277.2012.700820

Draganski, B., Gaser, C., Kempermann, G., Kuhn, H. G., Winkler, J., Buchel, C., & May, A. (2006). Temporal and spatial dynamics of brain structure changes during extensive learning. *Journal of Neuroscience*, *26*, 6314–6317. doi:10.1523/JNEUROSCI.4628-05.2006

Duff, S. J., & Hampson, E. (2001). A sex difference on a novel spatial working memory task in humans. *Brain and Cognition*, *47*, 470–493. doi:10.1006/brcg.2001.1326

Engle, R. W., Nations, J. K., & Cantor, J. (1990). Is "working memory capacity" just another name for word knowledge? *Journal of Educational Psychology*, *82*, 799–804.

Esposito, G., Kirkby, B. S., Van Horn, J. D., Ellmore, T. M., & Berman, K. F. (1999). Context-dependent, neural system-specific neurophysiological concomitants of ageing: Mapping PET correlates during cognitive activation. *Brain*, *122*, 963–979.

Evans, G. W., & Schamberg, M. A. (2009). Childhood poverty, chronic stress, and adult working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 6545–6549.

Evans, K. L., & Hampson, E. (2015). Sex differences on prefrontally-dependent cognitive tasks. *Brain and Cognition*, *93*, 42–53. doi:10.1016/j.bandc.2014.11.006

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *NeuroImage*, *26*, 471–479. doi:10.1016/j.neuroimage.2005.02.004

Federico, F., Delogu, F., & Raffone, A. (2014). Maintenance and manipulation of object sequences in working memory: A lifespan study. *Neurological Sciences*, *35*, 1883–1887. doi:10.1007/s10072-014-1851-0

Fournet, N., Roulin, J.-L., Vallet, F., Beaudoin, M., Agrigoroaei, S., Paignon, A., ... Desrichard, O. (2012). Evaluating short-term and working memory in older adults: French normative data. *Aging & Mental Health*, *16*, 922–930. doi:10.108 0/13607863.2012.674487

Fraundorf, S. H., & Jaeger, T. F. (2016). Readers generalize adaptation to newly-encountered dialectal structures to other unfamiliar structures. *Journal of Memory and Language*, *91*, 28–58.

Ginter, E., & Simko, V. (2013). Women live longer than men. *Bratislavske Lekarske Listy*, *114*(2), 45–49. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23331196

Goldman, N., Weinstein, M., Cornman, J., Singer, B., Seeman, T., Goldman, N., & Chang, M.-C. (2004). Sex differentials in biological risk factors for chronic disease: Estimates from population-based surveys. *Journal of Women's Health*, *13*, 393–403.

Goldstein, J. M., Jerram, M., Poldrack, R. A., Anagnoson, R., Breiter, H. C., Makris, N., ... Seidman, L. J. (2005). Sex differences in prefrontal cortical brain activity during fMRI of auditory verbal working memory. *Neuropsychology*, *19*, 509–519. doi:10.1037/0894-4105.19.4.509

Grady, C. L., & Craik, F. I. M. (2000). Changes in memory processing with age. *Current Opinion in Neurobiology*, *10*, 224–231.

Gregg, V. H., Freedman, C. M., & Smith, D. K. (1989). Word frequency, articulatory suppression and memory span. *British Journal of Psychology*, *80*, 363–374.

Grigorova, M., Sherwin, B. B., & Tulandi, T. (2006). Effects of treatment with leuprolide acetate depot on working memory and executive functions in young premenopausal women. *Psychoneuroendocrinology*, *31*, 935–947. doi:10.1016/j.psyneuen.2006.05.004

Hale, S., Rose, N. S., Myerson, J., Strube, M. J., Sommers, M., Tye-Murray, N., & Spehar, B. (2011). The structure of working memory abilities across the adult life span. *Psychology and Aging*, *26*, 92–110. doi:10.1037/a0021483

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation*, *22*, 193–225. doi:10.1016/S0079-7421(08)60041-9

Haut, M. W., Kuwabara, H., Moran, M. T., Leach, S., Arias, R., & Knight, D. (2005). The effect of education on age-related functional activation during working memory. *Aging, Neuropsychology, and Cognition*, *12*, 216–229. doi:10.1080/13825580590969325

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d′. *Behavior Research Methods*, *27*, 46–51.

Hedenius, M., Ullman, M. T., Alm, P. A., Jennische, M., & Persson, J. (2013). Enhanced recognition memory after incidental encoding in children with developmental dyslexia. *PLoS ONE*, *8*(5), e63998. doi:10.1371/annotation/925d6161-da3b-46f5-9946-e3cfec7a1681

Ingalhalikar, M., Smith, A., Parker, D., Satterthwaite, T. D., Elliott, M. A., Ruparel, K., ... Verma, R. (2014). Sex differences in the structural connectome of the human brain. *Proceedings of the National Academy of Sciences*, *111*, 823–828. doi:10.1073/pnas.1316909110

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

Janusik, L. A. (2007). Building listening theory: The validation of the conversational listening span. *Communication Studies*, *58*, 139–156.

Johnson, W., Logie, R. H., & Brockmole, J. R. (2010). Working memory tasks differ in factor structure across age cohorts: Implications for dedifferentiation. *Intelligence*, *38*, 513–528. doi:10.1016/j.intell.2010.06.005

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when. *Psychological Science*, *18*, 614–621. doi:10.1111/j.1467-9280.2007.01948.x

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

Karbach, J., & Verhaeghen, P. (2014). Making working memory work. *Psychological Science*, *25*, 2027–2037. doi:10.1177/0956797614548725

Kaufman, S. B. (2007). Sex differences in mental rotation and spatial visualization ability: Can they be accounted for by differences in working memory capacity? *Intelligence*, *35*, 211–223. doi:10.1016/j.intell.2006.07.009

Keenan, P. A., Ezzat, W. H., Ginsburg, K., & Moore, G. J. (2001). Prefrontal cortex as the site of estrogen's effect on cognition. *Psychoneuroendocrinology*, *26*, 577–590. doi:10.1016/S0306-4530(01)00013-0

Kumar, N., Priyadarshi, B., & Sah, S. U. (2017). Is there similar decline in visual and spatial domains of visuo-spatial working memory with ageing? *Indian Journal of Gerontology*, *31*, 152–168.

Lejbak, L., Crossley, M., & Vrbancic, M. (2011). A male advantage for spatial and object but not verbal working memory using the n-back task. *Brain and Cognition*, *76*, 191–196. doi:10.1016/j.bandc.2010.12.002

Levy, R. (2014). Using R formulae to test for main effects in the presence of higher-order interactions. arXiv preprint arXiv:1405.2094.

Li, S.-C., Schmiedek, F., Huxhold, O., Röcke, C., Smith, J., & Lindenberger, U. (2008). Working memory plasticity in old age: Practice gain, transfer, and maintenance. *Psychology and Aging*, *23*, 731–742. doi:10.1037/a0014343

Li, T., Luo, Q., & Gong, H. (2010). Gender-specific hemodynamics in prefrontal cortex during a verbal working memory task by near-infrared spectroscopy. *Behavioural Brain Research*, *209*, 148–153. doi:10.1016/j.bbr.2010.01.033

Loring-meier, S., & Halpern, D. F. (1999). Sex differences in visuospatial working memory: Components of cognitive processing. *Psychonomic Bulletin & Review*, *6*, 464–471. doi:10.3758/BF03210836

Lukács, Á., Kemény, F., Lum, J. A. G., & Ullman, M. T. (2017). Learning and overnight retention in declarative memory in specific language impairment. *PLoS ONE*, *12*(1), Article e0169474. doi:10.1371/journal.pone.0169474

Lupien, S. J., Maheu, F., Tu, M., Fiocco, A., & Schramek, T. E. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and Cognition*, *65*, 209–237. doi:10.1016/j.bandc.2007.02.007

Lynn, R., & Irwing, P. (2008). Sex differences in mental arithmetic, digit span, and g defined as working memory capacity. *Intelligence*, *36*, 226–235. doi:10.1016/j.intell.2007.06.002

Macmillan, N. A. (1993). Signal detection theory as data analysis method and psychological decision model. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 21–57). Hillsdale, NJ: Lawrence Erlbaum.

Marmot, M. (2005). Social determinants of health inequalities. *The Lancet*, *365*, 1099–1104. doi:10.1016/S0140-6736(05)71146-6

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*, 270–291. doi:10.1037/a0028228

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Montero-Melis, G., Jaeger, T. F., & Bylund, E. (2016). Thinking is modulated by recent linguistic experience: Second language priming affects perceived event similarity. *Language Learning*, *66*, 636–665.

Mooradian, A. D., & Korenman, S. G. (2006). Management of the cardinal features of andropause. *American Journal of Therapeutics*, *13*, 145–160. doi:10.1097/01.mjt.0000132252.80403.c9

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review*, *18*, 46–60. doi:10.3758/s13423-010-0034-0

Nagel, B., Ohannessian, A., & Cummins, K. (2007). Performance dissociation during verbal and spatial working memory tasks. *Perceptual and Motor Skills*, *105*, 243–250.

Nagel, I. E., Preuschhof, C., Li, S.-C., Nyberg, L., Bäckman, L., Lindenberger, U., & Heekeren, H. R. (2011). Load modulation of BOLD response and connectivity predicts working memory performance in younger and older adults. *Journal of Cognitive Neuroscience*, *23*, 2030–2045. doi:10.1162/jocn.2010.21560

Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., & Bäckman, L. (2012). Memory aging and brain maintenance. *Trends in Cognitive Sciences*, *16*, 292–305. doi:10.1016/j.tics.2012.04.005

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113. doi:10.1016/0028-3932(71)90067-4

Orsini, A., Chiacchio, L., Cinque, M., Cocchiaro, C., Schiappa, O., & Grossi, D. (1986). Effects of age, education and sex on two tests of immediate memory: A study of normal subjects from 20 to 99 years of age. *Perceptual and Motor Skills*, *63*, 727–732. doi:10.2466/pms.1986.63.2.727

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

Palacios, S., Henderson, V. W., Siseles, N., Tan, D., & Villaseca, P. (2010). Age of menopause and impact of climacteric symptoms by geographical region. *Climacteric*, *13*, 419–428. doi:10.3109/13697137.2010.507886

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*, 299–320. doi:10.1037/0882-7974.17.2.299

Pertzov, Y., Heider, M., Liang, Y., & Husain, M. (2015). Effects of healthy ageing on precision and binding of object location in visual short term memory. *Psychology and Aging*, *30*, 26–35. doi:10.1037/a0038396

Peterson, D. J., & Naveh-Benjamin, M. (2016). The role of aging in intra-item and item-context binding processes in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 1713–1730.

Pfeiffer, H. C. V., Løkkegaard, A., Zoetmulder, M., Friberg, L., & Werdelin, L. (2014). Cognitive impairment in early-stage non-demented Parkinson's disease patients. *Acta Neurologica Scandinavica*, *129*, 307–318. doi:10.1111/ane.12189

Phillips, D. R. (2002). *Ageing in the Asia-Pacific region: Issues, policies and future trends*. London: Routledge.

Pickering, S. J. (2006). *Working memory and education*. Burlington, MA: Academic Press.

Postma, A., Jager, G., Kessels, R. P., Koppeschaar, H. P., & van Honk, J. (2004). Sex differences for selective forms of spatial memory. *Brain and Cognition*, *54*, 24-34. doi:10.1016/S0278-2626(03)00238-0

Prado, E. L., & Ullman, M. T. (2009). Can imageability help us draw the line between storage and composition? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 849–866. doi:10.1037/a0015286

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*, 110–122. doi:10.1016/j.lindif.2009.10.005

Rechel, B., Grundy, E., Robine, J. M., Cylus, J., MacKenbach, J. P., Knai, C., & McKee, M. (2013). Ageing in the European union. *The Lancet*, *381*, 1312–1322. doi:10.1016/S0140-6736(12)62087-X

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

Reuter-Lorenz, P. A., Jonides, J., Smith, E. E., Hartley, A., Miller, A., Marshuetz, C., & Koeppe, R. A. (2000). Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. *Journal of Cognitive Neuroscience*, *12*, 174–187.

Reuter-Lorenz, P. A., & Sylvester, C.-Y. C. (2005). The cognitive neuroscience of working memory and aging. In R. Cabeza, L. Nyberg & D. C. Park (Eds.), *Cognitive neuroscience of aging: Linking cognitive and cerebral aging* (pp. 186–217) . New York, NY: Oxford University Press.

Ritchie, S. J., Bates, T. C., & Deary, I. J. (2015). Is education associated with improvements in general cognitive ability, or in specific skills? *Developmental Psychology*, *51*, 573–582. doi:10.1037/a0038981

Robert, M., & Savoie, N. (2006). Are there gender differences in verbal and visuospatial working-memory resources? *European Journal of Cognitive Psychology*, *18*, 378–397. doi:10.1080/09541440500234104

Ruggiero, G., Sergi, I., & Iachini, T. (2008). Gender differences in remembering and inferring spatial distances. *Memory*, *16*, 821–835. doi:10.1080/09658210802307695

Rypma, B., & D'Esposito, M. (2000). Isolating the neural mechanisms of age-related changes in human working memory. *Nature Neuroscience*, *3*, 509–515. doi:10.1038/74889

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428.

Schmidt, H., Jogia, J., Fast, K., Christodoulou, T., Haldane, M., Kumari, V., & Frangou, S. (2009). No gender differences in brain activation during the N-back task: An fMRI study in healthy individuals. *Human Brain Mapping*, *30*, 3609–3615. doi:10.1002/hbm.20783

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

Schmiedek, F., Li, S. C., & Lindenberger, U. (2009). Interference and facilitation in spatial working memory: Age-associated differences in lure effects in the N-back paradigm. *Psychology and Aging*, *24*, 203–210. doi:10.1037/a0014685

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2009). On the relation of mean reaction time and intraindividual reaction time variability. *Psychology and Aging*, *24*, 841–857. doi:10.1037/a0017799

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2014). A task is a task is a task: Putting complex span, n-back, and other working memory indicators in psychometric context. *Frontiers in Psychology*, *5*, Article 1475. doi:10.3389/fpsyg.2014.01475

Schneider, W., Eschman, A., & Zuccolotto, A. (2002a). *E-prime reference guide*. Pittsburgh, PA: Psychology Software Tools.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002b). *E-prime user's guide*. Pittsburgh. PA: Psychology Software Tools.

Scullin, M. K., & Bliwise, D. L. (2015). Sleep, cognition, and normal aging. *Perspectives on Psychological Science*, *10*, 97–137. doi:10.1177/1745691614556680

Shamosh, N. A., DeYoung, C. G., Green, A. E., Reis, D. L., Johnson, M. R., Conway, A. R. A., ... Gray, J. R. (2008). Individual differences in delay discounting: Relation to intelligence, working memory, and anterior prefrontal cortex. *Psychological Science*, *19*, 904–911.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society*, *31*, 137–149. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10495845

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, *8*, 448–460. doi:10.1017/S1355617702813248

Takahashi, T. A., & Johnson, K. M. (2015). Menopause. *Medical Clinics of North America*, *99*, 521–534. doi:10.1016/j.mcna.2015.01.006

Thornton, A., Chang, M.-C., & Sun, T.-H. (1984). Social and economic change, intergenerational relationships, and family formation in Taiwan. *Demography*, *21*, 475–499. doi:10.2307/2060911

Tsai, S.-L., Gates, H., & Chiu, H.-Y. (1994). Schooling Taiwan's women: Educational attainment in the mid-20th century. *Sociology of Education*, *67*, 243–263. doi:10.2307/2112815

Ullman, M. T., Miranda, R. A., & Travers, M. L. (2008). Sex differences in the neurocognition of language. In J. B. Becker, K. J. Berkley, N. Geary, E. Hampson, J. P. Herman & E. Young, (Eds.), *Sex differences in the brain: From genes to behavior* (pp. 291–309). New York, NY: Oxford University Press.

Ullman, M. T., & Pullman, M. Y. (2015). A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience & Biobehavioral Reviews*, *51*, 205–222. doi:10.1016/j.neubiorev.2015.01.008

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

van Gerven, P. W. M., Meijer, W. A., & Jolles, J. (2007). Education does not protect against age-related decline of switching focal attention in working memory. *Brain and Cognition*, *64*, 158–163. doi:10.1016/j.bandc.2007.02.005

van Gerven, P. W. M., Meijer, W. A., Prickaerts, J. H. M., & Van der Veen, F. M. (2008). Aging and focus switching in working memory: Excluding the potential role of memory load. *Experimental Aging Research*, *34*, 367–378. doi:10.1080/03610730802274165

Voyer, D., Voyer, S. D., & Saint-Aubin, J. (2017). Sex differences in visual-spatial working memory: A meta-analysis. *Psychonomic Bulletin & Review*, *24*, 307–334. doi:10.3758/s13423-016-1085-7

Weber, M. T., & Mapstone, M. (2009). Memory complaints and memory performance in the menopausal transition. *Menopause*, *16*, 694–700. doi:10.1097/gme.0b013e318196a0c9

Weber, M. T., Mapstone, M., Staskiewicz, J., & Maki, P. M. (2012). Reconciling subjective memory complaints with objective memory performance in the menopausal transition. *Menopause*, *19*, 735–741. doi:10.1097/gme.0b013e318241fd22

Weinstein, M., Goldman, N., Chang, M.-C., Lin, H.-S., Chuang, Y.-L., Peterson, C. E., ... Wu, S.-I. (2014). *Social Environment and Biomarkers of Aging Study (SEBAS) in Taiwan, 2000 and 2006*. doi:10.3886/ICPSR03792.v7

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*, 2020–2045. doi:10.1037/xge0000014

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, *91*, 461–481. doi:10.1037/0033-2909.91.3.461

Whitwell, J. L., Jones, D. T., Duffy, J. R., Strand, E. A., Machulda, M. M., Przybelski, S. A., ... Josephs, K. A. (2015). Working memory and language network dysfunctions in logopenic aphasia: A task-free fMRI comparison with Alzheimer's dementia. *Neurobiology of Aging*, *36*, 1245–1252. doi:10.1016/j.neurobiolaging.2014.12.013

Wurm, L. H., & Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language*, *72*, 37–48.

Zilles, D., Lewandowski, M., Vieker, H., Henseler, I., Diekhof, E., Melcher, T., ... Gruber, O. (2016). Gender differences in verbal and visuospatial working memory performance and networks. *Neuropsychobiology*, *73*, 52–63. doi:10.1159/000443174