# Effects of bilingualism on white matter integrity in older adults

**John A.E. Anderson**[a], **John G. Grundy**[a], **Jaisalmer De Frutos**[b], **Ryan M. Barker**[a], **Cheryl Grady**[c,d], and **Ellen Bialystok**[a,c,*]

[a]Department of Psychology, York University, Toronto, Ontario, Canada

[b]Laboratory of Cognitive & Computational Neuroscience, Autonomous University of Madrid Biological & Health Psychology Department, Canada

[c]Baycrest Rotman Research Institute, Toronto, ON, Canada

[d]University of Toronto, Departments of Psychology and Psychiatry, Canada

## Abstract

Bilingualism can delay the onset of dementia symptoms and has thus been characterized as a mechanism for cognitive or brain reserve, although the origin of this reserve is unknown. Studies with young adults generally show that bilingualism is associated with a strengthening of white matter, but there is conflicting evidence for how bilingualism affects white matter in older age. Given that bilingualism has been shown to help stave off the symptoms of dementia by up to four years, it is crucial that we clarify the mechanism underlying this reserve. The current study uses diffusion tensor imaging (DTI) to compare monolinguals and bilinguals while carefully controlling for potential confounds (e.g., I.Q., MMSE, and demographic variables). We show that group differences in Fractional Anisotropy (FA) and Radial Diffusivity (RD) arise from multivariable interactions not adequately controlled for by sequential bivariate testing. After matching and statistically controlling for confounds, bilinguals still had greater axial diffusivity (AD) in the left superior longitudinal fasciculus than monolingual peers, supporting a neural reserve account for healthy older bilinguals.

## Keywords

Bilingualism; DTI; White matter; Aging; Propensity score matching

## Introduction

Speaking two languages on a regular basis has been shown to lead to domain-general cognitive changes that persist across the lifespan (for recent reviews, see Bialystok, 2017; Grundy et al., 2017). However, it is unclear what neural mechanism might underlie these behavioral changes and whether this mechanism persists into old age. Uncovering such a mechanism is crucial in light of the increasing size of the elderly population. For example, in Canada the proportion of seniors aged 60–79 rose from 4.2% of the population in 2012 to

*Corresponding author. Department of Psychology, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada. ellenb@yorku.ca (E. Bialystok).

4.7% in 2016 (Statistics Canada, 2017). This rise in the size of the older adult population is associated with increases in the number of individuals suffering with dementia or cognitive decline. Importantly, there is converging evidence from multiple sources that symptoms of dementia and cognitive decline appear later in lifelong bilinguals than in comparable monolinguals. Older adult bilinguals are diagnosed with Alzheimer's disease (AD) on average four years later than their monolingual peers (Bialystok et al., 2007; Craik et al., 2010; Alladi et al., 2013). A study by Brookmeyer et al. (2007) demonstrated that a 1-year delay in symptoms would yield 11.8 million fewer cases of Alzheimer's disease worldwide by 2050. Clearly there is a need to expose the structural and functional brain differences that may underlie bilinguals' ability to protect cognitive function with aging and stave off dementia symptoms.

A consistent finding in the AD literature is a reduction in white matter integrity with disease progression. The anterior aspect of the corpus callosum and the superior longitudinal fasciculi are both sensitive to the progression of AD (Bartzokis et al., 2004; Rose et al., 2000; Bozzali et al., 2002). These white matter regions are also consistently remodeled by second-language experience in young adults. Structural magnetic resonance imaging (MRI) has revealed that young adult bilinguals have greater white matter volume than their monolingual peers. These differences are particularly reliable in the corpus callosum, and may allow bilinguals to exchange cross-hemispheric information more efficiently than monolinguals (e.g., Coggins et al., 2004; Felton et al., 2017).

More recently, the advent of diffusion tensor imaging (DTI) has allowed for a more detailed examination of water flow along gradients in the neurological pathways in the brain. This methodological development has allowed researchers to characterize white matter microstructural integrity using summary measures of the diffusion tensor (but see Jones et al., 2013; for an alternative interpretation). Anisotropic water diffusion along the primary eigenvector ($\lambda_1$), that is, parallel to a white matter tract is an index of axial diffusivity (AD) and has been shown to measure axon integrity, with higher values indicating better integrity. Isotropic water diffusion, largely influenced by increasing flow perpendicular to the primary diffusion gradient indicates radial diffusivity (RD: $\lambda_2, \lambda_3$) and is associated with demyelination such that higher values are generally associated with poorer integrity. The most widely reported measure, however, is the combination of the former two measures. This measure, called fractional anisotropy (FA), indexes the overall microstructural health of the white matter in a voxel and is calculated from a combination of the three eigenvalues, $\lambda_1, \lambda_2, \lambda_3$, by the following formula: $(3/2)* \; [(\lambda_1 - \lambda_{123})^2 + (\lambda_2 - \lambda_{123})^2 + (\lambda_3 - \lambda_{123})^2] / \; (\lambda_1^2 + \lambda_2^2 \; \lambda_3^2)$, where $\lambda_{123}$ is the mean of the eigenvalues. Therefore, FA is not a simple ratio of AD and RD but rather a complex summary of diffusion along the axon derived from the other two vectors. All three measures thus contribute meaningful information about white matter structure. Although greater FA is generally thought to index healthier white matter integrity, it is possible for changes to emerge in RD or AD without any effect on FA values. Accordingly, it is important to examine all three white matter components from the DTI analysis.

Studies using DTI to measure white matter integrity in young adults have revealed effects of bilingualism echoing the volumetric data. A recent study by Pliatsikas et al. (2015), for

example, showed that bilingual young adults expressed greater FA values than monolinguals in most regions of the corpus callosum, bilaterally in the inferior frontal occipital fasciculus, and external capsules. Training studies have also produced compelling evidence for white matter remodeling. Schlegel et al (2012) demonstrated that second-language training of Chinese by native English speakers over an eight-month period led to a linear increase in FA located predominantly in the anterior corpus callosum. To the degree that they successfully acquired their new language as measured by test scores, the students showed a steeper FA slope, indicating a more rapid remodeling of white matter. Parallels may also be drawn between how bilingualism and musicianship reshape the brain – and, in particular, the corpus callosum. As with bilinguals, musicians also appear to have larger corpus callosum volumes, an effect that is sensitive to the age at which the musician first acquired the skill (Schlaug et al., 1995; Wan & Schlaug, 2010). Echoing the arguments from the bilingual literature, the strengthening of the corpus callosum in musicians is also thought to reflect greater inter-hemispheric communication (e.g., Kraus et al., 2013).

Whether these increases in white matter integrity persist into the older adult years is still a matter of debate but essential for understanding the potential basis for cognitive reserve found for older bilinguals. Only two studies have examined how bilingualism impacts white matter integrity in the aging brain and these two studies report conflicting findings. The first study by Luk et al. (2011) showed that in a small but well-matched sample (N = 14 per group), bilingual older adults had higher FA values than monolinguals in the corpus callosum and bilateral superior and inferior longitudinal fasciculi, consistent with the young adult data. A second study by Gold et al. (2013) matched participants from a larger monolingual sample to a group of 20 bilinguals.[1] Whereas Luk et al. reported increased FA in corpus callosum and bilateral superior longitudinal fasciculi, Gold et al. reported the opposite: monolinguals were more likely to have higher FA values in a distributed set of regions including the corpus callosum, the inferior and superior fronto-occipital fasciculi, and the fornix. The authors noted that there were no regions in which bilinguals showed higher FA than monolinguals, but that bilinguals had higher RD values in most of these same regions. The latter finding that RD was higher for bilinguals was likely what drove the FA ratio, and led Gold and colleagues to conclude that their sample of bilinguals displayed remarkable cognitive reserve in the face of white matter atrophy relative to the monolingual sample.

One possible reason for the lack of consensus among group comparisons in neuroimaging studies is suboptimal matching. While many studies in neuroscience do attempt to rigorously match groups on behaviors and background variables to rule out the possibility that these other factors explain their findings, many others either do not, or simply present a subset of demographic variables without comment. Of those studies that do report matching groups, some indicate that they used *t*-tests to assess the (lack of) group differences in confounding variables, but often the matching procedure is not reported. More recently, techniques have

[1]Gold et al. (2013) matched participants for sex, education level age, and scores on ISP, Cattell IQ, MMSE, Vocabulary (PPVT), Digit span forward and backward, Spatial span forward and backward, Logical memory I and II and Task-switching RT & % errors. Luk et al. (2011) matched on age, gender, years of education, weekly hours of computer use, MMSE, Shipley English scores, Verbal fluency, Design Fluency, Stroop response time and Trail-Making response time. In both studies, matching success was assessed by a non-significant between-groups *p*-value for each measure.

been developed to carefully match groups on multiple variables simultaneously. One such technique, propensity score matching, fits a logistic regression to multiple confounds simultaneously and thus accounts for multivariate interactions among confounding variables that may differ between groups. We argue that there is a pressing need for more transparency about *how* participants are matched if we to assure that differences can be attributed to group characteristics and effects can be replicated. Propensity score matching is superior to sequential univariate group comparisons as it actively accounts for interactions between variables which may themselves differ by group.

Given the need to clarify the mechanism underlying bilinguals' ability to delay dementia symptoms, we investigated whether evidence for white matter differences following a lifetime of bilingual language use could be found in a large sample of older adults. We carefully matched monolingual and bilingual participants to control for multivariate interactions among potentially confounding variables, something previous studies have not done. Based on the evidence from younger adults, we expected to find greater white matter integrity for bilinguals than monolinguals in the corpus callosum, superior longitudinal fasciculi, and inferior fronto-occipital fasciculi. Such differences would contribute to our understanding of the factors responsible for neural reserve in general and the preserved cognitive function found for older bilinguals in particular.

## Method

### Participants

Sixty-one healthy older adults were recruited from the community. Thirty-one (11 men) of these participants were determined to be bilingual and 30 (8 men) were determined to be monolingual based on an extensive background questionnaire called the *Language and Social Background Questionnaire* (LSBQ; Anderson et al., 2017). Anderson et al. (2017) provide a method for calculating summary factor scores from which bilingual status can be determined, however validation of this method has not yet been extended to older adults. We therefore report English speaking and understanding and second-language speaking and understanding scores for each group (see Table 1). Importantly, English scores were equivalent for the two groups but second-language scores were significantly different. Screening for bilingual status was conducted via telephone interview and participants who could not be reliably categorized as monolingual or bilingual did not take part in the study. All participants were right handed and had no history of heart disease, psychological or neurological disease, or other MRI contraindications (see Table 1 for descriptive statistics). Bilinguals were lifelong bilinguals who were residents of Canada at the time of testing. We also asked participants "were any periods in your life when you did not use your second language?" If so, "how long?" The majority of the bilingual participants continually used their second language (64%) throughout their lives, a relationship that emerged even more strongly in the matched sample (72%).

### Data acquisition

Participants were scanned using a Siemens Trio 3T scanner using a 32-channel head coil. Head movement was constrained with foam padding. High-resolution T1-weighted

anatomical scans were acquired for registration purposes with a magnetized-prepared rapid gradient echo sequence using the following parameters: TR = 1.9 s, TE = 2.52 ms, FOV = 25.6 cm$^2$, 256 × 256 matrix, 192 slices of 1-mm thickness.

DTI scans were whole-brain 64-direction diffusion weighted images with the following parameters: TR = 9 200 s, TE = 86 s mm −2, 73 transverse slices with 2 mm thickness, FOV = 192 mm.

### Tract-Based-Spatial-Statistics (TBSS)

We performed a Voxelwise statistical analysis of the FA data employing Tract-Based Spatial Statistics (TBSS; Smith et al., 2006) included in FSL (Smith et al., 2004). Once FA images were generated by fitting a tensor model to the raw diffusion data using FDT, they were brain-extracted using Brain Extraction Toolbox (Smith, 2002). Following this, the nonlinear registration tool was applied to align the FA data from all subjects in a common space (Andersson et al., 2007a, 2007b), obtaining a b-spline representation of the registration warp field (Rueckert et al., 1999). Next the mean FA image was created and thinned so that a mean FA skeleton was obtained, representing the centers of tracts common to all participants. Finally aligned FA data from each participant was projected onto this skeleton and fed into voxelwise cross-subject statistics. We applied the same methodology to extract and compare RD and AD data, and tracts were identified post-hoc using the Johns Hopkins University DTI based probabilistic white matter atlas included with FSL (e.g., Mori et al., 2005).

### Propensity score analysis

All participants completed the D-KEFS battery (Delis et al., 2001). The D-KEFS battery was selected as a well-normed extensive battery covering a diverse array of frontal-lobe dependent cognitive processes including flexibility of thinking, inhibitory control, problem solving, planning, and impulse control (Homack et al., 2005). Data from the Trail-Making-Task (TMT), the Letter-Fluency-Task (LFT), and the Color-Word-Interference-Task (CWIT) are presented in Table 1 along with demographic and IQ information (Shipley verbal and nonverbal, Shipley, 1940). Between groups *t*-tests were computed for each set of scores (*p* values not corrected for multiple comparisons), and these are noted on the table as asterisks (significance < 0.05).

As shown in Table 1, neuropsychological performance was not equivalent for the two language groups in that monolinguals obtained better scores than bilinguals, a difference that confounds any interpretation of the brain data. Conclusions about differences between groups in white matter integrity require that cognitive level for the groups is equivalent; in the absence of such equivalence group differences could reflect simple differences in aging or cognitive decline rather than experience-dependent differences in white matter structure. Therefore, an explicit matching procedure was used. Several criteria were used to select variables for inclusion in the matching procedure, the first of which was a difference in mean performance on a neuropsychological sub-score. The TMT letter-number-switching score representing mental flexibility and the verbal and nonverbal components of the Shipley IQ test met this criterion, providing three matching variables. An additional four matching

criteria that were included were demographic scores routinely used for matching – age, education, gender, Mini-Mental State Examination (MMSE; Folstein et al., 1975) – producing 7 matching variables in total.

Rather than using sequential bivariate matching as is commonly reported in the literature (i.e., testing for an age difference using a *t*-test, reporting a null difference and moving on to the next potentially confounding variable), we used propensity-score matching to account for multivariate interactions. Briefly, propensity score matching uses logistic regression to predict group membership probability and then matches individuals from one group to those in the other based on the propensity (probability) scores (Rosenbaum and Rubin, 1983). This method is preferable to statistically controlling for multiple confounding variables in the typically smaller samples found in neuroimaging (Austin and Steyerberg, 2015).

A strength of propensity score matching is its ability to control for the interactions between variables which may differ by group. We conducted the equivalent of univariate matching as Gold et al. (2013) and Luk et al. (2011) did by matching bilinguals and monolinguals on each variable separately. Only Shipley IQ (verbal and nonverbal) yielded a loss of participants from either group suggesting that only two variables were unmatched from this perspective. This led us to suspect that *combinations* of demographic variables may yield group differences; that is, interactions between variables in multivariate space may reveal differences invisible to sequential bivariate testing. To illustrate this point, we matched the groups using MMSE and Age using a formulation identical to the one described above. Individually, neither variable yielded group differences; $t(51.57) = -1.37$, $p = 0.17$ for MMSE, and $t(57.99) = 1.5$, $p = 0.13$ for Age, but including them together led to the identification of 18 participants to be dropped, producing two groups of 21 participants each (see Fig. 2A). The interpretation is that the ***interaction*** between MMSE and Age is different for the two groups and it is the interaction that affects performance. This point is demonstrated in that the correlation between MMSE and Age was different for the two groups: for monolinguals, $r = 0.09$, $p = 0.61$; for bilinguals, $r = -0.40$, $p = 0.026$. A William's test for differences between correlations revealed that these correlations were significantly different from each other, $z = 1.93$, $p = 0.05$. Thus, bilinguals showed the expected negative relationship between age and MMSE scores whereas monolinguals did not. It is possible that monolinguals who showed declines in MMSE scores developed mild cognitive impairment (MCI) and were no longer part of the group of older adults considered to be experiencing healthy aging, leaving only more intact older monolinguals and undermining the correlation between Age and MMSE.

Bilinguals, in contrast, could cope longer with MCI symptoms before diagnosis (Bialystok et al., 2014) so remained in the sample of healthy older adults. Fig. 2B shows the bivariate relationships between each of the variables in the unmatched samples. Most between-group differences were eliminated using the propensity matching procedure.

If matching for neuropsychological performance eliminates effects in the DTI outcome measures, then group differences cannot be attributed to bilingualism. Conversely if matching enhances the between-group differences, we then can conclude that other factors were confounding the results and that the between-group differences are larger than might be

expected if careful matching were not conducted. Comparing data pre- and post-matching is novel in neuroimaging studies of bilingualism; although most studies claim to carefully match groups, none shows how this manipulation affects the data before and after matching. Finally, we compared the matched output from TBSS with analyses of the whole sample where these same variables were held constant via statistical control (i.e., were included in the linear model as covariates of no interest). We predicted that controlling for confounds using matching and adding these terms to the linear model would yield similar results.

As a first approach, we used propensity score analysis from the MatchIt R package to match groups of monolinguals and bilinguals. K-means nearest neighbor matching was then used to select two closely matched groups based on the propensity scores from the 7 selected variables. The formula used for matching was: matchit(formula = Group ~ TrailMaking-Task + MMSE + ShipBl + ShipV + Age + Gender + Education, data = TBSS, method = "nearest", discard = "treat"). The discard command removed bilingual individuals who were significantly different from the distribution of propensity scores of the monolingual participants. The remaining 23 bilinguals were then matched with the best subset of 23 monolinguals (see Fig. 1).

## Results

### Tract Based Spatial Statistics (TBSS)

Each of the following analyses was run using TBSS in FSL. Briefly, these between group analyses treat each voxel as independent and compare the group mean difference to a permuted null distribution to determine significance. Where the differences are adjusted for covariates, group differences above and beyond the influence of confounds were of interest.

The first analysis was run on the full sample prior to the matching procedure and revealed that the monolingual group had higher FA values than bilinguals, predominantly in the right hemisphere. This difference was found in the internal capsule, the anterior corpus callosum, the corona radiata and the inferior and superior longitudinal fasciculi. Bilinguals, in contrast, showed widespread RD at significantly greater levels than monolinguals in nearly all white matter brain regions. Bilinguals also had greater AD than monolinguals, particularly in the left hemisphere, likely contributing to the lack of significance of the FA contrast in that region. These results are shown in Fig. 3 Panel A, and coordinates are located in Table 2.

The second analysis was based on the propensity score matched samples and the results are shown in Fig. 3 Panel B. In this case, neither FA nor RD yielded significant clusters, but AD continued to reveal group differences in the same direction as found for the whole sample. Specifically, there were higher AD values for bilinguals than monolinguals in bilateral superior posterior corona radiata, the right external capsule, the midbody and splenium of the corpus callosum, the left superior temporal longitudinal fasciculus, and the anterior inferior frontal occipital fasciculus.

In the third analysis, conducted on the whole sample, the same analyses were used as previously but the 7 variables that had been used for propensity matching were entered as covariates in the analysis. These results revealed significant group differences only in AD in

the left superior temporal longitudinal fasciculus in a similar region also shown to be significant in the matched sample. These results are shown in Fig. 3 panel C. Using a covariate is a more stringent approach than matching because the group analysis is limited to examining residuals. In comparison, the matching procedure allows the group differences to examine the original variable space within the confines of a carefully matched sample. We suggest, therefore, that the matching procedure is more appropriate for neuroimaging studies of this sort but we report all the analyses here for completion.

## Discussion

The present study was designed to investigate conflicting findings regarding white matter integrity in older adult monolinguals and bilinguals. The results showed that when samples were unmatched, monolinguals displayed greater fractional anisotropy (FA) than bilinguals, and bilinguals displayed greater radial (RD) and axial (AD) diffusivity than monolinguals. However, when these groups were explicitly matched on seven background variables (Verbal and Spatial IQ, Age, Education, TMT, MMSE, and gender) using either a multivariate matching procedure (i.e., propensity score matching), or statistically controlled by entering the seven variables together as covariates, only the AD findings remained. Furthermore, sequential univariate techniques for matching (i.e., arguing for a lack of group differences based on *t*-tests for each variable) were insufficient, as they did not account for interactions between variables. These findings are discussed in the context of greater neural reserve for bilinguals than monolinguals and the importance of multivariate matching procedures in neuroimaging studies.

The idea that bilingualism leads to structural and functional brain adaptation is increasingly supported by evidence from studies of both grey matter volume (e.g., Abutalebi et al., 2015a, 2015b; Wei et al., 2015) and functional MRI (e.g., Rodríguez-Pujadas et al., 2014; Waldie et al., 2009). However, only two studies have examined white matter integrity in older adult monolinguals and bilinguals, and these studies yielded conflicting results. Consistent with the Gold et al. (2013) findings, the comparison of unmatched data in the present study showed greater FA and lower RD for monolinguals than bilinguals, a pattern associated with better white matter integrity for monolinguals. However, when a multivariate matching procedure was applied to match the samples on background measures, both the FA and the RD findings were eliminated, suggesting that confounds from these other measures were producing the differences. In contrast, bilinguals showed greater AD than monolinguals in both the matched and unmatched samples, a difference that could not be attributed to variation in the other background measures. This finding is consistent with the results of Luk et al. (2011) and fits with a neural reserve perspective in which lifelong bilingualism enhances white matter integrity in that AD is an index of diffusion along the primary gradient that is associated with positive cognitive outcomes (Urger et al., 2015). The idea of neural reserve is that some individuals, over a lifetime strengthen neural circuits and tissue providing a "cushion" against atrophy. Those without such protection decline at an accelerated rate and show symptoms of cognitive decline and dementia earlier. In contrast to neural reserve, cognitive reserve is thought to be resilience to neural insult. In this case, individuals with Alzheimer's pathology, for example, can remain symptom-free for longer than expected given the level of atrophy in their brains. It is thought that these individuals

have developed strategies that have strengthened alternative functional networks over a lifetime of practice.

It is important to note that evidence for the neural reserve hypothesis does not undermine a cognitive reserve perspective; the two accounts are not mutually exclusive. For example, proponents of the cognitive reserve perspective (e.g., Craik et al., 2010; Perani et al., 2017; Schweizer et al., 2012) usually include Alzheimer's disease patients in their studies whereas proponents of the neural reserve perspective typically recruit healthy older adults without disease progression (e.g., Abutalebi et al., 2015b; Li et al., 2017; Li et al., 2014; Olsen et al., 2015). Therefore, the theories largely describe different populations. It is also possible, as Gold et al. (2013) note, that some of the differences observed between the Luk et al. (2011) and Gold et al. (2013) studies arise due to a higher incidence of preclinical Alzheimer's disease in the bilingual sample. This is an interesting theory, and may be borne out by future replications. We note that previous studies have also shown that *immersion* may be important for explaining structural changes (e.g. Pliatsikas et al., 2015). While the majority of our bilingual participants were continuously immersed in both languages, it is possible that differences in immersion duration or characteristics between our sample and previously reported samples may account for some of the observed differences.

Neural reserve and cognitive reserve may work in tandem. There is some compelling evidence in the literature in line with the cognitive reserve hypothesis in which bilinguals are able to cope with more neurodegeneration than monolinguals. For example, Schweizer et al. (2012) showed that bilingual patients with Alzheimer's disease (AD) showed more brain atrophy in regions associated with the disease than monolinguals, despite equivalence on cognitive outcomes. More recently, Perani et al. (2017) examined monolingual and bilingual Alzheimer's disease patients that were matched for disease duration using positron emission tomography (PET). They showed that bilinguals were not only five years older than monolingual patients but also showed greater brain hypometabolism, which is a physiological index of the severity of Alzheimer's disease. These results suggest that bilinguals were able to cope with more diseased brains than monolinguals for longer periods of time before experiencing decline. It is possible that part of the adaptation allowing equivalent cognitive performance by bilinguals in the face of a greater degree of grey matter neurodegeneration than monolinguals is increased neural integrity in white matter tracts. Specifically, greater white matter integrity along the primary diffusion gradient (AD) might be a mechanism underlying reserve in bilinguals that facilitates communication between brain areas that are otherwise deteriorating. Thus, the combination of white matter integrity (Luk et al., 2011) and functional reorganization (Grady et al., 2015) might both contribute to a delay in cognitive decline for bilinguals relative to monolinguals.

The finding that consistently emerged across all the analyses was that bilinguals had greater AD in the left superior longitudinal fasciculus. The LSLF links the pars opercularis (Broca's area) with the receptive language areas in the temporal lobes. A case study of a tumor patient highlights the role the LSLF plays in language processing. This patient's tumor impinged on the LSLF, with symptoms manifesting as impairment in phonetic writing (Kana script). These symptoms resolved post surgery after the pressure was relieved (Shinoura et al., 2012). Corroborating evidence from DTI showed that the tract had been compressed by

the tumor. Greater mean diffusivity in LSLF has also been associated with a correspondingly more profound language deficit in autism spectrum disorder (Nagae et al., 2012). Given that this tract connects areas integral to the language network, it is not surprising that it can be remodeled by second-language experience. Notably, the LSLF is one of the tracts reported by Luk et al. (2011) as having greater FA for older bilinguals than older monolinguals. Similar findings were reported by Pliatsikas et al. (2015) who showed LSLF FA increases for bilinguals relative to monolinguals in a group of younger adults.

Somewhat surprisingly, after controlling for confounds, there were no group differences in the corpus callosum. Based on the literature, it was expected that bilinguals would have strengthened cross-hemispheric connections indexed by greater FA or AD in this region, although this was not the case. It is possible that this particular structure responds most plastically when a person is learning a second language, or using it in multiple contexts. Such a scenario would help to explain how children and young adults show remodeling of this region as language expertise develops, but once this expertise has reached a stable level, as in middle or older adult years, this callosal plasticity may no longer be evident.

Matching groups using a multivariate rather than a univariate method had a significant impact on the results, largely because matching on a single variable does not take into account possible interactions between the variables. This point was evident in the demonstration showing that entering MMSE scores or age into the model individually did not lead to the removal of any participants, but entering both variables into the same model led to the elimination of 9 participants from each group. This outcome suggests that the interaction of MMSE and age represented a significant confound in comparing the two groups, despite the inability to detect an influence of these variables when entered individually.

Matching on seven variables and their interactions revealed that only one aspect of the original results remained unchanged, namely, the finding that bilinguals had significantly higher AD than monolinguals. The original, unmatched findings that monolinguals had higher FA and lower RD values were eliminated. The outcome was confirmed through different statistical approaches. Using the seven variables as covariates and analyzing results from the whole sample produced similar results to those found in the propensity matched analysis, namely, higher AD for bilinguals than monolinguals with no other significant differences. However, lower thresholds were required in the latter method to see the full extent of overlap with the matched groups, suggesting that co-varying-out confounds likely requires more power. It is typically recommended that for each covariate in an analysis, N should be increased by 30 (e.g., Austin and Steyerberg, 2015); clearly it would be difficult to include 210 participants in most MRI studies. Therefore, propensity score analysis represents an excellent compromise in moderately sized studies such as those common in the neuroimaging literature where multiple covariates may affect the outcome, but it is statistically difficult to control for them.

In attempting to integrate our findings with the young adult white matter literature, we find that it converges on the finding that bilingualism is associated with *increased* white matter integrity. The spatial convergence of these beneficial effects is less clear, though a notable

exception is the anterior corpus callosum (see Grundy et al., 2017 for a recent review). One possible reason for the spatial inconsistency is the variety of methods that have been used to analyze white matter across studies. Some studies report volumetric data (e.g. voxel-based-morphometry) across the entire brain, others, including ourselves, report data restricted to a white-matter skeleton, and still other studies report significant group differences collapsed across entire tracts. It is hard to see how to directly compare findings across such widely differing methods. A second reason for spatial divergence is that while most studies include brain images, very few report spatial coordinates making it difficult to quantitatively synthesize the literature and answer questions about whether two studies reporting on similar regions are, in fact, referring to the same structure. In the present paper, we highlight a third possible reason for spatial inconsistencies: namely how matching – or lack of matching, between groups – affects outcomes.

In sum, we provide evidence that lifelong bilingualism leads to greater AD in healthy older bilinguals compared to monolinguals. This result persisted even after carefully controlling for multiple confounding variables and their interactions. These findings may help to explain why bilinguals show later cognitive decline than monolinguals in older age: second-language experience contributes to neural reserve.

## Acknowledgments

## References

Abutalebi J, Canini M, Della Rosa PA, Green DW, Weekes BS. The neuroprotective effects of bilingualism upon the inferior parietal lobule: a structural neuroimaging study in aging Chinese bilinguals. J Neurolinguistics. 2015a; 33:3–13.

Abutalebi J, Guidi L, Borsa V, Canini M, Della Rosa PA, Parris BA, Weekes BS. Bilingualism provides a neural reserve for aging populations. Neuropsychologia. 2015b; 69:201–210. [PubMed: 25637228]

Alladi S, Bak TH, Duggirala V, Surampudi B, Shailaja M, Shukla AK, et al. Bilingualism delays age at onset of dementia, independent of education and immigration status. Neurology. 2013; 81(22): 1938–1944. [PubMed: 24198291]

Anderson JAE, Mak L, Keyvani Chahi A, Bialystok E. The language and social background questionnaire: assessing degree of bilingualism in a diverse population. Behav Res Methods. 2017:1–14. [PubMed: 26660195]

Andersson, JL., Jenkinson, M., Smith, S. Non-linear Optimisation FMRIB technical report TR07JA1 University of Oxford. FMRIB Centre; Oxford, UK: 2007a.

Andersson, JL., Jenkinson, M., Smith, S. Non-linear Registration, Aka Spatial Normalisation FMRIB Technical Report TR07JA2. FMRIB Analysis Group of the University of Oxford; 2007b. p. 2

Austin PC, Steyerberg EW. The number of subjects per variable required in linear regression analyses. J Clin Epidemiol. 2015; 68(6):627–636. [PubMed: 25704724]

Barrett, Ashley, Strait & Kraus, (2013) with the following entry in the works cited:Barrett, KC., Ashley, R., Strait, DL., Kraus, N. Frontiers in psychology. 2013. Art and science: how musical training shapes the brain; p. 4Bartzokis G, Sultzer D, Lu PH, Nuechterlein KH, Mintz J, Cummings JL. Heterogeneous age-related breakdown of white matter structural integrity: implications for cortical "disconnection" in aging and Alzheimer's disease. Neurobiol Aging. 2004; 25(7):843–851. [PubMed: 15212838]

Bialystok E. The bilingual adaptation: how minds accommodate experience. Psychol Bull. 2017; 143(3):233–262. [PubMed: 28230411]

Bialystok E, Craik FI, Freedman M. Bilingualism as a protection against the onset of symptoms of dementia. Neuropsychologia. 2007; 45(2):459–464. [PubMed: 17125807]

Bialystok E, Craik FIM, Binns MA, Ossher L, Freedman M. Effects of bilingualism on the age of onset and progression of MCI and AD: evidence from executive function tests. Neuropsychology. 2014; 28:290–304. [PubMed: 24245925]

Bozzali M, Falini A, Franceschi M, Cercignani M, Zuffi M, Scotti G, et al. White matter damage in Alzheimer's disease assessed in vivo using diffusion tensor magnetic resonance imaging. J Neurology, Neurosurg Psychiatry. 2002; 72(6):742–746.

Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM. Forecasting the global burden of Alzheimer's disease. Alzheimers Dementia. 2007; 3:186–191.

Coggins PE III, Kennedy TJ, Armstrong TA. Bilingual corpus callosum variability. Brain Lang. 2004; 89:69–75. [PubMed: 15010238]

Craik FIM, Bialystok E, Freedman M. Delaying the onset of Alzheimer disease: bilingualism as a form of cognitive reserve. Neurology. 2010; 75(19):1726–1729. [PubMed: 21060095]

Delis, DC., Kaplan, E., Kramer, JH. Delis-Kaplan Executive Function System (D-KEFS). Psychological Corporation; 2001.

Felton A, Vazquez D, Ramos-Nunez AI, Greene MR, Macbeth A, Hernandez AE, Chiarello C. Bilingualism influences structural indices of interhemispheric organization. J Neurolinguistics. 2017; 42:1–11. [PubMed: 28579694]

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. J Psychiatric Res. 1975; 12(3):189–198.

Gold BT, Johnson NF, Powell DK. Lifelong bilingualism contributes to cognitive reserve against white matter integrity declines in aging. Neuropsychologia. 2013; 51:2841–2846. [PubMed: 24103400]

Grady CL, Luk G, Craik FIM, Bialystok E. Brain network activity in monolingual and bilingual older adults. Neuropsychologia. 2015; 66:170–181. [PubMed: 25445783]

Grundy JG, Anderson JAE, Bialystok E. Neural correlates of cognitive processing in monolinguals and bilinguals. Ann N Y Acad Sci. 2017

Homack S, Lee D, Riccio CA. Test review: Delis-Kaplan executive function system. J Clin Exp neuropsychology. 2005; 27(5):599–609.

Jones DK, Knosche TR, Turner R. White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. Neuroimage. 2013; 73:239–254. [PubMed: 22846632]

Li, L., Abutalebi, J., Emmorey, K., Gong, G., Yan, X., Feng, X., et al. How bilingualism protects the brain from aging: insights from bimodal bilinguals. Hum Brain Mapp Early View. 2017. https://doi.org/10.1002/hbm.23652

Li P, Legault J, Litcofsky KA. Neuroplasticity as a function of second language learning: anatomical changes in the human brain. Cortex. 2014; 58:301–324. [PubMed: 24996640]

Luk G, Bialystok E, Craik FIM, Grady CL. Lifelong bilingualism maintains white matter integrity in older adults. J Neurosci. 2011; 31:16808–16813. [PubMed: 22090506]

Mori, S., Wakana, S., Van Zijl, PC., Nagae-Poetscher, LM. MRI Atlas of Human White Matter. Elsevier; 2005.

Nagae LM, Zarnow DM, Blaskey L, Dell J, Khan SY, Qasmieh S, et al. Elevated mean diffusivity in the left hemisphere superior longitudinal fasciculus in autism spectrum disorders increases with more profound language impairment. Am J Neuroradiol. 2012; 33(9):1720–1725. [PubMed: 22492573]

Olsen RK, Pangelinan MM, Bogulski C, Chakravarty MM, Luk G, Grady CL, Bialystok E. The effect of lifelong bilingualism on regional grey and white matter volume. Brain Res. 2015; 1612:128–139. [PubMed: 25725380]

Perani D, Farsad M, Ballarini T, Lubian F, Malpetti M, Fracchetti A, et al. The impact of bilingualism on brain reserve and metabolic connectivity in Alzheimer's dementia. Proc Natl Acad Sci. 2017; 114(7):1690–1695. [PubMed: 28137833]

Pliatsikas C, Moschopoulou E, Saddy JD. The effects of bilingualism on the white matter structure of the brain. Proc Natl Acad Sci. 2015; 112(5):1334–1337. [PubMed: 25583505]

Rodríguez-Pujadas A, Sanjuán A, Fuentes P, Ventura-Campos N, Barrós-Loscertales A, Ávila C. Differential neural control in early bilinguals and monolinguals during response inhibition. Brain Lang. 2014; 132:43–51. [PubMed: 24735970]

Rose SE, Chen F, Chalk JB, Zelaya FO, Strugnell WE, Benson M, et al. Loss of connectivity in Alzheimer's disease: an evaluation of white matter tract integrity with colour coded MR diffusion tensor imaging. J Neurology, Neurosurg Psychiatry. 2000; 69(4):528–530.

Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 70:41–55.

Rueckert D, Sonoda LI, Hayes C, Hill DL, Leach MO, Hawkes DJ. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Trans Med Imaging. 1999; 18(8):712–721. [PubMed: 10534053]

Schlegel AA, Rudelson JJ, Peter UT. White matter structure changes as adults learn a second language. J Cognitive Neurosci. 2012; 24:1664–1670.

Schweizer TA, Ware J, Fischer CE, Craik FIM, Bialystok E. Bilingualism as a contributor to cognitive reserve: evidence from brain atrophy in Alzheimer's disease. Cortex. 2012; 48(8):991–996. [PubMed: 21596373]

Schlaug G, Jancke L, Huang Y, Steinmetz H. In vivo evidence of structural brain asymmetry in musicians. Sci - NY then Washingt. 1995:699.

Shinoura N, Midorikawa A, Onodera T, Yamada R, Tabei Y, Onda Y, et al. The left superior longitudinal fasciculus within the primary sensory area of inferior parietal lobe plays a role in dysgraphia of kana omission within sentences. Behav Neurol. 2012; 25(4):363–368. [PubMed: 22713399]

Shipley WC. A self-administering scale for measuring intellectual impairment and deterioration. J Psychol. 1940; 9(2):371–377.

Smith SM. Fast robust automated brain extraction. Hum Brain Mapp. 2002; 17(3):143–155. [PubMed: 12391568]

Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage. 2006; 31(4): 1487–1505. [PubMed: 16624579]

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, et al. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage. 2004; 23:S208–S219. [PubMed: 15501092]

Statistics Canada. Table 051-0001-Estimates of population, by age group and sex for July 1, Canada, provinces and territories, annual (persons unless otherwise noted), CANSIM (database). (accessed: June 6 2017).

Urger SE, De Bellis MD, Hooper SR, Woolley DP, Chen SD, Provenzale J. The superior longitudinal fasciculus in typically developing children and adolescents: diffusion tensor imaging and neuropsychological correlates. J Child Neurology. 2015; 30(1):9–20.

Waldie KE, Badzakova-Trajkov G, Miliivojevic B, Kirk IJ. Neural activity during Stroop colour-word task performance in late proficient bilinguals: a functional Magnetic Resonance Imaging study. Psychol Neurosci. 2009; 2(2):125–136.

Wan CY, Schlaug G. Music making as a tool for promoting brain plasticity across the life span. Neurosci. 2010; 16(5):566–577.

Wei M, Joshi AA, Zhang M, Mei L, Manis FR, He Q, et al. How age of acquisition influences brain architecture in bilinguals. J Neurolinguistics. 2015; 36:35–55. [PubMed: 27695193]
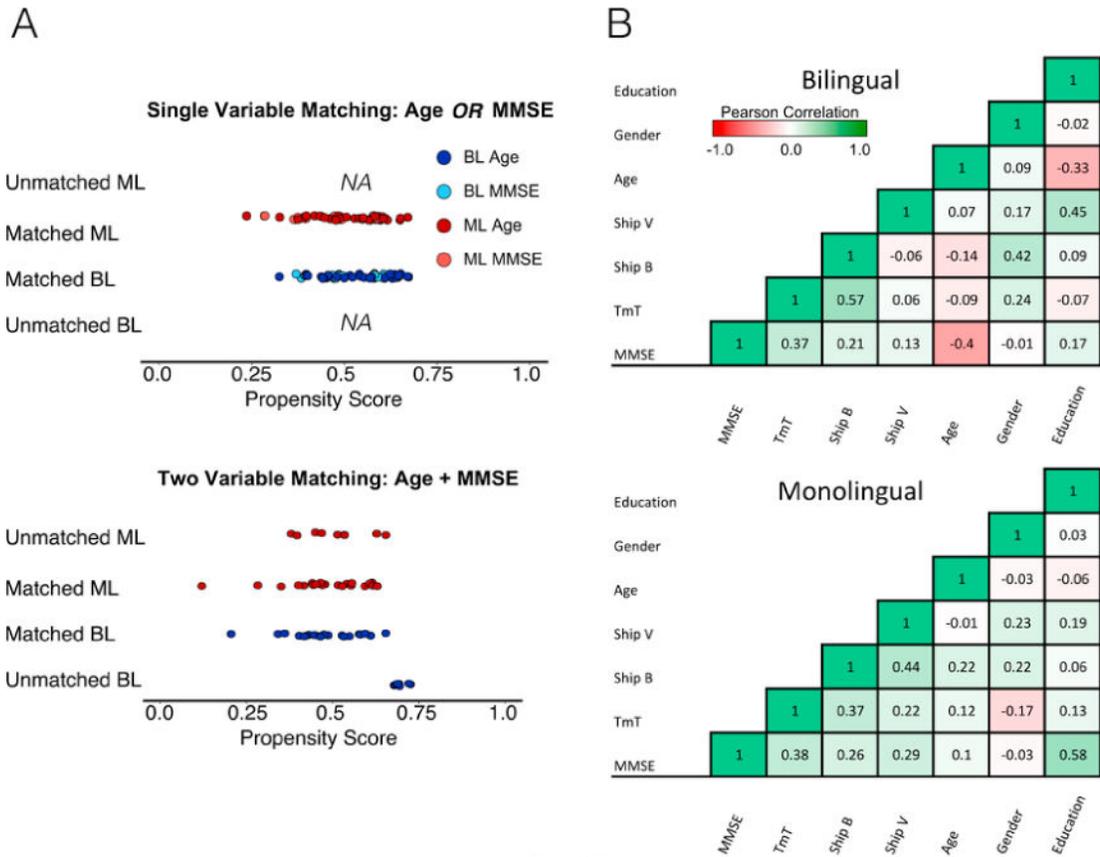
**Fig. 1.**
Effects of propensity score matching on single versus multiple variables. The left side of the Figure, Panel A, shows the effect of matching on either MMSE or age. The final propensity score analysis shows the effects of including both scores together. Only when multiple variables are entered into the matching procedure do participants start being removed. This effect is driven by the interactions between variables. Panel B shows the bivariate relationships between variables by group.
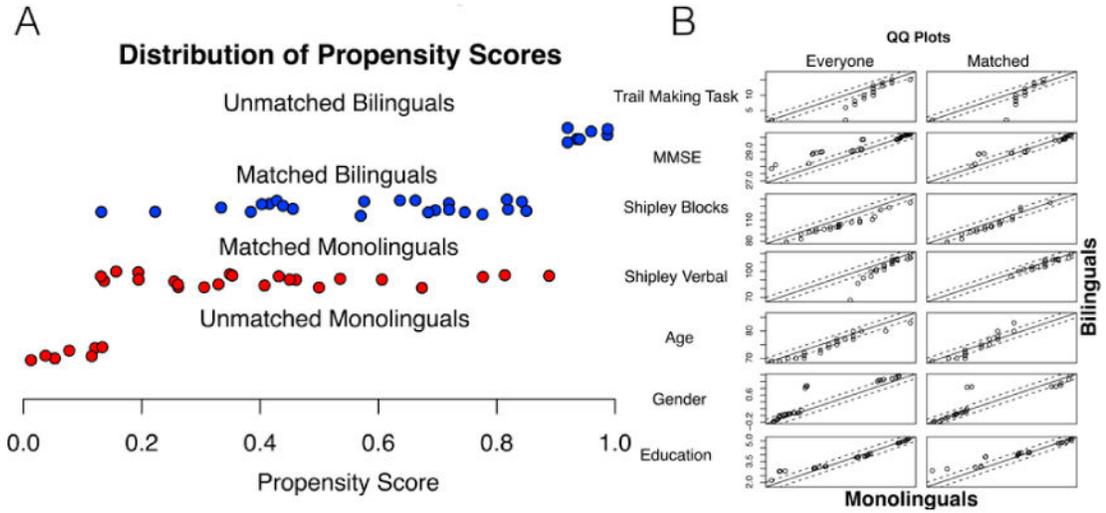
**Fig. 2.**
Distribution of Propensity Scores. Groups were matched using 7 measures (Trail-making [letter-number], MMSE, Shipley Verbal, Shipley Blocks, age, gender, and education) and k-means-nearest neighbors using the MatchIt package in R. Panel A shows the range of propensity scores, Panel B shows quantile-quantile plots for each of the measures in the unmatched and matched samples. Scores by quantile in the monolingual group were used to predict scores by quantile in the bilingual group. Deflections above or below the line correspond to systematic group differences, and a perfect relationship between the groups (i.e., no difference on this measure) is reflected by the degree to which the measure follows the line of union.
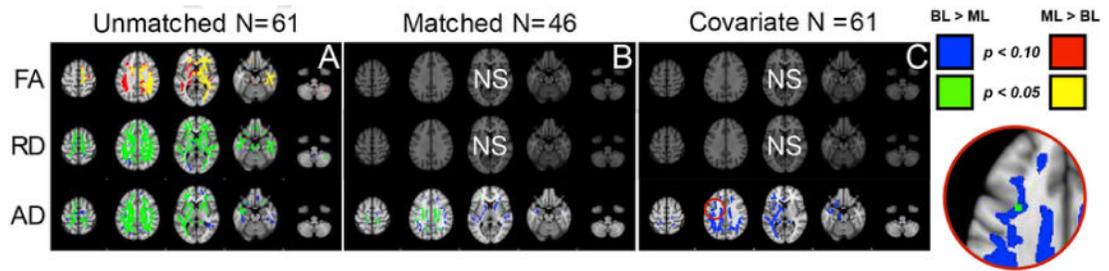
**Fig. 3.**
TBSS group comparisons for FA, RD, and AD. Panel A depicts the unmatched TBSS analysis, Panel B depicts results for the matched sample, and Panel C depicts results using seven covariates. Blue and red depict marginally significant group differences; green and yellow depict significant group differences (see legend for direction). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Anderson et al.

Page 17

**Table 1**

Demographic and Neuropsychological Measures. Means and SDs (in brackets) are displayed. Significant differences $p < 0.05$ between groups (uncorrected for multiple comparisons) are indicated by *. A superscript M next to a variable's name indicates it was used for propensity score matching.

| | Unmatched (Full Sample) | | | | | | Matched Sample | | | | | |
| | BL (N = 31) | | ML (N = 30) | | | | BL (N = 23) | | ML (N = 23) | | | |
| | M | SD | M | SD | p | | M | SD | M | SD | p |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Demographics** | | | | | | | | | | | |
| Age^M | 74.00 | 4.10 | 75.43 | 3.93 | 0.20 | | 74.09 | 4.28 | 74.43 | 2.95 | 0.75 |
| Education^M | 3.96 | 0.87 | 3.83 | 1.02 | 0.58 | | 4.13 | 0.76 | 3.96 | 0.98 | 0.50 |
| MMSE^M | 29.42 | 0.72 | 29.17 | 0.99 | 0.26 | | 29.43 | 0.66 | 29.26 | 0.86 | 0.45 |
| Gender (n Males)^M | 8 | | 11 | | 0.46 | | 8 | | 6 | | 0.53 |
| **LSBQ Scores** | | | | | | | | | | | |
| English Speaking | 9.23 | 1.02 | 9.55 | 0.59 | 0.13 | | 9.39 | 0.99 | 9.46 | 1.02 | 0.79 |
| English Understanding | 9.27 | 0.99 | 9.62 | 0.58 | 0.11 | | 9.37 | 1.07 | 9.54 | 0.99 | 0.50 |
| Second Language Speaking | 7.84 | 1.71 | 1.28 | 1.39 | 0.00* | | 7.61 | 1.76 | 1.21 | 1.71 | 0.00* |
| Second Language Understanding | 8.16 | 1.80 | 1.78005 | 1.78 | 0.00* | | 7.89 | 1.86 | 1.81 | 1.80 | 0.00* |
| Age Learned L2 | 3.03 | 4.88 | 8.75 | 6.21 | 0.00* | | 2.41 | 4.50 | 8.69 | 6.10 | 0.00* |
| Proportion of participants who did not use L2 for an extended period of time | 35.48% | 48.64% | 90.00% | 30.78% | 0.00* | | 27.27% | 45.58% | 87.50% | 34.16% | 0.00* |
| For participants not using a second language, for how long was this (in years)? | 12.82 | 12.98 | 64.31 | 16.13 | 0.00* | | 8.83 | 4.12 | 60.86 | 17.94 | 0.00* |
| **Shipley** | | | | | | | | | | | |
| Verbal^M | 106.20 | 10.55 | 111.27 | 5.53 | 0.02* | | 109.52 | 6.27 | 110.22 | 5.74 | 0.70 |
| Blocks^M | 101.90 | 12.10 | 110.00 | 13.73 | 0.02* | | 102.91 | 13.58 | 106.91 | 11.02 | 0.28 |
| Composite | 105.20 | 9.27 | 112.20 | 9.91 | 0.01* | | 107.78 | 8.01 | 109.87 | 8.37 | 0.39 |
| **Trail Making Task** | | | | | | | | | | | |
| Number Sequencing | 11.29 | 3.47 | 13.00 | 2.03 | 0.02* | | 11.61 | 3.24 | 12.87 | 2.24 | 0.13 |
| Letter Number Switching^M | 10.74 | 3.46 | 12.53 | 1.89 | 0.02* | | 10.48 | 3.89 | 12.43 | 2.04 | 0.04* |
| Switching errors | 11.12 | 1.80 | 11.67 | 0.66 | 0.12 | | 11.04 | 2.01 | 11.70 | 0.70 | 0.15 |
| **Verbal Fluency Task** | | | | | | | | | | | |
| Letter Fluency | 12.52 | 3.27 | 12.93 | 2.98 | 0.61 | | 12.65 | 3.52 | 12.65 | 2.89 | 1.00 |
| Category Fluency | 10.65 | 3.65 | 12.53 | 3.88 | 0.05* | | 11.39 | 3.55 | 12.17 | 3.71 | 0.47 |

| | Unmatched (Full Sample) | | | | | Matched Sample | | | | | |
| | BL (N = 31) | | ML (N = 30) | | | BL (N = 23) | | ML (N = 23) | | |
| | M | SD | M | SD | p | M | SD | M | SD | p |
|---|---|---|---|---|---|---|---|---|---|---|
| Category switching (total correct) | 10.03 | 2.94 | 11.87 | 3.31 | 0.03* | 10.30 | 2.95 | 11.43 | 3.38 | 0.23 |
| Category switching (total switching accuracy) | 10.74 | 2.73 | 12.00 | 2.77 | 0.08 | 11.13 | 2.70 | 11.70 | 2.88 | 0.50 |
| Percent set-loss errors | 11.58 | 1.57 | 11.30 | 1.82 | 0.52 | 11.17 | 1.61 | 11.04 | 1.94 | 0.81 |
| Percent repetition errors | 9.65 | 2.65 | 10.90 | 1.81 | 0.04* | 9.78 | 2.66 | 10.87 | 1.84 | 0.11 |
| Percent switching accuracy | 12.65 | 0.66 | 12.93 | 0.25 | 0.03* | 12.57 | 0.73 | 12.91 | 0.29 | 0.04* |
| **Stroop Task** | | | | | | | | | | |
| Color naming | 9.87 | 2.53 | 11.07 | 2.18 | 0.05* | 10.17 | 2.61 | 11.09 | 2.17 | 0.20 |
| Word reading | 10.71 | 2.88 | 11.07 | 2.27 | 0.59 | 11.17 | 2.27 | 10.96 | 2.38 | 0.75 |
| Inhibition | 13.32 | 10.71 | 12.20 | 2.38 | 0.58 | 11.57 | 2.15 | 12.00 | 2.49 | 0.53 |
| Inhibition/switching | 11.29 | 2.84 | 12.00 | 2.44 | 0.30 | 11.74 | 2.73 | 11.87 | 2.51 | 0.87 |
| Inibition/Switching vs. inhibition | 9.81 | 2.50 | 9.83 | 2.55 | 0.97 | 10.13 | 2.18 | 9.91 | 2.73 | 0.77 |
| Inhibition errors | 11.94 | 1.26 | 11.83 | 1.98 | 0.81 | 12.04 | 1.11 | 11.70 | 2.18 | 0.50 |
| Inhibition/Switching errors | 11.23 | 1.87 | 11.17 | 2.12 | 0.91 | 11.43 | 1.67 | 11.17 | 2.06 | 0.64 |

**Table 2**

Clusters exceeding threshold for significance for analyses controlling for confounds. CA = Covariate Analysis, MS = Matched Sample.

| Image | X | Y | Z | Hemisphere | Region |
|---|---|---|---|---|---|
| CA | −31 | 7 | 17 | L | Superior longitudinal fasciculus |
| MS | −27 | −30 | −4 | | Fornix (cres) |
| MS | −23 | −31 | −1 | | Fornix (cres) |
| MS | −14 | −1 | 36 | | Body of corpus callosum |
| MS | −12 | −40 | 28 | | Splenium of corpus callosum |
| MS | −10 | −5 | 32 | | Body of corpus callosum |
| MS | −7 | −21 | 25 | | Body of corpus callosum |
| MS | −7 | 25 | −2 | | Genu of corpus callosum |
| MS | 2 | 6 | 23 | | Body of corpus callosum |
| MS | 4 | −37 | 15 | | Splenium of corpus callosum |
| MS | 6 | −26 | 24 | | Body of corpus callosum |
| MS | 10 | 11 | 24 | | Body of corpus callosum |
| MS | 16 | 18 | 26 | | Body of corpus callosum |
| MS | −41 | −41 | −5 | L | Sagittal stratum (include inferior longitudinal fasciculus and inferior fronto-occipital fasciculus) |
| MS | −37 | −53 | 14 | L | Superior longitudinal fasciculus |
| MS | −37 | −45 | 2 | L | Posterior thalamic radiation (include optic radiation) |
| MS | −36 | −44 | 6 | L | Posterior thalamic radiation (include optic radiation) |
| MS | −35 | −42 | 7 | L | Posterior thalamic radiation (include optic radiation) |
| MS | −34 | −9 | −10 | L | External capsule |
| MS | −33 | −23 | 0 | L | Retrolenticular part of internal capsule |
| MS | −33 | −19 | −2 | L | External capsule |
| MS | −33 | −8 | −9 | L | External capsule |
| MS | −33 | −2 | 6 | L | External capsule |
| MS | −33 | −1 | 28 | L | Superior longitudinal fasciculus |
| MS | −32 | −22 | 1 | L | External capsule |
| MS | −32 | −5 | 24 | L | Superior longitudinal fasciculus |
| MS | −31 | −35 | 36 | L | Superior longitudinal fasciculus |
| MS | −30 | 10 | 4 | L | External capsule |

| Image | X | Y | Z | Hemisphere | Region |
|---|---|---|---|---|---|
| MS | −29 | −32 | 13 | L | Retrolenticular part of internal capsule |
| MS | −27 | −62 | 15 | L | Posterior thalamic radiation (include optic radiation) |
| MS | −27 | 11 | 27 | L | Superior corona radiata |
| MS | −26 | −27 | 17 | L | Retrolenticular part of internal capsule |
| MS | −26 | 13 | 27 | L | Superior corona radiata |
| MS | −25 | 15 | −10 | L | External capsule |
| MS | −23 | −40 | 35 | L | Posterior corona radiata |
| MS | −21 | 5 | 16 | L | Anterior limb of internal capsule |
| MS | −20 | −40 | 33 | L | Posterior corona radiata |
| MS | −20 | −28 | 38 | L | Posterior corona radiata |
| MS | −20 | 23 | −8 | L | Anterior corona radiata |
| MS | −18 | −7 | 39 | L | Superior corona radiata |
| MS | −18 | −5 | 37 | L | Superior corona radiata |
| MS | −18 | −4 | 8 | L | Posterior limb of internal capsule |
| MS | −18 | 26 | 27 | L | Anterior corona radiata |
| MS | −17 | −12 | −7 | L | Cerebral peduncle |
| MS | −16 | 15 | 1 | L | Anterior limb of internal capsule |
| MS | −11 | −24 | −11 | L | Cerebral peduncle |
| MS | 11 | 6 | 1 | L | Anterior limb of internal capsule |
| MS | 12 | 2 | 3 | L | Anterior limb of internal capsule |
| MS | 14 | −1 | 4 | L | Anterior limb of internal capsule |
| MS | 16 | −2 | 7 | L | Posterior limb of internal capsule |
| MS | 30 | −22 | 38 | L | Superior longitudinal fasciculus |
| MS | 33 | 6 | 20 | L | Superior longitudinal fasciculus |
| MS | 34 | −44 | 31 | L | Superior longitudinal fasciculus |
| MS | 34 | 0 | 29 | L | Superior longitudinal fasciculus |
| MS | 35 | −48 | 22 | L | Superior longitudinal fasciculus |
| MS | 35 | 6 | 21 | L | Superior longitudinal fasciculus |
| MS | 42 | −20 | 31 | L | Superior longitudinal fasciculus |
| MS | 11 | −47 | 23 | R | Cingulum (cingulate gyrus) |
| MS | 17 | 22 | 26 | R | Anterior corona radiata |

| Image | X | Y | Z | Hemisphere | Region |
|---|---|---|---|---|---|
| MS | 17 | 26 | 22 | R | Anterior corona radiata |
| MS | 18 | 21 | 26 | R | Anterior corona radiata |
| MS | 18 | 23 | 26 | R | Anterior corona radiata |
| MS | 19 | −28 | 35 | R | Posterior corona radiata |
| MS | 19 | 11 | 34 | R | Superior corona radiata |
| MS | 20 | 11 | 31 | R | Superior corona radiata |
| MS | 21 | −6 | 36 | R | Superior corona radiata |
| MS | 25 | 5 | 34 | R | Superior corona radiata |
| MS | 25 | 16 | 16 | R | Anterior corona radiata |
| MS | 27 | 32 | 4 | R | Anterior corona radiata |
| MS | 27 | 34 | 4 | R | Anterior corona radiata |
| MS | 32 | −37 | 15 | R | Retrolenticular part of internal capsule |
| MS | 34 | 5 | −7 | R | External capsule |
| MS | 35 | −10 | −4 | R | External capsule |