Published in final edited form as:
Psychol Aging. 2014 September ; 29(3): 696–705. doi:10.1037/a0037254.

Effects of Bilingualism and Aging on Executive Function and Working Memory

Ellen Bialystok[1], **Gregory Poarch**[1], **Lin Luo**[1], and **Fergus I.M. Craik**[2]

[1]York University

[2]Rotman Research Institute of Baycrest

Abstract

Two studies are reported in which younger and older monolingual and bilingual adults performed executive function tasks. In Study 1, 130 participants performed a Stroop task and bilinguals in both age groups showed less interference than monolinguals with a greater benefit for older adults. In Study 2, 108 participants performed a complex working memory task based on verbal or nonverbal stimuli. Bilinguals showed less interference than monolinguals, with a larger bilingual advantage in the older adult group and in the nonverbal task. Together, these results show that bilingual advantages in executive function depend on characteristics of the participants and features of the tasks, with larger effects found for older than younger adults and for complex tasks using nonverbal material.

Keywords

bilingualism; executive function; memory; working memory; Stroop task

---

Lifelong bilingualism has been shown to have a positive effect on the efficiency of the executive functioning (EF) system. Compared to monolinguals, bilinguals show superior performance in versions of such tasks as the Simon task (Bialystok, Craik, Klein, & Viswanathan, 2004; Poarch & Van Hell, 2012b), Stroop task (Bialystok, Craik, & Luk, 2008a; Blumenfeld & Marian, 2011), and flanker task (Costa, Hernández, & Sebastián-Gallés, 2008; Pelham & Abrams, 2013). All these tasks require resolving conflict from distracting cues, switching efficiently between types of trials, and maintaining rules in working memory, all components of EF. The enhancement of this EF system in bilinguals is thought to reflect its role in managing attention to competing linguistic representations (Bialystok & Craik, 2010; Kroll & Bialystok, 2013). Psycholinguistic research has demonstrated that both languages are activated during linguistic processing in bilinguals (Morford, Wilkinson, Villwock, Piñar, & Kroll, 2011; Thierry & Wu, 2007), creating potential competition between languages, even in strongly monolingual contexts for highly proficient bilinguals, and making linguistic processing more effortful (see Kroll, Dussias, Bogulski, & Valdes-Kroff, 2012, for a recent review). Thus there is a need for a control mechanism to select appropriately and avoid intrusions from the unwanted language (Green,

---

Address for Correspondence: Ellen Bialystok, Department of Psychology, York University, 4700 Keele Street, Toronto, Ontario, Canada M3J 1P3.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

1998; La Heij, 2005). Bilingualism thus entails extensive practice of this control mechanism because it is invoked every time language is used, strengthening it for functions beyond language control.

The case linking bilingualism to enhanced executive control is strengthened by findings from cognitive neuroscience. Abutalebi and Green (2007) have shown that a network of regions involving the anterior cingulate cortex, left prefrontal cortex, left inferior parietal lobule and caudate is activated in monolinguals performing tasks requiring resolution of conflict from competing responses. Strikingly, they show that bilinguals use the same network to control interference from the non-target language, and argue that this overlap in activated regions leads to strengthening of general control processes in bilinguals simply by virtue of knowing and using two languages.

In contrast to its enhancement of EF, bilingualism shows opposite effects for aspects of linguistic processing. Studies comparing monolinguals and bilinguals have shown smaller vocabulary size in bilinguals at all ages (Bialystok, Luk, Peets, & Yang, 2010; Bialystok & Luk, 2012; Portocarrero, Burright, & Donovick, 2007), slower picture naming times (Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Poarch & Van Hell, 2012a), more tip-of-the-tongue states (Gollan, Montoya, & Werner, 2002), and more interference in lexical decision tasks (Michael & Gollan, 2005). The combination of reduced lexical resources and the need to recruit EF to resolve competition makes linguistic processing more effortful for bilinguals than for monolinguals. Thus, bilinguals tend to perform more poorly than monolinguals in tasks that rely on vocabulary knowledge or lexical access.

The contrast between a nonverbal advantage and a verbal disadvantage for bilinguals in tasks involving control and interference suggests that monolinguals might outperform bilinguals on verbal control tasks, but that the bilinguals would be superior on similar tasks involving nonverbal materials. This result was reported by Luo, Craik, Moreno and Bialystok (2013) using simple and complex span tasks that were either verbal (word and alpha span) or spatial (Corsi forward and backward); a language by domain interaction was found in both younger and older adults. This contrast between advantages in nonverbal EF and disadvantages in lexical processing has also been shown in studies in which both linguistic tasks and nonverbal control tasks were administered to the same participants (Bialystok et al., 2008a; Pelham & Abrams, 2013).

Clearer evidence of the interplay between bilingual disadvantages in verbal processing and bilingual advantages in EF comes from comparing performance in two conditions of a single task, namely, the standard neuropsychological assessment of verbal fluency. The two conditions differ in their processing demands. Specifically, category fluency relies primarily on linguistic processing and performance is associated with grey matter density in the left inferior temporal cortex, a region relevant for linguistic ability, whereas letter fluency relies additionally on executive control and performance is associated with grey matter density in the pre-supplementary motor area and head of caudate, both part of the EF network (Grogan, Green, Ali, Crinion, & Price, 2009). Thus, category fluency reflects vocabulary and letter fluency reflects both vocabulary and EF. For unselected groups of monolinguals and bilinguals, bilingual participants typically perform more poorly than monolinguals on category fluency (Gollan et al., 2002; Portocarrero et al., 2007; Rosselli et al., 2000; Sandoval, Gollan, Ferreira, & Salmon, 2010), reflecting smaller vocabulary but results are more mixed for letter fluency (Kormi-Nouri et al., 2012; Portocarrero et al., 2007; Rosselli et al., 2000; Sandoval et al., 2010). However, if monolinguals and bilinguals are matched on vocabulary size, then both groups perform equivalently on category fluency but bilinguals produce more words than monolinguals in letter fluency (Bialystok, Craik, & Luk, 2008b; Luo, Luk, & Bialystok, 2010). Thus, with equivalent vocabulary and lexical access, bilinguals outperformed monolinguals on a verbal task in the condition that required EF. To summarize, the bilingual disadvantage in lexical tasks is related to reduced vocabulary and more effortful lexical access because of competition from the other jointly activated language. Resolution of this competition through recruitment of EF ultimately benefits the EF system in general.

Although the predictions regarding bilingual advantages in EF and disadvantages in lexical processing are clear in this literature, the bilingual advantage in EF has become somewhat elusive in recent studies with young adults (e.g., Paap & Greenberg, 2013). Therefore, one purpose of the present studies was to provide clarification of the conditions under which the effect is found and when it is not. From the results of studies reported above, it seems that the use of verbal as opposed to nonverbal materials is one factor that would reduce or even reverse the bilingual advantage in EF tasks. A second factor may be the age of participants, as several previous studies have found a larger bilingual advantage in older than in younger adult participants (e.g. Bialystok et al., 2004, 2008a; Gold, Kim, Johnson, Kriscio, & Smith, 2013), but studies failing to find such effects have all been conducted with only young adults. A third factor may be complexity of the task, with the bilingual advantage being strongest with complex tasks requiring substantial amounts of EF. This result was reported by Bialystok (2006) in a study of young adults using simple and complex versions of the flanker task. Therefore, on the basis of these previous studies it was predicted that the bilingual advantage would be most obvious in older adults performing a relatively complex nonverbal task.

In a previous study, Luo and colleagues (2013) found an interaction between adult age and processing domain in simple working memory (WM) span tasks, but no study to date has examined performance on complex WM tasks requiring EF. Therefore, Study 2 used such a task to address aspects of EF that pertain to WM processing, such as maintenance of relevant information but inhibition and deletion of information that is no longer relevant. Moreover, to cover a broader spectrum of EF processing, Study 1 used the Stroop paradigm, a well-established assessment of EF (Duncan & Owen, 2000; Stuss, Flodden, Alexander, Levine & Katz, 2001). Thus, Study 1 examined effects of aging and bilingualism in a task requiring online perceptual control and inhibition, whereas Study 2 explored the effects of these variables in a WM task requiring controlled attention to the current version of a series of displays while also inhibiting interference from previous versions of highly similar displays. Our predictions were that bilinguals would show more resistance to interference in both studies, and that this effect would be stronger in older adults.

In summary, the present research addressed three main questions. First, in tasks involving EF, are bilingual advantages more apparent at different points in the lifespan? To test this idea we administered the same EF tasks to younger and older adults who were comparable on general cognitive ability. The hypothesis was that language group differences would be more substantial in older than in younger adults. Second, we investigated whether the materials used in EF tasks, specifically verbal versus nonverbal materials, interacted with bilingualism in influencing performance outcomes. The hypothesis was that a bilingual advantage would be greater in nonverbal than in equivalent verbal tasks. Third, we asked whether a bilingual advantage would be found in a WM task involving a high degree of EF. Given that executive processes are generally considered to be an inherent part of WM (Baddeley, 2003), it would seem that such an effect should be present, but WM performance has not been well studied in bilinguals.

Study 1

The first experiment examined the effects of aging and bilingualism on a version of the Stroop paradigm. The hypothesis was that the ability to resist interference, as indicated by the difference between the time taken to name the color of non-matching color words and time to name color patches, would be greater in younger adults and in bilinguals than in their respective counterparts. Although the Stroop paradigm uses printed words as stimuli and so is essentially a verbal task, it is nonetheless appropriate for use with bilinguals if two conditions are met. First, if monolinguals and bilinguals read simple versions of the color words at the same speed indicating the same degree of automaticity in word reading, then it is likely that the degree of semantic interference is equivalent as well. In this case, differences in interference can be attributed to differences in executive control. Second, the processing “cost”, or Stroop effect, is calculated for each participant individually as the proportion increase in time needed to name the color in the incongruent Stroop condition than in the color patch condition. Thus, these possible differences in color naming time are controlled for individual naming times and costs can be compared across participants.

Method

Participants—There were 130 participants consisting of older and younger individuals who were monolingual or bilingual (see Table 1). The younger participants were undergraduate psychology students who received course credit for their participation; the older participants were volunteers from a senior research participant pool who received a small cash gift in appreciation for their participation. All older participants reported themselves to be in good health.

All participants completed the Language and Social Background Questionnaire (LSBQ; Luk & Bialystok, 2013) to assess language and social background. The non-English language of the bilingual participants included at least one of Arabic, Armenian, Cantonese, Creolese, Estonian, Farsi, Finnish, French, German, Hebrew, Italian, Malay, Maltese, Marathi, Punjabi, Russian, Spanish, Tagalog, Telugu, Ukrainian, or Urdu. For the younger bilinguals, 19 (43.2%) reported English to be their first language, with an average age of acquisition for the second language for all younger bilinguals of 5.02 years. For the older bilinguals, 10 (29.4%) reported English to be their first language, and the average age of second language acquisition was 8.04 years. All bilinguals reported high levels of fluency in both languages, regular use of both languages, and an average self-rated bilingualism level out of 5 as 4.2 (*SD* = 0.7) for younger participants and 4.6 (*SD* = 0.8) for older participants.

Tasks

Background measures: Participants were administered the Shipley Vocabulary Test (Zachary, 1986) and the Cattell Culture Fair Test (Cattell, 1957) to assess English receptive vocabulary and nonverbal fluid intelligence, respectively. Cattell raw scores were converted to standardized scores based on a population mean of 100 and a standard deviation of 15.

Stroop task: Participants were given three sheets of paper. The first sheet contained 100 color patches arranged in a 10 × 10 matrix and participants were asked to name the ink color by going across each row and saying the word aloud. The second sheet contained color words printed in black ink in the same arrangement on the page, and participants were asked to read the words aloud. This condition was included as a control to assure that there were no differences in the speed or automaticity with which participants in the two language groups read the word. On the third sheet, the interference condition, the color words were printed in incongruent ink colors and participants were required to name the ink color and ignore the word. Instructions were to complete each sheet as fast as possible without making errors. The time taken to complete each sheet was measured in seconds. Interference costs were expressed as the proportion increase in time to name the color of the font in the interference condition compared to the time taken to name the color patches, that is, (interference time – color naming time)/color naming time.

Results

Background data are presented in Table 1. There were no significant differences between groups on standardized Cattell scores. For Shipley vocabulary results, monolinguals, $F(1, 124) = 6.36$ $p < .02$, and older adults, $F(1, 124) = 134.72$, $p < .0001$, obtained higher scores than bilinguals and younger adults, respectively, with no interaction. These results are consistent with previous studies showing that older adults (Verhaeghen, 2003) and monolinguals (Bialystok & Luk, 2012) typically obtain higher vocabulary scores than their respective counterparts.

There were few errors in the Stroop task, with the mean number of errors ranging from 0.14 to 2.40 out of 100. These data were not examined further. Mean completion times in seconds for each condition are reported in Table 2. A 2-way ANOVA for age and language group was performed on each condition. For color naming, there were main effects of age group, $F(1, 123) = 5.77$, $\eta^2 = .05$, $p < .02$, with faster times for younger adults, and language group, $F(1, 123) = 6.18$, $\eta^2 = .05$, $p < .02$, with faster times for monolinguals, and no interaction, $F(1, 123) = 2.01$, n.s. For word reading, there were no main effects or interactions, all Fs < 1. The absence of any group difference in word reading rules out the possibility that there might be more interference for monolinguals because word reading was more automatic. For the interference condition, there was a main effect of age group, $F(1, 123) = 30.72$, $\eta^2 = .20$, $p < .001$, and an interaction of age group and language group, $F(1, 123) = 5.56$, $\eta^2 = .05$, $p < .02$. Older participants were generally slower, but for the older adults, there was a significant language group difference in which monolinguals were slower than bilinguals, F

(1, 55) = 4.38, $\eta^2 = .03$, $p < .04$, whereas there was no difference between language groups for younger adults, $F(1, 68) = 1.12$, n.s.

Interference costs are also reported in Table 2. Proportional costs for the young adults were 0.80 and 0.71 for monolinguals and bilinguals, respectively; corresponding costs for older adults were 1.16 and 0.85 for monolinguals and bilinguals, respectively. A 2-way ANOVA revealed a main effect of age group, $F(1, 123) = 20.11$, $\eta^2 = .14$, $p < .001$, indicating larger costs for older than younger adults, and a main effect of language group, $F(1, 123) = 11.87$, $\eta^2 = .08$, $p < .01$, indicating larger costs for monolinguals than for bilinguals. The interaction was also significant, $F(1, 123) = 3.78$, $\eta^2 = .03$, $p = .05$, with the difference between the language groups greater for the older adults (0.31) than for the young adults (0.09).

Discussion

The Stroop test is used as a standard assessment of EF and as a neuropsychological measure of frontal lobe functioning (Duncan & Owen, 2000; Stuss et al., 2001).The paper version used in the present study is convenient, especially for older adults, but has the disadvantage of less precise measurement than a computer-based version and lacks the ability to analyze correct response and errors separately. However, as reported above there were relatively few errors and the large differences between conditions shown in Table 2 gives us confidence that the results are valid. The important finding is that in spite of comparable performance on the simple conditions, older bilinguals performed the interference condition faster than older monolinguals, and bilinguals in both age groups demonstrated less interference than monolinguals. Moreover, the interaction of age and language for resolving interference was weak but significant, suggesting additional benefits for older bilinguals. These results support the conclusion of a bilingual advantage in EF even in younger adults and suggest that the difference may be larger in older age. It should be noted that a bilingual advantage was obtained in the Stroop paradigm despite the fact that it requires lexical access and processing. We argue that the bilingual disadvantage in verbal processing was again shown in the present experiment by the bilinguals' significantly slower naming times for color patches, but that the extra time required to name colors of the incongruent Stroop words reflects EF processing. Thus a bilingual advantage was found in interference costs, even although the bilinguals were slower at naming colors when no interference was present.

Study 2

The bilingual advantage in EF reported in Study 1 was based on a standard task used in this literature. The second study extends previous results to investigate whether older and younger bilinguals are also more able to employ EF to resolve interference in working memory and to compare performance on verbal and nonverbal versions of the same task. Verbal and nonverbal conditions of a recent-probe task (in which a short list of items is followed by a recognition-probe item) differed only in the memory items, namely, letters in the verbal task and visual 'stickman' figures in the nonverbal task (see Figure 1). The hypothesis was that both bilingual groups would show smaller interference costs than monolinguals in the figure task but that this advantage would be diminished in the verbal task because of reported disadvantages in verbal processing. Again, larger bilingual advantages were predicted for older adults.

Method

Participants—Participants included 108 individuals drawn from populations similar to those in Study 1 (Table 1). For the younger adults, 15 participants in each language group were formally part of a different study but completed the same working memory task so were included in these analyses. However, for these 30 participants, English vocabulary was assessed by the Peabody Picture Vocabulary Test (PPVT-III; Dunn & Dunn, 1997) rather than the Shipley. Like the Cattell, PPVT provides standard scores from age based on a mean of 100 and a standard deviation of 15. The monolinguals spoke only English. The non-English languages spoken by the bilinguals included Albanian, Arabic, Cambodian, Cantonese, Danish, Dari, Dutch, French, German, Greek, Hebrew, Hindi, Malayalam, Maltese, Polish, Portuguese, Punjabi, Romanian, Russian, Serbian, Singhalese, Spanish, Tagalog, Tamil, Turkish, Urdu, and Vietnamese. For the younger bilinguals, 11 (30.6%) reported English as their first language and their age of acquiring a second language as approximately 4.6 years. For the older bilinguals, 4 (22.2%) reported English to be the first language, with an average age of second language acquisition in the group as 6.6 years. Self-rated proficiency, indicated on scales from “unable to communicate” (0) to “native-like” (100), showed mean proficiency for English was 94 (*SD* = 13) and for the non-English language was 88 (*SD* = 14) for younger bilinguals, and 93 (*SD* = 13) for English and 98 (*SD* = 5) for the non-English language for older bilinguals.

Tasks

Background measures: These were the same as in Study 1 except for the inclusion of the PPVT-III for 30 of the young adults.

Working memory: Verbal (Figure 1A) and nonverbal (Figure 1B) versions of the recent-probe task (Jonides & Nee, 2006) were developed to assess proactive interference in working memory. In the verbal task, the trial began with a central fixation for 1000 ms, followed by the memory set containing 5 letters for 1000 ms. Next was a blank interval of 3000 ms, after which a probe appeared and the participant pressed a key to indicate whether the probe was one of the letters in the preceding memory set. In the nonverbal task, the memory set consisted of four stick figures (from a set of 26) presented on the screen for 2500 ms, followed by a blank interval of 1500 ms and then a probe figure. Pilot testing showed that the stick figures were more difficult to remember than the letters, so the timing parameters were adjusted to make the two tasks comparable in difficulty.

There are four trial types (Figure 1C) created by the combination of the correct response to the probe (yes or no) and whether the probe appeared in the previous trial (n-1): **Positive-baseline**: The probe appears in trial n so requires a “yes” response, and did not appear in trial n-1. **Positive-facilitation**: The probe appears in trial n so requires a “yes” response but also appeared in trial n-1 so potentially there is *facilitation*. **Negative-baseline**: The probe does not appear in trial n so requires a “no” response and did not appear in trial n-1. **Negative-interference**: The probe does not appear in trial n so requires a “no” response but did appear in trial n-1, potentially creating *interference*.

The verbal and nonverbal tasks were presented separately in counterbalanced order, each beginning with instructions and 12 practice trials, both of which were repeated if necessary. Because the figures were unfamiliar, the nonverbal task was preceded by a 3-minute familiarization phase in which participants were shown the 26 figures. Each task consisted of a block of 32 baseline trials (combined positive and negative), two blocks of 64 mixed trials each (16 trials for each of the 4 trial types), and finally another block of 32 baseline trials.

Results

Background measures are reported in Table 1. Two-way ANOVAs for standardized Cattell scores showed no effect of age group, $F < 1$, but a significant effect of language group, $F(1, 104) = 4.88$, $\eta^2 = .04$, $p < .03$, indicating higher scores by monolinguals than bilinguals. Although the interaction was not significant, $F(1, 104) = 2.31$, $p = .13$, it appears that the higher scores are largely in the group of older monolinguals, who were drawn from a volunteer pool containing many retired professional people. For participants who completed the Shipley test, older adults, $F(1, 74) = 31.80$, $\eta^2 = .25$, $p < .001$, and monolinguals, $F(1,74) = 14.13$, $\eta^2 = .12$, $p < .001$, obtained higher scores than their counterparts. There was also an interaction effect, $F(1, 74) = 5.65$, $\eta^2 = .05$, $p < .02$, in that the monolingual vocabulary advantage was larger in the older adult group than in the younger group. The PPVT scores were equivalent for the young monolinguals and bilinguals, $F(1, 28) = 1.35$, n.s.

Mean proportions correct for the letter and figures memory task are reported in Table 3. For simplicity, results from the two tasks were analyzed separately for both accuracy and RT. For the accuracy scores in the letter task, a 4-way mixed ANOVA for age, language, response (yes, no), and trial (baseline, alternative) indicated a main effect of age group, $F(1, 104) = 8.25$, $\eta^2 = .07$, $p < .005$, showing higher accuracy for younger participants, with no main effect of language group, $F(1, 104) = 1.91$, n.s., or age by language group interaction, $F < 1$. There was a significant main effect of response, $F(1, 104) = 39.08$, $\eta^2 = .27$, $p < .0001$, because accuracy was higher for negative trials, and an interaction of response and trial type, $F(1, 104) = 29.43$, $\eta^2 = .22$, $p < .0001$, because baseline trials were more accurate than the alternative (interference) for negative responses but less accurate than the alternative (facilitation) for positive responses.

A similar analysis was conducted on accuracy scores for the figures task. There was again a significant main effect of age, $F(1, 104) = 39.51$, $\eta^2 = .17$, $p < .0001$, with higher accuracy among younger participants than older ones. There was a main effect of response, $F(1, 104) = 57.74$, $\eta^2 = .36$, $p < .0001$, that interacted with each of age group, $F(1, 104) = 13.80$, $\eta^2 = .12$, $p < .0003$, and language group, $F(1, 104) = 4.28$, $\eta^2 = .04$, $p < .04$, although the three-way interaction was not significant, $F(1, 104) = 1.36$, n.s. For age, the difference between groups was in the positive trials, with young adults being more accurate than their counterparts. For language, bilinguals were more accurate than monolinguals for positive trials, but less accurate for negative trials. Overall, participants were more likely to respond “no” than “yes”, indicating that they did not recognize the probe as being part of the stimulus, inflating accuracy scores for the negative trials in which the probe was in fact not part of the stimulus set. Therefore, responses to the positive trials are more informative than

NIH-PA Author Manuscript
NIH-PA Author Manuscript

are responses to negative trials. As with the letter task, there was an interaction of response and trial type, F (1, 104) = 26.78, η^2= .20, $p < .0001$, in that the baseline trials were more accurate than interference trials for negative responses but less accurate than facilitation trials for positive responses.

The mean RTs for correct trials are also reported in Table 3. Trials with RTs shorter than 300 ms or longer than 3000 ms and below or above 1.5 interquartile range were excluded from the analysis. In the letter task, this procedure eliminated 6.8% of positive trials and 7.1% of negative trials for younger adults. For older adults in the letter task, 5.0% of positive trials and 8.3% of negative trials were excluded. In the figure task, 5.4% of positive trials and 5.8% of negative trials were excluded for younger adults, and 4.4% of positive trials and 5.7% of negative trials were excluded for older adults. As with accuracy, data were analyzed separately for each of the letter and figures tasks. For the letters task, a 4-way ANOVA for age, language, response, and trial showed main effects of age, F(1, 104) = 24.88, η^2= .19, $p < .0001$, with faster RTs for younger adults, and no effect of language group, F (1, 104) = 2.72, n.s., or the interaction of age and language, $F < 1$. For the response effects, positive trials were faster than negative trials, F (1, 104) = 15.83, η^2= .13, $p < .0001$, and baseline trials were different from their alternative, F (1, 104) = 23.31, η^2= .18, $p < .0001$, with these factors interacting, F (1, 104) = 22.64, η^2= .18, $p < .0001$. Baseline trials were faster than the alternative (interference) for negative responses but slower than the alternative (facilitation) for positive responses.

The same analysis was conducted on RT to the figures task. There was a main effect of age, F (1, 104) = 47.52, η^2= .32, $p < .0001$, with faster responses by younger participants, and a main effect of language group, F (1, 104) = 5.90, η^2= .05, $p < .01$, with faster responses by bilinguals. There was no interaction of age and language group, F (1, 104) = 2.33, n.s. For the response effects, positive trials were faster than negative trials, F (1, 104) = 17.22, η^2= .14, $p < .0001$. There was no main effect of trial type, F (1, 104) = 2.97, $p = .09$, but there was an interaction between response and trial type, F (1, 104) = 20.27, η^2= .16, $p < .0001$, again because baseline trials were faster than interference trials for "no" responses but slower than facilitation trials for "yes" responses. The difference between baseline and alternative trials interacted with language group, F (1, 104) = 17.29, η^2= .14, $p < .0001$, and entered a 3-way interaction with language group and response, F (1, 104) = 4.66, η^2= .04, $p < .03$, and a 4-way interaction with all the factors, F (1, 104) = 5.62, η^2= .05, $p < .02$. This interaction can be understood by considering the effect of age group and language group individually for each of the four conditions combining response (positive or negative) and trial type (baseline or alternative). All four conditions had significant main effects of age in which younger adults were faster than older adults, but the effect of language group was somewhat different. Bilinguals were faster than monolinguals for the positive baseline condition, F (1, 104) = 6.17, $p < .02$, the positive facilitation condition, F (1, 104) = 11.69, $p < .0009$, and the negative interference conditions, F (1, 104) = 11.13, $p < .001$. For the negative baseline condition, there was no main effect of language group, F (1, 104) = 1.85, n.s., but there was a significant interaction of age group and language group, F (1, 104) = 8.39, $p < .005$, in that the bilinguals were faster than monolinguals in the younger group but slower than monolinguals in the older group.

These effects become clearer by considering facilitation and interference cost scores indicating the difference between the baseline and alternative condition by age group and language group. These scores are shown in Figure 2 for each task and for each of the accuracy and RT data. These graphs show facilitation as the positive difference between baseline and facilitation trials and interference as the negative difference between baseline and interference trials. A bilingual advantage would thus be shown as more facilitation and less interference. Looking first at the patterns for young participants, there is little evidence of a consistent bilingual advantage. In the case of facilitation, two of the differences are in favor of the bilinguals and two favor the monolinguals; for interference, there is only one instance of bilingual advantage (Figure task RT) and two cases of equivalent performance. These language differences are also small. For older participants, however, three of the four tasks show a bilingual advantage for interference, and all four tasks show a bilingual advantage for facilitation. To understand these effects, the data in Figure 2 were examined in four 3-way ANOVAs for age group, language group, and measure (facilitation, interference). The main effect of measure is of little interest given that facilitation scores are necessarily more positive than interference costs, but interactions with this variable may yield meaningful results.

For letter accuracy the only result of interest is a marginally significant effect of language group, F (1, 104) = 3.64, η^2= .03, p < .06, showing an overall tendency for better performance (more facilitation, less interference) in the bilingual group (Figure 2a). For figure task accuracy there is a strong effect of age, F (1, 104) = 10.03, η^2= .09, p = .002, signifying better overall performance for the older group, and a marginally significant age by measure interaction, F (1, 104) = 3.10, η^2= .03, p = .08, indicating that this age-related advantage is stronger in facilitation than in interference (Figure 2c). For letter task RT (Figure 2b), the only effect to approach significance was the age by language group interaction, F (1, 104) = 3.18, η^2= .03, p < .08, which indicates a tendency for the bilingual participants to perform better in the old groups, but for the monolingual participants to do better in the young groups. Finally, in the figure task RT data (Figure 2d) there was a main effect of language group, F (1, 104) = 17.29, η^2= .14, p < .0001, indicating a large advantage for bilingual participants, that was qualified by a marginally significant age by language interaction, F (1, 104) = 3.32, η^2= .03, p = .07, showing that the bilingual advantage was somewhat larger in older participants. A language group by measure interaction, F (1, 104) = 4.66, η^2= .04, p = .03, showed that the bilingual advantage was more evident in smaller interference costs than in larger facilitation scores. Finally, a 3-way interaction of age group, language group, and measure, F (1, 104) = 5.62, η^2= .05, p < .02, indicated that the bilingual advantage was greatest in the older adults interference scores.

It is also illuminating to look at interactions with the two tasks (letters, figures), and these effects were assessed by carrying out four separate 3-way ANOVAs (age x language x task) on Accuracy-Facilitation, Accuracy-Interference, RT-Facilitation and RT-Interference. In summary, the Accuracy-Facilitation analysis found a significant effect of age, F (1, 104) = 16.03, p < .001, showing that older adults displayed more facilitation (Figure 2a, 2c, facilitation). Neither Accuracy-Interference nor RT-Facilitation analyses found any significant effects, but the RT-Interference comparisons found significant effects of language (bilinguals show less interference), F (1, 104) = 16.03, $p < .001$, and significant interactions between language and age (the bilingual advantage is shown by older adults only), F (1,104) = 14.22, $p < .001$, and between language and task (the bilingual advantage is on the figures task), F (1,104) = 9.61, p < .003. The 3-way interaction does not reach significance, F (1,104) = 1.43, p = .23, but Figures 2b and 2d (interference) show that the trend is for the bilingual advantage to be greatest for older adults in the figures task.

Discussion

The raw scores shown in Table 3 show strong effects of age on both accuracy and RT in both letter and figure tasks; in all cases young participants were faster and more accurate. A bilingual advantage was found in only one condition – RT in the figure task, arguably the more complex task paired with the more sensitive index of performance. This language group effect was found across both age groups with no interaction between them.

The more relevant data, however, are illustrated in Figure 2. This figure shows the differences between baseline conditions and facilitation trials (positive trials preceded by a trial containing the current target) and between baseline conditions and interference trials (negative trials preceded by a trial containing the current target). The data thus show the extent to which positive trials were aided by recent exposure to the current target for participants in the four groups and the extent to which the groups were able to resist interference from prior stimuli in negative trials. Effects of aging in these data were found in accuracy measures for the figure task in which older participants showed better performance (more facilitation, less interference) than younger participants, and there was a trend for this effect to be stronger in facilitation than in interference (Figure 2c). Main effects of bilingualism were found as a trend in accuracy measures in the letter task and as a strong effect in RT measures in the figure task. In the latter case the bilingual advantage tended to be larger in the older group, and was stronger for interference than for facilitation. Finally, there was a trend for an interaction between age group and language group in RT measures in the letter task; in older adults, bilinguals showed somewhat more facilitation and less interference than monolinguals, but in younger adults these effects were reversed.

In summary, the results reveal evidence for a bilingual advantage although it was restricted to specific conditions, particularly RT measures in the figure task (Figure 2d). Moreover, the bilingual advantage was found primarily in older participants – for facilitation in all four tasks, and for interference in 3 of the 4 tasks, although again most clearly in RT in the figure task. Thus, as predicted, the bilingual advantage was most evident in the nonverbal task and in older adults.

General Discussion

Bilingual participants outperformed monolinguals on a standard measure of EF (Study 1) and on a working memory task that requires EF to resolve interference from a previously presented stimulus (Study 2). These effects were found both in younger and older groups, with larger bilingual advantages in the older group, and in tasks based on different kinds of stimulus materials. These results provide some answers to the three questions raised in the introduction.

The first issue is the reliability of the bilingual advantage in EF. In Study 1, bilinguals experienced significantly less interference than monolinguals. The Stroop task is difficult and involves a range of EF and attentional components (MacLeod & MacDonald, 2000). In this task there was a reliable advantage for bilinguals even in a sample of young adults, with a greater advantage for bilinguals in the older adult groups. Studies that fail to find a bilingual advantage in EF typically investigate only younger participants and use very simple tasks (e.g., flanker task, switching task) for which reaction times are very fast (generally less than 500 ms) leaving little room for significant group differences to emerge above individual differences in performance. In contrast, tasks that involve slower and more effortful processing are more likely to show group differences, in this case, bilingual advantages.

The second issue is the role of verbal and nonverbal tasks in modulating the appearance of a bilingual advantage. Previous studies have typically reported bilingual advantages in nonverbal tasks but bilingual disadvantages in verbal processing, creating a sort of paradox: the consequences of a linguistic experience are to produce better nonverbal processing but poorer verbal processing. The results of Study 2 confirmed that bilingual advantages are more salient in nonverbal EF tasks. In the present study, there was no disadvantage for bilinguals on the letter task and clear advantages on the figure task. The finding that the reduced costs for bilinguals was greater in the figure task than in the letter task may reflect the fact that the figure task was more difficult (Table 3), but our preferred account is that it is a further example of the interaction between language group (bilingual/monolingual) and task (verbal/nonverbal) shown by Luo et al. (2013). Our argument is that a general enhancement of EF function in bilinguals is offset by their relative disadvantage in verbal processing (Bialystok et al., 2010; Gollan et al., 2005: Poarch & van Hell, 2012a).

Finally, the results of Study 2 provide an important extension of previous research into the crucial area of working memory. Using a version of the recent-probe task, the results showed age-related decrements in performance but no effects of language group in accuracy scores. For response latency, however, the bilingual participants showed reduced interference costs. Additionally, this bilingual advantage was greater in the nonverbal figure task and in older adults. We suggest that the bilingual advantage is typically greater in older adults and in children (see Barac et al., in press for review) because this advantage represents more efficient executive functions, and these EF abilities develop during childhood and adolescence but then decline during older adulthood. We further assume that since EF abilities are at their peak in young adults, they show a ‘functional ceiling’ in the sense that any further efficiencies associated with bilingualism have little effect, especially on relatively simple EF tasks that are performed quickly and accurately. On the other hand, the bilingual advantage has ‘room to emerge’ in children, older adults and on relatively complex tasks. The results of Study 2 show that the advantage is greatest in older adults performing an unfamiliar complex task (RTs on the Figures task).

Together, the two studies provide corroborating evidence for a bilingual advantage in EF and add to that literature by identifying factors associated with the strength of that advantage. Specifically, larger bilingual effects on EF performance are found in older adults than in younger adults, and in nonverbal material than in verbal material, with interactions between these factors. Moreover, Study 2 demonstrated these effects in a difficult working memory task, showing the breadth of these processing differences. EF is a complex set of processes, and bilingualism is a complex life experience, so it is not surprising that the relation between them is also complex. The present results shed light on those interactions.

Acknowledgments

This research was supported by grant R01HD052523 from the US National Institutes of Health and grant A2559 from the Natural Sciences and Engineering Council of Canada to EB, grant MOP57842 from the Canadian Institutes of Health Research to EB and FIMC, and grant A8261 from the Natural Sciences and Engineering Council of Canada to FIMC. We are grateful for the assistance of Sylvain Moreno in Study 2.

References

Abutalebi J, Green DW. Bilingual language production: The neurocognition of language representation and control. Journal of Neurolinguistics. 2007; 20:242–275.

Baddeley A. Working memory: Looking back and looking forward. Nature Reviews Neuroscience. 2003; 4:829–839.

Barac R, Bialystok E, Castro DC, Sanchez M. The cognitive development of young dual language learners: A critical review. Early Childhood Research Quarterly. in press.

Bialystok E. Effect of bilingualism and computer video game experience on the Simon task. Canadian Journal of Experimental Psychology. 2006; 60:68–79. [PubMed: 16615719]

Bialystok E, Craik FIM. Cognitive and linguistic processing in the bilingual mind. Current Directions in Psychological Science. 2010; 19:19–23.

Bialystok E, Luk G. Receptive vocabulary differences in monolingual and bilingual adults. Bilingualism: Language and Cognition. 2012; 15:397–401.

Bialystok E, Craik FIM, Klein R, Viswanathan M. Bilingualism, aging, and cognitive control: Evidence from the Simon task. Psychology and Aging. 2004; 19:290–303. [PubMed: 15222822]

Bialystok E, Craik FIM, Luk G. Cognitive control and lexical access in younger and older bilinguals. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2008a; 34:859–873.

Bialystok E, Craik FIM, Luk G. Lexical access in bilinguals: Effects of vocabulary size and executive control. Journal of Neurolinguistics. 2008b; 21:522–538.

Bialystok E, Luk G, Peets KF, Yang S. Receptive vocabulary differences in monolingual and bilingual children. Bilingualism: Language and Cognition. 2010; 13:525–531.

Blumenfeld HK, Marian V. Bilingualism influences inhibitory control in auditory comprehension. Cognition. 2011; 118:245–257. [PubMed: 21159332]

Cattell, RB. Culture fair intelligence test, a measure of “g”. Savoy, IL: Western Psychological Services; 1957.

Costa A, Hernández M, Sebastián-Gallés N. Bilingualism aids conflict resolution: Evidence from the ANT task. Cognition. 2008; 106:59–86. [PubMed: 17275801]

Duncan J, Owen AM. Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends in Neuroscience. 2000; 23:475–483.

Dunn, LM.; Dunn, LM. Peabody Picture Vocabulary Test–III. Circle Pines, MN: American Guidance Service; 1997.

Gold BT, Kim C, Johnson NF, Kriscio RJ, Smith CD. Lifelong bilingualism maintains neural efficiency for cognitive control in aging. Journal of Neuroscience. 2013; 33:387–396. [PubMed: 23303919]

Gollan TH, Montoya RI, Fennema-Notestine C, Morris SK. Bilingualism affects picture naming but not picture classification. Memory & Cognition. 2005; 33:1220–1234. [PubMed: 16532855]

Gollan TH, Montoya RI, Werner GA. Semantic and letter fluency in Spanish-English bilinguals. Neuropsychology. 2002; 16:562–576. [PubMed: 12382994]

Green DW. Mental control of the bilingual lexico-semantic system. Bilingualism: Language and Cognition. 1998; 1:67–81.

Grogan A, Green DW, Ali N, Crinion JT, Price C. Structural correlates of semantic and phonemic fluency ability in first and second languages. Cerebral Cortex. 2009; 19:2690–2698. [PubMed: 19293396]

Jonides J, Nee DE. Brain mechanisms of proactive interference in working memory. Neuroscience. 2006; 139:181–193. [PubMed: 16337090]

Kormi-Nouri R, Moradi AR, Moradi S, Akbari-Zardkhaneh S, Zahedian H. The effect of bilingualism on letter and category fluency tasks in primary school children: Advantage or disadvantage? Bilingualism: Language and Cognition. 2012; 15:351–364.

Kroll JF, Bialystok E. Understanding the consequences of bilingualism for language processing and cognition. Journal of Cognitive Psychology. 201310.1080/20445911.2013.799170

Kroll, JF.; Dussias, PE.; Bogulski, CA.; Valdes-Kroff, J. Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In: Ross, B., editor. The Psychology of Learning and Motivation. Vol. 56. San Diego: Academic Press; 2012. p. 229-262.

La Heij, W. Selection processes in monolingual and bilingual lexical access. In: Kroll, JF.; de Groot, AMB., editors. Handbook of bilingualism: Psycholinguistic approaches. New York: Oxford University Press; 2005. p. 289-307.

Luk G, Bialystok E. Bilingualism is not a categorical variable: Interaction between language proficiency and usage. Journal of Cognitive Psychology. 201310.1080/20445911.2013.795574

Luo L, Craik FIM, Moreno S, Bialystok E. Bilingualism interacts with domain in a working memory task: Evidence from aging. Psychology and Aging. 2013; 28:28–34. [PubMed: 23276212]

Luo L, Luk G, Bialystok E. Effect of language proficiency and executive control on verbal fluency performance in bilinguals. Cognition. 2010; 114:29–41. [PubMed: 19793584]

MacLeod CM, MacDonald PA. Interdimensional interference in the Stroop effect: uncovering the cognitive and neural anatomy of attention. Trends in Cognitive Sciences. 2000; 4:383–391. [PubMed: 11025281]

Michael, E.; Gollan, TH. Being and becoming bilingual: Individual differences and consequences for language production. In: Kroll, JF.; de Groot, AMB., editors. Handbook of bilingualism: Psycholinguistic approaches. New York: Oxford University Press; 2005. p. 389-407.

Morford JP, Wilkinson E, Villwock A, Piñar P, Kroll JF. When deaf signers read English: Do written words activate their sign translations? Cognition. 2011; 118:286–292. [PubMed: 21145047]

Paap KR, Greenberg ZI. There is no coherent evidence for a bilingual advantage in executive processing. Cognitive Psychology. 2013; 66:232–258. [PubMed: 23370226]

Pelham SD, Abrams L. Cognitive advantages and disadvantages in early and late bilinguals. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2013 Dec 2. Advance online publication. 10.1037/a0035224

Poarch GJ, Van Hell JG. Cross-language activation in children's speech production: Evidence from second language learners, bilinguals, and trilinguals. Journal of Experimental Child Psychology. 2012a; 111:419–438. [PubMed: 22138311]

Poarch GJ, Van Hell JG. Executive functions and inhibitory control in multilingual children: Evidence from second language learners, bilinguals, and trilinguals. Journal of Experimental Child Psychology. 2012b; 113:535–551. [PubMed: 22892367]

Portocarrero JS, Burright RG, Donovick PJ. Vocabulary and verbal fluency of bilingual and monolingual college students. Archives of Clinical Neuropsychology. 2007; 22:415–422. [PubMed: 17336036]

Rosselli M, Ardila A, Araujo K, Weekes VA, Caracciolo V, Padilla M, Ostrosky-Solis F. Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. Applied Neuropsychology. 2000; 7:17–24. [PubMed: 10800624]

Sandoval TC, Gollan TH, Ferreira VS, Salmon DP. What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. Bilingualism: Language & Cognition. 2010; 13:231–252.

Stuss DT, Flodden D, Alexander MP, Levine B, Katz D. Stroop performance in focal lesion patients: Dissociation of processes and frontal lobe lesion location. Neuropsychologia. 2001; 39:771–786. [PubMed: 11369401]

Thierry G, Wu YJ. Brain potentials reveal unconscious translation during foreign-language comprehension. Proceedings of the National Academy of Sciences. 2007; 104:12530–12535.

Verhaeghen P. Aging and vocabulary score: A meta-analysis. Psychology and Aging. 2003; 18:332–339. [PubMed: 12825780]

Zachary, RA. Shipley Institute of Living Scale – Revised. Los Angeles, CA: Western Psychological Services; 1986.

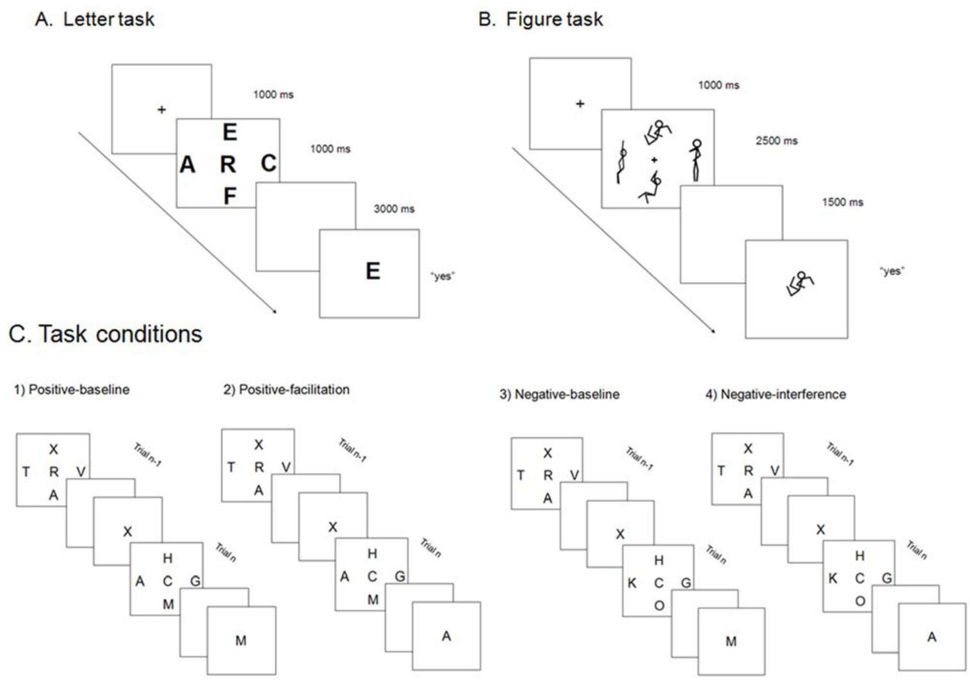

Figure 1.

Schematic illustration of the working memory interference tasks. A) A sample trial in the letter task; B) a sample trial in the figure task; and C) the four conditions in the tasks.

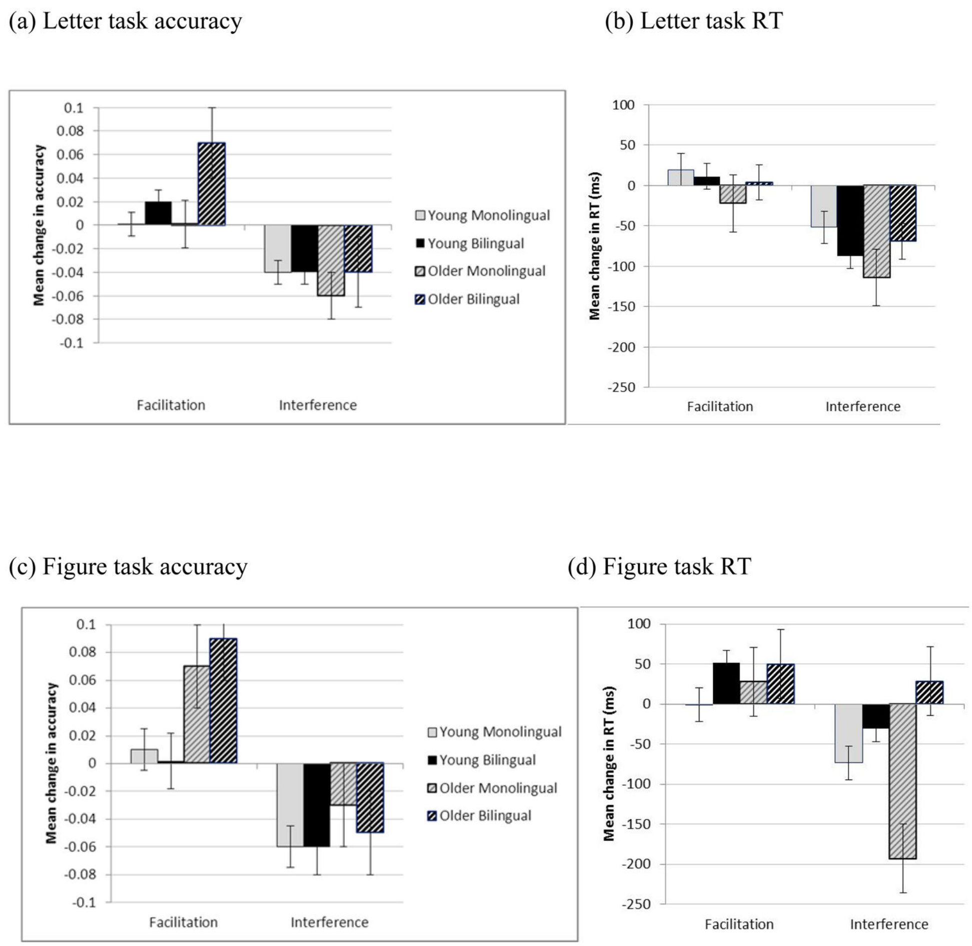

Figure 2.

Facilitation and interference scores for accuracy and reaction time in each of the letter and number tasks by age group and language group.

Table 1

Mean scores (standard deviations) of participant background measures by age group and language group for Studies 1 and 2.

	Younger adults		Older adults	
	Monolingual	**Bilingual**	**Monolingual**	**Bilingual**
STUDY 1				
N	27	44	25	34
Age	20.3 (2.5)	20.4 (2.2)	71.3 (5.1)	67.6 (4.5)
Cattell Std	106.2 (17.5)	105.6 (14.4)	113.0 (15.9)	107.9 (14.6)
Shipley	28.0 (3.7)	26.6 (4.1)	36.2 (2.7)	34.0 (4.3)
STUDY 2				
N	36	36	18	18
Age	21.4 (3.2)	20.2 (2.6)	72.4 (4.9)	69.1 (4.2)
Cattell Std	109.7 (13.5)	106.6 (15.7)	115.8 (11.0)	103.6 (16.8)
Shipley[1]	30.5 (2.8)	29.3 (3.0)	36.7 (1.9)	31.8 (5.1)
PPVT[2]	104.3 (8.2)	99.8 (12.4)	N/A	N/A

[1]For Young Monolinguals n = 21; for Young Bilinguals n = 21

[2]For Young Monolinguals n = 15; for Young Bilinguals n = 15

Cattell Std = Age-standardized scores on the Cattell culture fair intelligence test

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Mean reaction time in seconds (and standard deviation) for each condition of the Stroop task and proportion increase for Interference condition by age group and language group in Study 1.

	Younger adults		Older adults	
	Monolingual	**Bilingual**	**Monolingual**	**Bilingual**
Measure	**M (SD)**	**M (SD)**	**M (SD)**	**M (SD)**
Color Naming	58.5 (10.8)	67.3 (13.3)	68.2 (13.6)	70.4 (12.5)
Word Reading	48.7 (11.9)	47.3 (8.2)	47.6 (7.4)	46.7 (9.9)
Color Interference	106.0 (29.1)	112.3 (20.2)	143.6 (32.4)	129.0 (30.0)
Proportion Cost	0.80 (.29)	0.71 (.28)	1.16 (.34)	0.85 (.37)

Table 3

Mean accuracy rates as proportion correct (and standard deviation) and reaction times (standard deviation) for the letter and figure tasks by age group and language group in Study 2.

		Younger adults		Older adults	
	Condition	**Monolinguals**	**Bilinguals**	**Monolinguals**	**Bilinguals**
LETTER TASK ACCURACY					
Positive	Baseline	.87 (.11)	.83 (.13)	.81 (.14)	.77 (.14)
	Facilitation	.87 (.11)	.85 (.11)	.81 (.17)	.84 (.09)
Negative	Baseline	.95 (.06)	.95 (.06)	.95 (.06)	.88 (.12)
	Interference	.90 (.09)	.91 (.06)	.88 (.09)	.83 (.17)
LETTER TASK RT					
Positive	Baseline	937 (217)	828 (173)	1125 (286)	1046 (239)
	Facilitation	917 (213)	817 (164)	1147 (335)	1042 (270)
Negative	Baseline	945 (218)	841 (174)	1098 (226)	1113 (335)
	Interference	997 (230)	929 (219)	1212 (280)	1182 (312)
FIGURE TASK ACCURACY					
Positive	Baseline	.71 (.15)	.75 (.14)	.47 (.18)	.53 (.18)
	Facilitation	.72 (.17)	.74 (.16)	.54 (.22)	.62 (.15)
Negative	Baseline	.84 (.10)	.83 (.08)	.82 (.14)	.77 (.13)
	Interference	.78 (.10)	.78 (.13)	.79 (.18)	.71 (.15)
FIGURE TASK RT					
Positive	Baseline	1087 (208)	960 (180)	1332 (290)	1276 (216)
	Facilitation	1088 (198)	909 (159)	1304 (307)	1226 (270)
Negative	Baseline	1128 (210)	994 (171)	1257 (226)	1365 (232)
	Interference	1202 (224)	1025 (190)	1450 (305)	1336 (302)