Duration and extent of bilingual experience modulate neurocognitive outcomes

Article

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1016/j.neuroimage.2019.116222

Duration and extent of bilingual experience modulate neurocognitive outcomes

Vincent DeLuca[1]*, Jason Rothman[2,3], Ellen Bialystok[4], Christos Pliatsikas[3,5]

1. *School of Psychology, University of Birmingham, 52 Pritchatts Road, Birmingham, B15 2SA, UK*
2. *Department of Language and Culture, The University of Tromsø, Hansine Hansens veg 18, 9019 Tromsø, Norway*
3. *Facultad de Lenguas y Educación, Universidad Nebrija, Calle de Sta. Cruz de Marcenado, 27, 28015 Madrid, Spain*
4. *Department of Psychology, York University, 4700 Keele St, York, ON M3J 1P3, Canada*
5. *School of Psychology and Clinical Language Sciences, University of Reading, Harry Pitt Building, Whiteknights, Reading, RG6 6BE, UK*

*corresponding author

Abstract

The potential effects of bilingualism on executive control (EC) have been heavily debated. One possible source of discrepancy in the evidence may be that bilingualism tends to be treated as a monolithic category distinct from monolingualism. We address this possibility by examining the effects of different bilingual language experiences on brain activity related to EC performance. Participants were scanned (fMRI) while they performed a Flanker task. Behavioral data showed robust Flanker effects, not modulated by language experiences across participants. However, differences in duration of bilingual experience and extent of active language use predicted activation in distinct brain regions indicating differences in neural recruitment across conditions. This approach highlights the need to consider specific bilingual language experiences in assessing neurocognitive effects. It further underscores the utility and complementarity of neuroimaging evidence in this general line of research, contributing to a deeper understanding of the variability reported in the literature.

Keywords

Bilingualism, executive control, fMRI, individual differences

1. Introduction

The effect of bilingualism on domain-general cognition, particularly executive control (EC) has been the focus of significant debate in recent years (Antoniou, 2019; Bialystok, 2017; Donnelly, Brooks, & Homer, 2015; Lehtonen et al., 2018; Paap, Johnson, & Sawi, 2015; Valian, 2015). While several studies have provided evidence that bilingualism has a positive effect on various aspects of EC, including suppression of interfering information, cognitive cost of task switching, and use of facilitatory information in performing a task (Hernández, Costa, Fuentes, Vivas, & Sebastián-Gallés, 2010; Veroude, Norris, Shumskaya, Gullberg, & Indefrey, 2010; Zhou & Krott, 2018), others find restricted effects of bilingualism (Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009; Hernández, Martin, Barceló, & Costa, 2013) or none at all (Antón et al., 2014; Kirk, Fiala, Scott-Brown, & Kempe, 2014; Paap & Greenberg, 2013). This replication issue has led to recent claims that bilingualism has no meaningful effect on EC overall (Klein, 2014; Paap et al., 2015).

Variability between studies, however, is to be expected when one considers the multifarious nature of the bilingual experiences of individuals and even groups of individuals (Bak, 2016). Failure to replicate findings is not inherently a critical problem to the extent that (some) systematicity can reveal what conditions modulate effects. Bilingualism is a complex and dynamic process which encompasses a range of experiences that may drive neurocognitive adaptations (Bak, 2016; Bialystok, 2016; Li, Legault, & Litcofsky, 2014; Luk & Bialystok, 2013), however investigating how neurocognitive effects correlate with individual language experiences remains understudied. We confront this challenge in the present paper, seeking to identify specific language experiences within bilingualism that contribute to domain-general neurocognitive adaptations and assess how and why these adaptations manifest differentially. Importantly, the study focuses on specific aspects of bilingual experience that differ across individuals; no monolinguals are included. Given the potential for unpacking the basis for variability of cognitive adaptations across studies, the present approach can have significant impact for the field.

Given overlap in the networks that serve executive and language control, increased demands on the language control system in bilingualism have been argued to underlie the reported effects in brain structures and networks associated with domain-general EC (Anderson, Chung-Fat-Yim, Bellana, Luk, & Bialystok, 2018; De Baene, Duyck, Brass, & Carreiras, 2015; Garbin et al., 2010). Crucially, however, bilingualism-induced neurocognitive adaptations are often reported in the absence of commensurate task-performance effects. This suggests that monolinguals and bilinguals reach the same performance levels on EC tasks but do so by recruiting different underlying networks (Abutalebi et al., 2012; Ansaldo, Ghazi-Saidi, & Adrover-Roig, 2015; Bialystok et al., 2005; Costumero, Rodríguez-Pujadas, Fuentes-Claramonte, & Ávila, 2015). These language group differences in neural recruitment during EC tasks typically manifest in one of two ways. The first is a different spatial distribution of activations for the two language groups (Ansaldo et al., 2015; Bialystok et al., 2005; Costumero et al., 2015; Luk, Anderson, Craik, Grady, & Bialystok, 2010), indicating bilingual language control modulates the cognitive resources needed to handle the task demands. The second difference is decreased activation for bilinguals in the same regions used by monolinguals (Abutalebi et al., 2012), indicating decreased resources required for equivalent task performance.

These observations raise an important issue: most of the controversy in the literature on bilingual effects on cognition has arisen from behavioral studies utilizing binary comparisons

of monolinguals and bilinguals (Bak, 2016; Surrain & Luk, 2017). However, given the significant variation in quantity and quality of relevant experiences at the individual level, a binary approach collapses too much. A different approach, one that seeks to understand the relative contribution of distinct language experience to outcomes, is worth pursuing. The present study does just that by focusing on how bilingualism *itself* is defined for research purposes. Specifically, we endeavor to test the idea that bilingual effects on mind and brain are potentially conditioned, if not attenuated, by experiential factors related to the context of language exposure and opportunities for meaningful language use (DeLuca, Rothman, Bialystok, & Pliatsikas, 2019; Gullifer et al., 2018; Luo et al., 2019; Nichols & Joanisse, 2016).

General frameworks have been proposed to explain the nature of the neurocognitive accommodation to language experiences, dealing with the notion of changing brain network adaptations to accomplish the same behavioral task (Hernandez et al., 2019; Hernandez, Li, & MacWhinney, 2005; Hernandez & Li, 2007) and how these might manifest across the lifespan (Grant, Dennis, & Li, 2014). As a key example of this, the neuroemergentist perspective argues for the repurposing of brain regions and networks to most effectively and efficiently handle the cognitive demands imposed by bilingual language use (Hernandez, Claussenius-Kalman, Ronderos, & Vaughn, 2018).

Two models have been proposed which further this argument, by making more concrete predictions regarding the directionality and nature of these neurological effects. The Adaptive Control Hypothesis (ACH) (Abutalebi & Green, 2016; Calabria, Costa, Green, & Abutalebi, 2018; Green & Abutalebi, 2013) argues that the communicative context (single-language, dual-language, or dense code-switching) contributes to the recruitment of specific networks to handle language control demands imposed in that context. Situations of dense code switching, for example, where all languages one speaks are freely used, confer demands on the left inferior frontal gyrus (IFG) and cerebellum. A dual-language context in which all languages are available but only one is used at a given time imposes more control demands to select the correct language, necessitating a larger control network including the bilateral IFG, anterior cingulate cortex (ACC), thalamus, basal ganglia, cerebellum, and inferior parietal lobule. A single-language context, in which only one language is available, necessitates only selection and maintenance of the target language, placing demands on frontal control regions, specifically the left IFG. Finally, increased or continued engagement with one or more of these language use contexts will necessitate further engagement from the neural circuits required to handle them, leading to increased neural and cognitive adaptations.

A complementary approach is presented in the Bilingual Anterior to Posterior and Subcortical Shift (BAPSS) framework (Grundy, Anderson, & Bialystok, 2017). On that view, reliance on certain brain regions for control demands is modulated with continued second language (L2) exposure. Specifically, this reliance shifts from frontal regions involved in cognitive control, such as the regions within the dorsolateral prefrontal cortex (DLPFC) and ACC, to subcortical and posterior regions (Grant et al., 2014), such as the basal ganglia and occipital lobes, commensurate with increased efficiency and automation of bilingual language control.

The majority of existing evidence for the above models comes from research using between-group comparisons (monolingual/bilingual). Nonetheless, a few studies have directly examined potential neurocognitive effects of intensity and duration of bilingual language use, especially with respect to executive control, which is at the core of models such as the ACH and the BAPPS (Gullifer et al., 2018; Nichols & Joanisse, 2016; Yamasaki, Stocco, & Prat,

2018). For example, in a recent study Gullifer and colleagues scanned bilingual subjects (fMRI) at rest and performed a seed-based analysis focusing on regions described by the ACH (Gullifer et al., 2018). The authors utilized a measure of social diversity of language use (or *language entropy*) and reported that greater language entropy was associated with increased connectivity between the ACC and the putamen, regions with a prominent role in language control, as well as increased proactive control on an AX-Continuous Performance task (AX-CPT). Similarly, in a whole-brain resting state analysis, DeLuca et al. (2019) reported that earlier L2 AoA correlated with functional connectivity within the visual network, a finding that the authors interpreted as indicative of more efficient grapheme-to-phoneme mapping in L2 as a result of earlier L2 acquisition. While the results from these studies are encouraging, more research is needed regarding the specific effects of language experience on how the brain is recruited to handle EC demands.

Herein, we used fMRI to examine the impact of specific experience-based factors (EBFs) of bilinguals on several aspects of EC and its neural correlates. The participants resided in the UK where their L2, English, is the dominant language. Effects of specific language experiences on performance (accuracy and RTs) and neural recruitment were assessed in the MRI scanner while participants performed a Flanker task (Luk et al., 2010). This version of the flanker task included components aimed at assessing various aspects of EC claimed to be affected by bilingualism. These include the gating of non-target information (interference suppression) (Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002), the general cognitive load associated with switching between tasks (mixing cost) (Rubin & Meiran, 2005), and the use of information that assists in goal directed activity (facilitation) (Hedden & Gabrieli, 2010). Although facilitation is not as often studied as interference suppression or mixing cost, previous research has reported effects of bilingualism on facilitation both at the behavioral (Coderre, van Heuven, & Conklin, 2013) and neural (Luk et al., 2010) levels. The effects of bilingualism on facilitation are thought to stem from the recognition and use of contextual cues in conversation in order to facilitate more efficient language control. Facilitation is one of the cognitive processes defined within the ACH (salient cue detection) (Green & Abutalebi, 2013), and is, therefore, highly relevant to our investigation. The results will help us understand the effects of different bilingual experiences on the brain and provide arguments for the inclusion of EBFs in similar investigations. Because bilingualism is treated as a spectrum of experiences, the present approach will lead to an elaboration of theoretical proposals on bilingualism-induced adaptation with findings that could not emerge from a traditional bilingual/monolingual dichotomy.

1.1 The present study

The purpose was to investigate the role of specific EBFs in modulating neural recruitment for various EC demands. Following from a recent study (Deluca et al., 2019), we reduced our EBFs to two general domains: *duration* and *extent* of bilingual language use. *Duration* of L2 use was assessed with two variables: L2 age of acquisition (AoA), that is the absolute length of exposure to two languages, and the length of L2 immersion, that is the length of bilingual language use in settings where exposure to, and use of, the L2 is more frequent and robust (Linck, Kroll, & Sunderman, 2009). These factors allow for an examination of how prolonged bilingual language control demands affect domain-general EC. *Extent* of L2 use was also assessed using two variables. These were weighted factor scores derived from the Language and Social Background Questionnaire (LSBQ) (Anderson, Mak, Keyvani Chahi, &

Bialystok, 2018) which detail (a) extent of L2 engagement in home settings and (b) extent of L2 use in broader social/community settings. These allow for examination of whether recruitment patterns for EC are reinforced based on the specific context of L2 use. Although both measures potentially reflect a dual language context, L2 home may more specifically relate to this context. Lower scores on this scale would reflect engagement primarily with the L1 at home (e.g. partner, family), making broader social contexts a (potentially) L2 domain. Higher scores in L2 use in social settings, conversely, might indicate a dense code-switching context, especially in multilingual communities where language-switching and mixing is common, as in the UK. However, it is difficult to isolate a context of dense code-switching in the absence of a dual language context, especially in our sample of people who have migrated to the UK. Language proficiency was not included in the models, as it may be an outcome measure of bilingual experience in itself (for discussion on the suitability of proficiency measures as predictors of neural adaptation see DeLuca, Rothman, & Pliatsikas, 2018).

In addition to assessing the independent neurocognitive effects of duration and extent of bilingual language use, we also examined the combined effects of such factors. Given the dynamic nature of bilingual language use, adaptations to duration and extent of bilingual language occur together. Considering duration-based factors (e.g., L2 AoA) in isolation might be inadequate for meaningful cross-study comparisons precisely because they do not necessarily reflect the degree to which the second language was used. As such, we sought to combine the factors of duration and extent in a separate analysis, to assess the duration of *active* L2 use. We essentially converted the factors of AoA and Immersion into composite EBFs by accounting for reported active engagement with their L2 in the respective timeframes.

Based on previous findings (Abutalebi et al., 2012; Luk et al., 2010) and theoretical suggestions, we predicted neural recruitment for each EC aspect tapped by the Flanker task to be modulated by different EBFs. *Interference suppression* and *global switching costs* would relate to activations in fronto-parietal regions in the EC network, such as the ACC, MFG, IFG, and IPL (Abutalebi et al., 2012; Ansaldo et al., 2015). For these effects, longer duration of bilingual language use would relate to decreased activation in the fronto-parietal control network and increased activation in subcortical and posterior regions, including the occipital lobes and cerebellum (Filippi et al., 2011; Grundy et al., 2017; Pliatsikas, Johnstone, & Marinis, 2014), reflecting adaptations towards greater automation and/or efficiency in language control. Greater extent of L2 use, however, is expected to relate to increased activation in fronto-parietal control regions such as ACC and IPL (Abutalebi & Green, 2016; Green & Abutalebi, 2013), reflecting adaptations to increased control demands. Moreover, *facilitation* effects would manifest in regions such as the caudate nucleus, superior frontal gyrus, and occipital lobes (Luk et al., 2010). Longer duration of L2 use would result in decreased activations here or increased activations in posterior regions for this EC aspect, indicative of less active reliance on facilitative information in language contexts (Coderre et al., 2013; Grundy et al., 2017). Greater extent of L2 use would result in increased activations in fronto-parietal regions, reflecting adaptations towards language control and production demands (Abutalebi & Green, 2016). For our composite EBFs we predicted that the effects pertaining to duration of active bilingual language use will overlap with the absolute duration EBFs and relate to activation increases in posterior regions such as the cerebellum and occipital lobes, indicative of increased efficiency and automation in EC (Grundy et al., 2017). Finally, given findings from previous studies showing a disengagement between neural recruitment patterns and task performance

(Costumero et al., 2015; Luk et al., 2010), we predicted that the modulatory effects of language experience on neural outcomes would not have equivalents in behavioral outcomes.

2. Methods

2.1 Participants

Sixty-five bilingual adults (49 females, *m*age: 31.7yrs, SD: 7.24, range: 18-52) participated in the study. Inclusion criteria for the study included being right-handed, normal or corrected-to-normal visual acuity, and having no speech or language disorders, including dyslexia. All provided written informed consent and confirmed no contraindication to MRI scanning prior to participating in the study. Participants were native speakers of a variety of languages and all spoke English as one of their languages (*m*AoA: 8.5 yrs., SD: 4.9, range 0-22). The majority was born in other countries and moved to the UK at various ages, apart from three who were born in English speaking countries (UK and Ireland) to non-UK parents and moved to their parent's country of origin in childhood and then to the UK later in life. All participants were living in the UK at the point of testing, with varying lengths of residence (*m*length residence: 70.9 months, SD: 73.7, range 0.2-383.8). In terms of educational level, all participants reported holding at least a post-secondary degree or diploma apart from three who reported holding a high school diploma. Related to employment, all participants but one reported being either students in postgraduate education or professionals in a variety of sectors, including business, marketing, finance, health care and education. Minimal exclusion criteria were applied to recruit the widest possible range of linguistic experiences.

Approximately half of the participants (n=33) reported knowledge of additional languages beyond their native language and English. Of these participants, there was variability regarding the amount of current engagement with those languages. To control for potential ongoing effects of L3+ (Rothman, González Alonso, & Puig-Mayenco, 2019), engagement with these additional languages was included as a nuisance covariate in the analyses. This was calculated as a percentage of engagement (on a scale from 0-1) and was based on responses to four questions related to reading, writing, speaking, and listening for each additional language and then summed across all additional languages for each participant. Across the cohort we observed a current level of additional language exposure of 0.13 (SD: 0.26; range: 0-1.5).

2.2 Materials and Procedure

In addition to the LSBQ (Anderson et al., 2018), participants completed the Oxford Quick Placement Test (QPT) (Geranpayeh, 2003) for general English proficiency and Raven's Standard Progressive Matrices task (Ravens) (Raven, 1998) to control for intelligence/nonverbal spatial reasoning ability. Participants scored an average of 70.56% on the Ravens (SD: 8.38%, range 53.3%-90%) and were found to be high intermediate to high proficiency speakers of English via the QPT (avg. 88.4%, SD 10.8%, range: 52%-100%). The LSBQ documents language use in the participants' known languages in several life stages, from early childhood to the present day, and in a range of settings, both home/family settings and broader social and community settings.

Scores from the LSBQ were entered into a factor score calculator developed by Anderson and colleagues (Anderson, Mak, et al., 2018). The factor score calculator derives three individual factor scores and a composite factor score based on responses to questions in the LSBQ pertaining to language use in specific settings and time points. The data entered into

the calculations are numerated Likert scale scores detailing amount of L2 use, proficiency in each language, and frequency of use, respectively. These are standardized and summed into one of the three factor scores: 1) L2 use in home settings, 2) L2 use in Social/Community settings, 3) L1 proficiency. Additionally, there is a composite score of the three factor scores- labeled a "Bilingual Composite Score". Two of these factor scores were used as variables in our analyses to further isolate their respective contributions to neurocognitive adaptations. As referenced in previous sections, L2_Social, detailed L2 exposure and use in societal and community settings. The other, L2_Home, detailed the extent of L2 use in home settings. For both scores (L2_Home and L2_Social), a higher score indicates more L2 use, whereas a lower score indicates more engagement with the native language[i]. For these scores, a mean value of 51.5 was observed for L2_Social (SD: 11.36, range: 10.77-74.53), and a mean score of 2.38 for L2_Home (SD 5.25, range: -8.91-16.7).

For each of the behavioral and imaging analyses, two models were run to evaluate the role of the language factors and capture individual effects of duration- and extent-based language use as well as their combined effects on neurocognitive adaptation. Model 1 included four variables that assessed independent effects of (absolute) duration and extent of L2 exposure/use respectively. Duration of L2 use was measured via two variables: 1) L2 AoA and 2) length of L2 immersion (months). We log transformed the variables of L2 AoA and Immersion for two reasons: first, the data were not normally distributed (AoA: w=0.9521, p = 0.013; Immersion: w=0.81147, p<0.001) and second, we did not expect a linear adaptation over time (Kuhl et al., 2016). Extent of L2 use was assessed using the two factor scores derived from the LSBQ: L2_Home, and L2_Social. The four EBFs were included in the same model to allow us to control for their respective effects and isolate the individual neural effects of each type of language experience.

Model 2 assessed effects of the duration of *active* engagement with the additional language, thus the combined effects of duration and extent of L2 use. This was specified in two settings: 1) the total amount of time (of one's life) spent actively using the L2 (Active Duration) and 2) the length of time in immersion spent actively using the L2 (Active Immersion). Active Duration was determined by calculating the average percentage of English use in several stages, from the point the language was acquired through to the time of testing, as reported in the LSBQ. This percentage was multiplied by the total number of years spent using the L2. This resulted in a value per participant indicating the number of years spent actively using the L2 (English) (*m*length: 10.1yrs, SD: 5.1, range: 0.96-30.08). Active Immersion was determined by first calculating a percentage reflecting the regular use of English, including four questions related to reading, writing, speaking, and listening, respectively. This value was then multiplied by the number of months of immersion. This procedure produced values reflecting the amount of time actively engaged with English in immersion settings (*m*length active immersion: 58.4mo, SD: 60.85, range: 0.1-287.9). As neither of these predictor variables was normally distributed (Active Duration: w=0.907, p<0.001; Active Immersion: w=0.83128, p<0.001), both were log-transformed.

2.2.1 Flanker Task

Participants completed a version of the Flanker task (Eriksen & Eriksen, 1974; Luk et al., 2010) in the MRI scanner. The task was presented with E-Prime 2.0 Professional (Schneider, Eschman, & Zuccolotto, 2012). Participants were instructed to respond to the

direction of a red target arrow, surrounded by white colored symbols (Fig. 1), presented against a black background.

The task included 6 blocks of 72 trials each: three 'mixed', one 'congruent', one 'control' and one 'neutral'. The mixed blocks contained an equal number of congruent and incongruent trials, in which the flanking arrows pointed in the same or opposite direction as
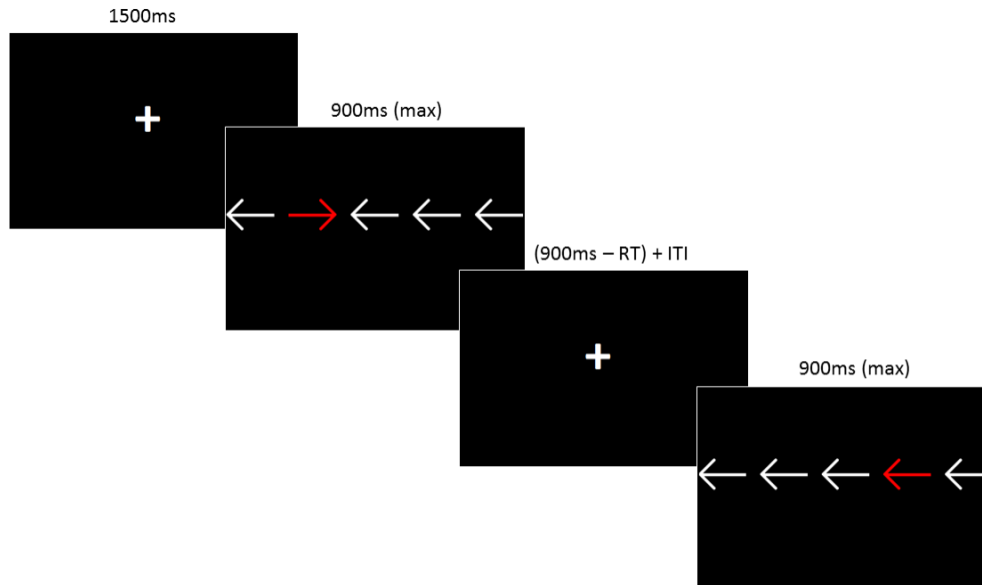
Figure 1: Presentation order and stimulus/ITI timings for the flanker task.

the target arrow. A congruent block contained only trials where the flanking arrows were in the same direction as the target arrow. In the neutral block, the arrow was surrounded by double-sided arrows, so the display had the same number of items (five) as the other blocks but contained no conflicting or facilitating information. Finally, in the control block a single arrow was presented in the middle of the screen. In all blocks, the order of presentation for individual trials was randomized. The mixed blocks were presented in an interspersed order with the other blocks, such that participants never saw the same block type sequentially. Blocks were presented on a Latin-square design to control for any potential effects of the order of block presentation. Target arrow direction was randomized across all trials within each block.

The first trial in each block began with a fixation cross, presented for 1500ms. The stimulus was then presented for up to 900ms. This screen was followed by a fixation cross which lasted for the remaining amount of time for the maximum trial length and the ITI (Fig 1). The remaining time for the trial was calculated as the difference between the trial reaction time and the maximum allowed time. As the mixed blocks contained two condition types (congruent and incongruent), stimuli were presented with a jittered inter-stimulus interval (ISI) of 1500 ± 500ms (minimum ISI 1000ms, maximum ISI 2000ms). The average trial length was 2400ms, but this varied from 1900 to 2900ms. As they contained only one condition, the neutral, control, and congruent blocks had a consistent ITI of 1500ms.

Breaks between blocks lasted 9 seconds during which two screens were shown. The first, lasting 3 seconds, gave instructions for participants to take a brief break and the second, lasting 6 seconds, instructed participants to get ready for the next block.

2.2.2 MRI data acquisition

Neuroimaging data were acquired with a 3T Siemens MAGNETOM Prisma_fit MRI scanner with a 32-channel Head Matrix coil and Syngo software. Whole-brain functional images were acquired during the Flanker task (735 volumes, FOV: 192 x 192, 68 transversal slices, 2.0mm slice thickness, voxel size 2.1x2.1x2.0mm, TR= 1500ms, TE= 30ms, flip angle 66°). A high-resolution anatomical scan using a MPRAGE sequence was carried out for purposes of registration (256 sagittal slices, 0.7 mm slice thickness, in-plane resolution 250 x 250, acquisition matrix of 246 x 256 mm, echo time (TE) = 2.41ms, repetition time (TR) = 2400ms, inversion time = 1140ms, flip angle = 8°).

2.3 fMRI data analysis

Due to an incidental finding within their structural scan, one participant was removed from the final imaging analysis. Neuroimaging data were processed and analyzed using the FEAT pipeline in FSL (Smith et al., 2004). The structural images were pre-processed with the fsl_anat software pipeline (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012). Functional data were motion-corrected using MCFLIRT, and slice-time corrected using Fourier-space time-series phase shifting. Non-brain tissue was removed using the brain extraction tool (BET) (Smith, 2002). Image distortion corrections were applied using field map-based echo-planar imaging (EPI) with PRELUDE+FUGUE (Jenkinson, 2003). The images were then registered to high resolution structural images using FLIRT (Jenkinson, Bannister, Brady, & Smith, 2002; Jenkinson & Smith, 2001). Registration from high resolution structural to standard space was then further refined using FNIRT nonlinear registration (Andersson, Jenkinson, & Smith, 2007). The images were also spatially smoothed using a Gaussian kernel with a Full Width and Half Maximum (FWHM) value of 4mm, and grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor was applied. Highpass temporal filtering was then applied (Gaussian-weighted least-squares straight line fitting, with sigma=50.0s).

The preprocessed FMRI data were first analyzed by task contrast at the subject level. Individual subject data were analyzed using the GLM package within FEAT (Woolrich, Ripley, Brady, & Smith, 2001). The task conditions (Congruent-Mixed Block, Incongruent-Mixed Block, Congruent, Neutral, and Control) were modelled as separate EVs. Incorrect and/or missing responses within the mixed blocks and breaks between blocks were modelled as covariates of no interest. Three task contrasts were specified to tap into specific cognitive demands. The first of these was the *Flanker effect*, which was run to assess neural correlates of interference suppression (Luk et al., 2010). This was assessed by contrasting incongruent against congruent trials and vice versa (incongruent>congruent and incongruent<congruent) within the mixed blocks. We also assessed neural correlates for the *facilitation effect* by contrasting the average activation for the congruent block against the average activation for the neutral block (congruent>neutral and congruent<neutral) (Luk et al., 2010). Finally, we examined global switching cost or *mixing cost* (Kray & Lindenberger, 2000; Rubin & Meiran, 2005) which was assessed by contrasting the congruent trials from the Mixed blocks with the average activation across the Congruent block (Congruent mixed>Congruent and Congruent mixed < Congruent). The resulting activation clusters (Gaussianized t-statistic images) for each task contrast were thresholded at Z>2.3 and a significance threshold of p=0.05.

Cross-subject analyses were carried out with mixed effects models using the FMRIB's Local Analysis of Mixed Effects (FLAME) pipeline in FSL (Woolrich, 2008; Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004). Contrast Parameter Estimates (COPEs) for

all the task contrasts (Flanker effect, mixing cost, and facilitation effect) from the subject-level analyses were entered into the models. The two cross-subject models specified group mean, age, sex, additional language use, and Ravens score as covariates of no interest, and then the EBFs as variables of interest. Model 1 (Duration and Extent) included AoA, Immersion, L2_Home, and L2_Social, Model 2 (Active Engagement) included Active Duration and Active Immersion. The same thresholding and correction were applied for the group level analyses as the subject-level analyses: the resulting statistic images were thresholded using a cluster-based threshold of Z>2.3 and a corrected cluster significance threshold of p=0.05.

3. Results

3.1 Behavioral

Accuracy rates were high across all conditions (Table 1), showing that our participants had no difficulty with the task. As such, we focused on the reaction time (RT) data. These data were submitted to a linear mixed-effect regression analysis using the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2014). Incorrect/missing trials and trials with RTs less than 200ms were excluded from the analysis. This resulted in a removal of 2.2% of the trials overall. Comparisons for model fit were assessed via an analysis of variance using the *anova()* function within the L (Kuznetsova, Brockhoff, & Christensen, 2017).

Table 1: Performance on the flanker task (accuracy and reaction time) globally and by condition.

		Whole task	Congruent (Mixed block)	Incongruent (Mixed block)	Neutral	Congruent	Control
Accuracy (%)	Mean	97.84	98.81	95.43	98.00	98.7	99.00
	SD	1.44	2.07	2.89	2.24	1.64	1.30
RT (ms)	Mean	456	460	538	474	424	384
	SD	40	43	43	43	48	43

The three task contrasts of interest—i.e., facilitation, mixing cost, and Flanker effect—were assessed separately. For the Flanker effect, Congruent (Mixed) condition was set as the reference level and was contrasted with Incongruent (Mixed) condition. For mixing cost, the Congruent block was set as the reference level and was contrasted against the Congruent-Mixed block. Finally, for facilitation effect, Neutral was set as the reference level and contrasted against the Congruent block. For all contrasts, a base model was first specified including fixed effects of condition and Ravens scores. Random intercepts of age, sex, and L3+ exposure were also added to the base model; however, they did not improve model fit and thus were not retained. The model failed to converge with addition of random slopes for L3+ exposure and thus these were not included. Random slopes for participants were not estimated for the models, as we wished to capture individual difference measures in fixed effects. Ravens scores significantly predicted RTs across all task contrasts, and thus were retained in the base model. The final base model thus contained fixed effects of condition and Ravens scores, and random intercepts of subjects.

For each task contrast, condition was found to significantly contribute to model fit (Flanker effect- $F(1,63)=1617.67$ $p<.0001$; Mixing cost- $F(1,63)=105.28$, $p<.0001$; Facilitation effect- $F(1,63)=125.88$, $p<.0001$). Model summaries showed that the incongruent condition was slower than congruent within the Mixed block (*est*=78.07, *SE*=1.941, p<.0001), the

congruent (Mixed) block was slower than the Congruent block (*est*=35.78, *SE*=3.4872, p<.0001), and Congruent block was faster than Neutral block (*est*=-49.64, *SE*=4.245, p<.0001). This confirmed that the expected task effects were present for all contrasts.

To assess if the EBFs modulated the above RT differences, two follow-up models were run for each task contrast. These models specified interaction terms between the EBFs and condition. The same EBFs were used for the RT analyses as for the neuroimaging analyses. Model 1 (duration and extent) included EBFs of AoA, Immersion, L2_Home, & L2_Social; Model 2 (Active engagement) included Active Duration and Active Immersion). The follow-up models did not improve model fit over the base model, indicating no modulatory effect of language experience on task performance.

3.2 Neuroimaging Results

Both Models 1 (Duration and Extent) and 2 (Active Engagement) showed modulations in brain activation across contrasts, which differed by EBF. Results are presented by model and planned task contrast, respectively. All clusters reported herein were corrected for multiple comparisons.

3.2.1 Model 1: Duration and Extent

Interference suppression (Flanker effect)

All four EBFs were related to distinct activation patterns (Table 2). AoA negatively correlated with activations across a range of regions, most in the right hemisphere, including the supramarginal/angular gyrus, middle frontal gyrus (MFG), and inferior frontal gyrus (IFG), and some in the left hemisphere including the cerebellum (Fig 2). Immersion also negatively correlated with activations across several regions including within the cerebellum, right inferior parietal lobule (IPL), thalamus, precentral gyrus, bilateral MFG, and thalamus (Table 2; Fig 2). L2_Home negatively correlated with activations in the cerebellum, MTG, and SMG. L2_Social positively correlated with activations across several regions including the cerebellum, anterior and posterior cingulate cortex (ACC/PCC), IPL and MFG (Table 2; Fig 2).

Table 2: Results for the Flanker effect for Model 1. Coordinates in MNI-space.

EBF	Hemisphere	Region	Direction	Voxels	Z score	X	Y	Z
AoA	L	Cerebellum- IX	-	542	3.86***	-4	-46	-46
		ITG	-	331	3.4**	-60	-58	-16
		Frontal pole	-	161	3.3*	-24	42	28
		Hippocampus	-	444	4.09***	-34	-34	2
	R	Occipital pole	-	493	3.82***	4	-88	18
		SFG	-	1093	3.87***	6	40	54
		SMG	-	991	4.13***	48	-38	8
		IPL	-	553	3.74***	54	-46	56
		Postcentral	-	245	3.35*	6	-36	78
		MFG	-	232	3.74*	44	28	32
		IFG	-	224	3.77*	58	24	12
		LOC	-	218	3.85*	40	-86	10
		Precentral	-	182	3.37*	62	-2	26
		Putamen	-	165	3.55*	26	2	-10
		ACC/PCC	-	408	4.01***	0	-16	38

Immersion	L	Cerebellum-crus I	-	1578	4.02***	-16	-84	-28
		IPL	-	605	3.77***	-50	-52	34
		LOC	-	255	3.79*	-34	-70	44
		Precentral	-	248	3.84*	-56	-6	24
		Opercular Cortex	-	237	3.62*	-60	-14	12
		MFG	-	328	3.72**	-34	8	60
		MFG	-	213	3.66*	-24	36	30
	R	Cerebellum- crus II	-	336	4.09**	44	-48	-44
		Thalamus	-	160	4.05*	8	-12	16
		Precuneus	-	554	3.98***	8	-64	-32
		MFG	-	2840	4.9***	50	16	40
		IPL	-	2651	4.52***	52	-50	40
		Precuneus	-	1635	4.02***	12	-66	28
		Frontal pole	-	402	3.64***	26	56	2
L2_Home	L	Cerebellum-VI	-	167	3.35*	-12	-62	-26
	R	MTG	-	299	3.6**	54	-38	0
		IPL	-	236	3.7*	46	-44	58
L2_Social	L	Precentral	+	1207	3.73***	-56	-26	50
		PCC	+	827	4.14***	-4	-28	38
		Cerebellum- crus II	+	230	3.4*	-34	-56	-44
		Postcentral	+	176	3.48*	-46	-36	58
		Frontal pole	+	160	3.62*	-42	52	2
	R	SMG	+	5894	5.08***	48	-40	8
		MFG	+	1438	4.55***	46	32	34
		ACC	+	1278	4.26***	4	26	18
		IPL	+	838	4.17***	50	-42	50
		Cerebellum-V/Brain stem	+	230	3.8*	6	-22	-30

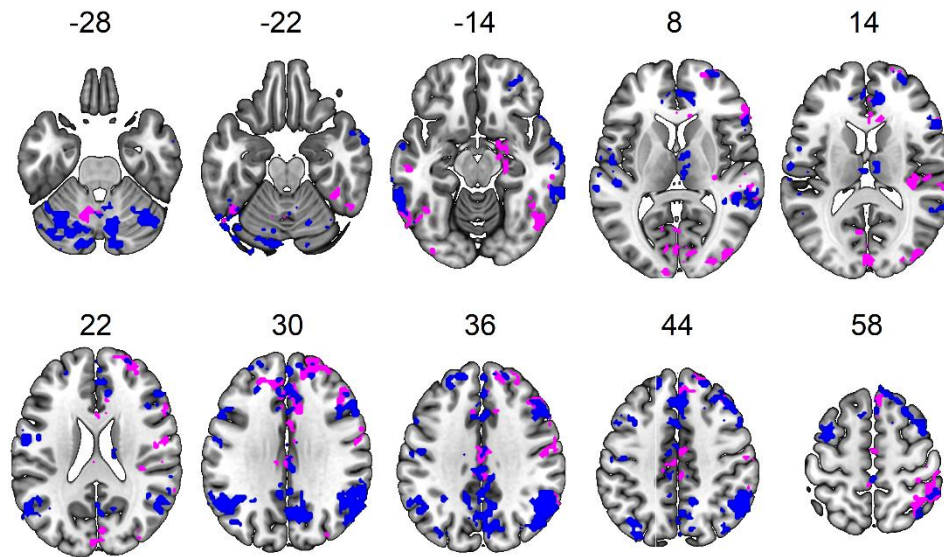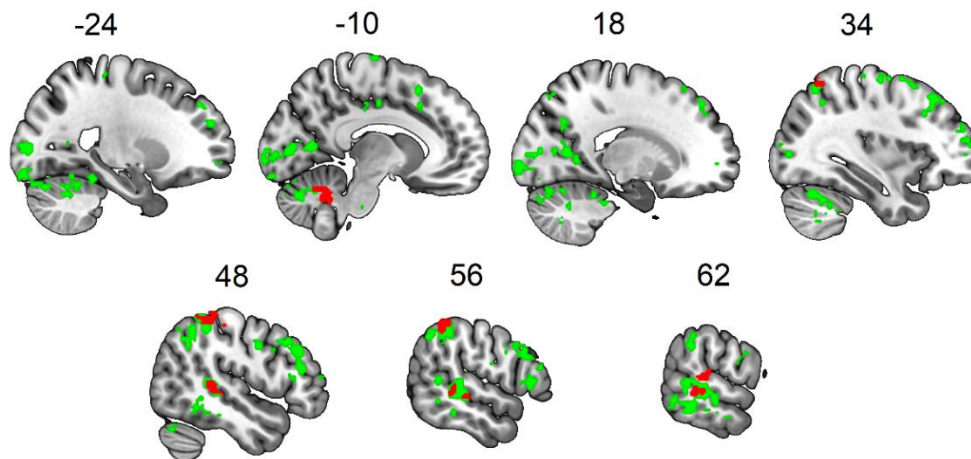* $p < .05$, ** $p < .001$, *** $p < .0001$; all p values corrected

Figure 2: Activations related to the Flanker effect. (above) Negative correlations with AoA (violet) and Immersion (blue). (below) Activations negatively correlated with L2_Home (red) and positively correlated with L2_Social (green).

Interim Summary- Interference suppression

AoA negatively correlated with activation in a range of regions implicated in EC. Recall, though, that this means that longer duration of L2 use relates to *greater* activation. Length of immersion and extent of L2 use at home both negatively correlated with activations in several regions involved in EC processes. Finally, L2_Social positively correlated with activation in a network of regions implicated in EC.

Mixing Cost

AoA negatively correlated with activations in the left lingual gyrus (332 voxels; Z score 3.3; -12, -52, 2). Immersion negatively correlated with activations in the left postcentral gyrus (633 voxels; Z score: 4.71; -52 -28, 60). Finally, L2_Home negatively correlated with activations in the precuneus (316 voxels; Z score: 3.67; -12, -70, 36) and left PCC (201 voxels; Z score 4.29; -6, -26, 30) (Fig. 3). L2_Social was not found to significantly correlate with any activations for mixing cost.
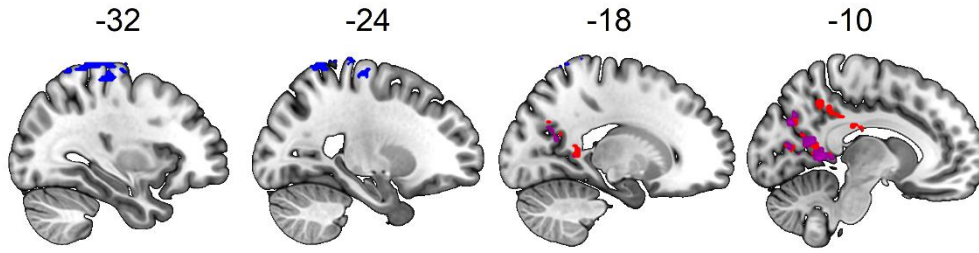
Figure 3: Activations negatively correlated with AoA (violet), immersion (blue), and L2_Home (red) for mixing cost.

Interim Summary- Mixing cost

For mixing cost, longer duration of L2 use (earlier AoA) related to increased reliance on posterior regions. Longer duration in immersion correlated with decreased reliance on superior parietal regions. More L2 use at home was related to less activation in posterior regions.

Facilitation Effect

AoA positively correlated with activations in the right STG and precentral gyrus. Immersion positively correlated with activations in the left MTG and SPL (Table 3, Fig 4). Neither L2_Home nor L2_Social were found to correlate with activations for this effect.

Table 3: Results for the facilitation effect for Model 1.

EBF	Hemisphere	Region	Direction	Voxels	Z score	X	Y	Z
AoA	R	STG	+	389	4.09**	64	-4	-4
		Precentral Gyrus	+	263	3.69*	22	-10	62
Immersion	L	MTG	+	237	3.93*	-50	-46	8
		SPL	+	216	3.68*	-28	-58	70

* p<.05, **p<.001, ***p<.0001; all p values corrected; MTG: middle temporal gyrus, STG: superior temporal gyrus, SPL: superior parietal lobule
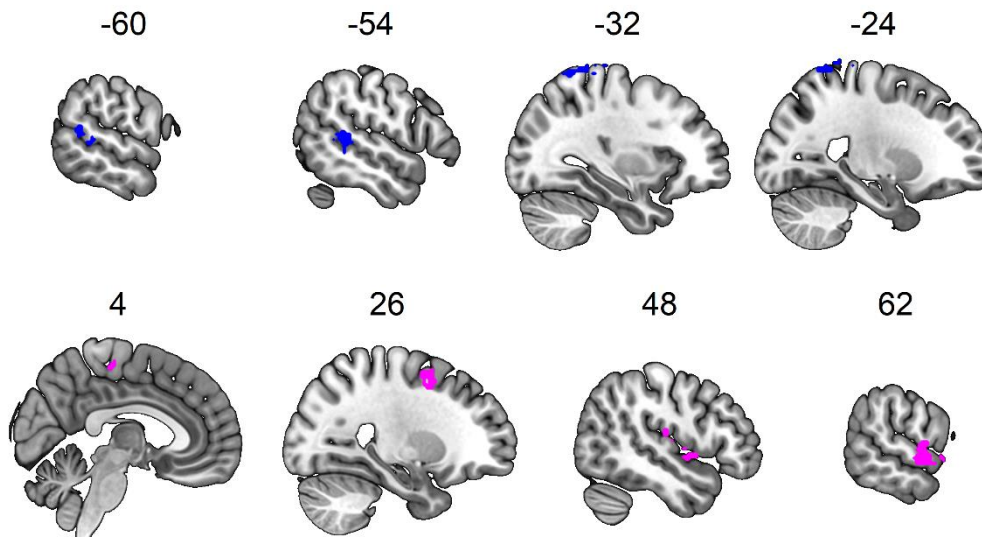
Figure 4: Activations positively correlated with immersion (blue) and AoA (violet) for the facilitation effect.

Interim Summary- Facilitation Effect

Only duration-based EBFs correlated with neural activation. Longer duration of L2 use (earlier AoA) was related to a decrease in recruitment of right hemisphere regions, whereas longer immersion correlated with increased recruitment within left hemisphere temporal/parietal regions.

3.2.2 Model 2: Active Engagement

Interference suppression (Flanker effect)

Active Duration positively correlated with activations in the left cerebellum (280 voxels; z score 3.67; -28, -82, -36). Active Immersion negatively correlated with activations in the right IPL (814 voxels; z score 3.93, p<.001; 58, -48, 30) and precuneus (354 voxels; z score 3.37; p<.001; 8, -76, 38) (Fig. 5).
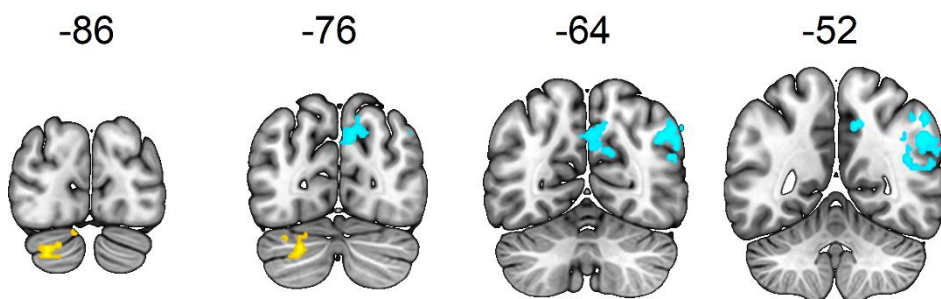

Figure 5: Activations positively correlated with Active Duration (yellow), and negatively correlated with Active Immersion (light blue) for the flanker effect.

Mixing Cost

Active Immersion negatively correlated with activations in the left postcentral gyrus (272 voxels; -42, -22, 58; Z score: 4.01). Active Duration did not significantly correlate with any activations for this effect.

Facilitation Effect
Neither EBF significantly predicted activations for this effect.

Interim Summary- Active Engagement
      Both EBFs related to active engagement displayed activation patterns similar to the duration-based counterparts (AoA and Immersion). However, not all of the regions implicated with the duration-based EBFs were found to correlate with the Active engagement EBFs.

4. Discussion
      The present study examined neurocognitive effects of experience-based factors (EBFs) within bilingualism, across several executive control processes. The results contribute to the contemporary debate regarding the extent to which bilingualism may be associated with adaptations to EC. Given its specific purpose, this study did not include monolingual controls. Nonetheless, the results have implications for the current controversy about whether bilingualism affects these processes at all, by specifically pointing to the importance of understanding potential neurocognitive adaptations related to different aspects of bilingual language use.
      Two findings from the present study are key. First, behavioral performance on the Flanker task was not modulated by participants' language experiences. Second, and in contrast to the behavioral results, the same language experiences did affect patterns of neural recruitment which also differed by each task contrast. Crucially, these neural effects were calibrated to the bilingual language experiences of the participants.
      Although participants showed the expected behavioral task effects for all measured contrasts (interference suppression, facilitation, and mixing cost), these were not modulated by their language experiences. This pattern supports two interrelated arguments. First, test-retest reliability of EC tasks of this type, and use of RT difference scores generally, are known to be low (Chan, Shum, Toulopoulou, & Chen, 2008; Draheim, Mashburn, Martin, & Engle, 2019). However, data from them are often used to make claims about certain neurocognitive differences in absolute terms, including, but not limited to, in the literature on bilingual cognition. The pattern of results in the present study provides a crucial example to this cautionary tale, precisely because they show that purely behavioral measures can have insufficient granularity to capture patterns in neurocognitive processing. Second, if there are behavioral constraints related to task granularity, for example in the motor responses involved in button pressing, or in the interaction between the motor and the EC systems, they do not necessarily apply to underlying neural recruitment. The disengagement of behavioral from neuroimaging results seen in the present study supports previous research showing differences in brain activity but *not* in task performance across adults with different language experiences (i.e. bilinguals vs. monolinguals) on executive function tasks (Ansaldo et al., 2015; Costumero et al., 2015; Luk et al., 2010). The present data strongly suggest a utility in corroborating what is found behaviorally with more granular (neuroimaging) modalities in future research, to more fully ascertain the nature of neurocognitive adaptations.
      Despite not being reflected in RTs, EBFs modulated neural recruitment patterns across each EC aspect. Furthermore, these patterns overlapped with previous work showing differences in brain activation between bilinguals and monolinguals in studies using similar tasks (Abutalebi et al., 2012; Ansaldo et al., 2015; Luk et al., 2010). That is, even in the absence

of a monolingual control group, the previously reported bilingualism-induced adaptations are documented *within* a bilingual group and are modulated by both the extent and duration of bilingual language experience. The present results fit the neuroemergentist framework in terms of differing regions implicated to accomplish the same task (e.g. Hernandez et al., 2019). Furthermore, this pattern of results fits with aspects of the BAPSS and ACH. As predicted under the BAPSS framework (Grundy et al., 2017), duration-based EBFs were shown to relate to neural recruitment patterns indicative of increased efficiency in handling EC demands. Similarly, consistent with aspects of the ACH (Abutalebi & Green, 2016), extent-based EBFs predicted neural recruitment patterns suggesting adaptation towards changing language control demands. Below we discuss the effects reported for each of the two types of EBFs.

EBFs related to *duration* of L2 use predicted distinct recruitment patterns for each aspect of EC examined. For interference suppression and mixing cost, length of immersion negatively correlated to degree of recruitment of several regions which have been implicated in selection and conflict monitoring processes (Abutalebi et al., 2012; Abutalebi & Green, 2016; Ansaldo et al., 2015), both of which are in line with our predictions (Table 4). These results likely reflect increased efficiency of language control, and thus EC, with prolonged intensive L2 exposure (Grundy et al., 2017; Linck et al., 2009). Longer overall duration of bilingual language use (earlier AoA) was related to increased recruitment of several regions involved in interference suppression and conflict monitoring including the IFG, ACC, and IPL (Abutalebi et al., 2012; Abutalebi & Green, 2016; Ansaldo et al., 2015). This finding did not follow our predictions, as we would have predicted more activation in subcortical and posterior regions, and less activation in fronto parietal regions with prolonged L2 exposure. Note, though, that as the duration-based factors were log-transformed, this pattern may reflect a plateau in the degree of recruitment of these regions for EC demands with prolonged bilingual experience. More evidence is required to assess the validity of that interpretation.

Regarding mixing cost, longer duration of L2 use was related to increased recruitment of posterior regions that have previously been implicated in switching demands (Anderson, Chung-Fat-Yim, et al., 2018). Longer immersion correlated to decreased activations in parietal regions which have previously been implicated in selection processes (Seo, Stocco, & Prat, 2018). Both patterns of results match our predictions (Table 4) and are in line with aspects of the BAPSS framework in which the shift in recruitment patterns would indicate improved efficiency in handling global switching demands with prolonged bilingual experience (Grundy et al., 2017).

Regarding the facilitation effect, the decreased right hemisphere activations, related to earlier AoA, may indicate decreased engagement with facilitation with prolonged bilingual language experience (Coderre et al., 2013). The increased left hemisphere activations for immersion may indicate an effect of facilitation in linguistic contexts from intensive L2 use in the immersive environment (Coderre et al., 2013; Costa, Caramazza, & Sebastián-Gallés, 2000).

EBFs related to *extent* of L2 use also predicted distinct activation patterns for interference suppression and mixing cost. Several of the regions which positively correlated with L2 use in social settings for interference suppression (including the ACC, MFG, and IPL) have been implicated in language and domain-general control processes (Abutalebi & Green, 2016; Ansaldo et al., 2015). These correlations with activations in fronto-parietal regions are in line with our predictions (Table 4). Interpreted within the ACH, these findings indicate adaptation towards increased language control demands of specific communicative contexts,

which in turn affect domain-general EC (Abutalebi & Green, 2016). The negative correlations found for L2 use in home settings for interference suppression and mixing cost did not match our specific predictions of greater activation in frontal regions involved in EC. However, as these correlations occurred in predominantly posterior and parietal regions (e.g. PCC, cerebellum, and IPL), this may indicate a transition in reliance from posterior regions (possibly towards frontal regions), to accommodate changing language control and switching demands at home.

Finally, the *duration of active* L2 use also predicted activation patterns, indicating increased neural efficiency and automation in handling non-linguistic cognitive control demands, as interpreted within the BAPSS framework (Grundy et al., 2017). These effects patterned with those derived from the absolute duration-based factors, albeit with a lesser spatial extent. The pattern of results for both EBFs is in line with our predictions of overlapping with duration-based variables, specifically in posterior regions such as the cerebellum (Table 4). Comparing length of immersion and its composite corollary (Active Immersion), the composite version overlapped with immersion, but was restricted to specific regions related to selection and control processes (Abutalebi & Green, 2016; Rossi, Newman, Kroll, & Diaz, 2018). The overlap is not surprising as these factors were highly correlated. The pattern of results for the composite EBF likely also indicate decreased cognitive demands for interference suppression and mixing cost. Similarly, when comparing AoA and its composite corollary (Active Duration), only activations in the cerebellum were found to correlate with interference suppression for the composite, also indicating increased automation (Filippi et al., 2011; Pliatsikas et al., 2014). A potential explanation for the differences between these EBFs is that regions correlated with the composite EBFs are where duration and extent of bilingual language use *converge* with respect to effects on domain-general EC. These results support the argument that the neurocognitive effects of absolute duration-based factors are modulated by what happens experientially within these timeframes.

Table 4: Summary of study predictions and findings.

Task Effect	EBFs	Prediction	Support for Predictions
Interference Suppression	Duration	Decreased activations in fronto-parietal regions, & increased activations in posterior regions	Supported- decreased activation for Immersion, increased activation for AoA across several regions
	Extent	Increased activation in fronto-parietal control regions	Partial support- increased activation for L2 use in social settings but decreased activation for home settings.
	Active Engagement	Overlap with Duration EBFs, specifically activation increases in posterior regions	Supported- overlapping activation with Duration EBFs but with lesser spatial extent
Mixing Cost	Duration	Decreased activations in fronto-parietal regions or increased activations in posterior regions	Supported- increased activation in posterior regions (AoA), decreased in fronto-parietal regions (Immersion)

	Extent	Increased activations in fronto-parietal regions	Not supported- decreased activation in posterior regions
	Active Engagement	Overlap with Duration EBFs	Supported- overlapping activation with Duration, but lesser spatial extent
Facilitation	Duration	Increased activations in posterior regions, decreased in fronto-parietal regions.	Not supported, immersion-increased activations in left temporal regions, AoA- decreased activations in right temporal regions
	Extent	Increased activation in fronto-parietal regions	Not supported- no significant effects
	Active Engagement	Overlap with Duration EBFs	Not supported- no significant effects

EBFs: Experience-Based Factors; Duration EBFs: Age of Acquisition (AoA) & months in UK (Immersion), Extent EBFs: L2 use at home (L2_Home) & in social/community settings (L2_Social), Active Engagement EBFs: Active Duration (years actively using the L2) & Active Immersion (time actively using the L2 in immersion settings. Task effects were Interference suppression was defined as incongruent (mixed block)-congruent (mixed block), Mixing cost: Congruent (mixed block)-Congruent block, Facilitation: congruent block- neutral block

Overall, our results indicate that increased duration of L2 use (measured by the proxies of L2 immersion and AoA) leads to more effective and efficient interference suppression processes, changing reliance on facilitation processes, and more efficient language switching. Moreover, it appears that increased L2 use in home and social contexts translates to the brain adapting to more effectively dealing with increased language control and mixing demands (Green & Abutalebi, 2013). These findings show that the continuum of experiences that comprise bilingualism leads to distinct, measurable neurocognitive outcomes that calibrate to the degree of language experience of the participants. Importantly, the EBFs examined in the present study are by no means an exhaustive list. Exploring other factors, such as language dominance (Yamasaki et al., 2018) and diversity of language use (Gullifer et al., 2018), and their relationships with the present ones, would better delineate the dynamic nature of bilingual language use and related neurocognitive adaptations. The present study provides both an example of how this can be done and evidence that it needs to be done moving forward.

It is worth noting that the specificity of adaptations in neural recruitment to individual EBFs reported in the present study is in line with the patterns reported for neuroanatomical adaptations with the same EBFs in a previous study (DeLuca et al., 2019). For example, in both studies the effects of immersion point towards dealing with language control with increasing efficiency as a consequence of longer immersion (i.e. greater opportunity and/or need to engage with the additional language). This manifested either as decreased recruitment of networks underlying conflict monitoring processes (the present study), or as renormalization of regions related to language control, such as the thalamus and the right caudate nucleus (DeLuca et al., 2019) as an effect of immersion. Similarly, increased exposure to the L2 in social settings causes increased language control demands; these manifested as increased reliance on networks underlying interference suppression (this study) and expansion of the regions that underlie language switching, like the left caudate nucleus (DeLuca et al., 2019). These patterns highlight

the need for a more generalized use of EBFs as predictors of functional and structural adaptations and the extent to which the two interact.

The difference in results between the absolute, duration-based EBFs (AoA and Immersion) and their composite corollaries also has consequences for comparisons of results across studies. If, for example, AoA can encompass any extent of usage—e.g. 10 years of L2 use in a group of 20 bilinguals could comprise 20 different levels of active engagement—then comparing adaptations for equivalent AoA across studies could be inadvertently misleading. It is possible, in the context of the replication debate, that studies showing no effects have greater variability between the AoA and active use corollary than other studies that show effects. Precisely because the present study and similar ones have recently shown that degree and quality of the bilingual experience matters, we should consider accounting for the degree of experience(s) within a given timeframe. A failure to replicate results between studies would then lead to an evaluation of the quantity/quality of the experience(s) of the individuals in the two cohorts to determine if such a comparison is justified. Even if absolute measures such as L2 AoA or immersion were equivalent in both cases, it may be that the patterns of language use during that time were different.

Bilingualism is a complex and dynamic process, comprising a range of experiences that contribute to distinct neurocognitive adaptations. The brain constantly adapts to be maximally effective at handling the cognitive load of the communicative environment. Modulations to that environment, specifically the language experiences it entails, will thus confer measurable and distinct outcomes for the mind and brain. Specific language experiences must be considered in more detail in future research examining the neurocognitive effects of bilingualism. Bilingualism is not a categorical label (Luk & Bialystok, 2013), and progress in understanding its impact on cognitive and brain systems will require investigating the effects of the *spectrum* of related factors that constitute this complex experience.

References

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernández, M., Scifo, P., Keim, R., … Costa, A. (2012). Bilingualism Tunes the Anterior Cingulate Cortex for Conflict Monitoring. *Cerebral Cortex*, *22*(9), 2076–2086. https://doi.org/10.1093/cercor/bhr287

Abutalebi, J., & Green, D. W. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Bilingualism: Language and Cognition*, *19*(04), 689–698. https://doi.org/10.1017/S1366728916000225

Anderson, J. A. E., Chung-Fat-Yim, A., Bellana, B., Luk, G., & Bialystok, E. (2018). Language and cognitive control networks in bilinguals and monolinguals. *Neuropsychologia*, *117*, 352–363. https://doi.org/10.1016/j.neuropsychologia.2018.06.023

Anderson, J. A. E., Mak, L., Keyvani Chahi, A., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, *50*(1), 250–263. https://doi.org/10.3758/s13428-017-0867-9

Andersson, J. L. R., Jenkinson, M., & Smith, S. M. (2007). Non-linear registration aka Spatial normalisation FMRIB Technial Report TR07JA2. *Oxford Centre for Functional Magnetic Resonance Imaging of the Brain, Department of Clinical Neurology, Oxford University, Oxford, UK*, 22.

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. https://doi.org/10.1080/13803395.2014.990359

Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., … Carreiras, M. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Frontiers in Psychology*, *5*, 1–12. https://doi.org/10.3389/fpsyg.2014.00398

Antoniou, M. (2019). The Advantages of Bilingualism Debate. *Annual Review of Linguistics*, *5*(1), 395–415. https://doi.org/10.1146/annurev-linguistics-011718-011820

Bak, T. H. (2016). Cooking pasta in La Paz. *Linguistic Approaches to Bilingualism*, *5*(2016), 1–19. https://doi.org/10.1075/lab.16002.bak

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bialystok, E. (2016). The signal and the noise: Finding the pattern in human behavior. *Linguistic Approaches to Bilingualism*, *6*(5), 517–534. https://doi.org/10.1075/lab.15040.bia

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, *143*(3), 233–262. https://doi.org/10.1037/bul0000099

Bialystok, E., Craik, F. I. M., Grady, C. L., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effect of bilingualism on cognitive control in the Simon task: Evidence from MEG. *NeuroImage*, *24*(1), 40–49. https://doi.org/10.1016/j.neuroimage.2004.09.044

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Immature Frontal Lobe Contributions to Cognitive Control in Children. *Neuron*, *33*(2), 301–311. https://doi.org/10.1016/S0896-6273(01)00583-9

Calabria, M., Costa, A., Green, D. W., & Abutalebi, J. (2018). Neural basis of bilingual language control. *Annals of the New York Academy of Sciences*, *1426*(1), 221–235. https://doi.org/10.1111/nyas.13879

Chan, R. C. K., Shum, D., Toulopoulou, T., & Chen, E. Y. H. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, *23*(2), 201–216. https://doi.org/10.1016/j.acn.2007.08.010

Coderre, E. L., van Heuven, W. J. B., & Conklin, K. (2013). The timing and magnitude of Stroop interference and facilitation in monolinguals and bilinguals. *Bilingualism: Language and Cognition*, *16*(02), 420–441. https://doi.org/10.1017/S1366728912000405

Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1283–1296. https://doi.org/10.1037/0278-7393.26.5.1283

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*(2), 135–149. https://doi.org/10.1016/j.cognition.2009.08.001

Costumero, V., Rodríguez-Pujadas, A., Fuentes-Claramonte, P., & Ávila, C. (2015). How

bilingualism shapes the functional architecture of the brain: A study on executive control in early bilinguals and monolinguals. *Human Brain Mapping*, *36*(12), 5101–5112. https://doi.org/10.1002/hbm.22996

De Baene, W., Duyck, W., Brass, M., & Carreiras, M. (2015). Brain Circuit for Cognitive Control Is Shared by Task and Language Switching. *Journal of Cognitive Neuroscience*, *27*(9), 1752–1765. https://doi.org/10.1162/jocn

DeLuca, V. F., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences*, *116*(15), 7565–7574. https://doi.org/10.1073/pnas.1811513116

DeLuca, V. F., Rothman, J., & Pliatsikas, C. (2018). Linguistic immersion and structural effects on the bilingual brain: a longitudinal study. *Bilingualism: Language and Cognition*, *1*, 1–16. https://doi.org/10.1017/S1366728918000883

Donnelly, S., Brooks, P. J., & Homer, B. D. (2015). Examining the Bilingual Advantage on Conflict Resolution Tasks: A Meta-Analysis. *37th Annual Conference of the Cognitive Science Society*, 596–601.

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, *145*(5), 508–535. https://doi.org/10.1037/bul0000192

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Filippi, R., Richardson, F. M., Dick, F., Leech, R., Green, D. W., Thomas, M. S. C., & Price, C. J. (2011). The Right Posterior Paravermis and the Control of Language Interference. *Journal of Neuroscience*, *31*(29), 10732–10740. https://doi.org/10.1523/JNEUROSCI.1783-11.2011

Garbin, G., Sanjuán, A., Forn, C., Bustamante, J. C., Rodríguez-Pujadas, A., Belloch, V., … Ávila, C. (2010). Bridging language and attention: Brain basis of the impact of bilingualism on cognitive control. *NeuroImage*, *53*(4), 1272–1278. https://doi.org/10.1016/j.neuroimage.2010.05.078

Geranpayeh, A. (2003). A quick review of the English Quick Placement Test. *UCLES Research Notes*, (12), 8–10.

Grant, A., Dennis, N. A., & Li, P. (2014). Cognitive control, cognitive reserve, and memory in the aging bilingual brain. *Frontiers in Psychology*, *5*, 1–10. https://doi.org/10.3389/fpsyg.2014.01401

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Grundy, J. G., Anderson, J. A. E., & Bialystok, E. (2017). Neural correlates of cognitive processing in monolinguals and bilinguals. *Annals of the New York Academy of Sciences*, *1396*(1), 183–201. https://doi.org/10.1111/nyas.13333

Gullifer, J. W., Chai, X. J., Whitford, V., Pivneva, I., Baum, S., Klein, D., & Titone, D.

(2018). Bilingual experience and resting-state brain connectivity: Impacts of L2 age of acquisition and social diversity of language use on control networks. *Neuropsychologia*, *117*(May), 123–134. https://doi.org/10.1016/j.neuropsychologia.2018.04.037

Hedden, T., & Gabrieli, J. D. E. (2010). Shared and selective neural correlates of inhibition, facilitation, and shifting processes during executive control. *NeuroImage*, *51*(1), 421–431. https://doi.org/10.1016/j.neuroimage.2010.01.089

Hernandez, A. E., Claussenius-Kalman, H. L., Ronderos, J., Castilla-earls, A. P., Sun, L., Weiss, S. D., & Young, D. R. (2019). Neuroemergentism: A framework for studying cognition and the brain. *Journal of Neurolinguistics*, *49*, 214–223. https://doi.org/10.1016/j.jneuroling.2017.12.010

Hernandez, A. E., Claussenius-Kalman, H. L., Ronderos, J., & Vaughn, K. A. (2018). Symbiosis, Parasitism and Bilingual Cognitive Control: A Neuroemergentist Perspective. *Frontiers in Psychology*, *9*, 1–7. https://doi.org/10.3389/fpsyg.2018.02171

Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin*, *133*(4), 638–650. https://doi.org/10.1037/0033-2909.133.4.638

Hernandez, A. E., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, *9*(5), 220–225. https://doi.org/10.1016/j.tics.2005.03.003

Hernández, M., Costa, A., Fuentes, L. J., Vivas, A. B., & Sebastián-Gallés, N. (2010). The impact of bilingualism on the executive control and orienting networks of attention. *Bilingualism: Language and Cognition*, *13*(3), 315–325. https://doi.org/10.1017/S1366728909990010

Hernández, M., Martin, C. D., Barceló, F., & Costa, A. (2013). Where is the bilingual advantage in task-switching? *Journal of Memory and Language*, *69*(3), 257–276. https://doi.org/10.1016/j.jml.2013.06.004

Jenkinson, M. (2003). Fast, automated, N-dimensional phase-unwrapping algorithm. *Magnetic Resonance in Medicine*, *49*(1), 193–197. https://doi.org/10.1002/mrm.10354

Jenkinson, M., Bannister, P. R., Brady, M., & Smith, S. M. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, *17*(2), 825–841. https://doi.org/10.1016/S1053-8119(02)91132-8

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Jenkinson, M., & Smith, S. M. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–156. https://doi.org/https://doi.org/10.1016/S1361-8415(01)00036-6

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 (Hove, England)*, *26*(6), 640–648. https://doi.org/10.1080/20445911.2014.929580

Klein, R. M. (2014). Is there a benefit of bilingualism for executive functioning? *Bilingualism: Language and Cognition*, *18*(01), 29–31. https://doi.org/10.1017/S1366728914000613

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging*, *15*(1), 126–147. https://doi.org/10.1037/0882-7974.15.1.126

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). **lmerTest** Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, *144*(4), 394–425. https://doi.org/10.1037/bul0000142

Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex*, *58*, 301–324. https://doi.org/10.1016/j.cortex.2014.05.001

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing Access to the Native Language While Immersed in a Second Language Evidence for the Role of Inhibition in Second-Language Learning. *Psychological Science*, *20*(12), 1507–1515. https://doi.org/10.1111/j.1467-9280.2009.02480.x.Losing

Luk, G., Anderson, J. A. E., Craik, F. I. M., Grady, C. L., & Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: Response inhibition versus interference suppression. *Brain and Cognition*, *74*(3), 347–357. https://doi.org/10.1016/j.bandc.2010.09.004

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. https://doi.org/10.1080/20445911.2013.795574

Luo, D., Kwok, V. P. Y., Liu, Q., Li, W., Yang, Y., Zhou, K., … Tan, L. H. (2019). Microstructural plasticity in the bilingual brain. *Brain and Language*, *196*, 104654. https://doi.org/10.1016/j.bandl.2019.104654

Nichols, E. S., & Joanisse, M. F. (2016). Functional activity and white matter microstructure reveal the independent effects of age of acquisition and proficiency on second-language learning. *NeuroImage*, *143*, 15–25. https://doi.org/10.1016/j.neuroimage.2016.08.053

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, *66*(2), 232–258. https://doi.org/10.1016/j.cogpsych.2012.12.002

Paap, K. R., Johnson, H. A., & Sawi, O. M. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265–278. https://doi.org/10.1016/j.cortex.2015.04.014

Pliatsikas, C., Johnstone, T., & Marinis, T. (2014). fMRI Evidence for the Involvement of the Procedural Memory System in Morphological Processing of a Second Language. *PLoS ONE*, *9*(5), e97298. https://doi.org/10.1371/journal.pone.0097298

R Core Team. (2014). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rossi, E., Newman, S., Kroll, J. F., & Diaz, M. T. (2018). Neural signatures of inhibitory control in bilingual spoken production. *Cortex*, *108*, 50–66. https://doi.org/10.1016/j.cortex.2018.07.009

Rothman, J., González Alonso, J., & Puig-Mayenco, E. (2019). *Third Language Acquisition and Linguistic Transfer*. Cambridge: Cambridge University Press. Retrieved from https://www.cambridge.org/gb/academic/subjects/languages-linguistics/psycholinguistics-and-neurolinguistics/third-language-acquisition-and-linguistic-transfer-volume-163?format=HB

Rubin, O., & Meiran, N. (2005). On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology: Learning Memory and Cognition*, *31*(6), 1477–1491. https://doi.org/10.1037/0278-7393.31.6.1477

Seo, R., Stocco, A., & Prat, C. S. (2018). The bilingual language network: Differential involvement of anterior cingulate, basal ganglia and prefrontal cortex in preparation, monitoring, and execution. *NeuroImage*, *174*, 44–56. https://doi.org/10.1016/j.neuroimage.2018.02.010

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143–155. https://doi.org/10.1002/hbm.10062

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., … Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*(SUPPL. 1), S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051

Surrain, S., & Luk, G. (2017). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*, 1–15. https://doi.org/10.1017/S1366728917000682

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, *18*(01), 3–24. https://doi.org/10.1017/S1366728914000522

Veroude, K., Norris, D. G., Shumskaya, E., Gullberg, M., & Indefrey, P. (2010). Functional connectivity between brain regions involved in learning words of a new language. *Brain and Language*, *113*(1), 21–27. https://doi.org/10.1016/j.bandl.2009.12.005

Woolrich, M. W. (2008). Robust group analysis using outlier inference. *NeuroImage*, *41*(2), 286–301. https://doi.org/10.1016/j.neuroimage.2008.02.042

Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*, *21*(4), 1732–1747. https://doi.org/10.1016/j.neuroimage.2003.12.023

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage*, *14*(6), 1370–1386. https://doi.org/10.1006/nimg.2001.0931

Yamasaki, B. L., Stocco, A., & Prat, C. S. (2018). Relating individual differences in bilingual language experiences to executive attention. *Language, Cognition and Neuroscience*, *33*(9), 1128–1151. https://doi.org/10.1080/23273798.2018.1448092

Zhou, B., & Krott, A. (2018). Bilingualism enhances attentional control in non-verbal conflict tasks – evidence from ex-Gaussian analyses. *Bilingualism: Language and Cognition*, *21*(01), 162–180. https://doi.org/10.1017/S1366728916000869

Acknowledgements

Ethics Statement

This research procedures in this study were approved by the University of Reading Research Ethics Committee. Prior to taking part in the experiment, participants gave written informed consent and confirmed no contraindication to MRI scanning.

Footnotes

---

[i] We used the version of the LSBQ (Version 1) available at the time of testing. Anderson and colleagues used a slightly updated version to create their factor score calculator (Version 3+). Thus, one of the questions included in the score L2_social ("Language use with Friends") could not be included the calculation of this factor score. In consultation with John Anderson to determine a good approximation, this question was removed from our factor score calculation such that the weighting automatically shifted for the remaining 19 of 20 questions.