Interference suppression in bilingualism

Citation for published version:
Xia, L, Bak, TH, Sorace, A & Vega-Mendoza, M 2022, 'Interference suppression in bilingualism: Stimulus-stimulus vs. stimulus-response conflict', *Bilingualism*, vol. 25, no. 2, pp. 256-268.
https://doi.org/10.1017/S1366728921000304

Digital Object Identifier (DOI):
10.1017/S1366728921000304

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Bilingualism

Running Head: Bilingualism and Inhibition: Stimulus-Stimulus vs. Stimulus-Response inhibition

Interference suppression in bilingualism: Stimulus-Stimulus vs. Stimulus-Response conflict

Lihua Xia[1], Thomas H. Bak[1], Antonella Sorace[1], Mariana Vega-Mendoza[1,2]

[1] School of Philosophy, Psychology & Language Sciences, The University of Edinburgh

[2] Department of Health, Education and Technology. Engineering Psychology, Luleå University of Technology, Luleå, Sweden

Address for correspondence

Lihua Xia, 7 George Square, Edinburgh, EH8 9JZ

S1464483@ed.ac.uk

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interests with respect to their authorship and/or the publication of this article.

Acknowledgement:

The authors thank the reviewers for their insightful comments on a previous version of this manuscript. This work received support from the School of Philosophy, Psychology & Language Sciences (PPLS) Research Support Grants, the University of Edinburgh.

Abstract

Studies examining the potential effects of bilingualism on interference suppression show inconsistent results. Our study approaches this topic by distinguishing two potential subcomponents within interference suppression (i.e., Stimulus-Stimulus and Stimulus-Response conflict). We investigated the two subcomponents through their operationalisation in different tasks and examined the role of language proficiency in modulating it. A sample of 111 young adult participants performed four non-linguistic cognitive tasks measuring both visual and auditory domains of cognitive control. Bilinguals outperformed monolinguals in tasks involving Stimulus-Stimulus conflict, but showed comparable performance in tasks involving Stimulus-Response conflict. Specific effects of language proficiency on cognitive control were observed: group differences in auditory inhibition and visual orienting were only observed between high-proficient bilinguals and monolinguals. Taken together, types of conflicts involved in interference tasks and language proficiency could differentially affect performance in monolinguals and bilinguals.

Keywords: Bilingualism, interference suppression, Stimulus-Stimulus inhibition, Stimulus-Response inhibition, language proficiency

1. Introduction

Studies across the lifespan have shown that bilingualism positively impacts performance on tasks tapping into multiple aspects of executive functions, particularly inhibitory control (children: Sorge, Toplak, & Bialystok, 2017; young adults: Costa, Hernández, & Sebastián-Gallés, 2008; and older adults: Bialystok, Craik, & Luk, 2008). The Inhibitory Control (IC) model (Green, 1998) and the Bilingual Interactive Activation Model (BIA+) (Dijkstra & Van Heuven, 2002), further refined in the Adaptive Control Hypothesis (Green & Abutalebi, 2013), suggest that bilinguals' two languages are simultaneously activated (Blumenfeld & Marian, 2013; Chen, Bobb, Hoshino & Marian, 2017; Thierry & Wu, 2007); thus, bilinguals have to select the target language and inhibit the non-target language, thereby posing increased cognitive demands on cognitive control (e.g., inhibitory control) for language processing (Bialystok, Craik, & Luk, 2012; Kroll, Dussias, Bice, & Perrotti, 2015). On the other hand, not all studies produced comparable results (Paap & Greenberg, 2013; Paap, Johnson, & Sawi, 2014), resulting in a debate as to whether bilingual experience benefits in terms of the nature and extent of possible bilingualism effects (Bak, 2016; Bialystok, 2017, 2020; Lehtonen et al., 2018; Kroll & Bialystok, 2013).

This conflicting evidence could be due to a variety of interacting variables, such as the definition of bilingualism, the type of selected tasks, and the populations under study (Cox et al., 2016). In terms of selected tasks, researchers have argued that the effects of bilingualism on executive control are dependent upon specific task conditions. Firstly, the demand levels of the selected tasks can vary. It appears that monolingual-bilingual group differences tend to emerge under sufficiently high cognitive control demands, such as a high task-demanding condition (Bialystok, 2006), a high-monitoring condition in the Flanker task (i.e., 75% congruent vs. 25% incongruent trials; Costa, Hernández, Costa-Faidella, & Sebastián- Gallés, 2009), or high working memory demands (Jiao, Liu, Wang, & Chen, 2019). Secondly, the bilingual effects might depend on the different types of inhibition involved in different tasks; that is, interference suppression (i.e., refers to the capacity to detect and filter out irrelevant information in the environment) and response inhibition (i.e., refers to the capacity to inhibit inappropriate, but prepotent, response tendencies) (Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002). Previous studies have found that bilingualism mainly influences interference suppression, but not response inhibition (Bialystok & Viswanathan, 2009; Esposito, Baker-Ward, & Mueller, 2013; Martin-Rhee & Bialystok, 2008). Luk, Anderson, Craik, Grady, and Bialystok (2010), for example, reported distinct neural correlates for the two types of inhibition in bilinguals. Specifically, they used partial least squares (PLS) to identify monolinguals and bilinguals' brain regions where activity covaried across flanker incongruent and non-go trials, representing interference suppression and response inhibition, respectively. The results showed that monolinguals and bilinguals activated similar regions for response inhibition but not for interference suppression.

However, even within interference suppression, the differences between monolinguals and bilinguals have not been consistently replicated, especially in young adult population. While some studies report better performance by bilinguals (Costa et al. 2009; Garraffa, Obregon, O'Rourke, & Sorace, 2020; Marzecová, Asanowicz, Krivá, & Wodniecka, 2012; 2013; Yow & Li, 2018), a substantial number of studies report no group differences (Kousaie & Phillips, 2012; Paap & Greenberg, 2013; Samuel, Roehr-Brackin, Pak, & Kim, 2018). One possibility is that the tasks typically used in bilingualism research (e.g., the Flanker, Stroop, and Simon tasks) are not sufficiently sensitive to detect group differences in performance for young adults, given that the null results tend to be observed in young adult population, who may be at their peak cognitive level and show ceiling performance (Grundy, 2020; Paap & Greenberg, 2013). An alternative possibility could be that the task performance might vary depending on the subcomponents within interference suppression; that is, Stimulus-Stimulus

(S-S) inhibition vs. Stimulus-Response (S-R) inhibition (Hilchey & Klein, 2011). Blumenfeld and Marian (2014), for instance, conducted two experiments comparing monolinguals to bilinguals on S-S and S-R conflict type of tasks and found that bilinguals showed more efficient S-S performance than S-R performance. They explained that S-S inhibition was more likely to be recruited for bilingual control as a result of parallel language activations (Blumenfeld & Marian, 2013) and competitions during language comprehension and production processes. On the other hand, S-R inhibition might be limited to production context where both languages remain active until a response is given. This could partly explain the lack of consistency across studies using different interference tasks. The current study aimed to investigate the role of conflict types (i.e., S-S and S-R conflict) in modulating interference suppression in relation to the bilingual effects in young adult bilinguals while accounting for the theoretical and methodological aspects of previous research (Bak, 2016; Bialystok, 2020).

1.1. The Stimulus-Stimulus and Stimulus-Response type of interference tasks

Interference tasks, such as the Flanker task/Attention Network Task (ANT), Simon task, and variations of Stroop tasks (e.g., colour-word vs. number version) have been frequently used in bilingualism research, in which bilinguals were reported to outperform monolinguals (Costa et al., 2008; Costa et al., 2009; Bialystok et al., 2004; Bialystok et al., 2008; also see Paap, Anders-Jefferson, Zimiga, Mason, & Mikulinsky, 2020 for counterarguments). Although all the tasks are assumed to measure the same executive control component (i.e., interference suppression), they may require different cognitive processes, as reflected in the different brain activation patterns (Hübner & Töbel, 2019; Liu, Banich, Jacobson, & Tanabe, 2004; Kousaie & Phillips, 2012, Pratte, Rouder, Morey, & Feng, 2010), as well as different sources of conflicts; that is, S-S conflict vs. S-R conflict (Hilchey & Klein, 2011; Paap, Anders-Jefferson, Mikulinsky, Masuda, & Mason, 2019).

Three non-linguistic interference tasks associated with different S-S and S-R conflict resolution, were included in the current study: the ANT (Costa et al., 2008), the Number Stroop task (adapted from Hernández, Costa, Fuentes, Vivas, & Sebastián-Gallés, 2010), and the Arrow Simon task (details of the three tasks were provided in the methodology section). Specifically, in the incongruent trials, the conflict resolutions were varied across the three tasks, reflecting different conflict sources (see Figure 1). While the S-R conflicts stem from competition between the task-relevant stimuli (S_r), the task-irrelevant stimuli (S_i), and response tendencies (R), the S-S conflicts stem from the same stimulus category (i.e., S_r – S_i)(Paap et al., 2020; Tiego, Bellgrove, Pantelis, & Whittle, 2018).

In the ANT, the pointing direction of the central arrow (S_r) is conflicted with that of the other arrows (S_i)(e.g., →→←→→), conforming a S-S conflict. In the Number Stroop task, the correct response (i.e., counting the number of digits; S_r) is conflicted with the numerical value of the digits (e.g., 222; S_i) that triggers the response tendency (R), reflecting a S-R conflict. The Number Stroop task could arguably reflect a S-S conflict, see Blumenfeld and Marian (2014). In the classic Colour-Word Stroop task, there are two stimulus dimensions: the colour of the word's ink (e.g., red) and the meaning of the word (e.g., green). In study conducted by Yow and Li (2015), for example, participants were instructed to respond according to the colour of the word by pressing a designated key on the keyboard (i.e., D, F, G, and H for red, yellow, green, and blue, respectively). There was no overlap between required response and key pressing tendency. Thus, the Colour-Word Stroop task was assumed to reflect a S-S conflict. In contrast, in the Number Stroop task, participants were instructed to count the number of digits by pressing corresponding 1, 2, or 3 keys on the keyboard. Given the natural tendency to react toward the value of digits, there was an overlap between required response and key pressing tendency; thus, reflecting a S-R conflict. Similarly, in the Arrow Simon task, the correct response (i.e., pointing direction of the arrow;

S_r) is conflicted with the location of the arrow (S_i) that triggers the response tendency (R), and thus conforms S-R type of conflict. More specifically, the Arrow Simon task used in the current study was a combination of Spatial and Vertical Stroop task with different response instructions (see more details in the methodology section) . The Spatial Stroop Task, in which participants identified the direction of an arrow (i.e., left and right) while ignoring its location (i.e., left and right) by pressing Left and Right key, could arguably reflect different conflict types: S-S conflict (Blumenfeld & Marian, 2014) and S-R conflict (Hilchey & Klein, 2011; Paap et al., 2020). The Vertical Stroop task, in which participants respond to the direction of an arrow (i.e., up and down) while ignoring its location (i.e., up and down) by pressing Left and Right key, was categorized as measuring the S-S conflict (Paap et al., 2020). In the current study, there was overlap between required response and key pressing tendency in the Arrow Simon task; thus, reflecting a S-R conflict.

1.2. The role of language proficiency in bilingualism

Language proficiency is one of the core aspects of bilingualism (Mishra, 2015). Given that the level of language proficiency may dynamically change with specific learning experiences (e.g., continuous language learning and daily language use), it is not clear when and how improvements in L2 proficiency might contribute to a change in executive control in bilingualism. Previous studies have shown that higher L2 proficiency predicts better executive control in bilinguals of various languages, including attentional monitoring abilities (Singh & Mishra, 2015), reactive inhibition (Khare, Verma, Kar, Srinivasan, & Brysbaert, 2013), and endogenous disengagement of attention in Hindi-English adults (Mishra, Hilchey, Singh, & Klein, 2012), conflict monitoring in Chinese-English adults (Xie & Pisano, 2018), and auditory inhibition and switching in English-Spanish adults (Vega-Mendoza, West, Sorace, & Bak, 2015). Some studies conducted with children also support this association. For instance, higher L2 proficiency was reported to be correlated with better performance in inhibition and shifting abilities in English-Hebrew children (Iluz-Cohen & Armon-Lotem, 2013), response inhibition in Chinese-English children (Chen, Zhou, Uchikoshi, & Bunge, 2014), and in conflict resolution and working memory capacity, but not in goal maintenance or task-set switching in Cantonese-English children (Tse & Altarriba, 2014).

Other studies have shown no effects of language proficiency on executive control in bilinguals (Dong & Xie, 2014; Rosselli, Ardila, Lalwani & Vélez-Uribe, 2016; Verreyt, Woumans, Vandelanotte, Szmalec & Duyck, 2016). For example, Dong and Xie (2014) reported no effects of language proficiency on either the inhibition or mental set shifting, but the language interpreting experience was strongly associated with better performance in mental set shifting. The authors explained that the small proficiency gap between high and low-proficiency level might be responsible for the inconsistent findings with regard to the effects associated with language proficiency. Rosselli et al., (2016) compared three groups of participants (i.e., balanced bilinguals vs. unbalanced bilinguals vs. monolinguals) performing both verbal and nonverbal executive tasks. The results demonstrated that nonverbal intelligence rather than language proficiency or bilingualism was a significant predictor of performance on these tasks. Similarly, Verreyt et al., (2016) compared three groups of bilinguals differing in language proficiency and switching experience. They found that language switching patterns rather than proficiency modulated executive control.

It has been suggested that a strong L2 (i.e., higher proficiency) is assumed to elicit a stronger interference on the L1 as compared to a weak L2 (i.e., lower proficiency), leading to increased cognitive demands (e.g., inhibitory control; Green,1998; Hui, Yuan, Fong, & Wang, 2020; Lehtonen et al., 2018). Given that bilingualism is a continuous and dynamic process, it could be possible that monolingual-bilingual differences could only be observed when bilinguals reach certain level of L2 proficiency (Vega-Mendoza et al., 2015). The current study aimed at investigating the extent to which level of proficiency might play a role in modulating executive control.

1.3. The current study

Previous research has shown group differences between monolinguals and bilinguals on a variety of tasks that have been described as tapping into executive functioning, but the precise mechanism for the cognitive effects associated with bilingualism is unknown and the executive functions are heterogeneous without a clear definition (Chun-Fat-Yim, Sorge, & Bialystok, 2017). The present study aimed to investigate the potential subcomponents (i.e., S-S vs. S-R inhibition) within interference suppression in relation to the bilingual effect while taking into account relevant theoretical and methodological issues, such as the role of language proficiency in modulating these effects.

The S-S and S-R inhibition were indexed by three different non-linguistic interference tasks: while the ANT taps into S-S conflicts, the Stroop and the Simon tasks tap into S-R conflicts. Following Blumenfeld and Marian (2014), we predicted that group differences between monolinguals and bilinguals (e.g., smaller conflict effects by bilinguals) would be observed in the ANT rather than the Stroop and Simon tasks.

The effects of language proficiency on executive control were examined by comparing performance on these abovementioned cognitive tasks of high and low-proficient bilinguals to monolinguals. We predicted that monolingual-bilingual differences would be more likely to be observed between high-proficient bilinguals and monolinguals, while low-proficient bilinguals took an intermediate position (Hui et al., 2020; Vega-Mendoza et al., 2015).

The selection of tasks takes into account the fact that the process of language learning involves multiple fundamental and interactive components, such as visual (i.e., reading and writing) and auditory ones (i.e., speaking and listening). While most of previous studies have been focusing on the visual domain, an increasing number of studies have demonstrated effects associated with bilingual experience in the auditory domain (Bak, Long, Vega-Mendoza & Sorace, 2016; Bak, Vega-Mendoza & Sorace, 2014; Garraffa et al., 2020; Ooi, Goh, Sorace, & Bak, 2018; Long, Vega-Mendoza, Rohde, Sorace & Bak, 2019; Vega-Mendoza et al., 2015). Hence, in addition to the visual domain (i.e., as measured by the three RT-based tasks), we adopted three Elevator subtests of the Test of Everyday Attention (TEA) from previous studies as a measure of executive control in the auditory domain (i.e., sustained attention, inhibitory control, and attentional switching). The combination of multiple tasks can avoid the task impurity problem (Miyake & Friedman, 2012) and limitations of the use of a single indicator to measure a single aspect of cognitive functions (von Bastian et al., 2016). It has been suggested that it would be difficult to establish whether the measured performance is a reflection of task-specific or ability-general effects (Shipstead, Redick & Engle, 2012). Moreover, the use of these tasks can provide a more varied approach in assessing attentional control: the ANT, the Stroop task, and the Simon task are motivated by a theoretical framework of attentional control and commonly used in an experimental setting, whereas the TEA was designed to stimulate daily activities using a humanised approach (Ooi et al., 2018).

2. Method

2.1. Participants

A total number of 111 University of Edinburgh students took part in this experiment. All were English native speakers. The participants were divided into three groups according to their second language (L2) proficiency: monolinguals (i.e., Mono group; n = 37), bilinguals with low proficiency (i.e., Bi-low group; n = 39), and bilinguals with relatively higher proficiency (i.e., Bi-high group; n = 35). As second language learning is a global communicative practice (Kramsch & Thorne, 2002), it is difficult to recruit pure "monolinguals" with no exposure to any other language than their mother tongue. Hence, the criterion for defining monolinguals and bilinguals was consistent with previous studies; as will be explained in more detail below, monolinguals were defined as those who were not functionally fluent in any other language (Luk & Bialystok, 2013; Vega-Mendoza et al., 2015).

Before attending the testing session, all participants were asked to self-report their overall L2 proficiency (the scale is 1-7 marked from "*poor*" to "*excellent*") and the ability to hold a conversation in an L2[1]. The grouping was based on the information collected from the background questionnaire that was adopted from a previous study (Ooi et al., 2018; see details in Appendix S1). Self-reported language proficiency was rated using 4-point scales (i.e., 1-4, marked from "*poor*" to "*excellent*"), participants rated their speaking, understanding, reading and writing skills in every language they had learnt. Moreover, participants were required to report whether they had actively used their L2s. Given the small language scale, we divided participants into three groups based on their L2 proficiency and the experience of active use of L2: monolinguals were those who could not hold a conversation in L2 and had never actively used their L2s, the Bi-low group were those who used to actively use their L2 but could only hold a basic conversation in L2 at the time of testing (e.g., this could be due to their lesser exposure to and usage of their L2s in their later life), the Bi-high group were those who could hold a conversation in L2 and actively use their L2. We also used an alternative splitting approach following Grundy, Pavlenko, and Bialystok (2020): low and high-proficient bilingual groups were created by averaging

[1] The 1-7 scale was used only at participants' recruitment, which was to get a general language profile of our participants before testing. To keep the consistency with a previous study (i.e., Ooi et al., 2018), we did not change the scale in the questionnaire (i.e., 1-4).

bilinguals' L2 proficiency and then doing a median split: participants whose L2 proficiency was below the median were assigned into the Bi-low group and the others were assigned into the Bi-high group. The two approaches to grouping high and low-proficient groups showed the same results.

All bilinguals had learned an Indo-European language (e.g., French, German, or Spanish) as their main L2s. One participant in the Bi-low group had Korean as the main L2. Excluding and including this participant did not change the results, thus their results were included in the full set of data. The information of the bilingual participants' other languages is presented in Table S1. Due to technical issues, some participants' data was missing in some tasks, as follows: in ANT, one in the monolingual group; in the Number Stroop task, three in each of the three groups; and in the Simon task, five in the Bi-low group.

2.2. Background measures

Raven's Advanced Progressive Matrices (APM)

The APM (Raven & Foulds, 1962) was used as a control of nonverbal general intelligence. Consistent with a previous study (Ooi et al., 2018), we adopted Set I (i.e., Item 5 or Item 7) as practice and Set II as the experimental test. The design of the Matrices is that the demand of the level gradually increases with the items. Participants were instructed to complete the matrices item by item in 10 minutes and were told that if they were having difficulty with a specific item, they could guess the answer and continue to the next one. They started with Item 1 and had to answer as many items as they could. The results were scored as the number of correct items for each participant.

Corsi Tapping Task (CTT)

The Corsi Tapping task (Wechsler Memory Scale-III, Wechsler, 1997) was administered as a measure of working memory (forward and backward conditions; Luo, Luk & Bialystok, 2010) and mental rotation (rotated condition; adapted from Keehner & Gathercole, 2007), in order to control fundamental cognitive abilities. This task was presented on a plastic whiteboard (27.5 cm × 21cm) with ten blue cube-shaped numbered blocks (3 cm × 3 cm; from 1 to 10). During the testing, the board was placed between the experimenter and the participant, and the numbers were only visible to the experimenter. In the forward condition, the experimenter tapped a sequence of blocks at a rate of approximately one second per block in a predetermined order and participants had to reproduce the tapping sequence with the same blocks and same order. In the backward condition, all the procedures were the same except that the participants had to reproduce the tapping in the reverse order. There were eight items in both the forward and backward conditions, varying from two 2-block to 9-block trials. In the rotated condition, there were two identical whiteboards, with numbers facing the experimenter only and one of the boards was rotated for 180° from the other one. The experimenter tapped on the blocks in a predetermined order and the participants had to reproduce the tapping sequence with the same blocks and the same order on the rotated board. There were five items, consisting of six trials, which varied from 1-block to 5-blocks. The task would stop when the participants made errors on all trials (i.e., two trials in the forward and backward conditions, and five trials in the rotated condition, respectively) in a given item. The rotated condition started with four 1-block trials practice, and there were no practice in the forward and backward conditions.

Background Questionnaire

Participants completed both a background questionnaire about their demographic and language related information. The demographic information included gender, age, nationality, and handedness. Factors that have previously been reported to affect cognitive performance were collected, including socioeconomic status (SES; Xie & Pisano, 2018), immigration status (Paap et al., 2015), musical experience (Bialystok & DePape, 2009), and video-gaming experience (Bialystok, 2006).

2.3. Experimental Tasks

Four non-linguistic cognitive tasks (i.e., Attention Network Task-ANT, Number Stroop task, Simon task, and Elevator subtests of the Test of Everyday Attention) were employed to measure different aspects of executive functions. In the computerised tasks (i.e., the ANT, Stroop, and Simon tasks), all stimuli were presented on E-Prime (version 2.0) on a 17-inch computer screen. The subtests of the TEA were displayed through media player. A schematic representation of each task is depicted in Figure 1.

<Insert Figure 1 about here>

Attention Network Task (ANT)

This task is a well-established assessment of attentional capacities (i.e., alerting, orienting, and inhibition) (Fan, McCandliss, Sommer, Raz, & Posner, 2002), and has been used to investigate the cognitive effects of bilingualism (e.g., Costa et al., 2008). We used it to measure S-S conflict effects in the current study. Participants were instructed to respond to the central arrow of the five horizontal arrows presented in the middle of the screen either below or above a fixation cross. In the congruent condition (e.g., →→→→→), the five arrows pointed in the same direction. In the incongruent condition (e.g., →→←→→), the central arrow was pointing opposite to the other arrows, creating a S-S conflict (see Figure 1).

Three attentional indices were obtained calculating the difference in RTs / Accuracy rate between the following trials: ANT conflict (congruent vs. incongruent trials), ANT alerting (double-cue vs. no-cue trials), ANT orienting (centre-cue vs. single-cue trials). Participants started with a practice block consisting of 24 trials and followed by three experimental blocks of 96 trials each. All stimuli were presented randomly at an equal number of times in each block. Feedback on performance was only provided in the practice block.

Number Stroop Task

To avoid any linguistic influence, we used a numerical version of the Stroop task adapted from Hernández et al. (2010). Participants were asked to count digits or symbols presented on the centre of the screen by pressing the keys 1, 2, or 3 on the keyboard while ignoring the numerical value of the digits. In the congruent condition (e.g., 22), the numerical value matched the number of digits. In the incongruent condition (e.g., 222), the correct response conflicted with the numerical value, creating a S-R conflict. The Stroop effect was assessed by the RTs/ accuracy differences between incongruent and congruent trials (Stroop, 1935) (see Figure 1). Participants were given a practice block with 18 trials which were followed by two experimental blocks of 90 trials each. Feedback on performance was only provided in the practice block.

Simon Task

We used an arrow version of the Simon task, which comprises a S-R conflict. In this task, an arrow appeared at one of four possible locations of the screen (left, right, up, or down), pointing at one of four possible directions (left, right, up, or down). Participants were instructed to respond to the direction of the arrow by pressing the correspondent button on the keyboard. There were three types of trials: baseline, congruent, and incongruent. Executive functioning was assessed by computing the RTs/ accuracy differences between incongruent and congruent trials (Simon Effect; Simon & Rudell, 1967). Participants started with a practice block consisting of 10 trials followed by three experimental blocks of 60 trials each in the order of baseline block, congruent block, and incongruent block.

Test of Everyday Attention (TEA)

The TEA (Robertson, Ward, Ridgeway, & Nimmo-Smith, 1994) is a well-established clinical assessment of attention. We selected three subtests of the Elevator tasks to measure executive control in the auditory domain (Bak et al., 2014; Vega-Mendoza et al., 2015). All tasks were presented through media player with a headset.

a) Elevator with Counting (EC: 7 trials): this task assesses sustained attention. Participants were asked to count tones of the same pitch presented at irregular intervals.
b) Elevator with Distraction (ED: 10 trials): this task assesses auditory selective attention/inhibition. Participants were asked to count low tones while ignoring interspersed high tones.
c) Elevator with Reversal (ER: 10 trials): this task assesses auditory attentional switching (auditory-verbal working memory). Participants were presented with high, middle, and low tones. The middle tones were to be counted while the high and low tones indicated the counting direction (upwards and downwards, respectively).

2.4. Statistical analyses

All analyses were conducted by fitting linear mixed-effect models (LMMs) from the lme4 packages (Bates, Mächler, Bolker & Walker, 2015) into R (Version 3.6.1, The R Foundation for Statistical Computing, 2019). The main motivation for using LMMs in favour of traditional analysis is that LMMs are robust against unbalanced datasets and based on by-subjects and by-items analyses (Baayen, Davidson & Bates, 2008). Given the advantages of using LMMs, it was recommended by Second Language Acquisition researchers (Cunnings 2012).

In the initial analysis, background measures with continuous variables were tested with the Shapiro-Wilk test for their normality. Normally distributed variables were analysed using the ANOVA test, and non-normally distributed variables (i.e., L2 proficiency, SES, age, AoA, and Raven's) were analysed using the Kruskal-Wallis test. Other experience measures on a nominal scale (i.e., gender, handedness, musical experience, and video-gaming experience) were using the chi-squared test. In the main analyses, LMMs were employed[2] Before the RT data analyses, following Ooi et al. (2018), we excluded the trials for which RTs were outside of 3-SD of each participant mean across all trial types and RTs associated with incorrect responses. Thus, the total trials excluded in each task were the following: ANT: 2.87% ; Stroop: 6.49%; Simon: 4.37%. In the LMMs for RT analyses, Group (i.e., Mono vs. High-Bi vs. Low-Bi) and Trial Type (e.g., congruent vs. incongruent trials in the model for Conflict effect) were fixed variables unless specified, and participants and items were random variables (e.g., including random intercepts for each participant and item). Fixed variables were allowed to interact with each other in a single model. For reasons of relevance to the research questions, background measures that indicated group differences were added into the model without interaction with the other fixed variables. All participants showed comparable performance and relatively high accuracy rate (i.e., ANT: 97.16%; Stroop: 94.98%; and Simon: 96.94%), therefore, we did not analyse it.

[2] As a reviewer suggested, we reanalysed the data with L2 proficiency as a continuous variable. When reanalysing the data, some of monolinguals did not report their L2 proficiency due to the fact that it was too low to be reported. The results showed that higher L2 proficiency predicted smaller conflict effects in the ANT and higher scores in the ED subtest, indicating better inhibitory control in both the visual and auditory domains. No correlations were found in the Stroop or Simon tasks.

As to the three subtests of the TEA, the accuracy rate was obtained based on the number of correct responses. Linear regression models were used with the accuracy rate as a dependent variable, and the Group as a fixed variable. Bonferroni corrections were applied with adjusted significance level of p = .0167. The outputs of main fixed effects of interest for each model are presented in Supporting Materials.

3. Results

3.1. Initial analyses

There were no group differences on age, AoA, gender distribution, video-gaming experience, and handedness distribution (all ps > .05); and no group differences on the APM and CTT (forward, backward, and rotated conditions) (all ps > .05) were found either, suggesting comparable basic cognitive abilities among the three groups.

Group differences were found on SES (p = .003) and musical experience (p = .03). Both SES and musical experience were put into the models as fixed variables, indicating that the main effects of Group and Trial Type were interpreted as controlling the effects of these background measures. As expected, the Bi-high group had higher self-reported L2 proficiency than that of the Bi-low group on overall proficiency, as well as four sub-components of language skills (all ps< .05; see Table1).

<Insert Table 1 about here>

3.2.Main analyses

ANT

Overall performance (i.e., RTs and accuracy rate) is shown in Figure 2. Mean RTs on the respective trial types are given in Table 2. The outputs of main fixed effects of interest are presented in Table S2.

<Insert Figure 2 about here>

<Insert Table 2 about here>

Overall performance

On overall RTs that collapsed across all conditions, the main effects of Group and SES were not significant (p = .23 and p = .44, respectively). The effect of musical experience was significant, with faster responses in those having musical experience than those who did not (p = .048).

Alerting Effect

The Alerting effect was significant, with faster responses on double-cue trials than on no-cue trials [F (1, 22.22) = 4.47, p = .046]. No other fixed effects or interaction effect were significant (all ps > .05).

Orienting effect

The main effects of Orienting, Group, and SES were not significant (all ps > .05). The main effect of musical experience was significant (p = .04), with faster responses in those having musical experience than those who did not. The interaction between Group and Orienting effect was significant [F (2, 106.89) = 3.84, p = .025]. Follow-up analysis showed that the Bi-high group displayed a smaller orienting effect than the monolinguals (β = 12.64, t = 2.65, p = .009), with no group differences between the Bi-low and monolingual groups (β = 2.94, t = .63, p = .53) or between the Bi-low and the Bi-high groups (β = 9.71, t = 2.071, p = .041) (p-value did not survive Bonferroni correction). Further analysis showed that the Bi-high and monolingual groups had similar response time on single-cue trials (p = .23), but the Bi-high group responded faster on centre-cue trials than the monolinguals did (p = .038), which led to a smaller Orienting effect in the Bi-high group relative to the monolinguals.

Conflict effect

The Conflict effect was significant, with faster responses on congruent trials than on incongruent trials [F (1, 33.02) = 90.76, p < .001]. The main effects of Group and SES were not significant (p = .15 and p = .53, respectively). The main effect of musical experience was significant (p = .043), with faster responses in those having musical experience than those who did not. The interaction between the Group and Conflict effect was significant [F (2, 106.61) = 4.16, p = .018]. Follow-up analysis found that both the Bi-high and Bi-low groups showed smaller conflict effects than the monolinguals did (β = 14.97, t = 2.56, p = .012 and β = 14.00, t = 2.45, p = .016, respectively), with no group difference was found between the Bi-high and Bi-low groups (β = .97, t = .17, p = .87).

Number Stroop Task

Overall performance is shown in Figure 3. Mean RTs on the respective trial types are given in Table 2. The outputs of main fixed effects of interest are presented in Table S3.

<Insert Figure 3 about here>

Overall performance

In the analysis of Global RTs, no fixed effects were significant (all ps > .05).

Stroop effect

The Stroop effect was significant, with faster responses on congruent trials than on incongruent trials [F (1, 7.31) = 19.36, p = .003]. No other fixed effects or interaction effect were significant (all ps > .05).

Simon Task

Overall performance (i.e., RTs and accuracy rate) is illustrated in Figure 4. Mean RTs on the respective trial types are given in Table 2. The outputs of main fixed effects of interest are presented in Table S4.

<Insert Figure 4 about here>

Overall performance

In the analysis of Global RTs, no fixed effects were significant (all *p*s > .05).

Simon effect

The Simon effect was significant, with faster responses on congruent trials than on incongruent trials [F (1, 9.88) = 52.74, p <. 001]. The main effect of SES was significant (p = .02): higher SES predicated faster response time. No other fixed effects or interaction effect were significant (all *p*s > .05).

TEA

In the EC and ER, there were no group differences (all *p*s > .05). In the ED, after Bonferroni correction (i.e., adjusted p = .0167), the Bi-high group showed a significant better performance than the monolinguals did [β = 13.49 [3.71, 23.26], t = 2.74, p = .007]; the difference between the Bi-low group and the monolinguals did not reach significance [β = 11.49 [1.89, 21.1], t = 2.37, p = .020](see Figure 5). There was no group difference between the Bi-high and Bi-low groups [β = 1.99 [-7.84, 11.83], t = 0.4, p = .689]. For background measures, musical experience was a significant predictor of performance in the ER, with better performance in those having musical experience than those who did not [β = 13.18 [1.17, 25.19], t = 2.18, p = .032]. The outputs of main fixed effects of interest are presented in Table S5.

<Insert Figure 5 about here>

4. Discussion

The main purpose of this study was to identify the effects of bilingualism on two potential subcomponents within interference suppression (i.e., S-S vs. S-R inhibition) in young adult monolinguals and bilinguals. Under the hypothesis that the bilingual experience affects specific inhibitory mechanisms, performance of young bilinguals and monolinguals was compared on three non-linguistic visual interference tasks (i.e., the ANT, Stroop, and Simon tasks) and one auditory inhibition task (i.e., the ED subtest in the TEA). We predicted superior performance in bilinguals relative to monolinguals in tasks tapping into inhibitory control in both the visual and auditory domains.

More specifically, based on previous research (Blumenfeld & Marian, 2014), better performance by bilinguals tends to be observed in tasks tapping into S-S inhibition (i.e., ANT), but not in tasks involving S-R inhibition (i.e., Stroop and Simon tasks) in the visual domain. The results of this study confirmed the prediction, showing better performance in the bilingual groups (i.e., the Bi-high and Bi-low groups) compared to the monolingual group in the ANT (i.e., smaller conflict effects), but comparable performance in the Stroop and Simon tasks between the groups. Our findings were consistent with previous research that bilingualism might engage S-S inhibition mechanisms more than S-R mechanisms due to cross-linguistic co-activation and competitions in language comprehension and competition processes. As Blumenfeld and Marian (2014) suggested, S-S and S-R inhibition might not arise together, although both of them are likely to present during bilingual production. Bilinguals who lived in their dominant-language linguistic context may co-activate their two languages up to the lemma level, with cross-linguistic competition resolved at this stage and with language selective processing at the response planning stages. Hence, such bilinguals may mainly experience S-S inhibition, but not S-R inhibition. This could explain the patterns of results in the current study, as the bilingual participants were native English speakers living in a unilingual context (i.e., English). Therefore, interference tasks (i.e., the ANT) involving S-S conflict, as compared to tasks involving S-R conflict (i.e., the Stroop and Simon tasks), are more sensitive to detect group differences between monolinguals and bilinguals in the current study. Moreover, the multiple alternative responses in the Stroop and Simon tasks (i.e., three and four alternative forced choices in the Stroop and Simon tasks, respectively) are more likely to result in slower responses and higher error rate (Albantaki, Branzi, Costa, & Deco, 2012). ERP measures have showed that different cognitive processes are required in the three different tasks; namely, conflict monitoring in the Stroop task, resource allocation in the Simon task, stimulus categorization and error-processing in the Flanker task (Kousaie & Phillips, 2012). However, in the present study, only behavioural measures were compared, so we do not know what the electrophysiological responses would have been to the three different types of conflict.

In the auditory domain, the results are consistent with previous studies (Bak et al., 2016; Bak et al., 2014; Long et al. 2019): only the Bi-high bilinguals outperformed the monolingual group in the ED subtest of the TEA (i.e., auditory inhibition). The Bi-low group took an intermediate position, not being significantly different from either group. This pattern was further confirmed by an additional linear trend analysis. These observations are consistent with previous research (Vega-Mendoza et al., 2015) that language proficiency might modify specific cognitive performance so that a group difference between the Bi-low group and the monolingual group might not emerge until bilinguals reach a considerable level of proficiency.

Interestingly, the Bi-high group showed a smaller orienting effect (centre-cue vs. single-cue trials) than the monolinguals in the ANT, while the Bi-low group and the monolinguals showed comparable performance. Further analysis showed that the Bi-high and the monolinguals had a similar response time on the single-cue trials, but that the Bi-high group responded faster on the centre-cue trials than the monolinguals did, thus leading to a smaller orienting effect. Previous research has shown that bilinguals show faster overall or global RTs than their monolingual counterparts (Bialystok et al., 2004; Costa et al., 2009;

Kousaie & Phillips, 2012), which was referred to as the Bilingual Executive Processing Advantage (BEPA) hypothesis (Hilchey & Klein, 2011). If so, the faster RTs in the centre-cue trials could be due to the general cognitive benefits associated with bilingualism.

With respect to background measures, our results suggest that musical experience is associated with better executive performance (e.g., auditory attentional switching and Global RTs in the ANT), while higher SES predicts faster response across the incongruent and congruent trials in the Simon task. Musical experience is one of the types of experience other than bilingualism that is associated with executive performance (Bialystok & DePape, 2009). It has been argued that musical performance is a complex task, which engages selective attention, monitoring, and shifting, particularly in the auditory domain. A growing number of studies have shown different cognitive benefits correlated with musical experience (Janus, Lee, Moreno & Bialystok, 2016; Moreno et al., 2011; Slevc, Davey, Buschkuehl & Jaeggi, 2016). As to SES, our results are consistent with previous research, in which SES has been reported to modulate cognitive performance (Xie & Pisano, 2018).

One limitation of the present study is the lack of a formal objective test of L2 proficiency. However, studies have confirmed the validity of self-reported language proficiency and usage (Luk & Bialystok, 2013; Marian, Blumenfeld, & Kaushanskaya, 2007). For instance, Marian et al. (2007) aimed to develop a valid and reliable tool for assessing language proficiency: the Language Experience and Proficiency Questionnaire (LEAP-Q). They conducted two studies to examine the internal validity and criterion-based validity of the LEAP-Q and suggested that self-reported language proficiency was reliable indicator of language performance. Moreover, there are studies that have shown that self-reports and objective measures tend to correlate. Vega-Mendoza et al. (2015), for example, found a correlation between self-reported L2 proficiency and accuracy to L2 words in a picture name verification task. Therefore, having used only one language measure (i.e., self-reports) in the current study is likely to be reliable.

Taken together, these results suggest that bilingualism benefits individuals in relation to executive functions even in young adults whose cognitive capacities are assumed to be at their highest level. Importantly, the current study separated two potential subcomponents within interference suppression (i.e., S-S vs. S-R inhibition), which could differently affect performance in monolinguals and bilinguals. Moreover, this study also investigates how language proficiency affect cognitive performance in both the visual and auditory domains: language proficiency turns out to modify cognitive performance and cause the emergence of group differences between monolinguals and bilinguals. The results may at least partly address the inconsistencies in the existing literature regarding the cognitive effects associated with bilingualism.

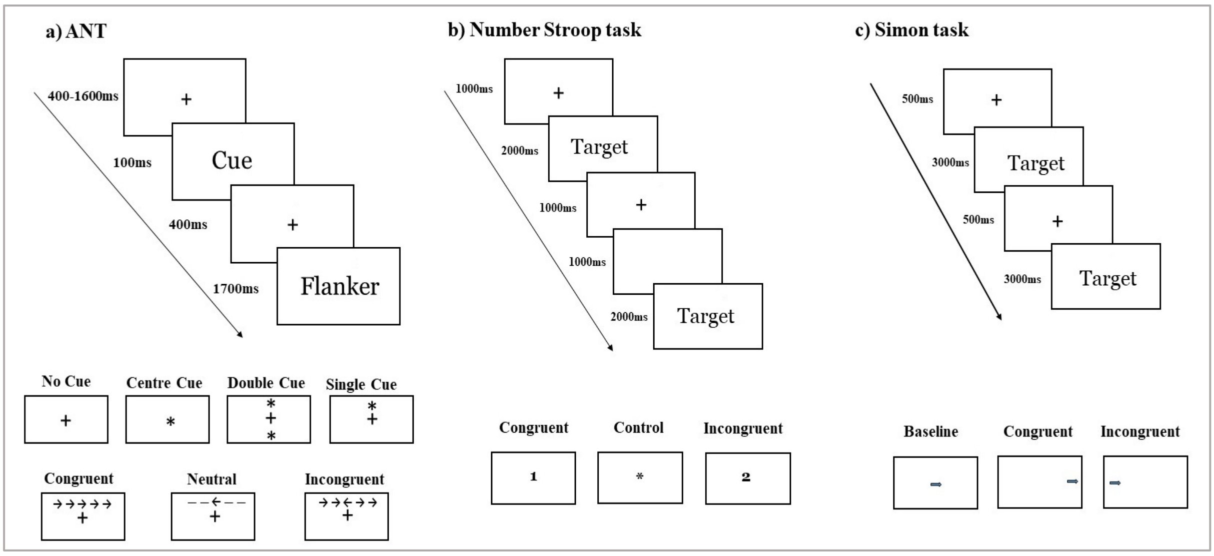

Figure 1. Schematic representation of the three RT-based interference tasks: a) ANT; b) the Number Stroop task; c) the Simon task

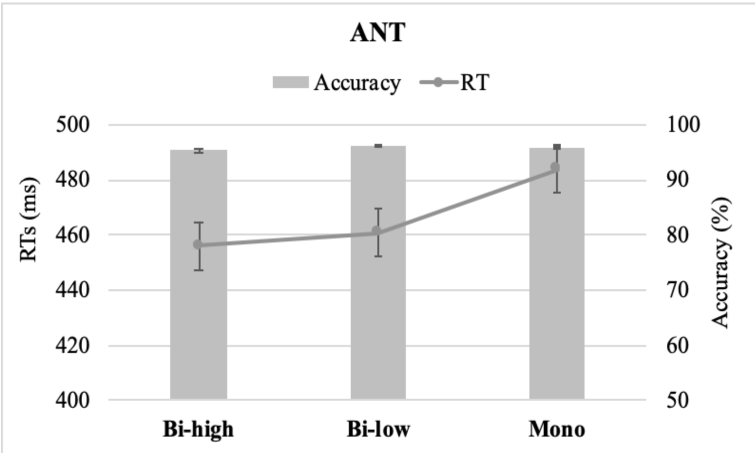

Figure 2. Overall performance in the ANT by group, collapsed across trial- and cue-type. Error bars represent ±1SE.

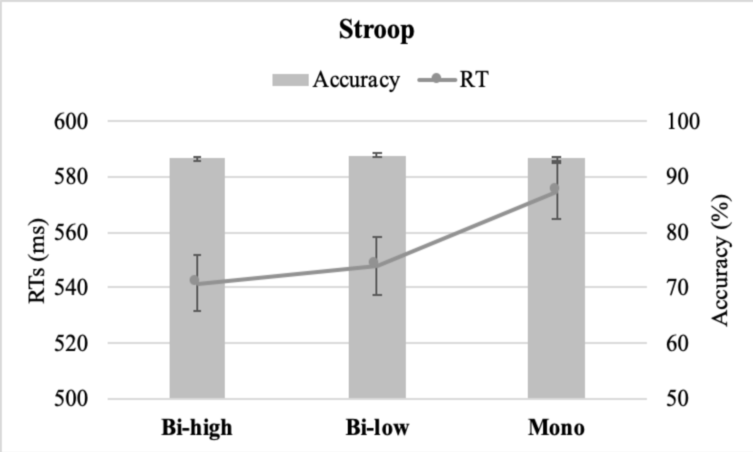

Figure 3. Overall performance in the Number Stroop task by group, collapsed across trial type. Error bars represent ±1SE.

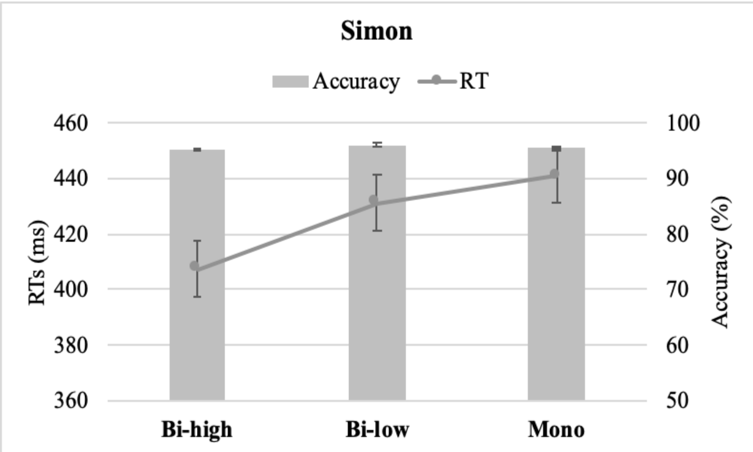

Figure 4. Overall performance in the Simon task by group, collapsed across trial type. Error bars represent ±1SE.

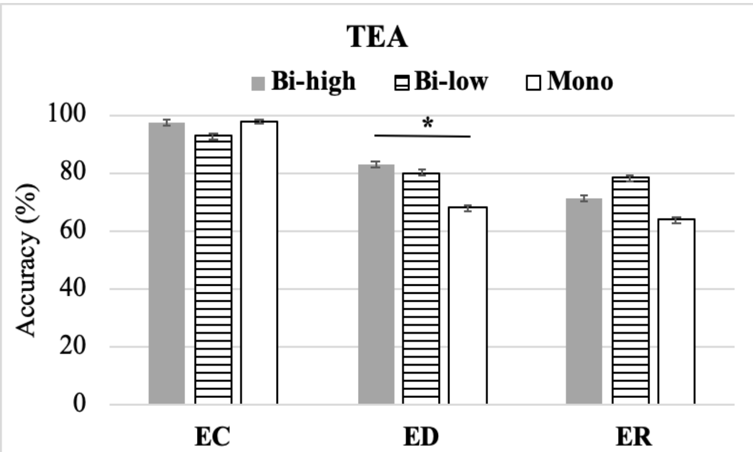

Figure 5. Performance in the respective TEA Elevator subtests by group. Error bars represent ±1SE.

Table1. *Participants' background measures. SDs are given in parentheses.*

	Mono	**Bi-low**	**Bi-high**
N /male	37 / 8	39 / 8	35 / 7
Age in year	19.86 (3.61)	18.71 (0.96)	18.88 (0.69)
SES indexed by Parents' Education [i]	3.89 (0.71)	3.73 (0.82)	4.26 (0.72) *
APM scores [ii]	17.43 (4.07)	17.51 (3.86)	17.2 (3.78)
Corsi Tapping Task			
Working memory span	52.03 (9.78)	53.04 (8.47)	51.89 (9.31)
Working memory	44.43 (10.80)	45.67 (9.20)	45.27 (11.05)
Mental rotation	36.46 (10.91)	36.41 (10.04)	36.36 (11.44)
Age of L2 acquisition in years	9.89 (3.93)	10.74 (3.13)	9.53 (4.05)
Self-reported L2 Proficiency [iii]	-	6.87 (1.72)	11.97 (1.93) *
Speaking	-	1.56 (0.6)	2.85 (0.7) *
Understanding	-	2.13 (0.73)	3.26 (0.45) *
Reading	-	1.85 (0.59)	3.12 (0.48) *
Writing	-	1.38 (0.49)	2.74 (0.71) *
Number of left-handed	7	5	5
Musical Experience Y%	64.86%	89.74%	80% *
Video-Gaming Experience Y%	24.32%	25.64%	34.29%

[i] This is an average score based on parental education level. The scale ranged from 1: *primary school*, 2: *O level or equivalent*, 3: *A level*, 4: *Bachelor's or equivalent to* 5: *postgraduate*, 6: *Ph.D.*

[ii] APM scores were the number of corrected items (the total number was 36)

[iii] The scale of self-reported L2 proficiency ranged from 1-4, marked by "*poor*" to "*excellent*"

*The difference between the groups was statistically significant

Table 2. *Interference tasks indices by Group. SDs are given in parentheses.*

	Mono	**Bi-low**	**Bi-high**
ANT			
Congruent	453 (59)	434 (52)	429 (44)
Incongruent	550 (72)	517 (58)	511 (56)
Conflict effect	**97 (30)**	**83 (20)**	**82 (23)**
Double	475 (59)	453 (56)	447 (44)
No	514 (66)	490 (58)	487 (52)
Alerting effect	**40 (19)**	**37 (20)**	**41 (21)**
Single	451 (60)	430 (49)	430 (48)
Centre	488 (64)	464 (55)	454 (47)
Orienting effect	**37 (21)**	**34 (21)**	**24 (17)**
Stroop			
Congruent	536 (76)	511 (82)	509 (84)
Incongruent	624 (95)	593 (99)	582 (110)
Stroop effect	**88 (45)**	**82 (28)**	**72 (45)**
Simon			
Congruent	385 (57)	377 (49)	367 (59)
Incongruent	513 (120)	505 (115)	457 (85)
Simon effect	**127 (75)**	**128 (91)**	**91 (41)**

Note: ANT Conflict: congruent trial vs. incongruent trial; Alerting: double-cue vs. no-cue; Orienting: centre-cue vs. single-cue. Stroop effect: congruent trial vs. incongruent trial. Simon effect: congruent trial vs. incongruent trial.

Table S1 Summary of bilinguals' other languages

	L2s (N)	L3s (N)	L4s (N)
Bi-low	French (18) Spanish (13) German (4) Greek (1) Italian (1) Korean (1) Russian (1)	French (5) Spanish (3) Gaelic (1) German (1) Mandarin (1) Dutch (1)	French (1) Italian (1) Swahili (1)
Bi-high	French (17) Spanish (11) German (3) Russian (2) Italian (1) Norwegian (1)	Spanish (6) German (3) French (2) Latin (2) Italian (1) Mandarin (1) Portuguese (1) Slovak (1)	Ancient Greek (1) Chinese (1) Finnish (1) French (1) Gaelic (1) German (1) Japanese (1) Mandarin (1) Spanish (1) Swahili (1) Thai (1)

Table S2. *Fixed effects of group comparisons for the ANT*

	Est	SE	df	t	Pr(>|t|)	
Global RTs						
Group Bi-high	-21.104	13.124	105.027	-1.608	0.1108	
Group Bi-low	-16.848	12.851	105.015	-1.311	0.1927	
Conflict effect						
Group Bi-high	-38.399	14.947	107.139	-2.569	0.0116	*
Group Bi-low	-32.968	14.551	107.086	-2.266	0.0255	*
Conflict effect	96.95	9.766	42.113	9.928	1.35E-12	***
Group Bi-high: Conflict effect	-14.96	5.87	106.75	-2.549	0.0122	*
Group Bi-low: Conflict effect	-13.989	5.71	106.393	-2.45	0.0159	*
Orienting effect						
Group Bi-high	-33.518	13.24	107.058	-2.532	0.0128	*
Group Bi-low	-24.144	12.89	107.036	-1.873	0.0638	.
Orienting effect	37.712	19.273	23.13	1.957	0.0626	.
Group Bi-high: Orienting effect	-12.643	4.78	107.13	-2.645	0.0094	**
Group Bi-low: Orienting effect	-2.938	4.649	106.64	-0.632	0.5288	
Alerting effect						
Group Bi-high	-28.123	12.753	107.105	-2.205	0.0296	*
Group Bi-low	-21.222	12.417	107.109	-1.709	0.0903	.
Alerting effect	38.988	18.324	23.264	2.128	0.0442	*
Group Bi-high: Alerting effect	1.091	4.82	106.052	0.226	0.8214	
Group Bi-low: Alerting effect	-3.235	4.687	105.551	-0.69	0.4916	

Significant codes: ‘***’ 0.001; ‘**’ 0.01; ‘*’ 0.05; ‘.’ 1

Note: Intercept is the Mono group; ANT Conflict: congruent trial vs. incongruent trial; Alerting: double-cue vs. no-cue; Orienting: centre-cue vs. single-cue

Table S3. *Fixed effects of group comparisons for the Number Stroop task*

	Est	SE	df	t	Pr(>|t|)	
Global RTs						
Group Bi-high	-22.11	21.44	103.02	-1.031	0.305	
Group Bi-low	-20.25	20.72	103.01	-0.977	0.331	
Stroop effect						
Group Bi-high	-41.504	24.657	105.179	-1.683	0.09529	.
Group Bi-low	-29.821	23.439	105.142	-1.272	0.20607	
Stroop effect	89.425	19.473	8.502	4.592	0.00151	**
Group Bi-high: Stroop effect	-15.475	9.594	104.41	-1.613	0.10977	
Group Bi-low: Stroop effect	-5.614	9.115	104.142	-0.616	0.53927	

Significant codes: ‘***’ 0.001; ‘**’ 0.01; ‘*’ 0.05; ‘.’ 1

Note: Intercept is the Mono group; Stroop effect: congruent trial vs. incongruent trial

Table S4. *Fixed effects of group comparisons for the Simon task*

	Est	SE	df	t	Pr(>|t|)	
Global RTs						
Group Bi-high	-27.462	17.171	95.01	-1.599	0.1131	
Group Bi-low	-4.389	17.167	94.996	-0.256	0.7988	
Simon effect						
Group Bi-high	-18.58	13.754	97.125	-1.351	0.1799	
Group Bi-low	-8.376	13.54	97.008	-0.619	0.5376	
Simon effect	127.094	19.149	20.226	6.637	1.73e-06	***
Group Bi-high: Simon effect	-36.443	19.02	96.79	-1.916	0.0583	.
Group Bi-low: Simon effect	0.342	18.727	96.737	0.018	0.9855	

Significant codes: '***' 0.001; '**' 0.01; '*' 0.05; '.' 1
Note: Intercept is the Mono group; Simon effect: congruent trial vs. incongruent trial

Table S5. *Fixed effects of group comparisons for the TEA*

	Est	SE	t	Pr(>|t|)	
EC (sustained attention)					
Group Bi-high	-0.04005	2.6477	-0.015	0.988	
Group Bi-low	-4.60306	2.60301	-1.768	0.0799	.
SES	-0.54673	1.40101	-0.39	0.6971	
Musical Experience	-1.98739	2.61039	-0.761	0.4481	
ED (selective attention/ inhibition)					
Group Bi-high	13.487	4.93	2.736	0.0073	**
Group Bi-low	11.494	4.847	2.372	0.0195	*
SES	2.43	2.609	0.931	0.3538	
Musical Experience	4.259	4.86	0.876	0.3829	
ER (attentional switching)					
Group Bi-high	4.388	6.144	0.714	0.476616	
Group Bi-low	11.699	6.04	1.937	0.055421	.
SES	3.454	3.251	1.062	0.290451	
Musical Experience	13.181	6.057	2.176	0.031763	*

Significant codes: '***' 0.001; '**' 0.01; '*' 0.05; '.' 1
Note: Intercept is the Mono group

Appendix S1 (questionnaire)

Supplementary material

For supplementary material accompanying this paper, visit www.cambridge.org/core/journals/bilingualism-language-and-cognition

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. https://doi.org/10.1016/j.cortex.2012.08.018

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A Systematic Review and Meta-Analysis of the Cognitive Correlates of Bilingualism. *Review of Educational Research*, *80*(2), 207–245. https://doi.org/10.3102/0034654310368803

Albantakis, L., Branzi, F. M., Costa, A., & Deco, G. (2012). A multiple-choice task with changes of mind. *PLoS ONE*, 7(8). https://doi.org/10.1371/journal.pone.0043131

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

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

Bak, T. H., Long, M. R., Vega-Mendoza, M., & Sorace, A. (2016). Novelty, Challenge, and Practice: The Impact of Intensive Language Learning on Attentional Functions. *PLOS ONE*, *11*(4), e0153485. https://doi.org/10.1371/journal.pone.0153485

Bak, T. H., Vega-Mendoza, M., & Sorace, A. (2014). Never too late? An advantage on tests of auditory attention extends to late bilinguals. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00485

Barac, R., Bialystok, E., Castro, D. C., & Sanchez, M. (2014). The cognitive development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, *29*(4), 699–714. https://doi.org/10.1016/j.ecresq.2014.02.003

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

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *60*(1), 68–79. https://doi.org/10.1037/cjep2006008

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

Bialystok, E. (2020). Null results in bilingualism research: What they tell us and what they don't. *Journal of Multilingual Theories and Practices*, 1(1), 8-22. https://doi.org/10.1558/jmtp.17104

Bialystok, E., & DePape, A.-M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 565–574. https://doi.org/10.1037/a0012735

Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition*, *112*(3), 494–500. https://doi.org/10.1016/j.cognition.2009.06.014

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, 16(4), 240–250. https://doi.org/10.1016/j.tics.2012.03.001

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

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 859–873. https://doi.org/10.1037/0278-7393.34.4.859

Blasi, G., Goldberg, T. E., Weickert, T., Das, S., Kohn, P., Zoltick, B., Bertolino, A., Callicott, J. H., Weinberger, D. R., & Mattay, V. S. (2006). Brain regions underlying response inhibition and interference monitoring and suppression: Brain regions underlying cognitive control. *European Journal of Neuroscience*, *23*(6), 1658–1664. https://doi.org/10.1111/j.1460-9568.2006.04680.x

Blumenfeld, H. K., & Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology*, *25*(5), 547–567. https://doi.org/10.1080/20445911.2013.812093

Blumenfeld, H. K., & Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology, 25*(5), 547–567. https://doi.org/10.1080/20445911.2013.812093

Blumenfeld, H. K., & Marian, V. (2014). Cognitive control in bilinguals: Advantages in Stimulus-Stimulus inhibition. *Bilingualism: Language and Cognition*, 17(3), 610–629. https://doi.org/10.1017/S1366728913000564

Brydges, C. R., Clunies-Ross, K., Clohessy, M., Lo, Z. L., Nguyen, A., Rousset, C., Whitelaw, P., Yeap, Y. J., & Fox, A. M. (2012). Dissociable Components of Cognitive Control: An Event-Related Potential (ERP) Study of Response Inhibition and Interference Suppression. *PLoS ONE*, *7*(3), e34482. https://doi.org/10.1371/journal.pone.0034482

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

Chen, P., Bobb, S. C., Hoshino, N., & Marian, V. (2017). Neural signatures of language co-activation and control in bilingual spoken word comprehension. *Brain Research*, *1665*, 50–64. https://doi.org/10.1016/j.brainres.2017.03.023

Chen, S. H., Zhou, Q., Uchikoshi, Y., & Bunge, S. A. (2014). Variations on the bilingual advantage? Links of Chinese and English proficiency to Chinese American children's self-regulation. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01069

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. *Quarterly Journal of Experimental Psychology, 70*(3), 366–372. https://doi.org/10.1080/17470218.2016.1221435

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, *106*(1), 59–86. https://doi.org/10.1016/j.cognition.2006.12.013

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

Cox, S. R., Bak, T. H., Allerhand, M., Redmond, P., Starr, J. M., Deary, I. J., & MacPherson, S. E. (2016). Bilingualism, social cognition and executive functions: A tale of chickens and eggs. *Neuropsychologia*, *91*, 299–306. https://doi.org/10.1016/j.neuropsychologia.2016.08.029

Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research*, *28*(3), 369–382. https://doi.org/10.1177/0267658312443651

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/s1366728902003012

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. https://doi.org/10.1080/20445911.2014.924951

Esposito, A. G., Baker-Ward, L., & Mueller, S. T. (2013). Interference suppression vs. response inhibition: An explanation for the absence of a bilingual advantage in preschoolers' Stroop task performance. *Cognitive Development*, *28*(4), 354–363. https://doi.org/10.1016/j.cogdev.2013.09.002

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), 340–347. https://doi.org/10.1162/089892902317361886

Garraffa, M., Obregon, M., O'Rourke, B., & Sorace, A. (2020). Language and Cognition in Gaelic-English Young Adult Bilingual Speakers: A Positive Effect of School Immersion

Program on Attentional and Grammatical Skills. *Frontiers in Psychology, 11*, 1–12. https://doi.org/10.3389/fpsyg.2020.570587

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition, 1*(2), 67-81. https://doi.org/10.1017/S1366728998000133

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

Grundy, J. G. (2020). The effects of bilingualism on executive functions: an updated quantitative analysis. *Journal of Cultural Cognitive Science*, 5. https://doi.org/10.1007/s41809-020-00062-5

Grundy, J. G., Pavlenko, E., & Bialystok, E. (2020). Bilingualism modifies disengagement of attention networks across the scalp: A multivariate ERP investigation of the IOR paradigm. *Journal of* Neurolinguistics, 56, 100933. https://doi.org/10.1016/j.jneuroling.2020.100933

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

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. https://doi.org/10.3758/s13423-011-0116-7

Hübner, R., & Töbel, L. (2019). Conflict resolution in the Eriksen flanker task: Similarities and differences to the Simon task. *PloS one*, 14(3), e0214203. https://doi.org/10.1371/journal.pone.0214203

Hui, N. Y., Yuan, M., Fong, M. C. M., & Wang, W. S. yuan. (2020). L2 proficiency predicts inhibitory ability in L1-dominant speakers. *International Journal of Bilingualism*. https://doi.org/10.1177/1367006920914399

Iluz-Cohen, P., & Armon-Lotem, S. (2013). Language proficiency and executive control in bilingual children. *Bilingualism: Language and Cognition*, 16(4), 884–899. https://doi.org/10.1017/S1366728912000788

Janus, M., Lee, Y., Moreno, S., & Bialystok, E. (2016). Effects of short-term music and second-language training on executive control. *Journal of Experimental Child Psychology*, *144*, 84–97. https://doi.org/10.1016/j.jecp.2015.11.009

Jiao, L., Liu, C., Wang, R., & Chen, B. (2019). Working memory demand of a task modulates bilingual advantage in executive functions. *International Journal of Bilingualism*, *23*(1), 102–117. https://doi.org/10.1177/1367006917709097

Keehner, M., Gathercole, S.E. (2007). Cognitive adaptations arising from nonnative experience of sign language in hearing adults. *Memory & Cognition*, 35(4), 752–761. https://doi-org.ezproxy.is.ed.ac.uk/10.3758/BF03193312

Khare, V., Verma, A., Kar, B., Srinivasan, N., & Brysbaert, M. (2013). Bilingualism and the increased attentional blink effect: Evidence that the difference between bilinguals and monolinguals generalizes to different levels of second language proficiency. *Psychological Research*, *77*(6), 728–737. https://doi.org/10.1007/s00426-012-0466-4

Kousaie, S., & Phillips, N. A. (2012). Conflict monitoring and resolution: Are two languages better than one? Evidence from reaction time and event-related brain potentials. *Brain Research*, 1446, 71–90. https://doi.org/10.1016/j.brainres.2012.01.052

Kramsch, C., & Thorne, S. (2002). Foreign language learning as global communicative practice. In D. Block & D. Cameron (Eds.), *Globalization and language teaching* (pp. 83–100). London: Routledge.

Kroll, J. F., & Bialystok, E. (2013). Understanding the Consequences of Bilingualism for Language Processing and Cognition. *Journal of Cognitive Psychology* (Hove, England), 25(5), 37–41. https://doi.org/10.1080/20445911.2013.799170

Kroll, J., Dussias, P., Bice, K., & Perrotti, L. (2015). Bilingualism, Mind, and Brain. *Annu. Rev. Linguist.*, 1(1), 377-394. https://doi.org/10.1146/annurev-linguist-030514-124937

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

Liu, X., Banich, M. T., Jacobson, B. L., & Tanabe, J. L. (2004). Common and distinct neural substrates of attentional control in an integrated Simon and spatial Stroop task as assessed by event-related fMRI. *NeuroImage*, 22(3), 1097–1106. https://doi.org/10.1016/j.neuroimage.2004.02.033

Long, M. R., Vega-Mendoza, M., Rohde, H., Sorace, A., & Bak, T. H. (2020). Understudied factors contributing to variability in cognitive performance related to language learning. *Bilingualism: Language and Cognition*, 23(4), 801-811. http://doi.org/10.1017/S1366728919000749

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

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

Luk, G., De Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? – ERRATUM. *Bilingualism: Language and Cognition*, *14*(4), 596–596. https://doi.org/10.1017/S1366728911000186

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. https://doi.org/10.1016/j.cognition.2009.08.014

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*. https://doi.org/10.1044/1092-4388(2007/067)

Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *11*(1), 81–93. https://doi.org/10.1017/S1366728907003227

Marzecová, A., Asanowicz, D., Krivá, L., & Wodniecka, Z. (2013). The effects of bilingualism on efficiency and lateralization of attentional networks. *Bilingualism: Language and cognition, 16*(3), 608–623. https://doi.org/10.1017/S1366728912000569

Mishra, R. K. (2015). Let's not forget about language proficiency and cultural variations while linking bilingualism to executive control. *Bilingualism: Language and Cognition*, 18(1), 39–40. https://doi.org/10.1017/S1366728914000455

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. *Quarterly Journal of Experimental Psychology*, *65*(8), 1502–1510. https://doi.org/10.1080/17470218.2012.657656

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

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-Term Music Training Enhances Verbal Intelligence and Executive Function. *Psychological Science*, *22*(11), 1425–1433. https://doi.org/10.1177/0956797611416999

Ooi, S. H., Goh, W. D., Sorace, A., & Bak, T. H. (2018). From Bilingualism to Bilingualisms: Bilingual experience in Edinburgh and Singapore affects attentional control differently. *Bilingualism: Language and Cognition*, *21*(4), 867–879. https://doi.org/10.1017/S1366728918000020

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

Paap, K. R., Anders-Jefferson, R., Mikulinsky, R., Masuda, S., & Mason, L. (2019). On the encapsulation of bilingual language control. *Journal of Memory and Language*, 105, 76–92. https://doi.org/10.1016/j.jml.2018.12.001

Paap, K. R., Anders-Jefferson, R., Zimiga, B., Mason, L., & Mikulinsky, R. (2020). Interference scores have inadequate concurrent and convergent validity: Should we stop using the flanker, Simon, and spatial Stroop tasks? *Cognitive Research: Principles and Implications*, 5(1), 7. https://doi.org/10.1186/s41235-020-0207-y

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. https://doi.org/10.1080/20445911.2014.944914

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. https://doi.org/10.1016/j.cortex.2015.04.014

Pratte, M. S., Rouder, J. N., Morey, R. D., & Feng, C. (2010). Exploring the differences in distributional properties between Stroop and Simon effects using delta plots. *Attention, Perception & Psychophysics*, 72(7), 2013–2025. https://doi.org/10.3758/APP

R Core Team. (2019). *R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing*. Retrieved from http://www.R-project.org/

Raven, J. C., & Foulds, G. A. (1962). Advanced Progressive Matrices. London: H.K. Lewis & Co. Ltd.

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). The test of everyday attention (TEA). Bury St. Edmunds, UK: Thames Valley Test Company, 197-221.

Rosselli, M., Ardila, A., Lalwani, L. N., & Vélez-Uribe, I. (2016). The effect of language proficiency on executive functions in balanced and unbalanced Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *19*(3), 489–503. https://doi.org/10.1017/S1366728915000309

Samuel, S., Roehr-Brackin, K., Pak, H., & Kim, H. (2018). Cultural Effects Rather Than a Bilingual Advantage in Cognition: A Review and an Empirical Study. *Cognitive Science, 42*(7), 2313–2341. https://doi.org/10.1111/cogs.12672

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*(4), 628–654. https://doi.org/10.1037/a0027473

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, 51(3), 300–304. https://doi.org/10.1037/h0020586

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*. https://doi.org/10.3389/fpsyg.2014.01481

Slevc, L. R., Davey, N. S., Buschkuehl, M., & Jaeggi, S. M. (2016). Tuning the mind: Exploring the connections between musical ability and executive functions. *Cognition*, *152*, 199–211. https://doi.org/10.1016/j.cognition.2016.03.017

Sorge, G. B., Toplak, M. E., & Bialystok, E. (2017). Interactions between levels of attention ability and levels of bilingualism in children's executive functioning. Developmental Science, 20(1), 1–16. https://doi.org/10.1111/desc.12408

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643–662. https://doi.org/10.1037/h0054651

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, *104*(30), 12530–12535. https://doi.org/10.1073/pnas.0609927104

Tiego, J., Testa, R., Bellgrove, M. A., Pantelis, C., & Whittle, S. (2018). A hierarchical model of inhibitory control. *Frontiers in Psychology*, 9, 1399. https://doi.org/10.3389/fpsyg.2018.01339

Tse, C.-S., & Altarriba, J. (2014). The relationship between language proficiency and attentional control in Cantonese-English bilingual children: Evidence from Simon, Simon switching, and working memory tasks. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00954

Vega-Mendoza, M., West, H., Sorace, A., & Bak, T. H. (2015). The impact of late, non-balanced bilingualism on cognitive performance. *Cognition*, *137*, 40–46. https://doi.org/10.1016/j.cognition.2014.12.008

Verreyt, N., Woumans, E., Vandelanotte, D., Szmalec, A., & Duyck, W. (2016). The influence of language-switching experience on the bilingual executive control advantage. *Bilingualism: Language and Cognition*, *19*(1), 181–190. https://doi.org/10.1017/S1366728914000352

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), 246–258. https://doi.org/10.1037/xge0000120

Vuillier, L., Bryce, D., Szücs, D., & Whitebread, D. (2016). The Maturation of Interference Suppression and Response Inhibition: ERP Analysis of a Cued Go/Nogo Task. *PLOS ONE*, *11*(11), e0165697. https://doi.org/10.1371/journal.pone.0165697

Wechsler D (1997) Wechsler Memory Scale (WMS-III) (Vol. 14). San Antonio: TX: Psychological Corporation.

Woumans, E., Surmont, J., Struys, E., & Duyck, W. (2016). The Longitudinal Effect of Bilingual Immersion Schooling on Cognitive Control and Intelligence*: Bilingualism

Makes You Smarter. *Language Learning*, *66*(S2), 76–91. https://doi.org/10.1111/lang.12171

Xie, Z., & Pisano, T. S. (2018). Second language (L2) proficiency, socioeconomic status (SES), and intelligence (IQ) are significant predictors of cognitive control differences among young adult unbalanced Chinese–English bilinguals. *Bilingualism: Language and Cognition*, *22*(04), 866–882. https://doi.org/10.1017/S1366728918000822

Yow, W. Q., & Li, X. (2015). Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: why variations in bilingual experiences matter. *Frontiers in Psychology*, 6, 164. https://doi.org/10.3389/fpsyg.2015.00164