Examining the effects of active versus inactive bilingualism on executive control in a carefully matched non-immigrant sample

Angela de Bruin¹, Thomas H. Bak¹, ², Sergio Della Sala¹, ²

1) Human Cognitive Neuroscience, Psychology, University of Edinburgh, Edinburgh, UK
2) Centre for Cognitive Ageing and Cognitive Epidemiology, Psychology, University of Edinburgh, Edinburgh, UK

Address correspondence to:
Angela de Bruin
Human Cognitive Neuroscience, Psychology
University of Edinburgh
7 George Square
Edinburgh EH8 9JZ - United Kingdom
Phone: +44 (0) 131 650 3426
Fax: +44 (0) 131 651 3230
Email: angela.debruin@ed.ac.uk
Abstract

Bilinguals have been argued to show a cognitive advantage over monolinguals, although this notion has recently been called into question. In many studies, bilinguals and monolinguals vary on background variables. Moreover, most studies do not distinguish between potential effects of language knowledge and language use. We examined the effects of bilingualism on executive control in older adults by comparing active and inactive bilinguals and monolinguals matched on lifestyle, socio-economic status, education, IQ, gender, and age. In the Simon arrow task, no effect of bilingualism was observed on overall RTs or the Simon effect. In the task-switching paradigm, although there was a difference between active (but not inactive) bilinguals and monolinguals on raw switching costs, the groups did not differ on overall RTs and proportional switching or mixing costs. Thus, our findings do not reveal an overall cognitive advantage of bilingualism on executive control tasks in groups matched on background variables.

Key words: Bilingualism, Language Use, Inhibitory Control, Task switching, Immigrant Status

Word count: 150
Introduction

Speaking two languages requires a constant control of both. Activating a word in the target language not only requires the speaker to activate that word, but also to inhibit the corresponding one from the non-target language (Green, 1998). This ongoing practice of language inhibition has been argued to lead to improved interference suppression in non-linguistic executive control tasks: the ability to suppress task-irrelevant information. Evidence for a bilingual advantage on inhibition tasks has been found for different age groups. Bilingual children have been found to outperform monolingual children on various inhibitory control tasks (e.g., Martin-Rhee & Bialystok, 2008; Bialystok & Viswanathan, 2009; Engel de Abreu, Cruz-Santos, Tourinho, Martin, & Bialystok, 2012). Although more inconsistently, similar inhibitory advantages have also been observed for younger adults (e.g., Costa, Hernández, & Sebastián-Gallés, 2008; Treccani, Argyri, Sorace, & Della Sala, 2009; Pelham & Abrams, 2014). Several studies comparing younger to older adults have furthermore suggested that bilingual advantages on inhibition tasks may be larger in older adults. Bialystok, Craik, Klein, & Viswanathan (2004) compared middle-aged to older bilinguals and monolinguals on a Simon task. Participants were presented with blue and red squares that were associated with a left or right button press. Stimuli appeared on the left or right side of the screen, thus leading to incongruent (e.g., left button, right side screen) and congruent (e.g., left button, left side screen) trials. Reaction times (RTs) are generally shorter for congruent trials than incongruent trials (Simon effect), but this difference was found to be smaller for bilinguals than monolinguals. This bilingual advantage was furthermore greater for older than middle-aged adults. Similarly, using a Simon arrow paradigm, Bialystok, Craik, and Luk (2008) found that older bilinguals were better at suppressing irrelevant information than monolinguals. This advantage was not found for the younger
adults tested in the same study. These advantages have also been found in verbal executive control tasks such as the Stroop task (e.g., Bialystok et al., 2008). However, to reduce the potentially confounding effects of lexical processing differences between bilinguals and monolinguals (cf., Bialystok, 2009), bilingual-monolingual differences have mainly been explored in non-verbal interference suppression tasks.

Other studies have challenged these findings. Testing participants across a range of inhibitory control tasks, several studies failed to observe a behavioural bilingual advantage in children (e.g., Antón et al., 2014; Duñabeitia et al., 2014), younger adults (e.g., Kousaie & Phillips, 2012a; Paap & Greenberg, 2013), and older adults (e.g., Kousaie & Phillips, 2012b; Kirk, Fiala, Scott-Brown, & Kempe, 2014).

Besides inhibitory control, bilingual advantages may also be related to task switching. Bilinguals were found to be faster at switching between non-verbal tasks than monolinguals in groups of children (e.g., Barac & Bialystok, 2012; Bialystok & Viswanathan, 2009), younger adults (e.g., Prior & MacWhinney, 2010), and older adults (e.g., Gold, Kim, Johnson, Kryscio & Smith, 2013). Prior and MacWhinney (2010) presented participants with stimuli that had to be responded to according to colour or shape. In the blocked condition, participants were presented with shape or colour stimuli only. In the mixed condition, participants had to switch between colour and shape decisions. The mixed condition consisted of both switch trials (switching between colour and shape) and non-switch trials (two consecutive colour or shape decisions). Bilingual participants were found to be better at switching than monolinguals. No difference was observed for the mixing costs: the difference between non-switch trials in the mixed condition and the blocked condition. These mixing costs have been argued to reflect more global mechanisms needed to maintain two competing tasks in a mixed condition (Rubin & Meiran, 2005). This suggests that the bilingual advantage is related
to switching specifically rather than more global task control. Comparing younger to older adults on a switching task, Gold et al. (2013) only found a bilingual switching advantage for the older but not the younger group. Yet bilingual switching advantages have been challenged too (e.g., Paap & Greenberg, 2013; Hernández, Martin, Barceló, & Costa, 2013).

Effects of bilingualism have predominantly been tested in inhibitory control and task-switching paradigms. However, if an advantage is found, its exact nature remains debated. Bilingual advantages have been found on inhibition costs such as the Simon costs (e.g., Bialystok et al., 2008), suggesting that the bilingual advantage concerns incongruent trials specifically. Yet bilingual advantages have not only been found on incongruent trials, but also on both congruent and incongruent trials (e.g., Bialystok, 2006). In 2011, Hilchey and Klein reviewed 31 experiments examining effects of bilingualism on executive control tasks. They concluded that there was hardly any evidence to support the hypothesis that bilinguals have an advantage on inhibitory control. Rather, they concluded, that bilinguals may have a more global advantage in monitoring conflict and regulating task demands. If bilinguals indeed have a more global monitoring advantage, this should be reflected in faster overall RTs on both congruent and incongruent trials in inhibition tasks and both switch and non-switch trials in switching tasks. Other studies have confirmed the suggestion that a bilingual advantage may be more widespread than just inhibition or switching, stating that the bilingual advantage may be found in conflict monitoring (Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009) or general mental flexibility (Kroll & Bialystok, 2013).

However, the idea of a bilingual advantage on either inhibitory control or switching specifically or on a more global level has been challenged in several recent studies. Paap, Johnson, and Sawi (2014) analysed 76 studies conducted after Hilchey and Klein’s review (2011). They included tasks measuring specific interference suppression and switching costs
or global monitoring effects and mixing costs. The majority of studies did not observe a significant effect of bilingualism neither on interference suppression and switching costs nor monitoring and mixing costs, especially for larger sample sizes. Similarly, in an update to their 2011 review, Hilchey, Saint-Aubin, and Klein (in press) conclude that the evidence for a bilingual advantage on inhibitory control is still weak. Contrary to the 2011 review, however, they now also argue that evidence for a more global bilingual advantage has evaporated since their initial review. Thus, whereas initial studies showed evidence supporting a bilingual advantage, more recent studies have challenged the reliability and even existence of this effect (a process that is known as the ‘decline effect’, see de Bruin & Della Sala, 2015).

A recent meta-analysis of studies on bilingualism and executive control (de Bruin, Treccani, & Della Sala, 2015) showed an average effect size of $d = .30$, suggesting a significant but small and inconsistent effect of bilingualism across studies. However, this meta-analysis was based on published studies only. The same study showed that the interpretation of the current literature may be distorted by a publication bias: The tendency that studies with positive findings are more likely to be published than papers with null or negative findings. An analysis of publication rates of conference abstracts found that 63% of results supporting a bilingual advantage were published in a scientific journal compared to only 36% of results challenging a bilingual advantage. The current literature of published studies may thus overestimate the actual effects of bilingualism on domain-general cognitive processes.

Thus, the current literature shows an inconsistent pattern of results (see Baum & Titone, 2014; Valian, 2015; Paap, Johnson, & Sawi for recent overviews). Whereas several studies have found a cognitive effect of bilingualism, this is challenged in more recent studies. Two key issues may affect the type of results found in studies on bilingualism and
executive control: the extent to which language groups are matched on potentially confounding variables, and the type of bilinguals that is tested. In the present study, we therefore matched the language groups on background variables. We also examined the effects of language use on executive control. We will first discuss the importance of background variables.

**The importance of confounding variables**

In order to examine effects of bilingualism on executive control, one has to ensure that bilingual and monolingual groups only differ in the number of languages that they speak. In most studies, however, this is not the only difference between language groups. Morton and Harper (2007) showed the importance of matching groups on Socio-Economic Status (SES). When groups were matched on SES, bilingual and monolingual children performed similarly on the Simon task, whereas children with a higher SES showed an advantage compared to children with a relatively low SES. Another important variable is immigration (cf., Fuller-Thomson, 2015). Many studies have compared bilingual immigrants to monolingual non-immigrants. Bialystok et al. (2004) compared Canadian monolinguals to Indian bilinguals. The majority of older bilingual participants (20 out of 24) in Bialystok et al. (2008) were immigrants, whereas all monolinguals were non-immigrants. In Gold et al. (2013), 85% of the older bilinguals were immigrants compared to 25% of the monolinguals. Similarly, two of the first studies on bilingualism and Alzheimer’s Disease (Bialystok, Craik, & Freedman, 2007; Craik, Bialystok, & Freedman, 2010) confounded bilingualism and immigration. In Craik et al. (2010), 79% of the bilinguals were immigrants, compared to 32% of the monolinguals. The issue of immigration, however, also plays a role in studies that have not found effects of bilingualism (e.g., Paap & Greenberg, 2013). Differences in immigrant status could lead to
differences in cultural, ethnical and genetic background, education, and lifestyle. Indeed, several studies have suggested that immigrants, independently of bilingualism, show increased cognitive control and less cognitive decline when compared to non-immigrants (e.g., Kopec, Williams, To, & Austin, 2001; Hill, Angel, Balistreri, & Herrera, 2012). This could be due to several factors. Immigrants for example, could be argued to be more practiced in adapting to new circumstances, which would then lead to better cognitive flexibility. Cultural background has also been shown to be related to executive control. Carlson and Choi (2009) demonstrated how the entanglement of bilingualism and cultural differences can lead to misunderstandings. They found a bilingual advantage for Korean-English bilinguals living in the USA when compared to ‘matched’ American monolinguals. When compared to Korean monolinguals, however, there was no bilingual advantage at all. Other differences between immigrants and non-immigrants might be due to the so-called ‘healthy immigrant effect’. Fuller-Thompson, Nuru-Jeter, Richardson, Raza, & Minkler (2013), for example, found that white Hispanic immigrants in the USA had a 26% lower chance of functional limitations than white non-immigrant inhabitants of the USA. It could be possible that healthy people are more likely to immigrate and thus to become bilingual (cf., Fuller-Thompson & Kuh, 2014). Immigrant status is thus an important, but often neglected, background variable that may explain part of the advantages that have been assigned to bilingualism.

Some studies have examined effects of bilingualism in a non-immigrant sample and have found positive effects for healthy adults (e.g., Costa et al., 2008; Bak, Nissan, Allerhand, & Deary, 2014; Woumans, Ceuleers, Van der Linden, Szmalec, & Duyck, 2015) and dementia patients (Alladi et al., 2013, Woumans et al., 2015). On the other hand, studies with non-immigrant samples did not observe a cognitive effect of bilingualism in children (e.g., Antón et al., 2014), younger adults (e.g., Kousaie & Philips, 2012a, although some differences were
found in the ERP data), healthy older adults (e.g., Kousaie & Philips, 2012b), and dementia patients (e.g., Chertkow et al., 2010; Lawton, Gasquoine, & Weimer, 2015).

In a recent study, Kirk, Fiala, Scott-Brown, and Kempe (2014), compared Gaelic-English bilinguals to monodialectals, bidialectals, monolinguals, and Asian-English bilinguals. All five language groups showed similar Simon costs. Crucially, the control groups were living in an urban environment, whereas the Gaelic-English bilinguals were living in a rural and more isolated environment. Lifestyle and environment thus might have been confounding factors. It is, however, not always the bilingual group that consists of immigrants. In a recent study by Clare et al. (2014), which found no evidence for a delaying effect of bilingualism on the onset of Alzheimer’s Disease, many monolinguals were people who migrated to Wales from other parts of the country, while the bilingual group consisted mainly of autochthonous population.

The importance of language use

Very little is known about the potential effects of language use on executive control. Yet bilinguals differ in language acquisition, proficiency, and use, and these differences could explain the inconsistent effects on executive control. Bilingual-monolingual differences may be due to the knowledge of different languages or to the actual use of two languages. Prior and Gollan (2011) suggested that task-switching performance may be affected by daily language switching. A group of frequently switching Spanish-English bilinguals showed smaller switching costs than monolinguals, whereas this advantage was not found for less frequently switching Mandarin-English bilinguals. Verreyt, Woumans, Vandelanotte, Szmalec, and Duyck (2015) tested balanced switching, balanced non-switching, and unbalanced bilinguals on a Simon and flanker task and found a relative advantage for the
balanced switching bilinguals. These two studies suggest that language use may affect executive control. On the other hand, Bak et al. (2014) found only small differences when comparing the impact on cognitive ageing of actively using two languages (active bilingualism) to low use of a second language (passive bilingualism). However, the interpretation of their results was limited by the fact that the vast majority of their participants used their second language only rarely. In the field of dementia studies, Freedman et al. (2014) argued that the differences between the results of studies conducted in Toronto, Montreal and Hyderabad could be due to different patterns of language use (in particular of switching) in these places. However, in previous studies, participants not only differed in language use, but also in language acquisition and language proficiency. To examine the effects of language use rather than competence more specifically, we compared participants with the same pattern of language acquisition but diverging patterns of language use.

**Present study**

The present study aimed, therefore, to address two issues. First, we wanted to examine effects of bilingualism on executive control in a sample carefully matched for potentially confounding variables such as lifestyle, education, SES, immigrant status, and IQ. We therefore tested groups of bilinguals and monolinguals on the Hebrides, a group of islands located in the Western part of Scotland. All participants were born and raised and at the time of testing living on these islands. Bilinguals acquired both Gaelic and English during their childhood, whereas monolinguals only spoke English. The isolated location of these islands has resulted in a relatively homogenous population with similar levels of education and SES.
Hence, the bilinguals and monolinguals only differed in their languages, but not in their background or country of origin. Furthermore, when effects of bilingualism are observed, they are mainly found in samples of children and older adults. In studies directly comparing older and younger adults, effects of bilingualism are largest or only found in the group of older adults (e.g., Bialystok et al., 2008; Gold et al., 2013). Consistent with these findings, we tested bilinguals and monolinguals who were aged 60 years or older.

Secondly, we examined not only the knowledge of a second language but also its active use. The specific environment of the Hebrides offers a unique opportunity to dissociate these two aspects of bilingualism. In the Outer Hebrides, most people over 65 years grew up in a Gaelic-speaking family and neighbourhood environment and acquired English with the beginning of their schooling (age 5 years). However, in their adulthood some Gaelic-English bilinguals continued to actively use both languages (active bilinguals), while others mainly used English in their later life (inactive bilinguals). This allowed us to examine the effects of knowing versus actually using two languages on executive functioning while keeping the Age of Acquisition (AoA) constant. Thus, active and inactive bilinguals acquired both languages in similar manners and up to similar proficiency levels. Language usage did not differ until adulthood. Two common, external, reasons were reported by inactive bilinguals for their low use of Gaelic in adulthood. Some inactive bilinguals married a non-Gaelic speaking spouse, which required them to speak English even with their Gaelic-speaking family members. Other inactive bilinguals stopped using Gaelic because of the general decline in Gaelic speakers and the increase of English speakers in their community.

Our monolingual controls came from the Inner Hebrides: an environment very similar in terms of landscape, living conditions and lifestyle to the Outer Hebrides. However, due to a closer proximity to the English-speaking mainland, Gaelic has been replaced by
English in this area already in the early 20th century and our monolinguals had either no or only minimal knowledge of Gaelic.

To examine effects of bilingualism and language use on executive functioning, we tested bilinguals’ and monolinguals’ performance on two executive tasks that have been used in previous studies of bilingualism. Firstly, we used a Simon arrow task, which has been linked to advantages for both young bilinguals (Bialystok, 2006) and older bilingual immigrants (Bialystok et al., 2008). If bilinguals are better at interference suppression than monolinguals, they should show smaller Simon costs. If bilinguals have a more global cognitive advantage, they should show faster RTs on both congruent and incongruent trials. Secondly, participants completed a task-switching paradigm. Prior and MacWhinney (2010) and Gold et al. (2013) found smaller switching costs for bilinguals as opposed to monolinguals and these effects have been associated with daily language switching frequency (Prior & Gollan, 2011). If bilinguals are better at switching than monolinguals, they should have smaller switching costs than monolinguals. If this is furthermore affected by language use, active bilinguals should have smaller switching costs than inactive bilinguals. If bilingualism affects general monitoring rather than switching specifically, bilinguals should have reduced mixing costs.

**Methods**

**Participants**

Seventy-six older adults (25 men) participated in the study (mean age = 70.91, SD = 6.82, range = 60 – 89 years). All participants were born and raised on the Hebrides and were living on the Isles of Harris, Islay, Lewis, Mull, or Skye. Twenty-eight participants were Gaelic-English bilinguals and still used both languages on a daily basis (‘active bilinguals’; 32% men).
Twenty-four participants were Gaelic-English bilinguals, but mainly used English ('inactive bilinguals'; 29% men). Twenty-four adults were English monolinguals (33% men). None of the participants were immigrants. All participants had normal or corrected-to-normal vision and none reported colour blindness. All took part in the experiment in their home or a community centre and gave informed consent. Participants received a gift card in return for participation. Data from one Simon task and one task-switching paradigm were not recorded due to equipment malfunction.

Participants first completed a questionnaire including questions about their language use and proficiency, education, and lifestyle. Following Hollingshead’s Four Factor Index of Social Status (1975), we calculated their socio-economic status (SES) score based on education and occupation. We also derived a lifestyle score from participation in 18 activities (Scarmeas et al., 2003). As further background measures, participants completed two non-verbal components (block design and matrix reasoning) of the Wechsler Abbreviated Scale of Intelligence (WASI), which were taken as a measure of IQ. As a screening for dementia, we also administered the Addenbrooke’s Cognitive Examination-III (ACE-III, Hsieh, Schubert, Hoon, Mioshi, & Hodges, 2013). The three language groups were matched on age, SES, years of education, lifestyle (including music practice), ACE-III score, and IQ (all Fs < 1). These background data are provided in Table 1.

Furthermore, participants completed two background measurements that were not expected to show a beneficial effect of bilingualism; the outcome of these measures was not used as an inclusion/exclusion criteria. In the Tower of London task (TOL; Shallice, 1982), participants had to move coloured balls on three pegs of different height to solve 12 problems increasing in difficulty (PEBL software, cf., Mueller & Piper, 2014). They received 3 points for solving the problem in the first attempt, 2 points for the second attempt, and 1
point for the third attempt (maximum score of 36). In a baseline processing speed task, participants saw 96 centrally presented arrows pointing to the left or the right and they had to press the corresponding button. Although participants were not matched on these two tasks, they did not differ significantly on TOL performance ($\chi^2(2) = .15, p = .93$) or speed processing ($F(2, 72) = .34, df = 2, p = .71$).

The questionnaire provided us with information about the participants’ AoA and language use. All active and inactive bilinguals acquired Gaelic as their first language. Most bilinguals only acquired English at the age of five, when they had to use this language at school. Monolingual participants were all native English speakers and did not speak any other language. Participants rated their language use on a scale from 1 (‘never’) to 10 (‘always’) for five time frames: childhood at home, childhood at school, work, later life at home, and after retirement (see Table 2 for mean ratings). Importantly, active and inactive bilinguals did not differ in English and Gaelic language use during their childhood at home or at school. During their later life, however, active bilinguals reported a much more frequent use of Gaelic than inactive bilinguals.

In terms of language switching, active bilinguals reported to switch often between English and Gaelic. On a scale from 1 (‘never’) to 4 (‘very often’), the mean rating for
switching on a daily basis was 3.82 (SD = .48), for switching in a conversation 3.54 (SD = 3.54), and for switching in a sentence 3.20 (SD = 1.26).

Language proficiency in English and Gaelic was measured using both self-ratings as well as a picture-word matching task. Participants rated their proficiency in both Gaelic and English on a 10-point scale ranging from 1 (‘no proficiency’) to 10 (‘excellent proficiency’) for speaking, understanding, writing, and reading (see Table 3 for mean ratings). Active bilinguals gave higher ratings for all Gaelic components than inactive bilinguals and monolinguals. In the picture-word matching task, participants were presented with a picture and a word and had to indicate whether the word and picture formed a match or mismatch. Active bilinguals had a higher accuracy than inactive bilinguals in Gaelic. English accuracy was similar for all three language groups (see Table 3).

[insert Table 3 about here]

Materials and procedure

Participants completed a Simon arrow task, and a task-switching paradigm1.

Simon arrow task

The Simon arrow task was adapted from the paradigm used by Bialystok et al. (2008), Bialystok (2006), and Bialystok and DePape (2009). Participants were presented with arrows pointing to the left or right and had to press the button corresponding with the direction of the arrow. The arrows were presented on the left or right side of the screen. In this way,

---

1 Our experiment also included the Test of Everyday Attention (TEA). Twenty-two participants, however, were not able to complete this task because of hearing difficulties (the task requires to distinguish between high- and low-pitched tones). Therefore, we did not include the full results of this task. Performance on the TEA did not differ significantly between the three groups (F(2, 51) = .23, p = .80).
stimulus presentation and arrow direction could be congruent (e.g., left side, left direction) or incongruent (e.g., left side, right direction). The difference between incongruent and congruent trials was defined as the Simon cost. There were an equal number of congruent and incongruent trials as well as an equal number of arrows pointing to the left and right. The experimental trials were preceded by ten practice trials; participants could only start the experiment if 80% of the practice trials were answered correctly. The conflict blocks consisted of two experimental conditions: a low and high switching task. Following Bialystok (2006), we manipulated the frequency of inter-trial response switching (i.e., how often the required response for the present trials differed from the response required on the preceding trial). In the high switch condition, the 96 trials included 67 inter-response switches. In the low switch condition, the 96 trials contained 36 inter-response switches. Each switching condition consisted of two blocks of 96 trials each. The order of high and low switching blocks was counterbalanced.

A trial started with the presentation of a fixation cross for 250 ms, followed by the presentation of the arrow that remained on the screen until a response was given. Arrows were 8 cm long and 4 cm wide at the widest point. The Simon task lasted approximately 10 minutes.

**Task-switching paradigm**

The task-switching procedure was adapted from Prior and MacWhinney (2010). Participants were presented with red or blue triangles or circles. The experiment consisted of a blocked and a mixed part. In the blocked condition, participants were asked to perform one task only (i.e., sort on colour or shape). In the mixed conditions, participants had to switch between colour and shape task according to a visual cue. These conditions consisted of both switch (tasks differs from previous trial) and non-switch (task is the same in two or more
subsequent trials) trials. The difference between mixed non-switch and blocked trials was defined as mixing costs; the difference between switch and non-switch trials as switching costs. Following Gold et al. (2013), who used a similar paradigm with older adults, we calculated proportional switching \( \left( \frac{\text{switch trials RT} - \text{non-switch RT}}{\text{non-switch RT}} \times 100 \right) \) and mixing costs \( \left( \frac{\text{non-switch trials RT} - \text{blocked RT}}{\text{blocked RT}} \times 100 \right) \) for each participant individually to correct for potential baseline differences.

The trials followed an unpredictable pattern of switch and non-switch trials. Each part of the experiment was preceded by a practice block. Participants could only start the experiment if they scored 80% correct on the practice trials. Participants completed two single task blocks (colour and shape were counterbalanced across participants), which consisted of 8 practice trials and 36 experimental trials each, and a mixed condition consisting of 16 practice trials and 144 experimental trials. In the mixed condition, half of the trials were switch trials and the other half non-switch trials, equally distributed across shape and colour. No more than three stimuli of the same trial type appeared in a row. In the mixed condition, the first trial after the short break was a dummy trial excluded from analysis.

A trial started with the presentation of a fixation cross for 350 ms, followed by a blank interval for 150 ms. Then, the visual cue appeared on the screen above the fixation cross for 250 ms. Next, the stimulus was presented while the visual cue remained present. The stimulus remained on the screen for 4000 ms or until a response was given. After a blank interval of 500 ms, the next trial started. The cue for the colour task was a colour gradient; the cue for the shape task was a row of small black shapes. Although participants did not need the cue to select the task in the blocked context, the cue was present in both the blocked and mixed condition to minimise differences between the two conditions.
Participants were instructed to perform one task using their left hand and the other one using their right hand (counterbalanced across participants). The specific responses were assigned to the middle or index finger. Stimuli were presented at the centre of a white background. Visual cues were presented 2.5 cm above the stimulus (4 x 4 cm).

**Analysis**

We analysed all data using both null hypothesis statistical testing (NHST) as well as Bayesian analysis. If the null is true, the $p$-value used in NHST only states that there is lack of evidence for an effect, but it cannot support the null hypothesis directly. Bayesian analysis, however, allows us to directly compare evidence favouring the null (‘no effect of bilingualism’) to evidence favouring the alternative (‘an effect of bilingualism’). In this way, it is possible to quantify evidence for the null hypothesis. The Bayes factor (BF) is the likelihood ratio of the probability of the data given the null hypothesis over the probability of the data given the alternative. For example, when $BF_{01} = 5$, the observed data are five times more likely to have occurred under the null hypothesis than under the alternative hypothesis. When $BF_{01} = .20$, the observed data are five times more likely to have occurred under the alternative hypothesis. The BF thus gives the likelihood of observing the data if there is no effect of bilingualism compared to if there is an effect of bilingualism. Contrary to $p$-values, a null effect can therefore be supported by statistical evidence. We used the Bayes Factor package in R (Morey & Rouder, 2011), with the default prior (Rouder, Morey, Speckman, & Province, 2012), and one million iterations to calculate the Bayes factors.
Results

Simon arrow task

Accuracy scores were close to ceiling (see Table 4) and were not analysed further. For RT analysis, we removed all incorrect answers as well as RTs more than 2.5 SD above the mean (calculated separately for each condition; an additional 2.06% of the correct trials). We carried out a three-way repeated ANOVA with language group as the between-subject variable, and congruency (incongruent, congruent), and condition (high switch, low switch) as the within-subject variables.

Analysing RTs, we observed a main effect of congruency ($F(2, 72) = 118.72, p < .001, \eta_p^2 = .63$), with incongruent trials being slower ($M = 946.44, SD = 241.76$) than congruent trials ($M = 872.49, SD = 220.06$). There was also a main effect of condition ($F(2, 72) = 5.78, p = .019, \eta_p^2 = .08$), with reaction times being faster in the low switch condition ($M = 892.56, SD = 220.37$) than high switch condition ($M = 926.05, SD = 251.77$). There was no main effect of language group ($F(2, 72) = .09, p = .92$), indicating that active bilinguals ($M = 896.78, SD = 170.91$), inactive bilinguals ($M = 912.91, SD = 244.46$), and monolinguals ($M = 919.12, SD = 274.64$) were equally fast. We observed an interaction between congruency and condition ($F(2, 72) = 8.27, p = .005, \eta_p^2 = .10$), suggesting that the Simon cost was larger in the low switch condition ($M = 85.55, SD = 58.07$) than the high switch condition ($M = 61.61, SD = 72.13$). There was no significant interaction between language and congruency ($F(2, 72) = .28, p = .75$), suggesting that Simon costs were similar for the three language groups. The absence of an interaction was also supported by the Bayesian analysis. Comparing the model with main effects only to the model with main effects + congruency*language, to examine the interaction of congruency and language specifically, provided strong evidence against
the model with the interaction. The model without an interaction fits the data better by a factor of 26.78 (± 1.09%), showing that bilingualism did not affect Simon costs (see Figure 1).

[insert Table 4 about here]

[insert Figure 1 about here]

**Task switching**

Accuracy scores were close to ceiling and were not analysed further (see Table 5). For RT analysis, we removed all incorrect answers and trials preceded by an incorrect trial. RTs more than 2.5 SD above the mean (calculated separately for each condition) were also excluded from analysis (an additional 1.95% of the correct trials). We carried out a two-way repeated ANOVA with language group as the between-subject variable, and trial type (switch, non-switch, or blocked trial) as the within subject variable. We observed a main effect of trial type, \(F(2, 144) = 208.516, p < .001, \eta_p^2 = .74\), with switch trials being slower \((M = 1352.91, SD = 306.37)\) than non-switch trials \((M = 1266.97, SD = 279.52)\) and blocked trials \((M = 857.21, SD = 185.41)\). There was no main effect of language group \((F(2, 72) = .290, p = .749)\), indicating that overall reaction times were equally fast for active bilinguals \((M = 1119.58, SD = 463.60)\), inactive bilinguals \((M = 1186.68, SD = 487.44)\), and monolinguals \((M = 1137.98, SD = 432.29)\). The interaction between trial type and language group was not significant \((F(2, 144) = 1.379, p = .244)\). Using BFss to compare the model with main effects only to the model with main effects + trial type*language, in order to examine the interaction of trial type and language specifically, provided strong evidence against the model with the interaction. The model without an interaction fits the data better by a factor
of 22.49 (± .86%), showing that bilingualism did not affect RT differences between blocked, non-switch, and switch trials (see Figure 1).

We also examined effects on switching and mixing costs specifically. Raw switching costs did show a significant effect of language group \( F(2, 72) = 3.51, p = .035, \eta_p^2 = .09, BF_{01} = .60 \pm .03\% \). A post hoc Tukey test showed that active bilinguals and monolinguals differed significantly \( (p = .032) \), but the group of inactive bilinguals did not differ significantly from either active bilinguals or monolinguals. Raw switching costs, however, were affected by (non-significant) slower RTs on non-switch trials for bilinguals. Using proportional costs to correct for these baseline differences, language group did not have a significant effect on switching costs \( (F(2, 72) = 2.99, p = .057, BF_{01} = .89 \pm .03\%) \). Language group also did not affect raw \( (F(2, 72) = 1.29, p = .28, BF_{01} = 3.17 \pm .03\%) \) or proportional mixing costs \( (F(2, 72) = 1.732, p = .184, BF_{01} = 2.32 \pm .04\%) \).

Discussion

This study examined the effects of bilingualism on executive control in older bilinguals and monolinguals matched on background variables such as immigrant status, education, IQ, SES, and lifestyle. Performance of active bilinguals, inactive bilinguals, and monolinguals was compared on two executive tasks. No consistent effects of bilingualism were observed on the task-switching paradigm or Simon arrow task. Accordingly, we did not reproduce the bilingualism effects described in previous studies in our population that was carefully matched on background variables.
Bilingual participants did not have an advantage over monolinguals on the Simon task. The Simon arrow task has shown an effect of ‘bilingualism’ in groups with bilingual immigrants (Bialystok, 2006; Bialystok et al., 2008; Bialystok & DePape, 2009). The number of participants per language group in our study was similar to the sample size in these three studies that showed an effect of language group (Bialystok, 2006: N per group = 40; Bialystok et al., 2008: N per group = 24; Bialystok & DePape, 2009: N per group = 24). Following Bialystok’s (2006) suggestion that the bilingual advantage may be strongest or only present in the most difficult executive control condition, we manipulated switching in the Simon task to modify task difficulty. No effect of bilingualism on the Simon effect was found in either the low-switching or more difficult high-switching Simon task. This is compatible with the view of Hilchey and Klein (2011), who argued that there is only limited evidence for a bilingual advantage on local inhibitory control processes. At the same time, they argued that bilinguals may have a more global conflict monitoring advantage. Our study does not support this idea, as we did not observe an effect of bilingualism on overall RTs on both congruent and incongruent trials. Our study, however, is not the first to find no difference between bilinguals and monolinguals on a Simon task. Several studies, including those that tested non-immigrants, also observed no bilingual advantage on Simon costs or overall RTs (e.g., Paap & Greenberg, 2013; Kousaie & Phillips, 2012a; Kirk et al., 2014).

On the task-switching paradigm, no effect of bilingualism was observed on either overall RTs or proportional switching costs. This is contrary to two studies with younger and older immigrants (Prior & MacWhinney, 2010; Gold et al., 2013) who found a bilingual advantage on switching trials. Again, our sample size was similar to the sample size in these two studies that found an effect (Prior & MacWhinney, 2010: N per group = 45; Gold et al., 2013: N per group experiment 1 = 15, N experiment 2 = 20). We did observe an effect of
bilingualism on raw switching costs, an effect that was only significant for active but not inactive bilinguals compared to monolinguals. This finding, however, is difficult to interpret as raw switching costs were not only based on differences in switch, but also non-switch trials. In the study conducted by Prior and MacWhinney (2010), the effect of bilingualism on switching costs was driven by switch trials only. Their two language groups had virtually identical RTs on non-switch trials. In our study, however, active bilinguals not only showed a small (but not significant) advantage on the switch trials, but also a disadvantage on the non-switch trials. When we corrected for differences on non-switch trials by analysing proportional switching costs, the effect of bilingualism was no longer significant. The pattern of the interaction thus does not support a bilingual advantage for switching specifically. However, it could suggest that active bilinguals use a different strategy in this task than monolinguals. If this is the case, further research will be needed to elucidate possible mechanisms of this phenomenon.

Hilchey and Klein (2011) suggested that bilinguals may have an advantage in more global conflict monitoring processes rather than conflict resolution. Bilinguals may not show a switching advantage, but they could still be better at monitoring task switching in a mixed block compared to monolinguals. However, the mixing costs in our study did not differ between the three language groups. This does not support the idea that bilinguals are better at global monitoring processes.

We aimed to replicate task-switching paradigms that have been linked to bilingual advantages (Prior & MacWhinney, 2010; Gold et al., 2013). To replicate the exact paradigm, we only used one cue for colours and one cue for shapes. Several task-switching studies (cf., Monsell & Mizon, 2006) have suggested that using a single cue confounds task changes with cue repetition, because the cue is repeated on non-switch trials but not on switch trials. The
use of single cues could thus affect the task-switching results. In future task-switching studies, two cues per task would therefore be recommendable.

Our study thus failed to replicate previous results that showed an effect of bilingualism on Simon tasks and task-switching paradigms. The data add to an increasing literature of studies that have found no effect of bilingualism or only in restricted circumstances. A commonly used argument when no bilingual advantage is found in groups of younger adults is that this age group has already reached a peak in executive control performance, thus masking potential effects of bilingualism (cf., Kroll & Bialystok, 2013). This is unlikely the case in our sample of older adults, who had not participated in cognitive tests before and had limited experience with computers. Similarly, in some studies, an effect of bilingualism may not be found because of the type of language groups tested (e.g., bilinguals with a low proficiency, or monolinguals that have some knowledge in a second language). The bilinguals in our study, however, had acquired both languages during childhood and up to a high proficiency level. The monolinguals had never spoken a second language. The unsuccessful attempt to replicate previous positive findings of bilingualism may be due to the different populations recruited. In our case, the populations were well-matched and exempt from the immigration effect. In many previous studies, this is not the case and group effects that have been attributed to bilingualism may have been the result of other background differences.

As a second question, we examined the effects of language use on executive control. The inconsistent findings in studies measuring the effects of bilingualism on executive control may partly be explained by the different types of bilinguals that have been tested. Our study shows that bilinguals cannot be treated as one homogenous group. Our active and inactive bilinguals did not differ in AoA of Gaelic and English or in language use during their
childhood at home or school. During their later life, however, inactive bilinguals reported a lower use of Gaelic than active bilinguals. Language use, and the changes across the life span, are often neglected in bilingualism studies. We did not find significant differences between active and inactive bilinguals on the Simon task or proportional task switching costs. We did find a significant difference between active bilinguals and monolinguals on raw switching costs, whereas this difference was absent for inactive bilinguals. Active bilinguals used both languages on a daily basis and also reported switching regularly between the two languages. This may suggest that bilingual-monolingual differences could be related to using and switching between two languages rather than purely knowing two languages. However, this effect was not found when correcting for baseline non-switch differences in the proportional switching costs and thus does not seem to be switching specifically. Furthermore, the difference between active and inactive bilinguals was not significant. Nevertheless, our study shows that bilingualism cannot be treated as a categorical variable and emphasizes the need to characterise bilingualism in terms of language use.

**Implications for executive control**

Studies on bilingualism and executive control have been used to examine the effects of bilingualism. Yet, this question can also be reversed: how is executive control affected by different types of training? Examining effects of training could be a valuable method to explore the system of executive functions. Although we often refer to ‘executive functions’ or ‘executive control’ as if it is one uniform system, it refers to many different subsystems (cf., MacPherson & Della Sala, 2015; Stuss & Knight, 2013). The unity and diversity of executive functions have been discussed widely. Is there one underlying ability that governs all subsystems (unity), or are these subsystems related but distinct processes (diversity)?
Evidence has been found for both (cf., Jurado & Rosselli, 2007, for an overview). Studies with patients have shown that not all patients with frontal lobe damage perform similarly on neuropsychological tests and that some patients perform well on some tasks but not on others, suggesting that the frontal lobes and executive processes are fractionated into separate subsystems (e.g., Godefroy et al., 1999; Baddeley & Della Sala, 1996). Yet, larger correlations have been found between individual frontal tests than between frontal and non-frontal tests in patients with a single frontal lesion (Della Sala, Gray, Spinnler, Trivelli, 1998), providing evidence for the existence of a common and unifying component (e.g., general intelligence; Duncan, 2010). Miyake et al. (2000) reconciled the unity and diversity accounts. Using a latent variable analysis across various executive control tasks, they argued that there are three subcomponents: updating, inhibition, and shifting (also called ‘task switching’). The three components showed moderate correlations but are clearly separable. In an updated model, Miyake and Friedman (2012) found that the inhibition component correlated perfectly with the common component that is shared by the different subcomponents. Because of the lack of unique variance for inhibition, this component is left out in the updated model.

Transfer effects of training, and especially examining the specificity of this transfer, could further elucidate the debate on the construct of executive functions. If executive functions are exercised by domain-general mechanisms, one would expect general improved executive control performance after cognitive training. However, if executive control tasks are governed by task-specific mechanisms, cognitive training should only enhance performance on trained tasks without automatic transfer to other tasks. Bilingualism could be argued to be one such type of cognitive training (cf., Valian, 2015, for a similar argument). Especially in older adults, this expertise has been trained for many years and often on a daily
basis. Controlling and using two languages requires continuous inhibition of the non-target language. Our own data and many other studies discussed in this paper suggest that this ongoing language inhibition training does not necessarily lead to advantages on non-verbal inhibition tasks.

Besides studies comparing bilinguals to monolinguals on executive control tasks, other studies have compared language- to task-switching performance directly within bilinguals. Although studies have suggested that domain-general mechanisms are involved in both language and task switching (e.g., Abutalebi & Green, 2007; de Bruin, Roelofs, Dijkstra, & FitzPatrick, 2014), others have shown that the two types of switching are also governed by their own, task-specific mechanisms. Evidence for this claim is predominantly based on comparisons of language and switching costs. Bilinguals have been found to show different patterns of symmetry in language- and task-switching costs (e.g., Calabria, Hernández, Branzi, & Costa, 2011) and diverging effects of ageing on language- versus task-switching performance (e.g., Weissberger, Wierenga, Bondi, & Gollan, 2012). In combination with the inconsistent and sometimes ambiguous effects of bilingualism and language use on task-switching performance, this suggests that language- and task-switching may share overlapping mechanisms but are not identical. Training on language-switching tasks furthermore does not necessarily lead to enhanced performance on task-switching paradigms. Prior and Gollan (2013) tested participants who received practice in a language- or task-switching paradigm and then completed both tasks one week later. Although both language- and task-switching showed within-task training effects, the cross-task transfer effects were limited. There was no transfer effect from language-switching to task-switching. Only task-switching before the language-switching task showed a transfer effect: Mixing costs in the non-dominant language were smaller after task-switching training.
Limited transfer has not only been found in studies testing effects of bilingualism on executive control. Playing video games has been found to lead to enhanced attentional control or working memory (e.g., Glass, Maddox, & Love, 2013) but has also been challenged in other studies (Unsworth et al., 2015). In a recent study, including tasks similar to those used in bilingualism research such as the flanker and Stroop task, Unsworth et al. (2015) concluded that the relation between video-game playing and cognitive abilities is weak or non-existent.

Studies examining cognitive training have also found that improved performance is often shown on trained tasks, but does not automatically transfer to untrained tasks (e.g., Dahlin, Neely, Larsson, Bäckman, & Nyberg, 2008; Redick et al., 2013). In a study with preschoolers, effects of inhibition training were found on the trained tasks, but not on the untrained tasks (Thorell, Lindqvist, Bergman, Bohlin, & Klingberg, 2008). They included inhibition tasks that are also commonly used in bilingualism research: The trained flanker and go/no-go task showed effects of training, whereas this did not extend to the non-trained Stroop task. This has not only been found with healthy and young participants, but also for patients with frontal lobe damage. Intervention studies generally reduce executive problems in patients, but this does not always transfer to untrained tasks (for a review, see Boelen, Spikman, & Fasotti, 2011). Therefore, different treatments according to different executive problems (e.g., Mateer, 1999) and multifaceted strategy trainings (e.g., Spikman, Boelen, Lamberts, Brouwer, & Fasotti, 2010) have been advocated.

Problematic for the account of a unified system of executive functions are the low correlations between executive control tasks observed in studies on bilingualism. Even executive control tasks that have been assumed to measure similar aspects of executive control do not always correlate. Paap and Greenberg (2013) only found a correlation of $r = -$
0.01 between the Simon and flanker effects, two tasks that have been argued to measure inhibitory control. Similarly, Duñabeitia et al. (2014) observed low correlations between their two types of verbal and numerical Stroop tasks ($r = .07$ for the Stroop effect, $r = .05$ for facilitation, and $r = .14$ for inhibition). Thus, even though we assume that two tasks measure the same mechanisms, the low correlations suggest that task-specific mechanisms may overrule potential domain-general effects. Although Simon and flanker tasks may appear to tap into a common mechanism and have a similar set-up in terms of incongruent and congruent trials, they also differ in many other aspects including the type of stimuli used. Possibly the most striking difference is the type of interference. In a flanker task, this interference is caused by flanker arrows surrounding a target arrow. The interference is thus found in the periphery. In the Simon task, however, there is only one stimulus and the interference is related to the match between presentation side and button side. This problem has also been referred to as task impurity: An executive control task does not only measure executive functions but also includes non-executive components (such as type of stimuli or processing speed) that affect performance (Jurado & Rosselli, 2007). In this light, it is perhaps also not surprising that effects of bilingualism are inconsistent across tasks.

Even tasks that are assumed to tap into similar processes may be largely affected by task-specific differences. Executive functioning may have an underlying domain-general mechanism but it also consists of different subcomponents that do not necessarily affect each other. This could partly explain why effects of bilingualism on executive control are inconsistent and may not always be found. Thus, both when studying executive control in general as well as the effects of bilingualism in particular, we should not aim to only identify

---

2 The absence of a correlation between two inhibitory control tasks seems to contradict the correlations found by Miyake et al. (2000). However, it should be noted that their overview of tasks did not include the Simon or flanker tasks, two paradigms that have often been studied in bilingualism research.
one general structure. Rather, we also need to establish the more specific mechanisms that are involved. Increasing our knowledge of the mechanisms behind executive control and the components that are involved in different tasks will also aid researchers to establish a more theory-driven approach when studying the effects of bilingualism (see Jared, 2015).

**Conclusion**

Our study is not the first to question the reliability of bilingual advantages in executive control tasks. Paap and Greenberg (2013), Paap and Sawi (2014), Gathercole et al. (2014), and Duñabeitia et al., (2014) all reported no effects of bilingualism across multiple tasks and using large sample sizes. This study, however, has been particularly thorough in matching participant groups on all background variables other than bilingualism itself. The Simon task and task-switching paradigm used in this study have been shown to be influenced by bilingualism in other studies comparing bilingual immigrants to monolingual non-immigrants (e.g., Bialystok et al., 2008; Gold et al., 2013). Our study shows the need to match bilingual and monolingual groups on background variables, including immigrant status. If we want to investigate the effects of *bilingualism*, we have to ensure that language groups only differ in terms of the languages that they speak. Moreover, we need to dissociate the effects of the knowledge of a language from the effects of its regular use.
References


Barac, R., & Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child Development, 83*(2), 413-422.


Fuller-Thomson E., & Kuh D. (2014). The Healthy Migrant Effect May Confound the Link between Bilingualism and Delayed Onset of Alzheimer’s Disease, *Cortex (52)*, 128-130.


