Bilingualism: Language and Cognition 22 (4), 2019, 733–751 © Cambridge University Press 2018 doi:10.1017/S1366728918000317

# Bilingual experience and executive control over the adult lifespan: The role of biological sex

SIVANIYA SUBRAMANIAPILLAI
*Department of Psychology, Faculty of Science, McGill University*
*Douglas Brain Imaging Centre, Douglas Institute Research Centre*
*Centre for Research on Brain, Language & Music, McGill University*
MARIA NATASHA RAJAH
*Department of Psychiatry, Faculty of Medicine, McGill UniversitDouglas Brain Imaging Centre, Douglas Institute Research Centre*
*Centre for Research on Brain, Language & Music, McGill University*
STAMATOULA PASVANIS
*Douglas Brain Imaging Centre, Douglas Institute Research Centre*
DEBRA TITONE
*Department of Psychology, Faculty of Science, McGill University*
*Centre for Research on Brain, Language & Music, McGill University*

(Received: January 12, 2017; final revision received: January 29, 2018; accepted: February 01, 2018)

*We investigated whether bilingual language experience over the lifespan impacts women and men in a manner that differentially buffers against age-related declines in executive control. To this end, we investigated whether executive control performance in a lifespan sample of adult women and men were differentially impacted by individual differences in bilingual language experience, assessed using an unspeeded measure of executive control: the Wisconsin Card Sort Test. The results suggested that women showed both the greatest degree of age-related decline across WCST measures, and a greater likelihood than men to express improved performance as a function of increased bilingual experience. We consider implications of this finding for advancing our understanding of the relation between bilingualism and cognition, and also the effects of biological sex on cognitive aging.*

Keywords: bilingualism, sex, aging, Wisconsin Card Sort Test, executive control

## Introduction

Executive control refers to the domain-general neurocognitive coordination of skills or habits that allow us to successfully implement goal-directed behaviour (e.g., Braver, 2012; Braver & West, 2008; Miyake, Friedman, Emerson, Witzki, Howerter & Wager, 2000; Miyake & Friedman, 2012). Healthy aging is typically associated with declines in executive control (e.g., Amer, Campbell & Hasher, 2016; Salthouse, 2009, 2010; Stuss, 2011); however, there is great variability in how age-related declines manifest across executive control tasks (Christensen, Mackinnon, Korten, Jorm, Henderson, Jacomb & Rodgers, 1999; Mungas, Beckett, Harvey, Farias, Reed, Carmichael, Olichney, Miller & DeCarli, 2010; Wilson, Beckett, Barnes, Schneider, Bach, Evans & Bennett, 2002; Zelinski, Gilewski & Schaie, 1993). Moreover, it is currently unknown how women and men vary in executive control as a function of biological and genetic differences across the adult lifespan. The demonstration of such differences due to biological sex may help to clarify prior conflicting findings in the aging literature on this point.

Of relevance here, recent work suggests that certain lifestyle factors (e.g., education level, socioeconomic status) might differentially modulate executive control performance over the adult lifespan (Mungas et al., 2010; Scarmeas & Stern, 2003; Stern, 2012). Given that the number of older adults will markedly increase over the next several decades (Alzheimer's Association, 2015; Beard, Officer & Cassels, 2015), it is crucial that we rigorously evaluate which potential lifestyle factors may promote successful aging so that we can optimize the

Address for correspondence:
Debra Titone, Ph.D., Department of Psychology, McGill University, 1205 Dr. Penfield Ave., Montreal, Quebec, H3A1B1, Canada
*dtitone@psych.mcgill.ca*

neurocognitive health of this population (Davis, Marra, Najafzadeh & Liu-Ambrose, 2010; Forte, Boreham, de Vito & Pesce, 2015; Vaughan & Giovanello, 2010).

One lifestyle factor that has received great attention in recent years is BILINGUALISM – that is, the knowledge, acquisition and regular use of two or more languages in daily life (Bialystok, Abutalebi, Bak, Burke & Kroll, 2016; Bialystok, Craik, Klein & Viswanathan, 2004; Coderre, Smith, van Heuven & Horwitz, 2016; Costa & Sebastián-Gallés, 2014; Kroll, Bobb & Hoshino, 2014; Kroll & Bialystok, 2013). Bilingualism is thought to protect against age-related executive control declines because people who regularly speak more than one language face unique cognitive challenges compared to people who speak only one language (i.e., monolinguals). For example, bilinguals must monitor and select the appropriate language to use within a given context. This may involve appropriately suppressing the impulse to use English in one's French-speaking workplace, but doing the reverse when conversing with exclusively Anglophone friends (Abutalebi, Tettamanti & Perani, 2009; Green, 1998; Guo, Liu, Misra & Kroll, 2011; Kroll, Bobb, Misra & Guo, 2008; Meuter & Allport, 1999; Misra, Guo, Bobb & Kroll, 2012; Pivneva, Palmer & Titone, 2012; von Studnitz & Green, 2002). Bilinguals must also suppress activation of specific sound patterns, words, or meanings from other known languages when speaking or reading (Blumenfeld & Marian, 2011; Christoffels, Firk & Schiller, 2007; Dijkstra, 2005; Green, 2011; Guo et al., 2011; Kroll et al., 2008; Macizo, Bajo & Martin, 2010; Martin, Macizo & Bajo, 2010; Mercier, Pivneva & Titone, 2014; Misra et al., 2012; Pivneva, Mercier & Titone, 2014). For example, by inhibiting the French meaning of *chat* when reading that word in an English context (Libben & Titone, 2009; Pivneva et al., 2014; Titone, Libben, Mercier, Whitford & Pivneva, 2011). Consequently, the need to routinely exercise executive control to suppress knowledge of a whole language, or to suppress specific representations within a particular language, is hypothesized to enhance executive control capacity over the lifespan (e.g., see reviews by Baum & Titone, 2014; Titone, Gullifer, Subramaniapillai, Rajah & Baum, 2017).

Enhanced executive control arising from greater bilingual experience could also contribute to greater COGNITIVE RESERVE in older adults (Barulli & Stern, 2013). Cognitive reserve refers to how individual differences in environment, genetics, and life experiences impact how well one's brain can qualitatively cope with age-related changes in neurocognitive function (Stern, 2009). Individuals with high cognitive reserve are able to maintain higher levels of cognitive function than would be predicted from their current measures of brain and/or neural reserve. Thus, having high cognitive reserve may slow the progression of both normal age-related cognitive decline and pathological aging (i.e., dementia) (Calvo, García, Manoiloff & Ibáñez, 2016; Scarmeas & Stern, 2003; Stern, 2002, 2009, 2012; Tucker & Stern, 2011). However, empirical support has been mixed for the idea that bilingualism may be a proxy measure of cognitive reserve: that it may help to mitigate the effects of age-related decline in domain-general executive control. Some studies offer clear support for a link between bilingualism and improved executive control in older adults (Bialystok et al., 2004, 2008; Gold, Kim, Johnson, Kryscio & Smith, 2013). Yet, other studies have obtained mixed support, in that bilingual effects occur for some executive function measures with age but not others (Ansaldo, Ghazi-Saidi & Adrover-Roig, 2015; Grady, Luk, Craik & Bialystok, 2015; Kousaie, Sheppard, Lemieux, Monetta & Taler, 2014; Zahodne, Schofield, Farrell, Stern & Manly, 2014), or have failed to support this view (de Bruin, Bak & Della Sala, 2015; Gathercole, Thomas, Kennedy, Prys, Young, Guasch, Roberts, Hughes & Jones, 2014; Kirk, Fiala, Scott-Brown & Kempe, 2014; Kousaie & Phillips, 2012).

Of relevance here, the emerging story may not be clear for a variety of potential reasons. These include: 1) a tendency in the literature to investigate heterogeneous dimensions of bilingual experience across studies (e.g., L2 AoA in one study, L2 proficiency in another); 2) the use of categorical age-group comparisons rather than a continuous lifespan approach; 3) the use of speeded executive control tasks that may be specifically problematic for older adults; and 4) not accounting for other variables that could modulate the impact of bilingual experience on executive control performance over the lifespan, such as biological sex. Thus, in an attempt to address these limitations, we investigate whether different kinds of bilingual experience modulate executive control performance across the adult lifespan (i.e., including younger, middle-aged, and older adults), as a function of biological sex, using a non-speeded executive control task (i.e., the Wisconsin Card Sorting Test).

The Wisconsin Card Sorting Test (WCST) is a well-established measure of executive control, in which participants are instructed to match a target card to one of four reference cards according to the colour, number, or shape of the stimuli on the cards (e.g., Miyake et al., 2000; Nyhus & Barceló, 2009). Participants must infer the operative sorting rule based on feedback given after each trial as to whether their match was correct or incorrect. After 10 consecutive successful matches (i.e., completion of one category), the sorting rule is changed without notice and participants must dynamically adjust to a new sorting classification. Accordingly, the WCST deconstructs participant performance into different measures that are thought to reflect different facets of executive control. Here, we focus on the following WCST measures that are sensitive to age-related decline, which are presumed to be enhanced by greater bilingual experience: the number of Perseverative Errors, the number of Non-Perseverative Errors, and the number of Categories Completed. Below we discuss each of the measures and the executive control processes they presumably reflect.

Perseverative Errors arise when participants fail to shift from a previously relevant sorting rule to a new one following an implicit rule change. Thus, these errors reflect a failure to use executive control to inhibit use of a no longer operative sorting rule in order to select from a new sorting category (e.g., when there is a switch in whether participants were expected to sort on the basis of colour, shape, or number). An age-related increase in Perseverative Errors is a robust finding in the aging literature, with this measure repeatedly linked with issues in inhibitory control (e.g., Head, Kennedy, Rodrigue & Raz, 2009; Fristoe, Salthouse & Woodard, 1997; Hartman, Bolton & Fehnel, 2001). With respect to bilingualism, older adults with greater bilingual experience are hypothesized to exercise greater executive control through constant practice inhibiting a non-target language while using a target language (e.g., Bialystok, Craik & Luk, 2008). Thus, we would expect people with greater bilingual experience to make fewer Perseverative Errors compared to people with less bilingual experience.

In contrast, Non-Perseverative Errors arise when a participant makes an error within a sorting category, when there was no implicit rule change. Thus, Non-Perseverative Errors reflect a failure to maintain attention within the same category in order to continue applying the same rule to subsequent trials within the same category. This might especially occur if people become distracted and are unable to inhibit interference from co-existing perceptual stimuli (e.g., Barceló, 1999; 2001). With respect to bilingualism, older adults with greater bilingual experience presumably pay greater attention in order to use a target language appropriately, and to switch when necessary, thus potentially exercising attentional control more rigorously than people with less bilingual experience, or even monolinguals who only use one language (e.g., Zhou & Krott, 2016; Tao, Marzecová, Taft, Asanowicz & Wodniecka, 2011).

Finally, the number of Categories Completed, which in this test could range between 0 and 9, increases when a person successfully makes 10 consecutive correct responses for a particular card sorting rule (i.e., colour, shape, or number). This measure is thus a reflection of global performance, where a low number of Categories Completed would reflect both an inability to maintain attention within a category AND a failure to shift set (e.g., Barcelo, 1999; 2001). To the extent that older adults with greater bilingual experience use executive control to more efficiently process a target and non-target language, bilinguals may show a greater number of Categories Completed.

Unlike the large number of studies using speeded executive control tasks to investigate the impact of bilingual language experience (e.g., Bialystok et al., 2004; Bialystok, Martin & Viswanathan, 2005; de Bruin et al., 2015; Antón, García, Carreiras & Duñabeitia, 2016), only two studies concerning older adults have made use of the non-speeded WCST, and both failed to find any positive impact of greater bilingual language experience (Kousaie et al., 2014; Gathercole et al., 2014). In one study, Kousaie and colleagues tested monolingual Francophones (30 young and 30 older adults), monolingual Anglophones (40 young and 31 older adults) and French–English bilinguals (51 young and 36 older adults), who were non-immigrants living in Ottawa or Quebec (Kousaie et al., 2014). In contrast to predictions of the bilingual advantage hypothesis, they found that monolingual Francophones achieved more Categories Completed than monolingual Anglophones and bilinguals. However, this study did not report results for other WCST measures (i.e., Perseverative and Non-Perseverative Errors), which may be more sensitive to the effects of aging (e.g., Rhodes, 2004).

In another study, Gathercole et al. (2014) used a card sorting task, similar to the WCST, to test how bilingualism influenced executive function in a lifespan sample of English–Welsh bilinguals in Wales (age 3 to older adults). However, their card sorting task involved explicit, experimenter-induced rule changes rather than implicit rule changes used in the WCST (Gathercole et al., 2014). The results showed a few significant bilingual effects on executive function across groups; however, the overall pattern of results were mixed, and several methodological constraints make it difficult to draw definitive conclusions. These include: the use of slightly different tasks over the different age ranges; the exclusive focus on difference scores for accuracy and RT instead of raw scores when assessing global performance; only assessing overall accuracy and RT rather than more specific performance measures (i.e., Perseverative, Non-Perseverative Errors, etc.); and finally, binning a large sample into seven smaller-N groups crossed with four language groups, ANOVA style, which may have reduced their ability to detect subtle effects.

In contrast with the ambiguous impact of bilingual effects on WCST performance across the lifespan, age-related decline on WCST performance has been consistently demonstrated (Axelrod & Henry, 1992; Daigneault, Braun & Whitaker, 1992; Fristoe et al., 1997; Kousaie et al., 2014; Rhodes, 2004). Moreover, an additional factor that has not been systematically addressed with respect to the bilingualism hypothesis, to our knowledge, is biological sex. While the literature on this topic is also variable (Geary, Saults, Liu & Hoard, 2000; Hyde, 1981; Li Zhang, Duann, Yan, Sinha & Mazure, 2009; Roivainen, 2011), some work supports the idea that biological sex modulates cognitive and executive control performance over the lifespan. For example, several studies have shown that women outperform men on tasks assessing memory, reasoning and verbal ability (Duff & Hampson, 2001; Finkel, Reynolds, McArdle, Gatz & Pedersen, 2003; Maitland, Herlitz, Nyberg, Bäckman & Nilsson, 2004; McCarrey, An, Kitner-Triolo, Ferrucci & Resnick, 2016), whereas men outperform women on tasks assessing spatial and visuospatial ability, inhibition, and task-switching (Bieri, Bradburn & Galinsky, 1958; Finkel et al., 2003; Karlsson, Thorvaldsson, Skoog, Gudmundsson & Johansson, 2015; Reimers & Maylor, 2005; Tun & Lachman, 2009; Voyer, Voyer & Bryden, 1995).

Prior work investigating the role of bilingualism on executive control tends to use non-verbal tasks, and thus, may favour greater performance in men vs. women (Clayson, Clawson & Larson, 2011; Colzato, Hertsig, van den Wildenberg & Hommel, 2010; Evans & Hampson, 2015; Halari, Hines, Kumari, Mehrotra, Wheeler, Ng & Sharma, 2005; Halari & Kumari, 2005; Stoet, 2010). Moreover, hormonal changes at puberty and in middle-age (i.e., menopause) may differentially influence the cognitive aging trajectory of women, compared to men (Janicki & Schupf, 2010; Keenan, Ezzat, Ginsburg & Moore, 2001; Shanmugan & Epperson, 2014). Thus, biological differences in men and women may impact executive control performance, in a manner that could also interact with other factors, such as bilingualism. If true, it would be difficult to interpret past findings about bilingualism and executive control as being generally applicable to all people, or more or less specific to men vs. women.

In the specific case of the WCST, a non-verbal task, there is no current consensus on the role of biological sex on task performance, which can be partly explained by the scarcity of the studies that have explicitly attempted to address this question (Ferland, Ramsay, Engeland & O'Hara, 1998; Yeudall, Fromm, Reddon & Stefanyk, 1986; Boone, Ghaffarian, Lesser, Hill-Gutierrez & Berman, 1993). In one study, Ferland et al. administered the WCST to a sample of young adults with traumatic brain injury and healthy controls, and found in their sub-analysis of healthy controls ($M$ = 20.82 years, $SD$ = 3.44) that men outperformed women, making fewer Perseverative Errors and responses (Ferland et al., 1998). In another study of individuals aged 15 to 40 years of age, Yeudall et al. (1986) found no sex differences in total number of errors (i.e., the sum of Perseverative, Non-Perseverative, and unique errors) between women and men. Finally, using a middle-aged and older adult sample (45 to 83 years of age), Boone et al. found that women completed more categories and had fewer Perseverative Errors compared to men (Boone et al., 1993). Overall, the results to date suggest that sex differences in WCST performance may be inconsistent across the lifespan, such that women and men might perform better or worse than one another depending on the particular period of life (i.e., young adulthood, midlife, older age).

### *The present study*

To summarize thus far, a growing body of work suggests that bilingual language experience modulates age-related declines in executive control, presumably by increasing cognitive reserve among older adults. However, the findings across studies are mixed, likely due to cross-study variability in individual differences among bilinguals, heterogeneity across tasks, the failure to take a lifespan approach, and the heretofore unstudied factor of biological sex. Thus, we investigate whether biological sex and individual differences in bilingual language experience relate to WCST performance in an adult lifespan sample.

Specifically, this study was a secondary analysis of a larger investigation of episodic memory function across the adult lifespan (e.g., Rajah, Wallace, Ankudowich, Yu, Swierkot, Patel, Chakravarty Naumova, Pruessner, Joober, Gauthier & Pasvanis, 2017; Ankudowich, Pasvanis & Rajah, 2017). Our goal was to investigate the potential impact of three different continuous measures of bilingual experience that were available to us in this dataset (no other language background measures were possible to derive that would enhance our ability to test the hypothesis). These measures included: the age of second language (L2) acquisition, the number of languages known, and the percentage of non-native (i.e., non-L1) language usage. Because the overwhelming majority of people in Montreal are bilingual or multilingual, we cannot rigorously investigate the impact of group differences between bilinguals/multilinguals vs. monolinguals. However, as previously argued (e.g., Baum & Titone, 2014; Titone & Baum, 2014; Titone et al., 2017), group comparisons of monolinguals vs. bilinguals presume that within group variability is negligible, an assumption that is highly likely to be untrue. Thus, here we investigated how the full continuum of bilingual experience (i.e., bilinguals with greater or lesser bilingual experience) impacts executive control performance, which may allow us to understand the impact of bilingual language experience in a more nuanced manner.

To this end, our specific predictions were:

1. Given sex differences in executive control processes (that presumably underlie the WCST), and changes specific to midlife that may affect cognitive performance of women to a greater extent than men, we predicted greater decline in WCST performance with aging in women compared to men.

2. Given past work suggesting that bilingual language experience ameliorates age-related cognitive decline,

we predicted that greater language experience (i.e., earlier age of L2 acquisition, greater non-native language usage, and greater number of languages known) will have a beneficial effect on WCST performance measures, particularly for women who may have more room to improve given prediction #1.

## Methods

### *Participants*

Participants were 152 bilingual and multilingual adults (98 women, 54 men; age range 19–76 years; $M_{age} = 48.23$, $SE_{age} = 1.31$) who had no self-reported history of psychiatric illness, neurological disorders, or substance abuse. Although we analyzed age continuously, a total of 41 participants could be classed as younger adults (19–35 years of age; $M_{age} = 25.59$, $SE_{age} = .61$), 66 as middle aged adults (40–58 years of age; $M_{age} = 49.68$, $SE_{age} = .66$), and 45 as older adults (60–76 years of age; $M_{age} = 66.73$, $SE_{age} = .59$). In addition to examining the effects of age continuously in all analyses, we also controlled for individual differences in participants' total number of years of education, a variable that served as a proxy of socioeconomic status and cognitive reserve. Participants were recruited through a variety of means, including advertisements (i.e., newspapers, magazines, etc.) and community and social engagement (i.e., TV and radio interviews), as part of a larger study investigating the role of episodic memory across the adult lifespan (Ankudowich, Pasvanis & Rajah, 2016; Ankudowich et al., 2017).

All participants had at least a high school education ($M_{education} = 15.68$, $SE_{education} = .16$), and were right-handed as measured by the Edinburgh Inventory for Handedness (Oldfield, 1971). In order to be eligible to participate in the study, all participants had to meet the eligibility criteria for the Mini-International Neuropsychiatric Interview (MINI; inclusion cutoff $\leq 2$); the Folstein Mini Mental State Examination (MMSE; exclusion cutoff $< 27$); and the Beck Depression Inventory (BDI; exclusion cutoff $< 15$). In addition, participants self-reported as having no history of diabetes, cardiovascular disease, neurological or psychological illness. All eligible participants performed a series of neuropsychological tests, in addition to the WCST, as part of the larger fMRI study (see Table 1).

All participants completed the Language and Social Background Questionnaire (LSBQ; Luk & Bialystok, 2013). Based on this measure, we determined that our sample contained 109 bilinguals, 39 trilinguals, and 4 multilinguals (knowing four languages). Participants spoke a variety of languages, including: Tamil, Spanish, Hungarian, Bulgarian, Greek, Polish, Arabic, Chinese, Italian, Gujarati, Creole, German, and Swedish; however, a large percentage were either English–French or French–English bilingual (68%). Thus, the first language (L1) was identified as English for 45 participants, French for 87 participants, and another language for 20 participants. The second language (L2) was identified as English for 96 participants, French for 50 participants, and another language for 6 participants. The age of L2 acquisition in the sample ranged from 0 to 22 years. Table 1 summarizes participant demographic information including bilingual characteristics of our sample.

### *Procedure*

Participants first signed a consent form approved by the ethics board of the Faculty of Medicine, McGill University. Participants then completed the WCST as part of a battery of neuropsychological tests. We specifically chose the WCST from this larger battery as it was the only task administered that would tap into executive control in a non-linguistic manner. Participants were given a computerized version of the WCST (Mueller & Piper, 2014) using the Psychology Experiment Building Language (PEBL) Version 0.13 (retrieved from http://pebl.sourceforge.net). The PEBL implementation of the WCST is cited as the Berg's Card Sorting Test, but this manuscript will use the more conventional and familiar task name (i.e., WCST). Each trial required participants to match a card to one of four cards (with no prior rule to match). For example, a participant might be given a card with four yellow circles that they must match to one of four cards that vary in terms of shape, colour, and number (e.g., one red triangle, two green stars, three yellow plus signs, four blue circles). Since the participants are given no rules to match, logically, they can match their card based on shape (i.e., circle), colour (i.e., yellow), or the number of shapes present (i.e., four). Thus, in this case, there are two possible cards they could reasonably pick (i.e., three yellow plus signs or four blue circles). Feedback is given after every trial by informing participants whether the match was correct or incorrect. If the correct way to match is by colour (i.e., yellow), then the participant must apply that rule for 10 consecutive trials to achieve a category, after which the rule changes (i.e., to shape or number) unbeknownst to the participant. Based on the negative feedback, participants must learn that the previously used rule to match no longer works and must determine and apply a new rule to match. Categories (i.e., rules) varied by colour, number, and form. The task is completed when the participant successfully sorts nine categories (three categories repeated three times) or alternatively, progresses through 128 card sort trials.

The dependent variables of interest were the number of Perseverative Errors, the number of Non-Perseverative Errors, and the number of Categories Completed. As previously discussed, Perseverative Errors are the total

Table 1. *Mean Background Measures (and Standard Errors) by Age Group.*

| | Younger Adults (YA) | | | Middle-Aged Adults (MA) | | | Older Adults (OA) | | |
|---|---|---|---|---|---|---|---|---|---|
| | YA Full Sample (n = 41) | Females (n = 25) | Males (n = 16) | MA Full Sample (n = 66) | Females (n = 45) | Males (n = 21) | OA Full Sample (n = 45) | Females (n = 28) | Males (n = 17) |
| Age (Years) | 25.59 (.61) | 25.44(0.76) | 25.81(1.05) | 49.68 (.66) | 50.38(0.76) | 48.19(1.25) | 66.73 (.59) | 66.68(0.75) | 66.82(0.98) |
| Education (Years) | 15.76 (.29) | 15.64(0.38) | 15.94(0.43) | 15.68 (.24) | 15.64(0.28) | 15.76(0.46) | 15.60 (.34) | 15.18(0.40) | 16.29(0.60) |
| L2 AoA (Years) | 5.90 (.58) | 6.04(0.76) | 5.69(0.93) | 8.26 (.59) | 7.58(0.66) | 9.71(1.13) | 9.04 (.68) | 8.57(0.83) | 9.82(1.20) |
| Non-Native Language Usage (%) | 36.48 (5.56) | 29.59(6.61) | 47.24(9.47) | 27.50 (3.61) | 29.36(4.74) | 23.49(5.08) | 19.04 (3.93) | 23.68(6.01) | 11.41(2.48) |
| Number of Languages Known | 2.37 (.08) | 2.28(0.09) | 2.5(0.13) | 2.27 (.06) | 2.33(0.08) | 2.14(0.08) | 2.31 (.09) | 2.36(0.12) | 2.24(0.14) |
| CVLT – LFR$^{*\%}$ | 13.61(.32) | 14.32(0.34) | 12.5(0.53) | 12.48(.33) | 13(0.33) | 11.38(0.71) | 12.22(.39) | 12.68(0.52) | 11.47(0.56) |
| CVLT – LCR$^{*\%}$ | 13.61(.32) | 14.32(0.35) | 12.5(0.51) | 12.79(.30) | 13.2(0.33) | 11.9(0.61) | 12.76(.34) | 13.14(0.43) | 12.12(0.52) |
| CVLT – RG | 15.27(0.16) | 15.40(0.15) | 15.06(0.34) | 14.88(0.18) | 14.91(0.22) | 14.81(0.33) | 15.11(0.14) | 15.14(0.19) | 15.06(0.18) |
| DKEFS – LF | 45.05(1.59) | 45.40(2.11) | 44.5(2.49) | 43.62(1.50) | 43.13(1.66) | 44.67(3.15) | 44.11(1.79) | 44.14(2.19) | 44.06(3.17) |
| DKEFS – CF | 43.49(1.38) | 43.92(1.84) | 42.81(2.13) | 41.97(1.14) | 42.00(1.36) | 41.9(2.14) | 42.96(1.01) | 43.57(1.27) | 41.94(1.67) |
| DKEFS – CS$^{*}$ | 15.8(0.48) | 16.08(0.64) | 15.38(0.72) | 15.53(0.37) | 16.27(0.41) | 13.95(0.67) | 15.58(0.37) | 15.82(0.45) | 15.18(0.64) |

Note: This table presents the group means and standard errors (SE) for the demographic, bilingual, and neuropsychological measures. CVLT = California Verbal Learning Task: LFR = Long Free Recall; LCR = Long Cued Recall; RG = Recognition; DKEFS = Delis-Kaplan Executive Function System: LF = Letter Fluency; CF = Category Fluency; CS = Category Switching. $^{*}$Denotes a significant effect of Sex ($p<0.05$). $^{\%}$Denotes a significant effect of Age ($p<0.05$).

number of errors made following a rule change (e.g., errors made immediately after an implicit rule change from colour to shape). Thus, this measure is hypothesized to assess successful task switching, as a low number of Perseverative Errors would indicate that participants were better able to inhibit a previous sorting rule that was no longer operative following an implicit rule change. Non-Perseverative Errors are the total number of errors made after a rule change that are not Perseverative Errors (e.g., the participant matches the card based on shape when they should be sorting based on colour and they have not received negative feedback to prompt them to try another card sorting rule). Thus, this type of error reflects a failure to maintain attention within the same category in order to continue applying a learned rule to subsequent trials within the same category. The number of Categories Completed can range from 0 to 9, where one category is achieved when a participant correctly responds to ten consecutive trials, after which the rule to match is changed without the participant's knowledge. This WCST measure reflects overall global performance (i.e., attention, set-shifting), where the greater number of Categories Completed corresponds to better performance on the WCST.

## Results

We performed a series of multiple linear regressions using robust regression with maximum-likelihood estimation. These regression models were implemented using iterated re-weighted least squares (IRLS) with Huber weights using the rlm of the MASS package, version 7.3-45 (Venables & Ripley, 2002) in R (R Development Core Team, 2016). Robust regression allows us to use all the observations present in the data, but attenuates the effect of large residuals (i.e., outliers or influential points). This allows us to include outlier participants (i.e., older participants) that deviate from the norm. An analysis of regression diagnostics showed homoscedasticity of the residuals and no outlier points having a large Cook's distance, demonstrating that model assumptions were met.

Regression models were constructed to assess the impact of specific measures of bilingualism (discussed below), Age, Sex, and their interaction, on the following critical WCST measures: Perseverative Errors, Non-Perseverative Errors, and Categories Completed. Our three bilingual language variables of interest were: MODEL 1 – L2 age of acquisition (L2 AoA, that is, the earliest age at which participants started to learn their L2 language either at home or school); MODEL 2 – the Number of Languages Known; and MODEL 3 – the percentage of Non-native Language Usage (calculated as 100% minus the percentage of L1 usage). Specifically, L1 usage was calculated as the percentage of the average time spent using the native language at home (i.e., speaking, listening, reading, writing, watching TV, listening to the radio) and at school/work (i.e., speaking, listening, reading, writing). Across models we tested the three-way interaction between Age, Sex and each bilingual language experience variable individually, while statistically controlling for the number of years of education and the two other language variables not included as part of the three-way interaction. Thus, in terms of R syntax, the specific models fitted were: Model 1: DV ~ Age*Sex*L2 AoA + Years of Education + Languages Known + Non-Native Language Usage; Model 2: DV ~ Age*Sex*Languages Known + Years of Education + L2 AoA + Non-Native Language Usage; Model 3: DV ~ Age*Sex*Non-Native Language Usage + Years of Education + Languages Known + L2 AoA.

All fixed effects variables were treated as continuous, with the exception of Sex, which was treated as categorical through deviation coding (-0.5, 0.5). All continuous fixed effects variables were standardized using a Z-score transformation, thus permitting comparisons in the effect size of the model regression coefficients. Tables 2–4 shows a summary of the regression output results for Models 1–3, respectively. Figures 1–4 present the partial effects plots of the results.

### *The number of perseverative errors*

Model 1, which examined the interaction of Age*Sex*L2 AoA, showed a significant Age*Sex interaction ($\beta = 3.13$, $SE = 1.28$, $t = 2.43$, $p = .01$), indicating that women made a greater number of Perseverative Errors than men beginning at midlife (see Figure 1). There was no main effect of L2 AoA on the number of Perseverative Errors, nor was there a significant interaction between L2 AoA and other independent variables.

Model 2, which examined the relationship between Age, Sex and Number of Languages Known on number of Perseverative Errors, also identified a significant Age*Sex interaction ($\beta = 3.20$, $SE = 1.19$, $t = 2.69$, $p = .01$). This interaction showed that women performed worse than men from midlife to older age. With respect to the Number of Languages Known, there were no significant main effects or interactions, similar to L2 AoA above.

Model 3, which examined the association between Age, Sex and Non-native Language Usage on number of Perseverative Errors, identified a significant interaction between Non-native Language Usage*Sex ($\beta = -3.26$, $SE = 1.64$, $t = -1.99$, $p < .05$), indicating that women who had a higher degree of Non-native Language Usage made fewer Perseverative Errors than men who had a lower degree of Non-native Language Usage (see Figure 2). Further, when sub-analyses were conducted separately for women ($n = 98$), there was a significant effect of Age ($\beta = 3.21$, $SE = .73$, $t = 4.41$, $p < .01$), a marginally significant effect of Non-native Language

Table 2. *Effect sizes ($\beta$), standard errors (SE), t values for Model 1, which examines the interaction between Age, Age of L2 Acquisition, and Sex.*

| | Perseverative Errors (PE) | | | Non-Perseverative Errors (NPE) | | | Categories Completed (CC) | | |
|---|---|---|---|---|---|---|---|---|---|
| Fixed effects | *β* | *SE* | *t value* | *β* | *SE* | *t value* | *β* | *SE* | *t value* |
| –Age | 1.52 | .63 | 2.39* | 2.00 | .53 | 3.75*** | –.76 | .18 | –4.15*** |
| –L2 AoA | –.02 | .62 | –.04 | .01 | .52 | .02 | .06 | .18 | .31 |
| –Sex | 2.50 | 1.26 | 1.99* | 1.89 | 1.06 | 1.78 | –.93 | .36 | –2.57* |
| –Age * L2 AoA | –.33 | .60 | –.54 | –.02 | .51 | –.05 | .02 | .17 | .09 |
| –Age * Sex | 3.13 | 1.28 | 2.43* | 1.27 | 1.08 | 1.17 | –1.15 | .37 | –3.11** |
| –L2 AoA * Sex | .12 | 1.27 | .09 | 3.78 | 1.07 | 3.54*** | –.62 | .37 | –1.68 |
| –Age * L2 AoA * Sex | .50 | 1.17 | .43 | 2.06 | .99 | 2.08* | –.47 | .34 | –1.39 |
| Control Predictors | *β* | *SE* | *t value* | *β* | *SE* | *t value* | *β* | *SE* | *t value* |
| Languages Known | –.65 | .66 | –.97 | –1.33 | .56 | –2.39** | .37 | .19 | 1.94 |
| Non-Native Usage | –.79 | .68 | –1.16 | .78 | .58 | 1.35 | –.11 | .20 | –.57 |
| Education Level | .06 | .59 | .10 | –.10 | .50 | –.19 | .34 | .17 | 2.03* |
| (Intercept) | 16.46 | .63 | 26.26*** | 10.08 | .53 | 19.11*** | 7.22 | .18 | 40.00*** |

* *p or Pr (>|z|) < .05;* ** *p or Pr (>|z|) < .01;* *** *p or Pr (>|z|) < .001*

Table 3. *Effect sizes ($\beta$), standard errors (SE), t values for Model 2, which examines the interaction between Age, Number of Languages Known, and Sex.*

| | Perseverative Errors (PE) | | | Non-Perseverative Errors (NPE) | | | Categories Completed (CC) | | |
|---|---|---|---|---|---|---|---|---|---|
| Fixed effects | *β* | *SE* | *t value* | *β* | *SE* | *t value* | *β* | *SE* | *t value* |
| –Age | 1.54 | .61 | 2.51* | 1.55 | .56 | 2.76** | –.69 | .19 | –3.59*** |
| – Languages Known | –.69 | .68 | –1.02 | –1.05 | .62 | –1.69 | .33 | .21 | 1.54 |
| –Sex | 2.50 | 1.21 | 2.07* | 3.03 | 1.11 | 2.72** | –1.05 | .38 | –2.74** |
| –Age * Languages Known | .48 | .62 | .77 | .28 | .57 | .50 | –.09 | .19 | –.47 |
| –Age * Sex | 3.20 | 1.19 | 2.69* | 2.50 | 1.09 | 2.29* | –1.29 | .38 | –3.45*** |
| –Languages Known * Sex | .03 | 1.26 | .03 | .15 | 1.16 | .13 | .01 | .40 | .03 |
| –Age * Languages Known * Sex | –.30 | 1.24 | –.24 | .23 | 1.14 | .20 | .20 | .39 | .52 |
| Control Predictors | *β* | *SE* | *t value* | *β* | *SE* | *t value* | *β* | *SE* | *t value* |
| L2 AoA | –.12 | .59 | –.20 | .04 | .55 | .08 | .06 | .19 | .34 |
| Non-Native Usage | –.75 | .65 | –1.15 | .59 | .60 | .99 | –.11 | .20 | –.55 |
| Education Level | .14 | .57 | .24 | –.40 | .53 | –.76 | .41 | .18 | 2.23* |
| (Intercept) | 16.43 | .60 | 27.56*** | 9.66 | .55 | 17.65*** | 7.23 | .19 | 38.41*** |

* *p or Pr (>|z|) < .05;* ** *p or Pr (>|z|) < .01;* *** *p or Pr (>|z|) < .001*

Usage ($\beta$ = -1.51, *SE* = .79, *t* = -1.90, *p* = .06), and overall significance in model fit, $F(6, 91) = 4.45, p < .01$. However, when sub-analyses were conducted separately for men, there were no significant predictors, and the overall model was not significant, $F(6,47) = .61, p = .72$.

Thus, to summarize the results for the number of Perseverative Errors: we observed an Age*Sex interaction such that women made more errors than men at midlife and older age. In terms of bilingual language experience, there was no significant effect of L2 AoA or Number of Languages Known on the number of Perseverative Errors; however, there was an interaction between Non-Native Language Usage*Sex, suggesting that women with greater Non-Native Language Usage made fewer Perseverative Errors than women with lower Non-Native Language Usage.

### *The number of non-perseverative errors*

Model 1, which examined the interaction of Age*Sex*L2 AoA, showed a significant three-way interaction between Age, L2 AoA, and Sex ($\beta$ = 2.06, *SE* = .99, *t* = 2.08,

Table 4. *Effect sizes ($\beta$), standard errors (SE), t values for Model 3, which examines the interaction between Age, Non-Native Usage, and Sex.*

| | Perseverative Errors (PE) | | | Non-Perseverative Errors (NPE) | | | Categories Completed (CC) | | |
|---|---|---|---|---|---|---|---|---|---|
| Fixed effects | $\beta$ | *SE* | *t value* | $\beta$ | *SE* | *t value* | $\beta$ | *SE* | *t value* |
| –Age | 2.08 | .63 | 3.28$^{**}$ | 1.48 | .61 | 2.42$^{*}$ | −.79 | .20 | −4.01$^{***}$ |
| –Non-Native Usage | .27 | .86 | .32 | .53 | .84 | .64 | −.29 | .27 | −1.06 |
| –Sex | 1.86 | 1.27 | 1.46 | 3.05 | 1.23 | 2.48$^{*}$ | −.99 | .40 | −2.50$^{*}$ |
| –Age * Non-Native Usage | .39 | .70 | .55 | .14 | .68 | .21 | −.09 | .22 | −.40 |
| –Age * Sex | 2.25 | 1.25 | 1.81 | 2.62 | 1.20 | 2.18$^{*}$ | −1.10 | .39 | −2.84$^{**}$ |
| –Non-Native Usage * Sex | −3.26 | 1.64 | −1.99$^{*}$ | .57 | 1.59 | .36 | .66 | .51 | 1.29 |
| –Age * Non-Native Usage * Sex | −1.93 | 1.43 | −1.35 | −.02 | 1.38 | −.02 | .37 | .44 | .84 |
| Control Predictors | $\beta$ | *SE* | *t value* | $\beta$ | *SE* | *t value* | $\beta$ | *SE* | *t value* |
| Languages Known | −.70 | .61 | −1.15 | −1.00 | .59 | −1.71 | .32 | .19 | 1.67 |
| L2 AoA | −.24 | .58 | −.41 | .15 | .56 | .27 | .05 | .18 | .30 |
| Education Level | −.19 | .56 | −.33 | −.41 | .55 | −.75 | .46 | .18 | 2.63$^{**}$ |
| (Intercept) | 16.83 | .63 | 26.82$^{***}$ | 9.71 | .61 | 16.01$^{***}$ | 7.22 | .19 | 37.07$^{***}$ |

* *p or Pr (>|z|) < .05 ; ** p or Pr (>|z|) < .01; *** p or Pr (>|z|) < .001*

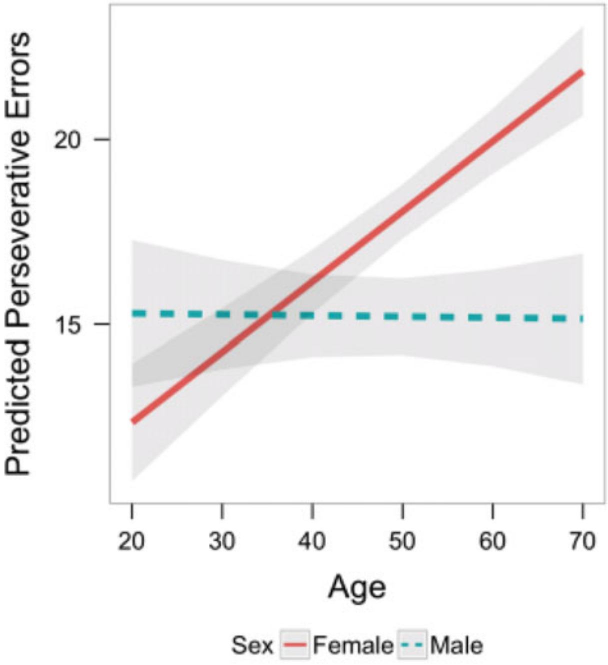

Figure 1. (Colour online) Partial effects plot demonstrating the Age x Sex interaction on the number of Perseverative Errors (Model 1).

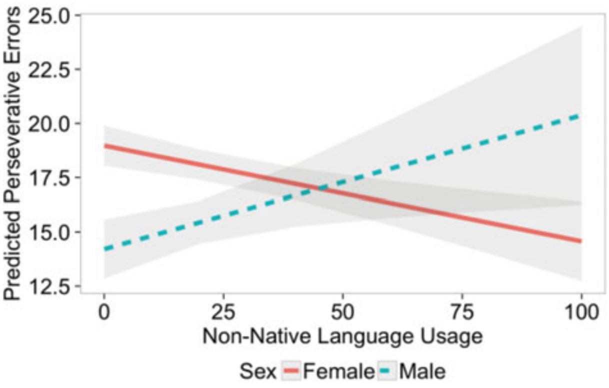

Figure 2. (Colour online) Partial effects plot demonstrating the Non-native Language Usage x Sex interaction on the number of Perseverative Errors (Model 3).

$p$ = .04), indicating that women made more Non-Perseverative Errors than men overall, although earlier L2 AoA among women was associated with fewer Non-Perseverative Errors (see Figure 3). Further, when sub-analyses were conducted on women alone, there were main effects of age ($\beta = 2.72$, $SE = .74$, $t = 3.70$, $p < .01$) in addition to L2 AoA ($\beta = 1.92$, $SE = .79$, $t = 2.42$, $p = .02$), and the overall model remained significant, $F(6, 91) = 4.35$, $p < .01$. Conversely, when sub-analyses were separately conducted for men alone, only a main effect of L2 AoA ($\beta = -1.59$, $SE = .64$, $t = -2.49$, $p = .02$) remained, although the overall model was marginally significant, $F(6, 47) = 1.95$, $p = .09$. This main effect among men indicated that earlier L2 AoA among men did not reduce Non-Perseverative Errors, but rather was associated with greater Non-Perseverative Errors.

Model 2, which examined the interaction of Age*Sex*the Number of Languages Known, also showed a significant Age*Sex interaction ($\beta = 2.50$, $SE = 1.09$, $t = 2.29$, $p = .02$). However, there were no significant

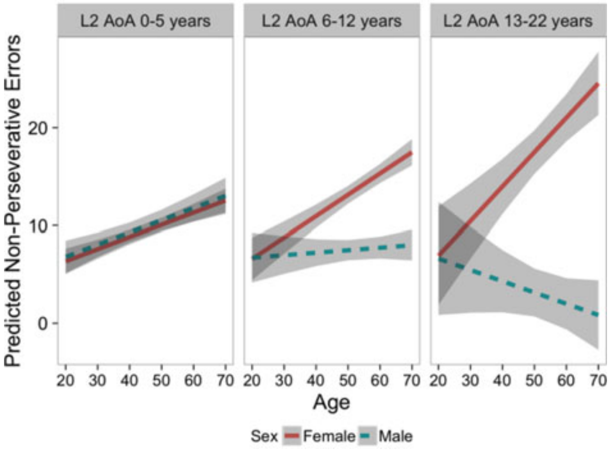

Figure 3. (Colour online) Partial effects plot demonstrating the three-way interaction between Age, Sex, and Age of L2 Acquisition on the number of Non-Perseverative Errors (Model 1).

main effects or interactions with respect to the Number of Languages Known.

Model 3, which examined the interaction of Age*Sex*Non-native Language Usage, also showed an Age*Sex interaction ($\beta$ = 2.62, *SE* = 1.20, *t* = 2.18, *p* = .03); however, there were no significant main effects or interactions involving Non-native Language Usage.

Thus, to summarize the results for Non-Perseverative Errors, we again found an interaction between Age and Sex such that women performed worse than men at around midlife: which progressed to older age, similar to the results observed for Perseverative Errors. With regard to the potential mitigating effects of bilingual language experience on task performance, we observed an Age*L2 AoA*Sex interaction where women and men showed an age-related increase in number of Non-Perseverative errors, with the exception of men with later AoA (i.e., 13–22 years) who made fewer errors with age. Moreover, women with earlier L2 AoA made fewer Non-Perseverative Errors than women with greater L2 AoA, an opposite pattern of effects than those observed in men (i.e., earlier L2 AoA was associated with greater errors).

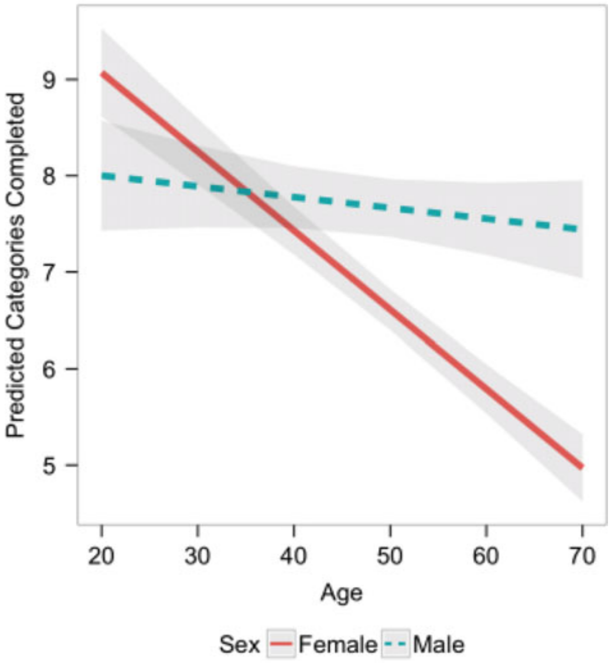

Figure 4. (Colour online) Partial effects plot demonstrating the Age x Sex interaction on the number of Categories Completed, where the number of possible categories can range from 0–9 (Model 1).

### *The number of categories completed*

Model 1, which examined the interaction of Age*Sex*L2 AoA, showed a significant Age*Sex interaction ($\beta$ = -1.15, *SE* = .37, *t* = -3.11, *p* < .01), but no main effects or interactions involving L2 AoA with respect to the number of Categories Completed. This Age*Sex interaction indicated that although women completed more categories than men at a younger age, this pattern reversed starting around midlife, with older women showing worse performance than older men (see Figure 4). Additionally, the covariate of education level was significant for L2 AoA ($\beta$ = 0.34, SE = 0.17, t = 2.03, p = 0.048), which indicated that greater education level was associated with greater number of Categories Completed.

Model 2, which examined the interaction of Age*Sex*the Number of Languages Known, also showed a significant Age*Sex interaction ($\beta$ = -1.29, *SE* = .38, *t* = -3.45, *p* < .01), demonstrating that, although men showed a steady state performance from young to older adulthood, women completed fewer categories with age, and also performed worse than men after late young adulthood. Finally, there were no main effects or interactions involving the Number of Languages Known, but the covariate of education level was significant for Number of Languages Known ($\beta$ = 0.41, SE = 0.18, t = 2.23, p = 0.03), which indicated that greater education level was associated with greater number of Categories Completed.

Model 3, which examined the interaction of Age*Sex*Non-native Language Usage, also showed a significant Age*Sex interaction ($\beta$ = -1.10, *SE* = .39, *t* = -2.84, *p* < .01). Similar to the previous Age*Sex interactions from past models, this interaction showed that participants achieved fewer categories with age, but that women showed a greater performance decrement compared to men starting as early as midlife. Lastly, the covariate of education level was significant for Non-native Language Usage ($\beta$ = 0.46, SE = 0.18, t = 2.63,

$p = 0.01$), which indicated that greater education level was associated with greater number of Categories Completed.

Thus, to summarize the results for the Number of Categories Completed, we consistently observed Age*Sex interactions across all the models, such that women completed fewer categories compared to men at midlife and older age. However, there was no modulating effect of bilingual language experience on this measure of the WCST. Interestingly, the effect of education was significant across all three models.

## Discussion

We investigated whether and how bilingual language experience, age and biological sex impacted executive control, as assessed by the WCST, in an adult lifespan sample. With respect to the past literature investigating the impact of bilingual language experience on domain-general cognition, this study had the following advantages: 1) we examined age continuously from young adulthood to older adulthood in a relatively large sample, thus enabling us to assess whether any effects would emerge in midlife; 2) we systematically investigated interactions between biological sex, age, and different measures of bilingualism on WCST performance; and 3) we used robust regression methods which enabled the inclusion of outlier participants to investigate three continuous measures of bilingual experience (i.e., L2 AoA, number of languages known, percentage of non-native language usage; see Von Bastian, Souza & Gade, 2016, for the use of continuous bilingual measures in a younger adult sample). Finally, all models controlled for number of years of education, a potential proxy of socioeconomic status and cognitive reserve.

Our findings showed different aspects of bilingual experience appeared to enhance or diminish WCST performance depending on biological sex. Generally, women had the greatest age-related cognitive decline across all WCST measures compared to men, but were more likely to show improved performance with increased bilingual experience. In the sections below we focus primarily on how age, sex and bilingualism related to Non-Perseverative and Perseverative Errors, since level of bilingualism did not impact the measure of Categories Completed.

### *The impact of age and biological sex on WCST performance*

Consistent with past work (Axelrod & Henry, 1992; Daigneault et al., 1992; Fristoe et al., 1997; Kousaie et al., 2014; Rhodes, 2004), we found that performance on the WCST declined with age. As people aged, they made more Perseverative and Non-Perseverative Errors. The measure of Perseverative Errors has been linked with reduced inhibitory control and cognitive flexibility in order to shift to a new sorting classification (e.g., Head et al., 2009; Fristoe et al., 1997; Hartman, Bolton & Fehnel, 2001; Miyake et al., 2000). In contrast, Non-Perseverative Errors reflect errors of distraction or failure to pay attention to the current sorting rule (e.g., Barceló, 1999; 2001). Therefore, our current findings corroborate previous studies indicating that healthy aging is associated with reduced performance on the WCST, and reductions in associated executive functions.

We also observed greater age-related cognitive decline in women vs. men at midlife and later life, a finding that is inconsistent with Boone et al. (1993) who investigated sex differences on WCST performance in middle-aged and older adults. In their study, Boone et al. found that women performed better than men on percent Perseverative Errors and Categories Completed (the measure of Non-Perseverative Errors was not analyzed). This inconsistency in findings is likely due to differences across studies in the way age-groups were defined, and the way the analyses testing for sex differences in WCST were conducted. For example, in our study, age was a continuous variable in our analyses exploring sex effects, whereas Boone et al. collapsed across ages 45–83 years to examine sex effects. Moreover, our study included different covariates when examining sex differences in WCST performance (e.g., presence of covariate bilingual measures). Therefore, these methodological differences could explain the dissimilar results we obtained, compared to Boone et al. (1993).

Sex differences in Non-Perseverative Errors may emerge in midlife (and remain into older age) because midlife is a sensitive period in adult development for women due to changes in hormonal levels associated with the transition into menopause. Indeed, the majority of our female sample in midlife included women in pre/peri- and post-menopausal stages. During menopausal transition, women experience a decline in endogenous estrogen levels, a primarily female sex hormone that plays a neuroprotective role in cognition (Brinton, Yao, Yin, Mack & Cadenas, 2015; Green & Simpkins, 2000; Janicki & Schupf, 2010). This decline may contribute to the steeper cognitive decline post-menopause. Moreover, this variability in hormonal levels at midlife in women, in addition to interactions with social and environmental factors (e.g., psychosocial stress and stress management, greater number of major life roles such as parenting, employee, caregiver to elderly parents), may contribute to observed sex differences in cognitive function with aging (e.g., Agrigoroaei & Lachman, 2011; Sullivan Mitchell & Fugate Woods, 2001). For example, given that Non-Perseverative Errors are thought to reflect declines in attention, it is possible that the accumulation of these sociocultural and biological factors in women, may contribute to their greater difficulties in concentrating on a task (e.g., Barling & MacEwen, 1991; Sullivan Mitchell & Fugate Woods, 2001; Xu, Lang & Rooney, 2014). It would be important for future research to target this transition more explicitly: as, unfortunately, we did not have a sufficient number of pre/peri-menopausal women vs. post-menopausal women to conduct such analyses here.

### *The impact of bilingual language experience on WCST performance*

#### *The effect of L2 AoA*

With respect to the impact of bilingual language experience, we found that middle-aged and older women with earlier L2 AoA made fewer Non-Perseverative Errors. Taken together with the above findings (sex differences in Non-Perseverative Errors), these results suggest that earlier L2 AoA mitigated the age-related increase in Non-Perseverative Errors in women, at midlife and older age. This finding is consistent with past work showing that earlier L2 AoA is important for forestalling age-related cognitive decline (general to both sexes) (Luk, De Sa & Bialystok, 2011; Perquin, Vaillant, Schuller, Pastore, Dartigues, Lair & Diederich, 2013), and thus reduces the number of Non-Perseverative Errors. The fact that this effect was only observed in middle-aged and older women may be due to the effect of menopause, as discussed above.

#### *The effect of non-native language usage*

We found that women with greater Non-native Language Usage exhibited fewer Perseverative Errors than men across the adult lifespan. Accordingly, the active use of non-L1 language(s), related to fewer Perseverative Errors in women, and that, unlike L2 AoA, this relationship was not moderated by age. Fewer Perseverative Errors reflects a greater ability to inhibit an irrelevant rule in order to apply a new rule (e.g., task-switching). Interestingly, younger and older adult women, compared to men, have been shown to have greater task-switching abilities (e.g., Kuptsova, Ivanova, Petrushevsky, Fedina & Zhavoronkova, 2015; Stoet, O'Connor, Conner & Laws, 2013), although no study to our knowledge has examined sex differences in task-switching in older adults. Thus, it is possible that women with greater non-native language usage have greater practice inhibiting the non-target, native, language(s), which might transfer to greater task-switching processes that may be reflected in this task as a reduction in the number of Perseverative Errors. This finding is also consistent with Kavé et al. (Kavé, Eyal, Shorek & Cohen-Mansfield, 2008), who found that bilinguals self-reporting a higher degree of non-native vs. native language fluency performed better on a cognitive screening test (assessing time orientation, memory and concentration) (Katzman, Brown, Fuld, Peck, Schechter & Schimmel, 1983) than those who self-reported a higher degree of native vs. non-native language fluency (Kavé et al., 2008). Similarly, Prior and Gollan (2013) report a significant link between executive control and bilingual language control, particularly for non-dominant language production and error monitoring. Thus, greater use of non-L1 languages may involve greater demand for language control that may be associated with more efficient task-switching processes.

However, greater non-native language usage (e.g., those with greater than 90% non-native usage) might also indicate that participants no longer use their L1, and perhaps more dominantly use their non-L1 language(s) (rather than balanced usage of two languages). This is conceivable in our sample where people may have learned their native language at home early in life, but may have started to more frequently use another language (i.e., French or English) in a new cultural context. This may attenuate the influence of non-native language usage on task performance (as 'greater' does not necessarily mean 'better', in terms of active use of one's languages).

The effect of non-native language use on Perseverative Errors, and L2 AoA on Non-Perseverative Errors, suggests that these two bilingual factors may tap into different cognitive processes. Indeed, there is a small but significant negative correlation between both bilingual factors ($r = -0.20$, $p = 0.01$), a pattern suggesting that earlier L2 AoA is related to greater non-native use. Accordingly, the earlier an individual learns their second language, the greater the opportunity they have to exercise the use of that non-native language. However, it is important to emphasize that non-native usage is an average of how much bilinguals and multilinguals use their non-L1 language(s). Thus, the non-native language usage measure for a trilingual may not be exactly comparable to a bilingual, for example – i.e., non-native usage in the bilingual only refers to the use of second language whereas this same measure for a trilingual is a composite of both their second and third language usage. Importantly, a higher level of non-native language use does not necessarily mean that an individual necessarily switches between their different languages more frequently. Thus, both bilingual measures can be different in that the non-native usage measure refers to active use of one's non-L1, whereas that is not necessarily the case for L2 AoA. That is, an individual can acquire their L2 early on but not actively use it, which can potentially explain the weak correlation between both bilingual measures and why they may separately contribute to different measures of the WCST.

Another possible explanation for why L2 AoA and non-native language usage may differentially contribute to the WCST measures may be related to how bilinguals uniquely use their languages in Montreal. In this city, bilinguals often mix their languages interchangeably, even within the same conversation. Thus, the need to exercise greater inhibitory control may be less necessary in Montreal than other bilingual cities where it may be more typical to observe bilinguals speaking one language in one setting and another language in another (i.e., work vs. home). In the latter case, which involves cross-language mixing, this may involve a greater need to exercise inhibitory control to suppress the non-target language(s). Thus, it is possible that earlier L2 AoA may not necessarily lead to greater inhibitory control. One study by Klein et al. (Klein, Mok, Chen & Watkins, 2014) highlights this point. The authors assessed cortical thickness in monolinguals and simultaneous (0-3 years L2 AoA) and late bilinguals (4-13 years L2 AoA) in Montreal and found no difference in cortical thickness between monolinguals and simultaneous bilinguals, but did find a difference between these two groups and late bilinguals (Klein et al., 2014). This suggests that simultaneous or early acquisition of multiple languages may not necessarily be distinct from what is experienced by monolinguals at a neural level, which may also manifest into a more complicated understanding of how L2 AoA may contribute to cognitive function.

### *Caveats*

When considering bilingual effects on WCST performance and the moderating role of biological sex, the results appear to show that, unlike women, men do not benefit from greater bilingual experience in terms of better task performance. Specifically, men perform worse with greater non-native usage and longer L2 AoA. In fact, when the models were separately conducted for women and men, the overall regression models are significant for women but were still marginally or non-significant for men. Moreover, post-hoc analyses using a matched subsample of $N = 54$ women and men (matched on age and education) revealed a similar pattern of results reported in our primary analyses. In other words, even in a matched sample, the effect of bilingualism on WCST was only observed in middle-aged and older women. Taken together, our results suggest that the null effect in men observed in our primary analyses was not due to the small sample size of men relative to women (54 men, 98 women in our overall sample).

It is also possible that the null effects reported for men may reflect the fact that the men tested in this study were cognitively higher performers than the sample of women tested, which may be suggestive of a selection bias in participant recruitment. However, all participants in the current study met the same inclusion criteria listed in the Methods. Moreover, as indicated in Table 1, there were no neuropsychological tests in which men outperformed women in the current study. In contrast, women were significantly better than men on some measures (see Table 1). This suggests that the men sampled in our study were not cognitively higher performers, compared to women. Therefore, the null effect in males may indeed reflect a sex-specific benefit of bilingualism on WCST performance in women, but not men, at midlife and later life. Future studies are needed to confirm this effect.

Interestingly, our results did not identify a main effect of the number of languages known, a dimension that has previously been reported to have a significant effect in delaying cognitive decline in older adults (Ihle, Oris, Fagot & Kliegel, 2016; Kavé et al., 2008), and in delaying symptom onset of dementia (Chertkow, Whitehead, Phillips, Wolfson, Atherton & Bergman, 2010). Other work also supports a more protective role of multilingualism (i.e., knowing three or more languages) compared to bilingualism in protecting cognition in older age, suggesting that the use and practice of more than two languages has a more significant effect on cognition, potentially through greater exercise of executive control mechanisms, and thereby acting as a more powerful source of cognitive reserve (Bak, Nissan, Allerhand & Deary, 2014; Chertkow et al., 2010; Kavé et al., 2008; Perquin et al., 2013). It is thus possible that the lack of variability in our study in terms of the number of languages known (with a smaller sample of multilinguals) possibly attenuated any effects related to this language experiential variable on cognitive performance of the WCST.

Variations in bilingual and multilingual patterns of use might also differentially exercise executive control (e.g., Bak, 2016; Baum & Titone, 2014; Titone et al., 2017). Prior work using the WCST in younger adult bilinguals shows differences in performance depending on how bilinguals use their languages [i.e., non-switchers vs. switchers (Festman & Münte, 2012), interpreters vs. bilinguals (Yudes, Macizo & Bajo, 2011)]. Thus, a hypothesis of greater bilingual language experience may not necessarily lead to greater cognitive outcome; however, differences in how these bilinguals distribute use of their languages in daily life (holding total amount of bilingual experience constant) might exert an effect on cognition. This is not a factor that we directly considered in our study, though we did repeat our analyses on a subset of our English–French bilinguals living in Montreal (where the study was conducted), the details of which we present below.

Given our diverse multilingual sample, it is possible that participants may not use their native languages in a way that is similar to the usage patterns of English and French in Montreal. For example, previous research suggests that inactive bilinguals (i.e., balanced bilinguals in early life actively using one language in later life) performed more like monolinguals than active bilinguals (de Bruin et al., 2015; de Bruin, Della Sala & Bak, 2016). Moreover, the heterogeneity of the linguistic backgrounds of our multilingual speaking sample might have had an impact on our study findings given the close association between language experience and other cultural factors that may exercise cognitive function, and which might be hard to tease apart in experimental methodology. In Montreal, however, French is the city's official language, but there is frequent switching between English and French across and within social contexts, suggesting similar patterns of language usage across participants. When we conducted our analyses exclusively on English–French bilinguals, the same general patterns reported for the larger, more heterogeneous, group were observed. This suggests that the results reported above for the whole sample are due to differences in bilingual language acquisition and usage rather than within-population variability.

Finally, a possible limitation of the present study is that we investigate the relationship between bilingual experience and cognitive performance without including a monolingual control group. This makes it difficult to generalize our findings to past studies that have compared bilinguals as a group to monolinguals as a group, or to make affirmative statements about how any form of bilingual experience (even in low doses) compares with no bilingual experience whatsoever. However, given the multilingual nature of Montreal, the city in which we conducted this study, it is challenging to recruit pure monolinguals or to even make the assumption that pure monolingualism is possible here, given ambient exposure to French in the linguistic landscape (e.g., Vingron, Gullifer, Hamill, Leimgruber & Titone, in press). As well, the practice of comparing bilinguals as a group to monolinguals as a group has its own limitations, most notably the incredible within group heterogeneity in terms of bilingual experience, other forms of heterogeneity among monolinguals, and unspecified differences among bilinguals and monolinguals that could also impact executive control (for detailed discussion of such issues, see Baum & Titone, 2014; Titone & Baum, 2014; Valian, 2015; Titone et al., 2017, Bak, 2017).

## Conclusion

To conclude, the results of this study suggest that increased bilingual language experience modulated performance in an unspeeded executive control task as a function of increasing age and biological sex. Specifically, women showed greater age-related declines in WCST performance compared to men; however, greater bilingual language experience (i.e., earlier age of L2 acquisition) was related to better WCST performance as they aged. Our findings show that a deeper consideration of sex differences in age-related executive control trajectories may be crucial when investigating the relative impact of other factors, such as bilingual language experience, on cognition.

## References

Abutalebi, J., Tettamanti, M., & Perani, D. (2009). The bilingual brain: Linguistic and non-linguistic skills. *Brain and language, 109*(*2-3*), 51–54. doi:10.1016/j.bandl.2009.04.001

Agrigoroaei, S., & Lachman, M. E. (2011). Cognitive functioning in midlife and old age: combined effects of psychosocial and behavioral factors. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *66*, i130–i140.

Alzheimer's Association. (2015). 2015 Alzheimer's disease facts and figures. *Alzheimer's and Dementia: The Journal of the Alzheimer's Association, 11*(*3*), 332–384.

Amer, T., Campbell, K. L., & Hasher, L. (2016). Cognitive Control As a Double-Edged Sword. *Trends in Cognitive Sciences*, *20*(*12*), 905–915. https://doi.org/10.1016/j.tics.2016.10.002

Ankudowich, E., Pasvanis, S., & Rajah, M. N. (2017). Changes in the correlation between spatial and temporal source memory performance and BOLD activity across the adult lifespan. *Cortex*, *91*, 234–249.

Ankudowich, E., Pasvanis, S., & Rajah, M. N. (2016). Changes in the modulation of brain activity during context encoding vs. context retrieval across the adult lifespan. *NeuroImage*, *139*, 103–113. http://doi.org/10.1016/j.neuroimage.2016.06.022

Ansaldo, A. I., Ghazi-Saidi, L., & Adrover-Roig, D. (2015). Interference Control In Elderly Bilinguals: Appearances Can Be Misleading. *Journal of Clinical and Experimental Neuropsychology*, *2*(*1*), 1–16. http://doi.org/10.1080/13803395.2014.990359

Antón, E., García, Y. F., Carreiras, M., & Duñabeitia, J. A. (2016). Does bilingualism shape inhibitory control in the elderly?. *Journal of Memory and Language*, *90*, 147–160.

Axelrod, B. N., & Henry, R. R. (1992). Age-related performance on the wisconsin card sorting, similarities, and controlled oral word association tests. *The Clinical Neuropsychologist*, *6*(*1*), 16–26. https://doi.org/10.1080/13854049208404113

Bak, T. H. (2017). The impact of bilingualism on cognitive ageing and dementia. *Growing Old with Two Languages: Effects of Bilingualism on Cognitive Aging*, *53*, 243.

Bak, T. H. (2016). Cooking pasta in La Paz. *Linguistic Approaches to Bilingualism*, *6*(*5*), 699–717. http://doi.org/10.1075/lab.16002.bak

Bak, T. H., Nissan, J. J., Allerhand, M. M., & Deary, I. J. (2014). Does bilingualism influence cognitive aging? *Annals of Neurology*, *75*(*6*), 959–63. http://doi.org/10.1002/ana.24158

Barceló, F. (1999). Electrophysiological evidence of two different types of error in the Wisconsin Card Sorting Test. *Neuroreport*, *10*(*6*), 1299–1303.

Barceló, F. (2001). Does the Wisconsin Card Sorting Test Measure Prefontral Function? *The Spanish journal of psychology*, *4*(*1*), 79–100.

Barling, J., & MacEwen, K. E. (1991). Maternal employment experiences, attention problems and behavioral performance: A mediational model. *Journal of Organizational Behavior*, *12*(*6*), 495–505.

Barulli, D., & Stern, Y. (2013). Efficiency, capacity, compensation, maintenance, plasticity: emerging concepts in cognitive reserve. *Trends in cognitive sciences*, *17*(*10*), 502–509.

Baum, S., & Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Applied Psycholinguistics*, *35*(*5*), 857–894. http://doi.org/10.1017/S0142716414000174

Beard, J., Officer, A., & Cassels, A. (2015). *World Report on Ageing and Health*. Luxembourg: World Health Organization Retrieved from http://apps.who.int/iris/bitstream/10665/186463/1/9789240694811_eng.pdf.

Bialystok, E., Abutalebi, J., Bak, T. H., Burke, D. M., & Kroll, J. F. (2016). Aging in two languages: Implications for public health. *Ageing Research Reviews*, *27*, 56–60. http://doi.org/10.1016/j.arr.2016.03.003

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, Aging, and Cognitive Control: Evidence From the Simon Task. *Psychology and Aging*, *19*(*2*), 290–303. http://doi.org/10.1037/0882-7974.19.2.290

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, memory, and cognition*, *34*(*4*), 859.

Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism*, *9*(*1*), 103–119.

Bieri, J., Bradburn, W. M., & Galinsky, M. D. (1958). Sex differences in perceptual behavior. *Journal of Personality*, *26*(*1*), 1–12. http://doi.org/10.1111/j.1467-6494.1958.tb01566.x

Blumenfeld, H. K., & Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition, 118*(*2*), 245–257. doi:10.1016/j.cognition.2010.10.012

Boone, K. B., Ghaffarian, S., Lesser, I. M., Hill-Gutierrez, E., & Berman, N. G. (1993).Wisconsin Card Sorting Test performance in healthy, older adults: Relationship to age, sex, education, and IQ. *Journal of Clinical Psychology*, *49*(*1*), 54–60.

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*. http://doi.org/10.1016/j.tics.2011.12.010

Braver, T. S., & West, R. (2008). Working memory, executive control, and aging. *The handbook of aging and cognition, 3*, 311–372.

Brinton, R. D., Yao, J., Yin, F., Mack, W. J., & Cadenas, E. (2015). Perimenopause as a neurological transition state. *Nature Reviews Endocrinology*, *11*(*7*), 393–405. http://doi.org/10.1038/nrendo.2015.82

Calvo, N., García, A. M., Manoiloff, L., & Ibáñez, A. (2016). Bilingualism and cognitive reserve: A critical overview and a plea for methodological innovations. *Frontiers in Aging Neuroscience, 7*, 249. http://doi.org/10.3389/fnagi.2015.00249

Chertkow, H., Whitehead, V., Phillips, N., Wolfson, C., Atherton, J., & Bergman, H. (2010). Multilingualism (but not always bilingualism) delays the onset of Alzheimer disease: Evidence from a bilingual community. *Alzheimer Disease & Associated Disorders*, *24*(*2*), 118–125. http://doi.org/10.1097/WAD.0b013e3181ca1221

Christensen, H., Mackinnon, A. J., Korten, A. E., Jorm, A. F., Henderson, A. S., Jacomb, P., & Rodgers, B. (1999). An analysis of diversity in the cognitive performance of elderly community dwellers: Individual differences in change scores as a function of age. *Psychology and Aging, 14*(*3*), 365–379. http://doi.org/10.1037/0882-7974.14.3.365

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research, 1147*, 192–208. doi:10.1016/j.brainres.2007.01.137

Clayson, P. E., Clawson, A., & Larson, M. J. (2011). Sex differences in electrophysiological indices of conflict monitoring. *Biological Psychology*, *87*(*2*), 282–289. http://doi.org/10.1016/j.biopsycho.2011.03.011

Coderre, E. L., Smith, J. F., van Heuven, W. J. B., & Horwitz, B. (2016). The Functional Overlap of Executive Control and Language Processing in Bilinguals. *Bilingualism: Language and Cognition*, *19*(*3*), 471–488. http://doi.org/10.1017/S1366728915000188

Colzato, L. S., Hertsig, G., van den Wildenberg, W. P. M., & Hommel, B. (2010). Estrogen modulates inhibitory control in healthy human females: Evidence from the stop-signal paradigm. *Neuroscience*, *167*(*3*), 709–715. http://doi.org/10.1016/j.neuroscience.2010.02.029

Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience*, *15*(*5*), 336–345. http://doi.org/10.1038/nrn3709

Daigneault, S., Braun, C. M. J., & Whitaker, H. A. (1992). Early effects of normal aging on perseverative and non-perseverative prefrontal measures. *Developmental Neuropsychology*, *8*(*1*), 99–114. http://doi.org/10.1080/87565649209540518

Davis, J. C., Marra, C. A., Najafzadeh, M., & Liu-Ambrose, T. (2010). The independent contribution of executive functions to health related quality of life in older women. *BMC Geriatrics*, *10*(*1*), 16. http://doi.org/1471-2318-10-16[pii]\r10.1186/1471-2318-10-16

de Bruin, A., Bak, T. H., & Della Sala, S. (2015). Examining the effects of active versus inactive bilingualism on executive control in a carefully matched non-immigrant sample. *Journal of Memory and Language*, *85*, 15–26. http://doi.org/10.1016/j.jml.2015.07.001

de Bruin, A., Della Sala, S., & Bak, T. H. (2016). The effects of language use on lexical processing in bilinguals. *Language, Cognition and Neuroscience*, *31*(*8*), 967–974. http://doi.org/10.1080/23273798.2016.1190024

Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). New York: Oxford University Press.

Duff, S. J., & Hampson, E. (2001). A sex difference on a novel spatial working memory task in humans. *Brain and Cognition*, *47*(*3*), 470–93. http://doi.org/10.1006/brcg.2001.1326

Evans, K. L., & Hampson, E. (2015). Sex-dependent effects on tasks assessing reinforcement learning and interference inhibition. *Frontiers in Psychology*, *6*, 1044. http://doi.org/10.3389/fpsyg.2015.01044

Ferland, M. B., Ramsay, J., Engeland, C., & O'Hara, P. (1998). Comparison of the performance of normal individuals and survivors of traumatic brain injury on repeat administrations of the Wisconsin Card Sorting Test. *Journal of clinical and experimental neuropsychology*, *20*(*4*), 473–482.

Festman, J., & Münte, T. F. (2012). Cognitive control in Russian-German bilinguals. *Frontiers in Psychology*, *3*(*115*). http://doi.org/10.3389/fpsyg.2012.00115

Finkel, D., Reynolds, C. A, McArdle, J. J., Gatz, M., & Pedersen, N. L. (2003). Latent growth curve analyses of accelerating decline in cognitive abilities in late adulthood. *Developmental Psychology*, *39*(*3*), 535–550. http://doi.org/10.1037/0012-1649.39.3.535

Forte, R., Boreham, C. A. G., de Vito, G., & Pesce, C. (2015). Health and quality of life perception in older adults: The joint role of cognitive efficiency and functional mobility. *International Journal of Environmental Research and Public Health*, *12*(*9*), 11328–11344. http://doi.org/10.3390/ijerph120911328

Fristoe, N. M., Salthouse, T. A., & Woodard, J. L. (1997). Examination of age-related deficits on the Wisconsin Card Sorting Test. *Neuropsychology*, *11*(*3*), 428–436. http://doi.org/10.1037/0894-4105.11.3.428

Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Guasch, N. V., Roberts, E.J., Hughes, E.K., & Jones, L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology, 5*(*11*). http://doi.org/10.3389/fpsyg.2014.00011

Geary, D. C., Saults, S. J., Liu, F., & Hoard, M. K. (2000). Sex differences in spatial cognition, computational fluency, and arithmetical reasoning. *Journal of Experimental Child Psychology*, *77*(*4*), 337–353. http://doi.org/10.1006/jecp.2000.2594

Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., & Smith, C. D. (2013). Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *Journal of Neuroscience*, *33*(*2*), 387–396.

Grady, C. L., Luk, G., Craik, F. I. M., & Bialystok, E. (2015). Brain network activity in monolingual and bilingual older adults. *Neuropsychologia*, *66*, 170–81. http://doi.org/10.1016/j.neuropsychologia.2014.10.042

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 67–81.

Green, D. W. (2011). Language control in different contexts: the behavioral ecology of bilingual speakers. *Frontiers in Psychology, 2*. doi:10.3389/fpsyg.2011.00103

Green, P. S., & Simpkins, J. W. (2000). Neuroprotective effects of estrogens: potential mechanisms of action. *International Journal of Developmental Neuroscience*, *18*(*4*), 347–358. http://doi.org/S0736-5748(00)00017-4[pii]

Guo, T., Liu, H., Misra, M., & Kroll, J. F. (2011). Local and global inhibition in bilingual word production: fMRI evidence from Chinese-English bilinguals. *NeuroImage, 56*(*4*), 2300–2309. doi:10.1016/j.neuroimage.2011.03.049

Halari, R., Hines, M., Kumari, V., Mehrotra, R., Wheeler, M., Ng, V., & Sharma, T. (2005). Sex differences and individual differences in cognitive performance and their relationship to endogenous gonadal hormones and gonadotropins. *Behavioral Neuroscience*, *119*(*1*), 104–117. http://doi.org/10.1037/0735-7044.119.1.104

Halari, R., & Kumari, V. (2005). Comparable cortical activation with inferior performance in women during a novel cognitive inhibition task. *Behavioural Brain Research*, *158*(*1*), 167–173. http://doi.org/10.1016/j.bbr.2004.08.022

Hartman, M., Bolton, E., & Fehnel, S. E. (2001). Accounting for age differences on the Wisconsin Card Sorting Test: decreased working memory, not inflexibility. *Psychology and Aging*, *16*(*3*), 385–399. http://doi.org/10.1037/0882-7974.16.3.385

Head, D., Kennedy, K. M., Rodrigue, K. M., & Raz, N. (2009). Age differences in perseveration: Cognitive and neuroanatomical mediators of performance on the Wisconsin Card Sorting Test. *Neuropsychologia*, *47*(*4*), 1200–1203. http://doi.org/10.1016/j.neuropsychologia.2009.01.003

Hyde, J. S. (1981). How large are cognitive gender differences? A meta-analysis using!w$^2$ and d.. *American Psychologist*, *36*(*8*), 892–901. http://doi.org/10.1037/0003-066X.36.8.892

Ihle, A., Oris, M., Fagot, D., & Kliegel, M. (2016). The relation of the number of languages spoken to performance in different cognitive abilities in old age. *Journal of Clinical and Experimental Neuropsychology*, *38*(*10*), 1103–1114. http://doi.org/10.1080/13803395.2016.1197184

Janicki, S. C., & Schupf, N. (2010). Hormonal influences on cognition and risk for Alzheimer's disease. *Current Neurology and Neuroscience Reports*. http://doi.org/10.1007/s11910-010-0122-6

Karlsson, P., Thorvaldsson, V., Skoog, I., Gudmundsson, P., & Johansson, B. (2015). Birth cohort differences in fluid cognition in old age: Comparisons of trends in levels and change trajectories over 30 years in three population-based samples. *Psychology and Aging*. http://doi.org/10.1037/a0038643

Katzman, R., Brown, T., Fuld, P., Peck, A., Schechter, R., & Schimmel, H. (1983). Validation of a short orientation-memory-concentration test of cognitive impairment. *The American Journal of Psychiatry*, *140*(*6*), 734–739. http://doi.org/10.1176/ajp.140.6.734

Kavé, G., Eyal, N., Shorek, A., & Cohen-Mansfield, J. (2008). Multilingualism and cognitive state in the oldest old. *Psychology and Aging*, *23*(*1*), 70–78. http://doi.org/10.1037/0882-7974.23.1.70

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*(*6*), 577–590.

Kirk, N. W., Fiala, L., Scott-Brown, K. C., & Kempe, V. (2014). No evidence for reduced Simon cost in elderly bilinguals and bidialectals. *Journal of Cognitive Psychology*, *26*(*6*), 640–648. http://doi.org/10.1080/20445911.2014.929580

Klein, D., Mok, K., Chen, J. K., & Watkins, K. E. (2014). Age of language learning shapes brain structure: a cortical thickness study of bilingual and monolingual individuals. *Brain and language*, *131*, 20–24.

Kousaie, S., & Phillips, N. A. (2012). Ageing and bilingualism: Absence of a "bilingual advantage" in Stroop interference in a nonimmigrant sample. *The Quarterly*

*Journal of Experimental Psychology*, *65*(*2*), 356–369. http://doi.org/10.1080/17470218.2011.604788

Kousaie, S., Sheppard, C., Lemieux, M., Monetta, L., & Taler, V. (2014). Executive function and bilingualism in young and older adults. *Frontiers in Behavioral Neuroscience*, *8*(*250*). http://doi.org/10.3389/fnbeh.2014.00250

Kroll, J. F., & Bialystok, E. (2013). Understanding the Consequences of Bilingualism for Language Processing and Cognition. *Journal of Cognitive Psychology*, *25*(*5*), 497–514. http://doi.org/10.1080/20445911.2013.799170

Kroll, J. F., Bobb, S. C., & Hoshino, N. (2014). Two Languages in Mind: Bilingualism as a Tool to Investigate Language, Cognition, and the Brain. *Current Directions in Psychological Science*, *23*(*3*), 159–163. http://doi.org/10.1177/0963721414528511

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica, 128*(*3*), 416–430. doi:10.1016/j.actpsy.2008.02.001

Kuptsova, S. V., Ivanova, M. V., Petrushevsky, A. G., Fedina, O. N., & Zhavoronkova, L. A.(2015). Sex-related differences in task switching: An fMRI study. *Human Physiology*, *41*(*6*), 611–624.

Li, C. S., Zhang, S., Duann, J. R., Yan, P., Sinha, R., & Mazure, C. M. (2009). Gender differences in cognitive control: An extended investigation of the stop signal task. *Brain Imaging and Behavior*, *3*(*3*), 262–276. http://doi.org/10.1007/s11682-009-9068-1

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(*2*), 381–390.

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*(*5*), 605–621. http://doi.org/10.1080/20445911.2013.795574

Luk, G., De Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control?. *Bilingualism: Language and Cognition*, *14*(*4*), 588–595. http://doi.org/10.1017/S1366728911000186

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

Maitland, S. B., Herlitz, A., Nyberg, L., Bäckman, L., & Nilsson, L.-G. (2004). Selective sex differences in declarative memory. *Memory & Cognition*, *32*(*7*), 1160–1169. http://doi.org/10.3758/BF03196889

Martin, M. C., Macizo, P., & Bajo, T. (2010). Time course of inhibitory processes in bilingual language processing. *British Journal of Psychology, 101*, 679–693. doi:10.1348/000712609x480571

McCarrey, A. C., An, Y., Kitner-Triolo, M. H., Ferrucci, L., & Resnick, S. M. (2016). Sex Differences in Cognitive Trajectories in Clinically Normal Older Adults. *Psychology and Aging*, *31*(*1*), 166–175. http://doi.org/10.1037/pag0000070

Mercier, J., Pivneva, I., & Titone, D. (2014). Individual differences in inhibitory control relate to bilingual spoken word processing. *Bilingualism: Language and Cognition*, *17*(*1*), 89–117.

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40*(*1*), 25–40.

Misra, M., Guo, T., Bobb, S. C., & Kroll, J. F. (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language, 67*(*1*), 224–237. doi:10.1016/j.jml.2012.05.001

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science*, *21*(*1*), 8–14. http://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(*1*), 49–100. http://doi.org/10.1006/cogp.1999.0734

Mueller, S. T., & Piper, B. J. (2014). The Psychology Experiment Building Language (PEBL) and PEBL Test Battery. *Journal of Neuroscience Methods*, *222*, 250–259. http://doi.org/10.1016/j.jneumeth.2013.10.024

Mungas, D., Beckett, L., Harvey, D., Farias, S. T., Reed, B., Carmichael, O., Olichney, J., Miller, J., & DeCarli, C. (2010). Heterogeneity of Cognitive Trajectories in Diverse Older Person. *Psychology and Aging*, *25*(*3*), 606–619. http://doi.org/10.1037/a0019502

Nyhus, E., & Barceló, F. (2009). The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: a critical update. *Brain and cognition*, *71*(*3*), 437–451.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(*1*), 97–113.

Perquin, M., Vaillant, M., Schuller, A. M., Pastore, J., Dartigues, J. F., Lair, M. L., & Diederich, N. (2013). Lifelong Exposure to Multilingualism: New Evidence to Support Cognitive Reserve Hypothesis. *PLoS ONE*, *8*(*4*). http://doi.org/10.1371/journal.pone.0062030

Pivneva, I., Mercier, J., & Titone, D. (2014). Executive control modulates cross-language lexical activation during L2 reading: evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 787–796. doi:10.1037/a0035583

Pivneva, I., Palmer, C., & Titone, D. (2012). Inhibitory control and L2 proficiency modulate bilingual language production: evidence from spontaneous monologue and dialogue speech. *Frontiers in Psychology, 3*. doi:ARTN 57 10.3389/fpsyg.2012.00057

Prior, A., & Gollan, T. H. (2013). The elusive link between language control and executive control: A case of limited transfer. *Journal of Cognitive Psychology*, *25*(*5*), 622–645. http://doi.org/10.1080/20445911.2013.821993

R Development Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rajah, M. N., Wallace, L. M., Ankudowich, E., Yu, E. H., Swierkot, A., Patel, R., Chakravarty, M. M, Naumova, D., Pruessner, J., Joober, R., Gauthier, S., & Pasvanis, S. (2017). Family history and APOE4 risk for Alzheimer's disease impact the neural correlates of episodic memory by early midlife. *NeuroImage: Clinical*, *14*, 760–774.

Reimers, S., & Maylor, E. A. (2005). Task switching across the life span: effects of age on general and specific switch costs. *Developmental Psychology*, *41*(*4*), 661–671. http://doi.org/10.1037/0012-1649.41.4.661

Rhodes, M. G. (2004). Age-Related Differences in Performance on the Wisconsin Card Sorting Test: A Meta-Analytic Review. *Psychology and Aging*, *19*(*3*), 482–494. http://doi.org/10.1037/0882-7974.19.3.482

Roivainen, E. (2011). Gender differences in processing speed: A review of recent research. *Learning and Individual Differences*, *21*(*2*), 145–149. http://doi.org/10.1016/j.lindif.2010.11.021

Salthouse, T. A. (2009). Decomposing age correlations on neuropsychological and cognitive variables. *Journal of the International Neuropsychological Society: JINS*, *15*(*5*), 650–661. http://doi.org/10.1017/S1355617709990385

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society: JINS*, *16*(*5*), 754–760. http://doi.org/10.1017/S1355617710000706

Scarmeas, N., & Stern, Y. (2003). Cognitive Reserve and Lifestyle. *Journal of Clinical and Experimental Neuropsychology*. http://doi.org/10.1076/jcen.25.5.625.14576

Shanmugan, S., & Epperson, C. N. (2014). Estrogen and the prefrontal cortex: Towards a new understanding of estrogen's effects on executive functions in the menopause transition. *Human Brain Mapping*, *35*(*3*), 847–865. http://doi.org/10.1002/hbm.22218

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, *8*(*3*), 448–460. http://doi.org/10.1017/S1355617702813248

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia, 47*(*10*), 2015–2028. http://doi.org/10.1016/j.neuropsychologia.2009.03.004

Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *The Lancet Neurology, 11*(*11*), 1006–1012. http://doi.org/10.1016/S1474-4422(12)70191-6

Stoet, G. (2010). Sex differences in the processing of flankers. *Quarterly Journal of Experimental Psychology (2006)*, *63*(*4*), 633–8. http://doi.org/10.1080/17470210903464253

Stoet, G., O'Connor, D. B., Conner, M., & Laws, K. R. (2013). Are women better than men at multi-tasking?. *BMC Psychology*, *1*(*1*), 18.

Stuss, D. T. (2011). Functions of the frontal lobes: relation to executive functions. *Journal of the International Neuropsychological Society: JINS*, *17*(*5*), 759–65. http://doi.org/10.1017/S1355617711000695

Sullivan Mitchell, E., & Fugate Woods, N. (2001). Midlife women's attributions about perceived memory changes: observations from the Seattle Midlife Women's Health Study. *Journal of women's health & gender-based medicine*, *10*(*4*), 351–362.

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

Titone, D., & Baum, S. (2014). The future of bilingualism research: Insufferably optimistic and replete with new questions. *Applied Psycholinguistics*, *35*(*5*), 933–942. http://doi.org/10.1017/S0142716414000289

Titone, D., Gullifer, J., Subramaniapillai, S., Rajah, N., & Baum, S. (2017). History-inspired reflections on the Bilingual Advantages Hypothesis. *Growing Old with Two Languages: Effects of Bilingualism on Cognitive Aging*, *53*, 265.

Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: The effects of L2 knowledge, semantic constraint, and L1–L2 intermixing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(*6*), 1412–1431.

Tucker, A. M., & Stern, Y. (2011). Cognitive reserve in aging. *Current Alzheimer Research*, *8*(*4*), 354–60. http://doi.org/10.2174/156721112225912050

Tun, P. A., & Lachman, M. E. (2009). Age Differences in Reaction time and Attention in a National Telephone Sample of Adults: Education, Sex, and Task Complexity Matter. *Developmental Psychology*, *44*(*5*), 1421–1429. http://doi.org/10.1037/a0012845.Age

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, *18*(*1*), 3–24. http://doi.org/10.1017/S1366728914000522

Vaughan, L., & Giovanello, K. (2010). Executive function in daily life: Age-related influences of executive processes on instrumental activities of daily living. *Psychology and Aging*, *25*(*2*), 343–355. http://doi.org/10.1037/a0017729

Venables, W. N., & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

Vingron, N., Gullifer, J., Hamill, J., Leimgruber, J. R. E., & Titone, D. (in press). Developing Eye Tracking Methods to Investigate Linguistic Landscape Processing: A Selective Review and Preliminary Data from Bilinguals Viewing Unilingual and Bilingual Signs. Linguistic Landscapes.

Von Bastian, C. C., Souza, A. S., & Gade, M. (2016). No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of Experimental Psychology: General*, *145*(*2*), 246.

von Studnitz, R. E., & Green, D. W. (2002). The cost of switching language in a semantic categorization task. *Bilingualism: Language and Cognition, 5*(*3*), 241–251. doi:10.1017/s1366728902003036

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of Sex Differences in Spatial Abilities: A Meta-Analysis and Consideration of Critical Variables. *Psychological Bulletin*, *117*(*2*), 250–270. http://doi.org/10.1037/0033-2909.117.2.250

Wilson, R. S., Beckett, L. A., Barnes, L. L., Schneider, J. A., Bach, J., Evans, D. A., & Bennett, D. A. (2002). Individual differences in rates of change in cognitive abilities of older persons. *Psychology and Aging, 17*(*2*), 179–193. http://doi.org/10.1037/0882-7974.17.2.179

Xu, Q., Lang, C. P., & Rooney, N. (2014). A systematic review of the longitudinal relationships between subjective sleep disturbance and menopausal stage. *Maturitas*, *79*(*4*), 401–412.

Yeudall, L. T., Fromm, D., Reddon, J. R., & Stefanyk, W. O. (1986). Normative data stratified by age and sex for 12 neuropsychological tests. *Journal of Clinical Psychology*, *42*(*6*), 918–946.

Yudes, C., Macizo, P., & Bajo, T. (2011). The influence of expertise in simultaneous interpreting on non-verbal executive processes. *Frontiers in Psychology*, *2*(*309*). http://doi.org/10.3389/fpsyg.2011.00309

Zahodne, L. B., Schofield, P. W., Farrell, M. T., Stern, Y., & Manly, J. J. (2014). Bilingualism Does Not Alter Cognitive Decline or Dementia Risk among Spanish-Speaking Immigrants. *Neuropsychology*, *28*(*2*), 238–246. http://doi.org/10.1037/neu0000014

Zelinski, E. M., Gilewski, M. J., & Schaie, K. W. (1993). Individual differences in cross-sectional and 3-year longitudinal memory performance across the adult life span. *Psychology and Aging, 8*(*2*), 176–186. http://doi.org/10.1037/0882-7974.8.2.176

Zhou, B., & Krott, A. (2016). Bilingualism enhances attentional control in non-verbal conflict tasks–evidence from ex-Gaussian analyses. *Bilingualism: Language and Cognition*, 1–19.