1

2

3   1

4

5

6   Running head: Executive control in older bilinguals

7

8

9

10

Language proficiency does not modulate executive control in older

bilinguals

Ramesh Kumar Mishra [#,a], Manasa Padmanabhuni[a], Pratik Bhandari[b],

Shiji Viswambharan[a], & Seema Gorur Prasad [#,a]

[a]Centre for Neural and Cognitive Sciences, University of Hyderabad, India

[b]Basque Center on Cognition, Brain and Language, Spain

[#]Ramesh Kumar Mishra and Seema Gorur Prasad contributed equally to the manuscript.

Correspondence regarding this manuscript should be addressed to:

Ramesh Kumar Mishra, Centre for Neural and Cognitive Sciences, Science Complex,

University of Hyderabad, Hyderabad, Telangana 500046

Ph No: 09451872007

Email id: rkmishra@uohyd.ac.in

2

Abstract

We examined if language proficiency modulates performance in tasks that measure executive control in older Telugu-English bilinguals (n = 50, Mean age = 57.15 years). We administered numerical Stroop task, ANT, DCCS task and stop-signal task that are known to tap into different aspects of executive functioning on healthy ageing Telugu-English bilinguals. Second language (English) proficiency was calculated as a cumulative score that considered both subjective and objective measures of L2 fluency and use. Bilinguals were divided into two groups based on the cumulative score and compared on each task. We did not find any effect of language proficiency on any of the executive control measures. The additional Bayesian analysis also supported these findings. Therefore, the results do not support the claim that bilingual language proficiency modulates executive control, at least in the elderly population. We discuss the results with regard to the issue of bilingual advantage in executive control and the role of age and language use.

3

Introduction

Whether bilingualism should have any positive effect on one's cognitive and executive control has been controversial lately and remains unsettled (for reviews see Kroll & Bialystok, 2013; Paap, Johnson, & Sawi, 2015). Those who have obtained positive correlations between bilingualism and measures of executive functioning have attributed them to different mechanisms such as superior inhibitory control (Green, 1998), attentional disengagement (Mishra, Hilchey, Singh, & Klein, 2012), monitoring (Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009), executive processing (Hilchey & Klein, 2011), and selective attention (Chung-Fat-Yim, Sorge, & Bialystok, 2017). However, others have found null results in conceptual replications (Anton et al., 2014; Anton, García, Carreiras, & Duñabeitia, 2016; Paap & Greenberg, 2013). Apart from the question of which component of executive functions bilingualism should improve, in whom these improvements should be seen is critical. It's intuitive to expect that older people should show benefits of bilingualism more than younger people given their extensive use of and experience with bilingualism. In this study, we examined if such extensive practice with language management over the years makes older bilinguals perform better at executive control tasks. Many studies have shown that bilingualism provides a cognitive reserve that pushes the onset age of degenerative diseases (Alladi et al., 2016, 2017; Bialystok, Anderson, & Grundy, 2018; Bialystok, Craik, & Freedman, 2007; Filippi, D'Souza, & Bright, 2018; Woumans et al., 2015). But, most such reports have relied on hospital records and indirect methods of data acquisition on language use. It is necessary to determine if bilingualism modulates executive control in older, healthy bilinguals even before the onset of these diseases. We examined Telugu-English bilinguals on

4

a range of executive control tasks to explore if their second language (L2) proficiency predicts performance on such tasks.

The few studies that have examined bilingual advantage in older bilinguals have obtained conflicting results (See Figure 1 for a summary of some studies). In one of the earliest studies, Bialystok, Craig, Klein, and Viswanatha (2004) found reduced Simon effects in older bilinguals compared to monolinguals. Since then some behavioural and brain imaging evidence have been found for the positive effects of bilingualism on task-switching (e.g., Houtzager, Lowie, Sprenger, & De Bot, 2017; Gold et al., 2013), inhibitory control (eg., Bialystok et al., 2006, 2008), episodic memory (Schroeder & Marian, 2012) and selective attention (eg., Soveri et al., 2011) in healthy, older bilinguals. However, not all have observed such positive correlations between bilingualism and executive control in older participants (Anton et al., 2016; Incera & McLennan, 2017; Ivanova et al., 2016; Papageorgiou et al., 2018; Struys, Duyck, & Woumans, 2018). For instance, Ivanova et al. (2016) found that older bilinguals are not better in inhibiting intrusions during language switching. Recently, Borsa et al. (2018) compared bilinguals and monolinguals on the Attention Network Task (ANT) in a fMRI study. While they found evidence of a "neural reserve" for bilinguals compared to monolinguals in terms of ACC activation, the two groups did not differ on the behavioural measures.

(Figure 1 about here)

Another strand of research suggests that lifelong bilingualism prevents early onset of neurodegenerative diseases and offers protection against neural tissue loss (Alladi et al., 2013,

5

2016, 2017; Bialystok et al., 2007; Woumans et al., 2015). It is assumed that the extensive use

of two languages across the lifespan enhances the executive control system which in turn

provides a protective cover against such diseases (Abutalebi & Green, 2016; Bialystok,

Poarch, Luo, & Craik, 2014; Bialystok et al., 2016). But, it is to be noted that several others

following a similar methodology of using hospital records have not found any difference in

the onset age of dementia between bilinguals and monolinguals (eg., Clare et al., 2016;

Lawton, Gasquoine & Weimer, 2015). For instance, Clare et al. (2016) did not find any

significant differences in the age of onset of Alzheimer's disease bilingual older adults

compared to the monolinguals. Gollan et al. (2011) found that the age of onset of Alzheimer's

disease in Spanish-English bilinguals was later than in monolinguals, but only for the low

educated bilinguals.

Bilingualism induced cognitive reserve against diseases is mediated through

bilingualism's effect on executive control. Among many casual factors, L2 proficiency seems

critical as a proxy for bilingual fluency and competence. L2 proficiency (particularly for the

unbalanced, sequential type bilinguals) has been linked to language non-selective activation

(Mishra & Singh, 2016) and also to executive control (Mishra et al., 2012; Singh & Mishra,

2012, 2013, 2014, 2015). Similar findings have also been reported by other groups (Dong &

Xie, 2014; Festman & Münte, 2012; Iluz-Chen & Armon-Lotem, 2014; Tao et al., 2011; Tse

& Altarriba, 2012). Further, brain imaging studies have shown that higher L2 proficiency

modulates activation of the Anterior Cingulate Cortex (Abutalebi et al., 2013). Many

developmental studies have also shown that an increase in L2 proficiency in children alters

brain networks that are critical in conflict resolution, which presumably is also recruited for

6

language control in bilinguals (Green, 1998). Luk and Bialystok (2013) have suggested that language proficiency as a continuous measure can be a good variable instead of the traditional comparisons between bilinguals and monolinguals. Considering this, we correlated composite scores on L2 proficiency with executive control measures in older bilinguals.

Current study

Given this background, where some have found evidence for superior cognitive reserve in older, healthy bilinguals compared to monolinguals, and others have found null results, it is important to seek more cross-cultural data from healthy older bilinguals. The bilinguals in Hyderabad (a large cosmopolitan city in Southern India) live and work in a densely bilingual society. Hyderabad is a metropolitan city with Telugu as its primary language. Telugu is a Dravidian language largely spoken in the states of Telangana and Seemandhra. Telugu has its script which is derived from Brahmi. This script is alphasyllabic or more recently people have termed it "Akshara" (Nag, 2011). In contrast, English is a Roman language and has an alphabetic script. Telugu and English also have different phonological systems and there are no cognates between the languages. English is often spoken as the lingua franca in both work and home environments. We tested a group of Telugu-English bilinguals in Hyderabad on a range of executive control tasks. These bilinguals were fluent in English which they had acquired early in life.

In India, there are English-medium schools that have English as the main medium of education or state-run schools in different provinces where English is only taught as L2. However, all post-school education in India is in English. Therefore, although individuals

7

who have studied in English-medium schools tend to have higher fluency in L2, others

develop reasonable fluency in L2 when they enter college and university. There are few

studies that have directly examined how the medium of instruction at school or college may

modulate executive control directly. Our participants were all exposed to English during their

college years since all of them had at least completed graduate education. Many were in

professions where the demand for English was much higher than other professions. Since

these people are older bilinguals, the degree to which English was used at home vs. outside

could have differed. There is scant data on the quantum of English used by educated older

Indians at home and work. Many writers have indicated that in the last several decades the use

of English has grown significantly in India (Mohanty, 2010). In sum, our older bilinguals

having studied English either at school or more thoroughly in English medium schools had

experience in using both English and Telugu over a long period. Our primary independent

variable was language proficiency for which we developed a cumulative score (Ma, Chen,

Guo, & Kroll, 2017). This score gave an overall index of language ability and was developed

to treat L2 proficiency as a continuous variable. We reasoned that older bilinguals with better

cumulative scores should do better on executive control tasks.

Bilingual advantages have been mostly found across four domains (Bastian & Gade,

2016; Klein, 2016): an inhibitory control advantage (Green, 1998), an executive processing

advantage (Hilchey & Klein, 2011), advantage in switching or mental flexibility (Bialystok et

al., 2012) and more recently an advantage in attentional engagement (Bialystok, 2015).

Bilinguals activate both their languages simultaneously and have to routinely deal with

competition between these two languages to select the appropriate language. The control

8

mechanisms involved in this selection are said to transfer to other nonlinguistic domains as well. But the exact nature of this domain-general control process is still debated. The inhibitory control (IC) model was proposed with the assumption that selecting the appropriate language involves inhibiting the irrelevant language which then enhances inhibitory control in general. However, Hilchey and Klein (2011) argued that in several studies that apparently showed IC advantage, participants were faster on both congruent and incongruent trials which suggests an overall speed advantage and not necessarily only an advantage in inhibition. Subsequently, several studies argued that bilinguals show an advantage on tasks requiring mental flexibility such as switching since they have to routinely switch between two languages (Prior & MacWhinney, 2010). Most recently, the attentional system has been proposed as the core cognitive system responsible for bilingual advantage (Bialystok, 2017). Attending to two languages since infancy is said to enhance executive attention which then manifests in different tasks. There have been studies in support of all the above models. Thus, it is still not clear exactly which specific aspect of cognitive control is influenced by bilingualism or L2 proficiency.

We chose four tasks that each tap into one or more of the executive functions mentioned above: Attention Network Test (ANT), numerical Stroop task, Dimensional Change Card Sorting task (DCCS) and stop-signal task. The executive processing advantage is commonly observed as an overall RT advantage on a task due to enhanced performance. The attentional advantage was assessed through the ANT task. The task consists of three independent attentional networks: alerting, orienting and executive control network. Participants respond to a central directional arrow on the screen flanked by other arrows on

9

either side. A cue (central/spatial/double/no) is presented before the target phase. Bilinguals and monolinguals have shown differences in alerting (Costa et al., 2006), orienting (Poarch & van Hell, 2012) and conflict effect (Costa et al., 2006, 2009; Tao et al., 2011). Tao et al. (2011) found that early bilinguals (with higher L2 proficiency) showed reduced conflict effect compared to late bilinguals on a modified ANT task. In contrast, Pelham and Abrams (2014) did not observe such differences between early and late adult bilinguals. Anton et al. (2014) also failed to observe an advantage in the ANT task in bilingual children compared to monolingual children.

Numerical Stroop task and stop-signal tasks were used to examine the inhibitory control advantage. Numerical Stroop task is said to tap into the inhibitory control ability of bilinguals. Participants respond to the physical size of a number and ignore the numerical magnitude of the number. While some studies have found that young bilingual adults perform better (larger facilitation and lesser interference) than monolinguals on this task (Hernandez et al., 2010), studies with children and elderly population have sometimes failed to observe such differences (Anton et al., 2016; Dunabeitia et al., 2014). The stop-signal task has been used to measure reactive inhibition in participants. We chose this task to examine whether lifelong bilingualism leads to a benefit in the ability to reactively inhibit. In this task, the participants are required to quickly abort a response when presented with a stop-signal. The evidence for bilingual advantage in this task has been mixed. While Colzato et al. (2008) found no differences between monolinguals and bilinguals in the performance on the stop-signal task, using a similar task Antoniou et al. (2016) found that children speaking two closely related languages have better inhibitory control than monolinguals. The dimensional card sorting

10

(DCCS) task was administered to examine the set-shifting ability and mental flexibility in bilinguals. The task involves categorising bivalent stimuli based on either the shape of the object or the colour of the object. Bilinguals have been found to perform better in DCCS task compared to monolinguals indicating that they are better at shifting between two representations (Qu, Low, Zhang, & Zelazo, 2016; Yang, Hartanto & Yang, 2017). We divided the participants into high and low proficient based on the L2 cumulative score and compared the performance on all the tasks using both traditional and Bayesian analyses.

Methods

Participants

Fifty healthy ageing adults from the city of Hyderabad (34 male, Mean age = 57.2 years, SD = 4.12, Age range = 50 - 65 years) participated in the experiment (see Table 1 for demographic details). All participants reported normal or corrected-to-normal vision. All participants had Telugu as their L1 (native language). The mean age of acquisition (AoA) of L1 was 1.9 years (SD = 1.2) which was calculated based on the self-reported age at which the participants remember first using L1 words in speech. AoA of L2 was 9.46 years (SD = 2.91). All participants provided written consent for their participation. The Institutional Ethics Committee of the University of Hyderabad approved the study.

(Table 1 about here)

11

One crucial step in bilingual advantage research is finding an appropriate way to quantify second language proficiency. Researchers have traditionally used either subjective measures such as self-rated proficiency or standardised objective tests such as lexical decision tasks or verbal fluency tasks to assess language proficiency. In a comprehensive analysis of the different factors contributing to bilingualism, Luk and Bialystok (2013) found that bilingual experience was best characterised by a combination of multiple factors – assessed both subjectively and objectively. The authors thus suggest that "a composite of self-rated English proficiency in listening and speaking in conjunction with objective testing is more informative than either method alone" (p. 9). Thus, in this study, participants were divided into high and low L2 proficiency (HP and LP) groups using a median split on L2 proficiency composite score. Many recent studies have used such a composite score to denote bilinguals with different language proficiency (e.g., Ma et al., 2017; McMurray et al., 2010).

L2 proficiency composite score for each participant was calculated by first adding z-scores of all the objective L2 proficiency measures (i.e., score on WordORnot vocabulary test, L2 semantic fluency score, and L2 naming latency† in object naming task) and the L2 self-report score from the questionnaire data, then dividing the sum by square root of sum of variances and covariances of these individual values. Self-assessment has been found to provide a reliable measure of proficiency in bilinguals (Marian, Blumenfeld, & Kaushanskaya, 2007). Thus, to further validate the accuracy of the composite score, we performed correlation between L2 composite score and L2 self-reported proficiency (averaged over reading, speaking and comprehension). This correlation was found to be significant, $r = 0.73$, $p < 0.001$.

The mean L2 proficiency composite score for HP bilinguals was significantly higher ($M = 0.05$, $SD = 0.01$, t (36) = 8.38, $p < 0.001$) in comparison to the LP bilinguals ($M = 0.03$, $SD = 0.01$). No such differences were observed for the L1 proficiency composite score, t (48) = 0.65, $p = 0.52$. The scores on all the control tasks (reported in Table 1) were then compared between the two groups.

Control tasks

WordORnot. An online vocabulary test (WordORnot, centre for Reading Research, Ghent University) was administered to all the participants as one of the measures of their L2 proficiency. Participants were instructed to judge strings of English letters as a "word" or a "non-word". The total score was the difference between the percentage of correct and incorrect responses. The score for HP bilinguals ($M = 54.52$ %, $SD = 15.48$) was significantly greater than the score for LP bilinguals ($M = 33.44$ %, $SD = 17.46$), t (48) = 4.52, $p < 0.001$

Semantic Fluency Test. We administered semantic fluency tests in both Telugu and English (please Table 1 for the cores). Participants were asked to generate as many names as they could of everyday objects of 4 semantic categories ("birds", "vegetables", "animals", and "fruits") within 1 minute for each category. Language mapping with the categories was counterbalanced across participants. Semantic fluency scores (i.e., the average number of words produced per language per minute) was calculated. The L2 semantic fluency score was significantly higher for HP bilinguals ($M = 11.38$, $SD = 2.91$) compared to LP bilinguals ($M =$

13

7.58, $SD = 2.72$), $t (48) = 4.77$, $p < 0.001$. The two groups also differed on L1 semantic

fluency score (HP: $M = 13.48$, $SD = 3.9$, LP: $M = 11.4$, $SD = 2.8$), $t (48) = 2.1$, $p = 0.03$.

Language Questionnaire. Participants filled a language background questionnaire

that had questions on the native language, languages known, the AoA of L1 and L2, and

percentage of time exposed currently to L1, L2, and other language(s) if any. They also

provided their L1 and L2 proficiency self-report score for speaking, understanding and

reading ability on a scale of 1 to 10, 10 being the highest. The average self-reported L2

proficiency was higher for the HP bilinguals compared to the LP bilinguals, $t (48) = 4.6$, $p <$

0.001. No difference was found in the average self-reported L1 proficiency, $t (48) = 0.56$, $p =$

0.58. The self-report scores for each category (in both L1 and L2) along with significance

values (wherever found) are mentioned in Table 1.

Socio-Economic Status (SES) Questionnaire. Participants were asked to fill in a

three-point SES questionnaire where they had to indicate if they belonged to upper, middle or

lower socio-economic class. Alongside, they also indicated their monthly income. There was

no significant difference, $t (48) = 1.12$, $p = 0.27$ in SES of HP bilinguals ($M = 2.6$, $SD = 0.5$)

and LP bilinguals ($M = 2.44$, $SD = 0.51$).

Voluntary Object naming. A voluntary object naming task was administered to

further obtain an objective measure of proficiency in Telugu and English. The procedure for

the selection of black and white line drawings used in the experiment was similar to the one

used in the previous studies of Mishra and colleagues (Roychoudhury, Prasad, & Mishra,

14

2016; Bhatia, Prasad, Sake, & Mishra, 2016). Before the naming task, participants were exposed to the standardised 147 images with their names written both in English and in Telugu. The stimuli in the object naming task were presented using DMDX (Forster & Forster, 2003) version 5.1.1.3 with DirectX 9.0 on a laptop with 1366 x 768 pixel resolution and 60 Hz refresh rate. Participants were seated on a chair at a distance of 60 cm from the monitor. An i-ball M-27 table microphone was used to collect responses. Response latencies were recorded by software DMDX through a voice trigger. A fixation cross appeared at the centre of the screen for 1000 ms at the beginning of every trial. Then after a blank screen of 250 ms following the disappearance of the fixation cross, a black-and-white image to be named was displayed for a maximum duration of 2500 ms or till the participant named the picture. Participants were instructed to name the pictures "in any language that comes to their mind first" as quickly and as accurately as possible. Participants were asked to maintain a balance between the two languages when naming the pictures. We used this method to see if voluntarily these bilinguals are able to choose and name in both the languages.

The audio responses of the object-naming task were recorded by software Audacity-win2.0.5 using which naming errors were manually coded. Object naming latencies on correct trials in L1 and L2 were calculated after discarding trials with latencies less than 150 ms, trials with no or incorrect responses. There was no significant difference in L2 naming latency between HP (M = 1156.82 ms, SD = 201.95, t (48) = - 0.260, p = 0.8) and LP (M = 1172.39 ms, SD = 220.56) bilinguals. Participants did not differ in L1 naming latency as well, t (48) = - 0.9, p = 0.4 (HP: M = 1142 ms, SD = 237, LP: M = 1202 ms, SD = 229). L2 naming latency

15

for the two groups was also taken as one of the components for calculation of L2 proficiency composite score.

Ravens Standard Progressive Matrices. Non-verbal IQ was assessed through Ravens Standard Progressive Matrices (Raven, 1958). The task consisted of 60 items divided into 5 types (A to E), each type with 12 items. Each item was a pattern with a missing part and participants had to fill in the missing part from the possible pieces that could fit in the missing part. The order of types and items were arranged in the order of difficulty. Correct answers were counted as IQ score. HP bilinguals ($M = 36.04$, $SD = 7.04$) had higher IQ compared to the LP bilinguals ($M = 30.42$, $SD = 7.46$), $t (48) = 2.69$, $p = 0.01$

Mini-Mental State Examination (MMSE). MMSE (Folstein, Folstein, & McHugh, 1975) was administered to all participants to assess their global cognitive status. MMSE is a test that is widely used by clinicians and researchers to test cognitive function, especially in the elderly population. It is a 30-point questionnaire that examines basic functions under different categories such as recall ("Can you remember the three words you have just said?"), spatial and temporal orientation ("which city are we in?"), ability to follow simple commands ("fold this piece of paper") and so on. Each category carried different points, and the maximum possible score in MMSE was 30. HP bilinguals ($M = 28.92$, $SD = 1.53$) and LP bilinguals ($M = 28.72$, $SD = 1.28$) did not differ in their cognitive status as indicated by the scores on Mini-Mental State Examination, $t (48) = 0.5$, $p = 0.62$.

16

Visual Digit Span. The digit span task was administered using Inquisit 4.0.0 (2004) which adapted scripts from Woods et al (2011). Five participants (2 HP and 3 LP bilinguals) refused to complete the task. So, it was administered only to 45 participants. In the forward digit span task, there was no significant difference in the number of digits recalled between HP bilinguals ($n = 23$, $M = 6.43$, $SD = 0.95$) and LP bilinguals ($n = 22$, $M = 5.91$, $SD = 1.41$), t (43) = 1.47, $p = 0.15$ However, in backward digit span task, the number of digits recalled by HP bilinguals ($n = 23$, $M = 5.87$, $SD = 1.058$) was more compared to LP bilinguals ($n = 22$, $M = 5.18$, $SD = 1.37$) and this effect was marginally significant, t (43) = 1.89, $p = 0.07$.

Stimuli and procedure

Attentional Network Test (ANT)

Attentional Network Test (Fan et al., 2002; Costa et al., 2008, 2009) was administered to all the participants. Each trial began with a fixation cross at the centre of the screen for 400 ms. Then a cue appeared for 100 ms followed by a fixation cross for 400 ms. There were four cue conditions: centre cue, double cue, spatial cue (above or below the central fixation) and no-cue. In the centre cue condition, an asterisk appeared at the centre of the screen. In the double cue condition, two asterisks appeared simultaneously 1.06 degree above and below the central fixation cross. In the spatial cue condition, an asterisk appeared 1.06 degree either above or below the central fixation cross. In the no cue condition, no asterisk appeared. In all, except the centre cue condition, fixation cross remained on the screen throughout the cue presentation duration of 100ms. Following this, a target arrow with flankers was presented for

17

1700 ms or till the participant made a response. The target stimuli were presented as a set of

five horizontal black lines with an arrowhead pointing either right or left on a white

background screen. Each line or arrow subtended a visual angle of 0.55 degree, and the

distance of separation between two arrows/lines in a set was 0.06 degree. The set of arrows

appeared 1.06 degree above or below the central fixation cross. Three different types of target

and flanker combinations were used: congruent (→→→→→ or ←←←←←), incongruent

(←←→←← or →→←→→) or neutral (— — ← — — or — — → — —). Participants were

instructed to respond to the central arrow. If the central arrow was pointing towards left, they

had to press "A" and, if it was pointing towards right, they had to press "L" on the keyboard.

Participants were instructed to "try to be as quick and as accurate as possible".

Two within-subjects factors: cue type (central cue, double cue, no cue, spatial cue) and

flanker type (congruent, incongruent, neutral) interacted leading to 12 experimental

conditions. There was a total of 96 trials (12 experimental conditions repeated eight times).

All trials were randomised and presented in one block. There were total 32 congruent trials,

32 incongruent trials, and 32 neutral trials. There was a break for 2 minutes after 48 trials.

Twelve practise trials were administered before the experimental trials.

Dimensional Change Card Sort (DCCS) task

Standard condition of DCCS task (Qu et al., 2013; Wilson et al., 2008) was adopted

and administered to all participants. Three cards were simultaneously presented on the

computer screen along with a cue word "COLOUR" or "SHAPE". Two target cards were

18

presented on the top right and on the top left and a test card was presented on the bottom centre of the screen. The target cards were of different shape and colour (e.g., red star and a blue square). Participants were instructed to match the test card with one of the target cards according to the cue word ("SHAPE" or "COLOUR"). They had to press "A" if the target card on the left side matched with the test card, and press "L" if the target card on the right side matched with the test card. Participants were instructed to "try to be as quick and as accurate as possible".

The task consisted of two blocks: dominant block (non-switch block) and secondary block (switch block). There were 40 trials in the dominant block in which participants had to match target stimuli with the test stimuli by a dominant dimension. There were two types of dominant block: Colour dominant and Shape dominant. The dominant blocks were non-switch blocks because all the trials within a dominant block were to be matched by the same dimension (colour only, or shape only). The trials in a dominant block are considered baseline trials (discussed in 'Results' section).

In the secondary block, there were total 78 trials – 68 dominant trials of dominant dimension and 10 secondary trials of secondary dimension. There were two types of secondary block: one with "Colour" as dominant dimension and "Shape" as a secondary dimension, and the other with "Shape" as dominant dimension and "Colour" as a secondary dimension. If the dominant block was "Colour" then dominant dimension in the secondary block would be "Colour" and secondary dimension would be "Shape". If the dominant block was "Shape" then dominant dimension in the secondary block would be "Shape" and

19

secondary dimension would be "Colour". In the secondary block, the dominant trials and

secondary trials were randomised across the participants. The secondary block was switch

block as participants had to switch in some trials between dimensions (shape and colour).

There were two versions of the DCCS task, one with "Colour" as a dominant

dimension (i.e., Colour dominant version), and the other with "Shape" as a dominant

dimension (i.e., Shape dominant version). Each participant performed both the two versions

of this experiment.

Stop-signal task

Standard visual two choice stop-signal task (Colzato et al., 2008) with visual go signal

and an auditory stop signal was administered to all participants. The interval between two

choice signals varied from 1250 ms to 1750 ms randomly in steps of 125 ms (i.e., 1250 ms,

1375 ms, 1500 ms, 1675 ms and 1750 ms). During this interval, a fixation cross was presented

at the centre of the screen. The go signal was an arrow pointing towards right or left, for 2000

ms or till a response was made. Participants were instructed to press "A" on the keyboard if

the arrow was pointing towards left and "L" if the arrow was pointing towards the right. The

"stop" signal was a short auditory signal (a beep sound) presented at a short delay after the

presentation of the go signal. Participants were instructed to withhold their response in such

stop-signal trials. The difficulty of the stop-signal trials was manipulated by varying the stop-

signal delay (SSD), i.e., the time between the presentation of go signal and the presentation of

the auditory stop signal.

20

There were total 110 experimental trials out of which 30% (i.e., 33 trials) were stop-signal trials and 70% (i.e., 77 trials) were go trials or choice trials. The SSD was varied systematically in a simple staircase or a tracking method by gradual increase or decrease from the previous SSD by 35ms. If a previous stop-signal trial was successfully inhibited, then SSD was increased by 35ms to increase the difficulty in the next trial, and if the inhibition was not successful, then in the next trial, the SSD was decreased to decrease the difficulty. There were 12 practise trials before the experimental trials.

Numerical Stroop task

Numerical Stroop task (Anton et al., 2016) was administered to all the participants. Each trial began with a fixation cross at the centre of the screen for 300 ms. Then a pair of digits in Courier New black font appeared on the screen, for 4000 ms or till the participant made a response. The digits were of two different sizes, font sizes 24 and 48. There were three different combinations of digit-pair: congruent (when the bigger digit in size had a bigger numerical value), incongruent (when the bigger digit in size had a smaller numerical value) and neutral trial condition (when both the bigger and the smaller digit had the same numerical value). Participants were instructed to respond to the bigger-sized digit in the digit pair, regardless of its numerical value. They had to press "A" on the keyboard if the digit on the left side was bigger and press "L" if the digit on the right side was bigger. For half of the trials, digits on the right side were bigger and for another half, the digits on the left side were bigger. Participants were instructed to try to be "as quick and as accurate as possible".

21

Trial types were the within-subjects factor: congruent trials, incongruent trials, and neutral trials. There were total 96 experimental trials created using eight digits (1, 2, 3, 4, 6, 7, 8 and 9). Each digit was presented the same number of times (4 times) in each condition (congruent, incongruent, and neutral) and in total 24 times. There was an equal number of trials in all the three conditions: 32 congruent trials, 32 incongruent trials, and 32 neutral trials. In all the conditions, responses were counterbalanced across the sides: in half of the trials responses were on the right side and in the other half they were on the left side, and trials were randomised across participants. Twelve practise trials were given before the main experiment.

Participants first completed the 4 main cognitive control tests followed by the control tasks and the language questionnaire. The task order was randomized across participants to prevent the effects of fatigue and practice. The experiment was about two hours long. Participants were offered a 10-minute break half-way through the session, if they needed to take rest. All the tasks were administered in a quiet room with no distractions on a Dell Vostro machine with Intel Core i5 processor running Windows 7 connected to a LCD monitor with resolution 1024 * 768 pixels and refresh rate of 60 Hz. Responses were collected using a Logitech wireless keyboard.

Data analyses and Results

In each task, trials with RT above and below 2 standard deviations from each participant's mean were excluded as outliers. The resulting data was analysed through

22

ANOVA using SPSS (IBM SPSS Statistics for Windows). The specific factors and the

dependent variables are mentioned under each task. The descriptive statistics for all the tasks

are given in Appendix A. The results are shown in Figure 2. Some researchers have argued

that bilingualism must be treated as a continuous variable (Luk & Bialystok, 2013). Thus, we

also performed regression analyses on all the important task measures while treating

composite L2 proficiency score as a continuous variable.

Bayesian hypothesis testing and estimation

Since we did not find any group difference with traditional analyses, we additionally

performed Bayesian analysis because the traditional hypothesis testing only provides evidence

to reject a null hypothesis. Researchers have advocated for the additional use of Bayesian

statistics along with traditional statistics especially in the case of nonsignificant results

(Dienes, 2014). Two methods are primarily used to assess null values using the Bayesian

framework: Bayesian hypothesis testing using Bayes factor and Bayesian parameter

estimation (Kruschke, 2011). Bayesian hypothesis testing involves setting up two models with

two contrasting hypothesis and adjusting the likelihood of each model based on the evidence

(data obtained). We tested for model H_0 defined as no differences between the two groups

against model H_1 defined as better performance for HP bilinguals compared to LP

bilinguals.We performed independent sample t-tests using JASP comparing both the groups

on the key dependent measure in each task. Bayes factors (BF_{01}) were computed which

quantify the relative evidence for the two competing hypothesis. According to a commonly

accepted convention (Jefffreys, 1961; Kruschke, 2011), $3 > BF_{01} > 1$ denotes anecodtal

evidence, $10 > BF_{01} > 3$ denotes moderate evidence and $BF_{01} > 10$ denotes strong evidence for

the null hypothesis. Similarly, $1/3 < BF_{01} < 1$ denotes anecdotal evidence, $1/10 < BF_{01} < 1/3$

23

denotes moderate evidence and $BF_{01} < 1/10$ denotes strong evidence for the alternate

hypothesis. $BF_{01} = 1$ suggests that the data is inconclusive.

Bayesian parameter estimation for comparing two groups is a method that provides

complete information about the posterior distribution (that is, the prior distribution updated by

the current data) of the two groups. This estimation method can be used to assess the

credibility of differences in means using the HDI and a ROPE (Region of Practical

Equivalence) value. Highest density interval (HDI) refers to the 95% of the most credible

values of the posterior distribution. If 95 % of the HDI (for the distribution of the difference

in group means) contains zero, then the data is suggestive of a lack of group difference

because it indicates that there is not much difference in the individual distributions of each

group. A more conservative decision criterion to estimate the likelihood of group differences

using the ROPE has also been suggested (Kruschke, 2011). ROPE is a range of values around

0 that is statistically equivalent to the null value. Thus, the ROPE essentially gives the

minimum difference in group means that is practically equivalent to no difference between the

groups. The ROPE is generally decided based on previous research findings. If 95% of the

HDI falls within the ROPE, then the null effect is accepted. If 95% of the HDI falls outside

the ROPE, then the alternate hypothesis is accepted. The posterior distribution of the

difference in group means for each executive control measure along with the ROPE limits

was plotted (Figure 2) and analysed using the BEST package in R (Kruschke, Meredith &

Meredith, 2018).

There is no consensus in the literature regarding which of the two Bayesian methods is

more appropriate. Both the methods involve different interpretations of the Bayesian

inference (Kruschke & Liddell, 2018). We decided to test our data with both the approaches

24

to assess the likelihood of the null effects of L2 proficiency on cognitive control tasks in our

sample.

(Figure 2 about here)

(Figure 3 about here)

Attentional Network Test

Trials in which participants did not respond were excluded from further analyses

(3.83% of the total trials). Three participants' data had more than 80% outliers and were

excluded from further analyses. 4.55% trials in HP group and 4.6 % trials in LP group were

discarded as outliers. Incorrect trials (1.1%) were also filtered out before analysing the

response time data.

We performed separate analyses on each of the attention networks (Executive

network, Alerting network and Orienting network). For Executive network analysis, repeated

measures ANOVA was conducted on RT with group (HP and LP bilinguals) as a between-

subjects factor and flanker type (congruent, incongruent and neutral) as a within-subjects

factor. Similarly, for Alerting network analysis, ANOVA was performed on RT with cue type

(no cue, double cue) as within-subjects factor and group (HP and LP bilinguals) as between-

subjects factor. For Orienting network analysis ANOVA was performed on RT with cue type

(centre cue and spatial cue) as within-subjects factors and group (HP and LP bilinguals) as

between-subjects factor.

Executive Network. Repeated measures ANOVA on executive network showed a

significant effect of flanker type, $F(2, 44) = 168.65$, $p < 0.001$, $\eta^2 = 0.79$. Participants

25

responded faster ($p < 0.001$) on congruent trials (754.54 ms) compared to incongruent trials (880.95 ms). Responses on neutral trials (755.7 ms) were also significantly faster ($p < 0.001$) compared to incongruent trials. There was no difference ($p = 0.81$) between congruent RT and neutral RT. There was no main effect of group, $F(2, 44) = 0.02$, $p = 0.89$, $\eta^2 < 0.001$. The interactions between flanker type and group also did not reach the level of significance ($F < 1$).

We further explored whether L2 cumulative score modulated conflict effect. Linear regression analysis was used to test if the L2 cumulative score significantly predicted conflict effect. The results of the regression analysis indicated that the L2 cumulative score is not a significant predictor of conflict effect, $r = 0.07$, $F(1, 45) = 0.26$, $p = 0.61$.

These findings were further confirmed by Bayesian t-test on conflict effect comparing the two groups which resulted in a $BF_{01} = 2.79$ suggesting that the null hypothesis was slightly more likely than the alternative one.

Alerting Network. ANOVA on the alerting network showed that the main effect of cue type was significant, $F(1, 45) = 5.41$, $p = 0.03$, $\eta^2 = 0.11$. RT on trials with double cue were faster (792.88 ms) than those with no cue (805.31 ms). Main effect of group was not observed, $F(1, 45) = 0.02$, $p = 0.89$, $\eta^2 < 0.001$. Interaction between flanker type and group was also not significant ($F < 1$). The correlation between alerting effect and L2 proficiency was not found to be significant either, $r = 0.07$, $F(1, 45) = 0.22$, $p = 0.64$. The lack of group difference was also confirmed by Bayesian t-test on alerting effect ($BF_{01} = 2.46$) indicating that the null hypothesis was slightly more likely than the alternative hypothesis.

26

Orienting network. ANOVA on orienting network revealed that the main effect of cue type was not significant, $F(1, 45) = 0.24$, $p = 0.63$, $\eta^2 = 0.005$. The main effect of group was also absent, $F(1, 45) = 0.18$, $p = 0.67$, $\eta^2 = 0.004$, and the interaction between cue type and group was not significant ($F < 1$). Interestingly, linear regression showed a significant positive correlation between orienting effect and L2 proficiency, $r = 0.3$, $F(1, 45) = 4.43$, $p = 0.04$ indicating that older bilinguals with higher L2 proficiency are better at orienting to cues. Bayesian t-test on orienting effect favoured the null hypothesis ($BF_{01} = 5.97$) confirming the lack of group difference on orienting effect.

A ROPE of 20 ms around 0 was set for all the three effects based on previous studies on bilingual advantage using the ANT (Anton et al., 2014; Costa et al., 2006, 2009). Specifically -20 to +20 ms was the range used in Anton et al. (2014) where the authors used a similar estimation method to evaluate the null results. Bayesian estimation method revealed that for alerting effect, 91% of the HDI was within the ROPE. For orienting and conflict effect, 77% and 61% of the HDI was within the ROPE (Figure 3A, 3B and 3C).

Numerical Stroop task

Trials in which participants did not respond were excluded (0.39% of total trials). 3.52% trials in HP group and 3.51% trials in LP group were excluded as outliers. Incorrect trials (0.44%) were also filtered out before further analyses.

We analysed response time data for numerical Stroop task using repeated-measures ANOVA with condition (congruent, incongruent, neutral) as within-subjects factors and group as between-subjects factor. ANOVA on RT showed no significant main effect of condition, $F(2, 48) = 0.68$, $p = 0.51$, $\eta^2 = 0.01$) indicating lack of Stroop effect in these

27

participants. There was no difference between the RT on congruent (620 ms), incongruent (624 ms) and neutral trials (617 ms). Main effect of group was also not observed, $F(1, 48) = 0.12$, $p = 0.732$, $\eta^2 = 0.002$. Interaction between condition and group was also not significant ($F < 1$).

The lack of Stroop facilitation effect was confirmed by Bayesian t-test comparing RT on congruent trials with RT on neutral trials (Stroop facilitation) for all participants ($BF_{01} = 9.21$). RT on neutral trials was also compared with RT on incongruent trials (Stroop interference, $BF_{01} = 1.69$) which suggested the absence of Stroop interference effect, although the evidence was not very strong. Bayesian t-test comparing the two groups on Stroop interference effect suggested that null hypothesis was slightly more likely than the alternative one ($BF_{01} = 5.54$). We did not perform regression analysis, Bayesian t-test or Bayesian estimation to examine the influence of L2 proficiency on Stroop facilitation/interference as there was no evidence for the basic effects.

Stop-signal task

We used filtering criteria only for the choice trials before calculating Median RT. Trials in which participants did not respond were excluded from further analyses (3.3% of the choice trials). 4.24% of trials in HP group and 4.15% trials in LP group were rejected as outliers. Trials with error in discriminating the go signal (1.5%) were also filtered out before analysing the RT data. Trials with failed inhibition (responding to stop signal) are generally not excluded from analysis as the dynamic adjustment of the onset of stop signal is designed to generate about 50% stopping efficiency (Wildenberg et al., 2006). The HP and LP bilinguals were able to stop their responses efficiently on about 85% of the trials which is

much higher compared to the stopping rates observed normally. It is possible that the older

bilinguals prioritises accuracy over speed thereby successfully inhibiting most responses.

First, we analysed Median RT on choice trials. There was no significant difference ($t =$

-0.166, $p = 0.87$) in median choice RT between HP (978.06 ms) and LP bilinguals (987.28

ms). Bayesian t-test also supported this result with a $BF_{01} = 3.13$. Linear regression analysis

showed no significant correlation between Median RT and L2 proficiency, $r = 0.07$, $F(1, 48)$

$= 0.27$, $p = 0.6$. We performed ANOVA on Stop Signal Reaction Time (SSRT) to verify if L2

proficiency could modulate inhibition efficiency. SSRT was calculated by subtracting the

mean SSD from the median choice RT. SSRTs were estimated separately for each participant

and for each group. The main effect of group was not observed ($t = -0.486$, $p = 0.63$) which

reveals that the LP (492.54 ms) and HP groups (467.63 ms) did not differ significantly in

SSRT.

This result was also confirmed by a Bayesian t-test on SSRT ($BF_{01} = 2.4$) and a

regression analysis, $r = 0.02$, $F(1, 48) = 0.01$, $p = 0.91$. A ROPE of (-40, 40) was set around

0 to assess group differences in SSRT using the estimation method. There have been very few

studies examining cognitive advantage in bilinguals using the stop-signal task and a bilingual

advantage has rarely been found with 10 – 20 ms group differences (Colzato et al., 2008;

Morales et al., 2013). Thus, we set a slightly larger ROPE of (-40, 40) for this measure. The

posterior distribution showed that 53% of the HDI was within the ROPE limits (Figure 3D).

Dimensional Change Card Sort (DCCS) task

The first four trials in the switch block were removed before the analysis (see Qu et

al., 2013 for similar analysis). Incorrect responses and the trials immediately following an

29

incorrect response were filtered out (1.45% in dominant block and 10.86 % in secondary

block in Shape dominant version, and 5.9% in dominant and 10.7 % in secondary block in

Colour dominant version). Further, trials with RT less than 100 ms and trials with RT above

and below 2 standard deviations from each participant's mean were excluded as outliers.

Based on this criteria, 3.4% trials in HP group and 3.5 % trials in LP group were excluded in

the dominant block of the Shape dominant version. 2.7% trials in HP group and 2.6% trials in

LP group were discarded in the secondary block. In the colour dominant version, 3.12% in the

HP group and 3.16% of trials in the LP group in the dominant block were discarded. 2.46%

trials in HP group and 2.47% of trials in LP group were discarded in the secondary block.

We compared the participants' performance on the switch trials during the switch (or

secondary) blocks with the performance on the baseline trials during the non-switch

(dominant) blocks. Local switch cost for RT were calculated by subtracting the median RT on

the baseline trials from the median RT on the switch trials during the switch block, then

dividing the resultant value by the median RT on the baseline trials. We also calculated global

switch cost by subtracting the median RT on the baseline trials from the median RT on the

repeat (stay) trials during the switch block then dividing the resultant value by the median RT

on the baseline trials (Qu et al., 2013).

We compared the local switch cost and global switch cost in colour task (colour

dominant version) and shape task (shape dominant version) separately.

Colour Task. Independent-samples t-test was conducted to compare local switch cost

across the two groups. There was no significant difference in local switch cost between LP

bilinguals (1.0) and HP bilinguals (0.85), $t(48) = 0.27$, $p = 0.79$. There was no significance

30

difference in the global switch cost between the LP bilinguals (0.33) and HP bilinguals (0.31), t (48) = -0.24, p = 0.82. We also did regression analysis on local switch cost and global switch cost with L2 proficiency cumulative score as an independent variable. The results of the regression analysis indicated that the L2 cumulative score is not a significant predictor of either local switch cost (r = 0.09, F (1, 48) = 0.38, p = 0.54) or global switch cost (r = 0.03, F (1, 48) = 0.05, p = 0.83).

Bayesian t-tests on local and global switch costs both provided moderate evidence for the null hypothesis (BF_{01} = 1.27 and BF_{01} = 2.96 respectively). For the estimation method, ROPE was set at (-0.1, 0.1) around 0 based on studies that have compared relative switch costs between bilinguals and monolinguals on colour-shape switching task (Prior & Gollan, 2011). 0.1 is still a conservative estimate because it indicates a 10% increase in switch cost between the two groups. The posterior distribution showed that 25% of the HDI was within ROPE for local switch costs and 80% of the HDI was within ROPE for global switch costs (Figure 3E and 3F). This suggests that there is some evidence to accept the null hypothesis for global switch costs while there is very little evidence to support the null difference on local switch costs.

Shape Task. There was no significant difference in local switch cost between LP (0.95) and HP bilinguals (0.89), t (48) = - 0.45, p = 0.66. LP (0.44) and HP bilinguals (0.48) did not differ on global switch cost as well, t (48) = 0.40, p = 0.69. Linear regression analysis was used to verify if the L2 cumulative score significantly predicted switch costs. The results of the regression analysis indicated that the L2 cumulative score is not a significant predictor of local switch cost (r = 0.02, F (1, 48) = 0.02, p = 0.88.), and global switch cost (r = 0.04, F (1, 48) = 0.06, p = 0.8).

31

Bayesian t-tests on local and global switch costs also provided moderate to strong evidence that the null hypothesis was more likely ($BF_{01} = 2.48$ and $BF_{01} = 4.6$ respectively). With a similar ROPE value as used for the shape dominant blocks, the estimation method showed that 53% of the HDI and 68% of the HDI were within the ROPE for local and global switch costs, respectively (Figure 3G and 3H).

We failed to find any significant effects of L2 proficiency on cognitive control abilities through multiple analyses. To test whether the quantum of engagement with L2 (and not just proficiency in L2) modulates performance on executive control tasks, we performed correlations between percentage of exposure to L2 (self-reported) and the executive control measures (the results are reported in Appendices B - E). We observed that higher exposure to L2 was linked to higher alerting and orienting effect ($r = 0.35$, $p < 0.05$ and $r = 0.4$, $p < 0.05$ respectively) as well as to lower switch costs in the DCCS task (shape, $r = -0.26$, $p < 0.1$). To test this relationship further, we divided the participants into two groups based on exposure to L2 (HE: high exposure, LE: low exposure). We performed an analysis of covariance (ANCOVA) on each executive control measure with exposure as the independent variable and L2 proficiency as the covariate (the group means, and F values are reported in Appendix F). After controlling for L2 proficiency, high exposure group showed a greater orienting effect, $F (1, 44) = 4$, $p = 0.05$, $\eta^2 = 0.08$ and lesser Stroop effect, $F (1, 47) = 6.14$, $p = 0.02$, $\eta^2 = 0.11$ compared to low exposure group. The percentage of exposure did not have a significant effect on any other measure.

Discussion

32

We examined whether L2 proficiency improves cognitive functioning in older

bilinguals. We administered a set of tasks measuring various aspects of executive control on

older, healthy, Telugu-English bilinguals in the age range of 50 – 65 years. The ANT,

numerical Stroop task, stop-signal task, and DCCS task were administered to all the

participants along with a battery of other tasks to measure L2 proficiency, memory capacity

and IQ. Bilinguals were divided into HP and LP based on a composite score derived from

various objective and subjective measures of L2 proficiency. The two groups' performance

was not found to differ on any of the tasks. The null effects were further confirmed through

Bayesian analysis. Based on our study, we conclude that L2 proficiency does not lead to

cognitive advantages in well educated, healthy, ageing Telugu-English bilinguals. Although

our group has earlier shown that young HP bilinguals (e.g., Singh & Mishra, 2012, 2013,

2015, 2016) are better than LP bilinguals on executive control tasks, we did not observe this

advantage in the older population. These results are in line with a growing number of

studies/reviews that have reported null effects of bilingualism on cognitive functioning (e.g.,

Lehtonen et al., 2018; Paap & Greenberg, 2013; Saint-Aubin et al., 2018). In a recent meta-

analysis of studies done on the protective role of bilingualism in cognitive decline, Mukudam,

Sommerlad and Livingston (2017) failed to find consistent evidence for a bilingual advantage

in cognitive decline. They argued that retrospective studies on patient records are susceptible

to several confounding factors making the findings difficult to interpret. Our study takes care

of these confounding factors since we experimentally tested a homogeneous group of

bilinguals with similar levels of education and socioeconomic background, differing only in

L2 proficiency.

33

What could be the reason for this lack of group difference? It is possible that language proficiency can only boost executive control if the bilinguals are practising bilingualism every day. Unlike university students who function within a bilingual setup most of their time, the older bilinguals spend most of the time at home. Self-report data from our participants show that the percentage of exposure to L2 did not differ ($p = 0.28$) between HP (26.3 %) and LP (20.1 %) participants. The current switching frequency also did not differ ($p = 0.61$) between the two groups. Additionally, the percentage of exposure to L2 was significantly lesser ($p < 0.001$) than L1 for both the groups. These measures suggest that although there was a difference in the level of proficiency between the two groups, the exposure to L2 did not differ significantly between the two groups. As a result, such bilinguals may not often use the skills like shifting, switching, inhibition or monitoring. And without the constant practice of these skills through bilingualism, executive control may not be modulated.

In a recent study, Borsa et al. (2018) found evidence for such a link between language use and executive control. On an ANT task, they observed that exposure to L2 was a significant predictor of interference effects. This clearly indicates that the quantum of engagement with L2 influences general executive control. Greater engagement of L2 provides greater opportunity for exercising language control resulting in executive control benefits. To test this hypothesis with our current dataset, we performed Analysis of Co-variance (ANCOVA) with the percentage of exposure to L2 as the independent variable and L2 proficiency as the covariate. This was done to examine whether exposure to L2 modulates cognitive functioning, independent of differences in proficiency. If exposure to L2 (and not L2 proficiency) modulates executive functioning in bilinguals, then we should observe better

executive control performance in the high exposure group. However, this was not observed in any of the measures except the ANT orienting effect and the Stroop effect. The reduction in Stroop effect is not straightforward to interpret as there was no significant evidence for the existence of the Stroop effect itself. Additionally, the increased orienting effect was only marginally significant. Nevertheless, these results underscore the importance of appropriately quantifying the context of a bilingual and considering it as an independent variable in studies examining cognitive advantage in bilinguals. Further, it is to be noted that we recorded the subjective report of participants' *exposure* to L2, but not their *usage* of L2. Although there might be correlations between the two, the assumption is that active usage of the two languages (not just passive exposure to them) requires some of the core cognitive mechanisms such as inhibition, monitoring etc. which then translates into better performance on tasks measuring these components (eg., Yang, Hartanto, & Yang, 2016).

Another possible reason for the lack of effect of L2 proficiency could be that there is a ceiling effect with respect to bilingual advantage and it is possible that both HP and LP groups might have shown an advantage compared to a monolingual group. For instance, Luo et al. (2010) found that high and low vocabulary bilinguals differed from monolinguals in executive function measures in a letter fluency task, but no difference was found within the two types of bilinguals. This shows that executive control (indexed by performance on the letter fluency task) could depend on the existence of bilingualism but not on the levels of it. We are not in a position to evaluate this since our study did not have a monolingual control group for the reasons mentioned in the introduction section. But it is to be noted that the two groups of bilinguals in Luo et al. (2010) differed in vocabulary size but not in proficiency. A

35

lot of research has shown that bilingual language control may be linked to language proficiency. Therefore, in the absence of a monolingual control group, one can expect a difference between groups differentiated by L2 proficiency. In line with this, Luk and Bialystok (2013) note that, bilingualism should not be treated as a categorical variable (in terms of presence vs. absence of bilingualism) and that multiple factors such as language use and proficiency must be taken into consideration while examining bilingualism.

Our data are in contrast to several earlier findings of bilingual advantage in the elderly population (e.g., Bialystok et al., 2006, 2007, 2008). It is to be noted that a vast majority of studies reporting differences in executive control in elderly bilinguals have compared bilinguals with monolinguals (eg., Alladi et al., 2013, 2015; Bialystok et al., 2006, 2007, 2008; Houtzager et al., 2017). This could be problematic because of the difficulty in matching two different groups on a host of other confounding factors. This is particularly true in the Indian context, in a large metropolitan city like Hyderabad, where monolinguals are mostly illiterates/low educated with no additional knowledge of any language. Mishra and colleagues have done many studies on Indian illiterates who were largely monolinguals knowing only their native language (Huettig, Singh, & Mishra, 2011; Mishra, Singh, Pandey, & Huettig, 2012; Olivers, Huettig, Singh, & Mishra, 2014). This could bring serious confounds into the study since the level of education has been shown to play a significant role in modulating bilingual advantage (Gollan et al., 2011). Thus, our data raises questions on studies that have argued for a cognitive reserve in bilinguals based on data from hospital records from a population similar to the one we have used in this study (Alladi et al., 2013, 2015). For instance, Alladi et al. (2013) find the literacy level to be higher in bilinguals compared to

36

monolinguals. They "correct" for this effect statistically through a univariate GLM analysis. However, such statistical correction of confounding factors has been criticised (Miller & Chapman, 2001; Paap, Johnson, & Sawi, 2014). It is also difficult to determine the extent to which the patients used both the languages when they were healthy based on hospital records or data from caretakers. As a result, it is hard to know whether bilingualism was indeed responsible for the later onset of such neurodegenerative diseases.

It is possible to question whether our data collection methods were sensitive enough to observe any subtle differences, considering the prevalence of null effects in our study. To examine this, we performed correlations between some demographic variables (such as age, gender, IQ), L2 measures (percentage of exposure to L2, L2 proficiency score etc.) and task performance (correlation tables are given in Appendices B - E). We observed that the percentage of exposure to L2 significantly predicted an increase in the scores on IQ and MMSE. Higher scores on MMSE also led to lower switch costs and lower Stroop effects. While there was no consistent pattern to these correlations, it nevertheless provides some evidence that our methods were sensitive enough to detect group differences, if any. Importantly, age did not correlate significantly with any of the executive control measures. This is surprising since cognitive capacities generally decline with age (Craik & Salthouse, 2011; Douaud et al., 2014). For instance, in a large sample of 484 participants, Douaud et al. (2014) found that age was a significant predictor of cognitive decline. In an analysis of the brain structures of healthy participants ranging from 8 years to 85 years, the authors observed that with old age, brain areas associated with disorders such as Alzheimer's degenerated faster compared to rest of the brain areas. Interestingly, Borsa et al. (2018) also found that age

37

correlated significantly with conflict effect (on ANT), but only in monolinguals and not in bilinguals. This was taken as evidence that bilingualism acted as a reserve against the natural age-related cognitive decline. While monolinguals experienced a decline in their cognitive capacities as they aged, the practice of two languages in bilinguals stopped or slowed down such decline. It is possible that this null correlation in our study between age and any of the executive control measures also suggests a positive influence of bilingualism in healthy older bilinguals. We can't be confident of this explanation since we did not have a monolingual group for comparison. However, this perspective of bilingualism as a factor that slows down cognitive decline in healthy older bilinguals is worthy of more serious enquiry.

We confirmed the lack of group differences on the cognitive control measures through multiple analyses methods. In addition to the traditional frequentist statistics, we also performed two different types of Bayesian analysis – a model comparison using Bayes factors and the Bayesian estimation method - to assess the likelihood of the null effect. Bayesian analysis is important in this context because the traditional null hypothesis testing (NHST) does not have provisions for accepting a null hypothesis. The results from both Bayes factors and estimation method failed to show any strong evidence for bilingual advantage. The Bayes factors for all the tasks indicated that there was moderate to strong evidence for accepting the null hypothesis. The parameter estimation method also some provided some support for the null hypothesis. According to this method, the alternative hypothesis can be accepted only if 95 % of the HDI falls *outside* the ROPE. This was not observed for any of the tasks suggesting that we don't have sufficient evidence in support of the bilingual advantage claim. Importantly, our data is more suggestive of a lack of group difference since the posterior

distribution of the difference of group means for most of the tasks was centered around zero suggesting a high probability of no difference between the groups. The only exception is the local switch cost on the DCCS task shape-dominant block which indicates a possible group difference. But it may not be a very reliable effect since is not supported by either the Bayes factor analysis or the traditional t-tests. Finally, it can be pointed out that the percentage of HDI within the ROPE for each task was always below 95% (the null hypothesis can be accepted when at least 95% of the HDI is *within* the ROPE) suggesting that maybe there wasn't enough evidence to conclusively accept the null hypothesis based on this method. However, it is to be noted that due to the paucity of experimental research examining the role of L2 proficiency on cognitive control in older bilinguals, specifically in India, it is difficult to decide on how big of an effect counts as a "true" effect. So, the ROPE limits set in this study have to be treated as a rough estimation. Secondly, the Bayesian estimation method, as the name suggests, is meant to be useful for estimating the extent of the difference in group means rather than providing strict criteria for binary decisions on whether an effect is significant or not.

One of the limitations of this study is the small sample size (n). Previous scholarship on Indian ageing bilinguals on this question (eg., Alladi et al., 2013) was based on hospital records and not experimental investigations. Because this was the first experimental study on an ageing population in India, there were several limitations in recruiting a large group of participants. It is true that a very large n in any experimental investigation induces confidence in the results - positive or negative (Paap, 2014). However, there have been other studies with low sample sizes that have shown a null effect of bilingualism (Anton et al., 2016; Kousaie &

39

Phillips, 2012; Papageorgiou et al., 2018; Ramos et al., 2017). And these findings have been used in other meta-analysis or reviews for claiming the no-advantage position (De Bruin, Treccani, & Della Sala, 2015; Hilchey & Klein, 2011; Paap, Johnson & Sawi, 2015). Further, to address this limitation as best as possible, we performed Bayesian analysis. The null hypothesis significance testing (NHST) is susceptible to false conclusions, especially with small sample sizes. One of the reasons is the dependence of p values on stopping intentions (the decision to stop collecting data). But, this problem is not as severe with the Bayesian method which is considered to be quite robust as "Bayesian inference does not depend on stopping and testing intentions" (Kruschke & Liddell, 2018). This effectively means that the n at which we stopped collecting data (big or small) should not drastically modify our Bayesian inference. However, it is true that large-scale studies (whenever possible) will be able to throw more light on the links between L2 proficiency and executive control.

Whether knowing more languages leads to higher executive functioning advantage is not clear from available studies (Chertkow, Whitehead, Philips et al. 2010). We are arguing that since older bilinguals may not indulge in many bilingual conversations, they may not find opportunities to switch/inhibit/monitor often. Ours is a performance-based account of neuroplasticity that bilingualism might have bestowed. However, some researchers (e.g., Bak et al., 2014) have argued that quantum of language use may not matter for executive control. Bak et al. (2014) argue that since bilinguals unintentionally activate the two languages and inhibit one for selection, their executive control gets exercised irrespective of language use. It is true that many studies have shown that bilinguals activate two languages in parallel, but we do not know from such studies how the competition between the two languages is resolved.

40

Additionally, language proficiency has been shown to influence such non-selective parallel language activation (Singh & Mishra, 2016). It seems problematic to assume that the overt practice of a skill plays a minor role in inducing plasticity. Unless bilingualism is in use in a societal scenario with actual partners, just the knowledge of two languages and passive activation of them may not be enough to induce neuroplasticity. In line with this argument, many studies have shown that bilinguals who use language more show better performance on executive control tasks (Hartanto & Yang, 2016).

In conclusion, our study indicates that differences in language proficiency may not be sufficient to induce an advantage in cognitive functioning in healthy, elderly bilinguals. It is the first study to examine the influence of L2 proficiency in Indian bilinguals using an experimental paradigm. However, we are not suggesting that there are no cognitive benefits of bilingualism at all. Rather, it is important to consider the socio-linguistic context of a bilingual population before taking either a null effect or an advantage into serious consideration. Bilingualism is a social phenomenon which cannot be studied in isolation and expected to generalise across different contexts. Future studies examining the role of bilingualism should consider including language use and the overall linguistic context of the bilinguals as a potential factor. Comparing two groups of bilinguals with similar levels of proficiency but differing in the amount of L2 use will provide useful evidence to understand this issue better. For example, one may find a difference between sedentary home-bound older bilinguals who just speak one language at home and other active older bilinguals who use two languages at work. In the Indian context, language use is often differentiated between work and home domains. Future studies with such variables may throw new light on the boundary

41

conditions of the phenomenon which is much required to resolve the debate surrounding this issue.

References

Abutalebi, J., Della Rosa, P. A., Ding, G., Weekes, B., Costa, A., & Green, D. W. (2013). Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex*, *49*(3), 905-911.

Abutalebi, J., & Green, D. W. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Bilingualism: Language and cognition*, *19*(4), 689-698.

Alladi, S., Bak, T. H., Duggirala, V., Surampudi, B., Shailaja, M., Shukla, A. K., ... & Kaul, S. (2013). Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurology*, *81*(22), 1938-1944.

Alladi, S., Bak, T. H., Mekala, S., Rajan, A., Chaudhuri, J. R., Mioshi, E., ... & Kaul, S. (2016). Impact of bilingualism on cognitive outcome after stroke. *Stroke*, *47*(1), 258-261.

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*, 398.

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

Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of clinical neuropsychology*, *15*(1), 31-36.

Bak, T. H., Nissan, J. J., Allerhand, M. M., & Deary, I. J. (2014). Does bilingualism influence cognitive aging?. *Annals of neurology*, *75*(6), 959-963.

Bhatia, D., Prasad, S. G., Sake, K., & Mishra, R. K. (2017). Task Irrelevant External Cues Can Influence Language Selection in Voluntary Object Naming: Evidence from Hindi-English Bilinguals. *PloS one*, *12*(1).

Bialystok, E., Craik, F. I., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia*, *45*(2), 459-464.

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

Bialystok, E., & Feng, X. (2009). Language proficiency and executive control in proactive interference: Evidence from monolingual and bilingual children and adults. *Brain and language*, *109*(2), 93-100.

43

Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010). Word mapping and executive

functioning in young monolingual and bilingual children. *Journal of Cognition and*

Development, *11*(4), 485-508.

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: consequences for mind and

brain. *Trends in cognitive sciences*, *16*(4), 240-250.

Bialystok, E., Poarch, G., Luo, L., & Craik, F. I. (2014). Effects of bilingualism and aging on

executive function and working memory. *Psychology and aging*, *29*(3), 696.

Bialystok, E. (2015). Bilingualism and the development of executive function: The role of

attention. *Child Development Perspectives*, *9*(2), 117-121.

Bialystok, E., Abutalebi, J., Bak, T. H., Burke, D. M., & Kroll, J. F. (2016). Aging in two

languages: Implications for public health. *Ageing research reviews*, *27*, 56-60.

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate

experience. *Psychological bulletin*, *143*(3), 233.

Bialystok, E., Anderson, J. A. E., & Grundy, J. G. (2018). Interpreting cognitive decline in the

face of cognitive reserve. Linguistic Approaches to Bilingualism, Online First. Retrieved from

https://benjamins.com/catalog/lab.18040.bia

44

Chertkow, H., Whitehead, V., Phillips, N., Wolfson, C., Atherton, J., & Bergman, H. (2010).

Multilingualism (but not always bilingualism) delays the onset of Alzheimer disease:

evidence from a bilingual community. *Alzheimer Disease & Associated Disorders*, *24*(2),

118-125.

Chung-Fat-Yim, A., Sorge, G. B., & Bialystok, E. (2017). The relationship between

bilingualism and selective attention in young adults: Evidence from an ambiguous figures

task. *The Quarterly Journal of Experimental Psychology*, *70*(3), 366-372.

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij,

W., & Hommel, B. (2008). How does bilingualism improve executive control? A comparison

of active and reactive inhibition mechanisms. *Journal of Experimental Psychology. Learning,*

Memory, and Cognition, *34*(2), 302–12. https://doi.org/10.1037/0278-7393.34.2.302

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict

resolution: Evidence from the ANT task. *Cognition*, *106*(1), 59-86.

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.

Craik, F. I., & Salthouse, T. A. (Eds.). (2011). *The handbook of aging and cognition*.

Psychology Press.

45

De Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An

example of publication bias?. *Psychological science*, *26*(1), 99-107.

Dong, Y., & Xie, Z. (2014). Contributions of second language proficiency and interpreting

experience to cognitive control differences among young adult bilinguals. *Journal of

Cognitive Psychology*, *26*(5), 506-519.

Douaud, G., Groves, A. R., Tamnes, C. K., Westlye, L. T., Duff, E. P., Engvig, A., ... &

Matthews, P. M. (2014). A common brain network links development, aging, and

vulnerability to disease. *Proceedings of the National Academy of Sciences*, *111*(49), 17648-

17653.

Draine, S. C. (2004). Inquisit 2.0. 50401 [Computer software]. *Seattle, WA: Millisecond

Software*

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the

Efficiency and Independence of Attentional Networks. *Journal of Cognitive Neuroscience*,

14(3), 340–7.

Festman, J., & Münte, T. F. (2012). Cognitive control in Russian–German

bilinguals. *Frontiers in Psychology*, *3*, 115.

46

Filippi, R., D'Souza, D., & Bright, P. (2018). A developmental approach to bilingual research: The effects of multi-language experience from early infancy to old age. *International Journal of Bilingualism*, Online First. Retrieved from https://journals.sagepub.com/doi/abs/10.1177/1367006917749061

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). ''Mini-mental state'': A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35, 116–124.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological science, 17*(2), 172-179.

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

Gray, J. R., & Thompson, P. M. (2004). Neurobiology of intelligence: science and ethics. *Nature Reviews Neuroscience, 5*(6), 471-482.

47

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism:*

Language and cognition, *1*(2), 67-81.

Hartanto, A., & Yang, H. (2016). Disparate bilingual experiences modulate task-switching

advantages: A diffusion-model analysis of the effects of interactional context on switch

costs. *Cognition*, *150*, 10-19.

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.

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic

interference tasks? Implications for the plasticity of executive control processes. *Psychonomic*

bulletin & review, *18*(4), 625-658.

Houtzager, N., Lowie, W., Sprenger, S., & De Bot, K. (2017). A bilingual advantage in task

switching? Age-related differences between German monolinguals and Dutch-Frisian

bilinguals. *Bilingualism: Language and Cognition*, *20*(1), 69-79.

Huettig, F., Singh, N., & Mishra, R. K. (2011). Language-mediated visual orienting behavior

in low and high literates. *Frontiers in psychology*, *2*, 285..

48

Incera, S., & McLennan, C. T. (2018). The time course of within and between-language interference in bilinguals. *International Journal of Bilingualism*, *22*(1), 88-99.

Ivanova, I., Murillo, M., Montoya, R. I., & Gollan, T. H. (2016). Does bilingual language control decline in older age?. *Linguistic approaches to bilingualism*, *6*(1), 86-118.

Jeffreys, H. (1961). *Theory of probability*, 3rd edn. Oxford: Oxford University Press

Klein, R. M. (2016). What cognitive processes are likely to be exercised by bilingualism and does this exercise lead to extra-linguistic cognitive benefits?. *Linguistic Approaches to Bilingualism*, *6*(5), 549-564.

Kousaie, S., & Phillips, N. A. (2017). A behavioural and electrophysiological investigation of the effect of bilingualism on aging and cognitive control. *Neuropsychologia*, *94*, 23-35.

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497-514.

Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, *6*(3), 299-312.

Kruschke, J. K., & Liddell, T. M. (2018). Bayesian data analysis for newcomers. *Psychonomic bulletin & review*, *25*(1), 155-177.

49

Kruschke, J. K., Meredith, M., & Meredith, M. M. (2018). Package 'BEST'.

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

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control

on verbal fluency performance in bilinguals. *Cognition*, *114*(1), 29-41.

Ma, F., Chen, P., Guo, T., & Kroll, J. F. (2017). When late second language learners access

the meaning of L2 words: Using ERPs to investigate the role of the L1 translation equivalent.

Journal of Neurolinguistics, *41*, 50–69.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and

Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and

multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940-967.

McMurray, B., Samelson, V. M., Lee, S. H., & Tomblin, J. B. (2010). Individual differences

in online spoken word recognition: Implications for SLI. *Cognitive Psychology*, 60(1), 1e39.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of

abnormal psychology*, *110*(1), 40.

50

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye-movements are modulated by reading ability-Evidence from Indian low and high literates. *Journal of Eye Movement Research*, *5*(1).

Mishra, R. K., Hilchey, M. D., Singh, N., & Klein, R. M. (2012). On the time course of exogenous cueing effects in bilinguals: higher proficiency in a second language is associated with more rapid endogenous disengagement. *The Quarterly journal of experimental psychology*, *65*(8), 1502-1510.

Mishra, R. K., & Singh, N. (2016). The influence of second language proficiency on bilingual parallel language activation in Hindi–English bilinguals. *Journal of Cognitive Psychology*, *28*(4), 396-411.

Mukadam, N., Sommerlad, A., & Livingston, G. (2017). The relationship of bilingualism compared to monolingualism to the risk of cognitive decline or dementia: a systematic review and meta-analysis. *Journal of Alzheimer's Disease*, (Preprint), 1-10.

Nag, S. (2011). The akshara languages: what do they tell us about children's literacy learning?. *Language-cognition: state of the art*, 291-310.

Olivers, C. N. L., Huettig, F., Singh, J. P., & Mishra, R. K. (2014). The influence of literacy on visual search. *Visual Cognition*, *22*(1), 74-101.

51

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.

Paap, K. R., Johnson, H. A., & Sawi, O. (2014). Are bilingual advantages dependent upon specific tasks or specific bilingual experiences?. *Journal of Cognitive Psychology*, *26*(6), 615-639.

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265-278.

Papageorgiou, A., Bright, P., Periche Tomas, E., & Filippi, R. (2018). Author accepted manuscript: Evidence against a cognitive advantage in the older bilingual population. *Quarterly Journal of Experimental Psychology*, 1747021818796475.

Pelham, S. D., & Abrams, L. (2014). Cognitive advantages and disadvantages in early and late bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 313.

Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and cognition*, *13*(2), 253-262..

52

Qu, L., Low, J. J. W., Zhang, T., Li, H., & Zelazo, P. D. (2016). Bilingual advantage in executive control when task demands are considered. *Bilingualism: Language and Cognition*, *19*(2), 277-293.

Ramos, S., García, Y. F., Antón, E., Casaponsa, A., & Duñabeitia, J. A. (2017). Does learning a language in the elderly enhance switching ability?. *Journal of Neurolinguistics*, *43*, 39-48.

Raven, J. C. (1958). Standard Progressive Matrices. *London: Lewis*

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic bulletin & review*, *16*(2), 225-237

Roychoudhuri, K. S., Prasad, S. G., & Mishra, R. K. (2016). Iconic Native Culture Cues Inhibit Second Language Production in a Non-immigrant Population: Evidence from Bengali-English Bilinguals. *Frontiers in Psychology*, *7*(October), 1–8.

Saint-Aubin, J., Hilchey, M. D., Mishra, R., Singh, N., Savoie, D., Guitard, D., & Klein, R. M. (2018). Does the relation between the control of attention and second language proficiency generalize from India to Canada?. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *72*(3), 208.

53

Singh, N., & Mishra, R. K. (2012). Does language proficiency modulate oculomotor control? Evidence from Hindi–English bilinguals. *Bilingualism: Language and Cognition*, *15*(4), 771-781.

Singh, N., & Mishra, R. K. (2013). Second language proficiency modulates conflict-monitoring in an oculomotor Stroop task: evidence from Hindi-English bilinguals. *Frontiers in psychology*, *4*, 322.

Singh, N., & Mishra, R. K. (2015). The modulatory role of second language proficiency on performance monitoring: evidence from a saccadic countermanding task in high and low proficient bilinguals. *Frontiers in psychology*, *5*, 1481.

Singh, J. P., & Mishra, R. K. (2016). Effect of bilingualism on anticipatory oculomotor control. *International Journal of Bilingualism*, *20*(5), 550-562.

Struys, E., Duyck, W., & Woumans, E. A. (2018). The role of cognitive development and strategic task tendencies in the bilingual advantage controversy. *Frontiers in psychology*, *9*, 1790.

Tao, L., Marzecová, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Frontiers in psychology*, *2*, 123.

54

Tse, C. S., & Altarriba, J. (2012). The effects of first-and second-language proficiency on

conflict resolution and goal maintenance in bilinguals: Evidence from reaction time

distributional analyses in a Stroop task. *Bilingualism: Language and Cognition*, *15*(3), 663-

676.

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, *18*(1),

3-24.

Von Bastian, C. C., Souza, A. S., & Gade, M. (2016). No evidence for bilingual cognitive

advantages: A test of four hypotheses. *Journal of Experimental Psychology: General*, *145*(2),

246.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J.

(2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t

tests. *Perspectives on Psychological Science*, *6*(3), 291-298.

Wilson, C. M., Christensen, B. K., King, J. P., Li, Q., & Zelazo, P. D. (2008). Decomposing

perseverative errors among undergraduates scoring high on the Schizotypal Personality

Questionnaire. *Schizophrenia Research*, *106*(1), 3–12.

Woods et al. (2011). Improving digit span assessment of short-term verbal memory. *Journal

of Clinical and Experimental Neuropsychology*, 33, 101-111

55

Woumans, E., Santens, P., Sieben, A., Versijpt, J. A. N., Stevens, M., & Duyck, W. (2015). Bilingualism delays clinical manifestation of Alzheimer's disease. *Bilingualism: Language and Cognition*, *18*(3), 568-574.

Yang, H., Hartanto, A., & Yang, S. (2016). The importance of bilingual experience in assessing bilingual advantages in executive functions. *Cortex*, *75*, 237.

Yang, H., Hartanto, A., & Yang, S. (2017). Bilingualism confers advantages in task switching: Evidence from the dimensional change card sort task. *Bilingualism: Language and Cognition*, 1-19.

Appendix A. Descriptive statistics of all the tasks.

Task	High proficient bilinguals		Low proficient bilinguals	
	Mean	**SD**	**Mean**	**SD**
Numerical Stroop				
Congruent	627. 24	(134.14)	613.02	(139.1)
Incongruent	632.06	(114.74)	616.86	(129.82)
Stroop Effect	4.83	(46.21)	3.83	(53.16)
Attentional Network Task (ANT)				
Executive Network				
Congruent	754.26	(104.01)	754.81	(80.76)
Neutral	752.3	(126.45)	759	(86)
Incongruent	877.99	(125.95)	883.90	(97.00)

56

Conflict Effect	123.74	(74.63)	129.09	(55.53)
Alerting Network				
No Cue	806.26	(125.10)	804.35	(86.92)
Double Cue	796.13	(133.96)	789.62	(83.20)
Alerting Effect	10.13	(36.37)	14.73	(36.91)
Orienting Network				
Central Cue	787.36	(121.12)	795.34	(82.63)
Spatial Cue	779.69	(103.20)	797.73	(97.88)
Orienting Effect	7.67	(32.30)	-2.39	(41.11)
Stop-signal task				
Stop Signal Reaction Time(SSRT)	467.63	(142.20)	492.54	(213.31)
Dimensional Change Card Sort (DCCS) task				
Colour Task				
Local Switch Cost	0.85	(0.59)	1.02	(0.39)
Global Switch Cost	0.31	(0.25)	0.33	(0.23)
Shape Task				
Local Switch Cost	0.89	(0.41)	0.95	(0.52)
Global Switch Cost	0.48	(0 .28)	0.44	(0.32)

The correlation tables for all the tasks are listed below. Only the values in the last column change substantially for each task.

Appendix B. Correlational analysis for ANT

	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. Gender	**0.37**[*]	0.23	0.2	-0.13	0.1	0.06	**0.3**[*]	-0.1	0.2	-0.03
2. Age		0.005	0.05	0.05	-0.07	-0.06	**0.8**[***]	-0.17	-0.05	-0.11
3. MMSE			**0.36**[*]	0.13	0.19	**0.34**[*]	0.1	0.06	0.1	0.04

57

	5.	6.	7.	8.	9.	10.	11.
4. IQ	0.19	0.28	**0.34***	0.09	0.11	-0.03	-0.12
5. SES		0.19	0.1	-0.09	-0.15	0.08	-0.13
6. L2 proficiency[a]			0.08	0.10	-0.07	**0.3***	-0.07
7. L2 exposure[b]				-0.05	**0.35***	**0.4****	-0.24
8. YOB[c]					**-0.29***	-0.03	0.04
9. Alerting effect						0.18	-0.07
10. Orienting effect							**-0.4****
11. Conflict effect							

Note: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, #$p < 0.1$

[a]L2 proficiency refers to the cumulative L2 proficiency score used to divide the participants into two groups

[b]L2 exposure refers to the percentage of exposure to L2 extracted from the questionnaire as an answer to the question: "Please list what percentage of the time you are currently, on average, exposed to each language. (Your percentages should add up to 100%)")

[c]YOB: Years of Bilingualism = Current age – AoA of L2

Appendix C. Correlational analysis for Numerical Stroop task

	2.	3.	4.	5.	6.	7.	8.	9.
1. Gender	**0.34***	0.22	0.2	-0.14	0.12	0.06	**0.3***	0.07
2. Age		0.002	0.05	0.06	-0.04	-0.07	**0.8*****	0.16
3. MMSE			**0.37***	0.11	0.19	**0.35***	0.09	**-0.33***
4. IQ				0.13	**0.32***	**0.36***	0.11	-0.08
5. SES					0.12	0.06	-0.11	-0.03
6. L2 proficiency						0.11	0.14	0.13
7. L2 exposure							-0.04	-0.23
8. YOB								0.13
9. Stroop effect								

Note: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, #$p < 0.1$

Appendix D. Correlational analysis for DCCS task

	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1. Gender	**0.34***	0.22	0.2	-0.14	0.12	0.06	**0.3***	-0.08	-0.15	-0.15	-0.004
2. Age		0.002	0.05	0.06	-0.04	-0.07	**0.8*****	0.02	0.04	-0.09	-0.02
3. MMSE			**0.37***	0.11	0.19	**0.35***	0.09	-0.07	**-0.24#**	**-0.4****	-

58

	5.	6.	7.	8.	9.	10.	11.	12.
3.								0.39**
4. IQ	0.13	0.32*	0.36*	0.11	-0.11	-0.18	-0.36*	-0.38**
5. SES		0.12	0.06	-0.11	-0.02	-0.04	-0.04	-0.21
6. L2 proficiency			0.11	0.14	0.03	-0.09	0.04	-0.02
7. L2 exposure				-0.04	-0.18	-0.26#	0.2	0.09
8. YOB					-0.05	-0.12	-0.12	0.02
9. Global SC (shape)[a]						0.66***	0.32*	0.23
10. Local SC (shape)[a]							0.34*	0.29*
11. Global SC (colour)[b]								0.82***
12. Local SC (colour)[b]								

Note: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, #$p < 0.1$

[a]Shape-dominant block, [b]Colour-dominant block

Appendix E. Correlational analysis for Stop-signal task

	2.	3.	4.	5.	6.	7.	8.	9.
1. Gender	0.34*	0.22	0.2	-0.14	0.12	0.06	0.3*	0.02
2. Age		0.002	0.05	0.06	-0.04	-0.07	0.8***	-0.22
3. MMSE			0.37*	0.11	0.19	0.35*	0.09	0.01
4. IQ				0.13	0.32*	0.36*	0.11	-0.09
5. SES					0.12	0.06	-0.11	0.03
6. L2 proficiency						0.11	0.14	0.02
7. L2 exposure							-0.04	-0.07
8. YOB								-0.16
9. SSRT[a]								

Note: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, #$p < 0.1$

[a]Stop-signal reaction time

Appendix F. ANCOVA results

Results from analysis of co-variance with cumulative L2 proficiency as the covariate.

59

	F	p	η^2	Group means (in ms)[a]
ANT alerting effect				
Exposure to L2	1.13	0.29	0.02	HE: M = 16.84, SD = 43.3 LE: M = 8.3, SD = 28.4
L2 proficiency	0.71	0.4	0.02	
ANT orienting effect				
Exposure to L2	4	0.05[#]	0.08	HE: M = 16.4, SD = 32.4 LE: M = -10.7, SD = 36.9
L2 proficiency	1.5	0.2	0.03	
ANT conflict effect				
Exposure to L2	0.81	0.37	0.02	HE: M = 116.5, SD = 41.6 LE: M = 136.1, SD = 81.1
L2 proficiency	0.01	0.9	< 0.001	
Stroop effect				
Exposure to L2	6.1	0.02[*]	0.1	HE: M = -7.9, SD = 43.8 LE: M = 16.5, SD = 52.3
L2 proficiency	3.72	0.06[#]	0.06	
Global SC (shape)				
Exposure to L2	0.54	0.47	0.01	HE: M = 0.5, SD = 0.27 LE: M = 0.43, SD = 0.33
L2 proficiency	0.006	0.94	<0.001	
Local SC (shape)				
Exposure to L2	0.00004	0.9	<0.001	HE: M = 0.92, SD = 0.4 LE: M = 0.93, SD = 0.5
L2 proficiency	0.02	0.89	<0.001	
Global SC (colour)				
Exposure to L2	1.15	0.29	0.02	HE: M = 0.29, SD = 0.28 LE: M = 0.35, SD = 0.2
L2 proficiency	0.42	0.52	0.009	
Local SC (colour)				
Exposure to L2	1.2	0.27	0.02	HE: M = 0.85, SD = 0.57 LE: M = 1.02, SD = 0.4
L2 proficiency	0.01	0.92	<0.001	
SSRT				

60

Exposure to L2	0.26	0.6	0.006	HE: M = 469.2, SD = 136.3 LE: M = 491, SD = 217.3
L2 proficiency	0.1	0.76	0.002	

[a]Group means obtained from ANCOVA after controlling for L2 proficiency scores.
HE: high exposure, LE: low exposure

Table 1: Demographic data of participants

	High-proficiency bilinguals (n=25)		Low-proficiency bilinguals (n=25)	
	Mean	**SD**	**Mean**	**SD**
Age (in years)	57.44	4.32	56.9	3.98
Age of Acquisition (in years) of:				
L1	2.20	1.41	1.62	0.90
L2	9.04	2.84	9.88	2.98

61

Number of years of formal education in:				
L1	8.32	4.63	10.1	3.89
L2	3.28	4.31	1.92	3.89
Percentage exposure to[a]:				
L1	66.16	18.48	75.52[#]	18.68
L2	26.32	16.09	20.1	18.07
Self-reported proficiency score of L1 in:				
Reading[b]	9.44	0.768	9.40	1.04
Speaking[b]	9.12	0.917	9.10	1.10
Understanding spoken language[b]	9.48	0.714	9.16	1.21
Self-reported proficiency score of L2 in:				
Reading[b]	8.92	0.909	7.68[*]	1.49
Speaking[b]	8.04	1.37	6.16[*]	1.68
Understanding spoken language[b]	8.68	1.11	7.16[*]	1.68
Naming latency (in ms) in:				
L1	1142.45	237.45	1202.17	229.04
L2	1156.82	201.95	1172.39	220.56
Semantic fluency score in:				

62

L1	13.48	3.94	11.4*	2.82
L2	11.38	2.91	7.58*	2.72
L2 vocabulary test score	54.52	15.48	33.44*	17.46
Composite score[c]:				
L1	0.000713	0.00832	-0.000712	0.0071
L2	0.0471	0.00538	0.0271*	0.01063
Current switching frequency[d]	4.48	1.58	4.20	1.98
Preferred switching frequency with a balanced bilingual[e]	3.96	1.95	3.04	1.84
MMSE[f]	28.92	1.52534	28.72	1.27541
Non-verbal IQ	36.0417	7.04322	30.4167[#]	7.45955
Working memory				
Forward digit span task	6.43 (n=23)	0.945	5.91 (n=22)	1.41
Backward digit span task	5.87 (n=23)	1.058	5.18 (n=22)	1.37
SES	2.6	0.5	2.44	0.507

*$p < 0.001$, #$p < 0.05$

[a]Question: "Please list what percentage of the time you are currently, on average, exposed to each language. (Your percentages should add up to 100%)"

[b] Self-rating of language proficiency by each participant on a scale of 1 to 10 (10 being the highest).

63

ᶜ Composite scores for L1 and L2 consisted of scores from object naming test, semantic

fluency test, self-ratings on proficiency and vocabulary test (only for L2 score).

ᵈ Question: "How often are you in a situation when you switch between the languages Telugu

and English?" – 1 (never) to 7 (very often).

ᵉ Question: "When choosing a language to speak with a person who is equally fluent in all

your languages, how often would you switch between languages?" – 1 (never) to 7 (very

often).

† As shorter naming latencies reflects higher language proficiency, naming latencies were

multiplied by -1 so that a larger resulting value corresponded with higher composite L2 score

(Ma, Chen, Guo, & Kroll, 2017).

Figure captions

Figure 1. A representative sample of studies conducted to examine bilingual advantage in the

elderly population since 2004. Blue circles indicate studies on healthy older adults measuring

performance on cognitive tasks and red circles indicate studies that have examined the role of

bilingualism in the onset of a neurological disease based on patient records. The double

circles indicate studies that have treated bilingualism as a continuous variable and tested

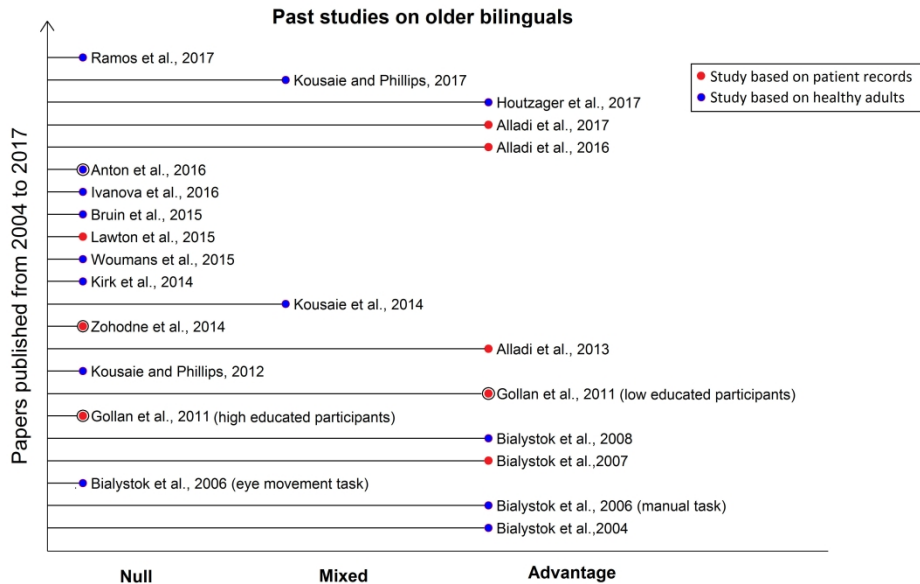bilinguals based on their second language proficiency.

Figure 2. Comparison of ANT (**A**), Alerting effect (**B**), Orienting effect (**C**) conflict effect

across L2 proficiency groups. (**D**) Group comparison of Stroop effect from numerical Stroop

task. (**E**) Local switch cost and Global switch cost comparison of DCCS task for Colour and

64

Shape dominant blocks across L2 proficiency groups. **(F)** Comparison of Stop-signal

Reaction Time(SSRT) across L2 proficiency groups.

Note: Error bars indicate +1 SE

Figure 3. The posterior distributions of the difference in group means for all the tasks. The y

axis is the probability density and the x axis is the parameter value (difference in group means

in ms). The ROPE limits are marked by red vertical lines. The criteria for setting the ROPE

limits for each task are described in the text.

Note: μ_1: high proficient group, μ_2: low proficient group

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

Past studies on older bilinguals

A representative sample of studies conducted to examine bilingual advantage in the elderly population since 2004. Blue circles indicate studies on healthy older adults measuring performance on cognitive tasks and red circles indicate studies that have examined the role of bilingualism in the onset of a neurological disease based on patient records. The double circles indicate studies that have treated bilingualism as a continuous variable and tested bilinguals based on their second language proficiency.

1272x781mm (96 x 96 DPI)

26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
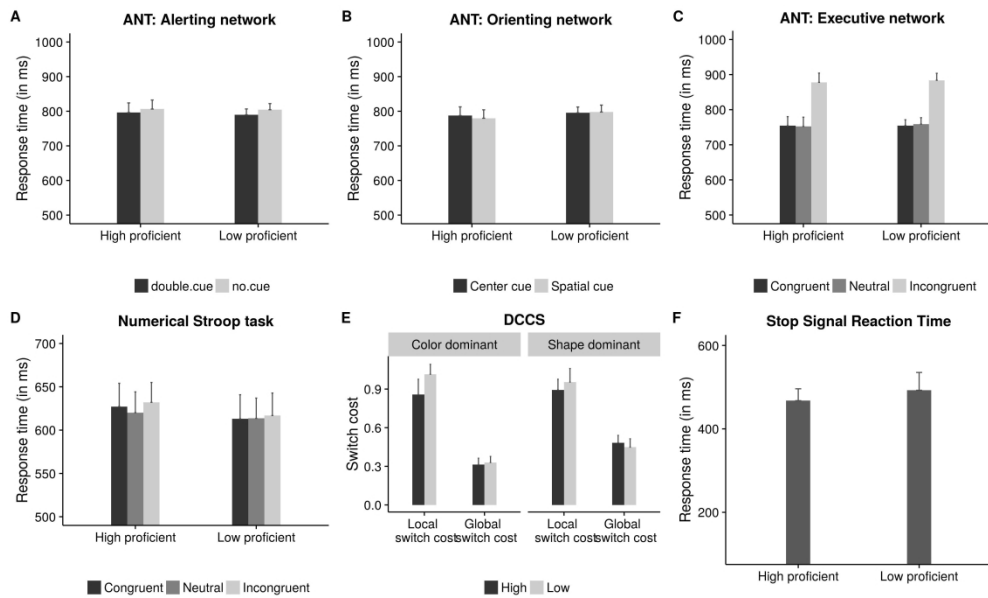49
50
51
52
53
54
55
56
57
58
59
60

Figure 2. Comparison of ANT (A), Alerting effect (B), Orienting effect (C) conflict effect across L2 proficiency groups. (D) Group comparison of stroop effect from numerical stroop task. (E) Local switch cost and Global switch cost comparison of DCCS for Colour and Shape dominant tasks across L2 proficiency groups. (F) Comparison of Stop-signal Reaction Time(SSRT) across L2 proficiency groups.
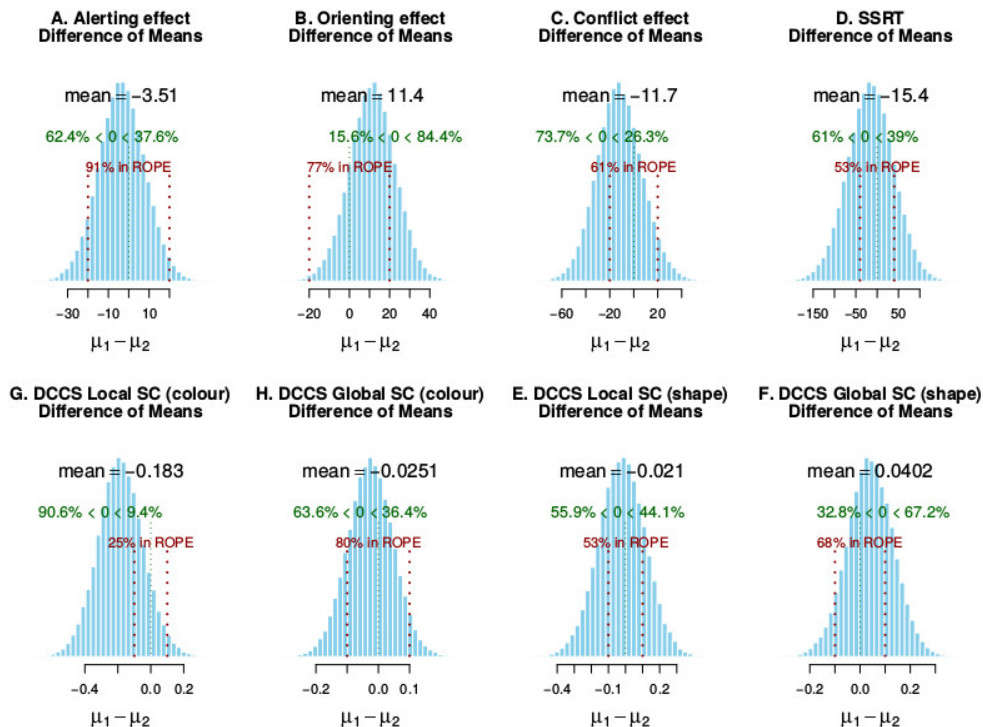Note: Error bars indicate +1 SE

335x203mm (300 x 300 DPI)

Figure 3. The posterior distributions of the difference in group means for all the tasks. The y axis is the probability density and the x axis is the parameter value (difference in group means in ms). The ROPE limits are marked by red vertical lines. The criteria for setting the ROPE limits for each task are described in the text.
Note: μ1: high proficient group, μ2: low proficient group

203x152mm (100 x 100 DPI)