Executive control processes in verbal and nonverbal working memory

The role of aging and bilingualism

Margot D. Sullivan, Yolanda Prescott, Devora Goldberg and Ellen Bialystok
York University

Studies across the lifespan have revealed modifications in executive control (EC) from bilingualism, but studies of working memory (WM), a key aspect of EC, have produced varied results. Healthy older (M = 71.0 years) and younger participants (M = 21.1 years) who were monolingual or bilingual, performed working memory tasks that varied in their demands for EC. Tasks included a star counting task, a flanker task, and a nonverbal recent probe memory task. Bilinguals performed similarly to monolinguals on the star counting task after controlling for differences in vocabulary. Monolinguals were faster than bilinguals on the flanker task with only age group differences significant for the WM manipulation. Bilinguals were faster than monolinguals on the nonverbal recent probe memory task, particularly for the condition that included proactive interference. The interpretation is that better bilingual performance in nonverbal working memory tasks is linked to the need for executive control.

Keywords: bilingualism, aging, working memory, executive control, interference

1. Introduction

The importance of attention in working memory (WM) is clear when one considers the cognitive operations needed to successfully encode, avoid irrelevant information/interference, maintain as well as retrieve information (Eriksson, Vogel, Lansner, Bergström, & Nyberg, 2015). This becomes relevant for how bilingualism impacts WM performance as there is much evidence for the parallel activation of bilinguals' two languages (for a recent review see Kroll, Dussias, Bolgulski, & Valdes-Kroff, 2012), and it has been put forth that bilinguals attend to both

Linguistic Approaches to Bilingualism 6:1/2 (2016), 147–170. DOI 10.1075/lab.15056.sul
ISSN 1879–9264 / E-ISSN 1879–9272

languages and therefore require the use of executive control (EC) to avoid interference (Bialystok, 2015). However, in contrast to a substantial amount of research showing that bilinguals outperform monolinguals on many nonverbal EC tasks, with the largest effects seen in older adults and children (Bialystok, Craik, Green, & Gollan, 2009), studies examining the effect of bilingualism on WM have produced more varied results (e.g., Bialystok, Craik, & Luk, 2008; Luo, Craik, Moreno, & Bialystok, 2013). The purpose of the present study was to investigate the effect of bilingualism on working memory by considering the level of executive control required, presence or absence of interference, and verbal demands on the WM task and the possible differences in these effects that might be found for younger and older adults. As Adesope, Lavin, Thompson, and Ungerleider (2010) point out, dual language activation in bilinguals could either increase cognitive load, making WM processing more cumbersome (cf., van Merrienboer & Sweller, 2005) or the intense practice with attentional control that is required for bilingual language production may benefit WM processing through practice (Luo, Craik, Moreno, & Bialystok, 2013). In both cases, these outcomes implicate the broader notion of EC of which WM is a component. Continual and lifelong experience with dual language control and, crucially, managing interference between languages suggests that better WM performance by bilinguals should be linked to the need for EC (Kroll et al., 2012); however such an effect may have been obscured in previous work due to either low level EC requirements of the task and/or the impact of verbal processing in the tasks due to the type of materials used.

Verbal memory performance in bilinguals is often confounded with language proficiency because receptive vocabulary is generally lower in both bilingual children (Bialystok, Luk, Peets, & Yang, 2010) and adults (Bialystok & Luk, 2012) than it is for their monolingual peers. Bilinguals across the lifespan also show reduced lexical access and slower retrieval in production tasks, including more tip of the tongue states (Gollan & Silverberg, 2001), slower picture naming in the L1 especially for low frequency words (Gollan, Montoya, Cera, & Sandoval, 2008; Ivanova & Costa, 2008), and reduced verbal fluency, all of which impacts tests of free recall. For instance, in a study by Fernandes, Craik, Bialystok, and Kreuger (2007), older and younger adult bilinguals and monolinguals performed a free recall test of semantically-related word lists under normal or divided attention conditions. Both bilinguals and older adults recalled fewer words, but the proportion of attention decrement in the divided attention conditions compared to the full attention condition was equivalent to monolinguals and younger adults. These findings suggest that both older adults and bilinguals were not further hindered by interference effects, despite recalling fewer words. Importantly, the effect of bilingualism but not aging disappeared after controlling for vocabulary and nonverbal fluid intelligence, indicating that although both older adults and bilinguals have initial difficulty with lexical retrieval, only age-related detriments on verbal recall remain once language proficiency is accounted for. Further analyses revealed that worse bilingual performance was tied to the full attention condition and only one of the four divided attention conditions, a pattern that the authors interpreted as utilization of a more efficient EC system to improve performance on the effortful conditions making it comparable to that of monolinguals.

In another study with young adults, a similar influence of verbal ability on verbal memory and interference was shown by having bilinguals and monolinguals recall four sequential word lists, with the first three lists containing words from the same semantic category to accumulate proactive interference (i.e., interference from previously relevant material, see Jonides & Nee, 2006) and the fourth list containing words from a separate sematic category to erase proactive interference. There were no language group differences in verbal recall or number of intrusions, but when differences in vocabulary knowledge were controlled, the bilinguals showed better verbal recall than monolinguals (Bialystok & Feng, 2009). The essential point from both of these studies is that bilinguals perform equivalently to or even better than monolinguals after verbal ability is considered, and importantly better WM performance is seen in bilinguals for the conditions involving controlled processing and interference.

To avoid the potential confound of language proficiency on WM performance, Bialystok, Craik, & Luk (2008) asked younger and older bilingual and monolingual adults to perform nonverbal span tasks -- forward and backward Corsi blocks and a self-ordered pointing task. Younger bilinguals recalled more than younger monolinguals on the Corsi blocks task, but there were no language group differences in older adults. On the self-ordered pointing task, there were age-related differences in performance but no effect of language group. The lack of language group differences seen with aging for these two tasks may be because they are simple span tasks with little need for EC. The suggestion is that bilinguals outperform monolinguals only on tasks that make substantial demands on EC, and the implication is that WM tasks that depend on EC are more likely to show better performance by bilinguals than by monolinguals. An example of this comes from the Simon task (Lu & Proctor, 1995) that was adapted to create a WM manipulation (Bialystok, Craik, Klein, and Viswanathan, 2004). The task included conditions in which participants responded to the color of either two stimulus options (red or green) or four stimulus options (red, blue, orange, or brown) that were placed on the left or right side of the screen in compatible or incompatible positions with the correct response key. Independent of differences in the Simon effect, both younger and older bilinguals outperformed their monolingual counterparts in the more demanding 4-color conditions. Therefore, bilinguals appear to show better WM performance on tasks involving nonverbal, speeded responses that require the use of EC.

The interaction between domain-specific ability and domain-general executive control on memory performance can be shown by testing monolinguals and bilinguals on the same task using different stimuli. Luo, Craik, Moreno, and Bialystok (2013) gave younger and older bilinguals and monolinguals simple and complex verbal and spatial span tasks. An interaction between language group and domain showed that monolinguals were better than bilinguals on verbal tasks and bilinguals were better than monolinguals on spatial tasks, differences that were stable across age groups. These effects remained after controlling for vocabulary and nonverbal intelligence. Although these data are in line with the current predictions, language differences are not always seen for span tasks (e.g. Bialystok, Craik, & Luk, 2008), suggesting that the EC requirements in these tasks may not be sufficient to elicit such differences.

Two additional studies have used the same task across different domains with older and younger adult bilinguals but have focused more on isolating controlled processing components in memory performance. First, Wodniecka, Craik, Luo, and Bialystok (2010) used a process dissociation paradigm (PDP; Jacoby, 1991) that was designed to assess the effects of aging and bilingualism on measures of familiarity and recollection. Familiarity is an automatic process but recollection requires controlled processing, or EC. The typical age-related declines were found for recollection, but language group effects depended on both domain and pre-existing ability. Older adult bilinguals for whom English was an L2 displayed worse performance for verbal recollection but better performance for nonverbal recollection compared to older adult monolinguals. In contrast, bilinguals with stronger English proficiency showed the reversed effect in which older adult bilinguals displayed better performance in verbal recollection. Bilinguals of both age groups outperformed monolinguals for nonverbal recollection. These studies provide some evidence for better bilingual performance in the controlled processes of recollection but not automatic familiarity, with clearer differences for older adult bilinguals and nonverbal tasks.

Finally, evidence that language group differences in nonverbal memory may be tied to the need to resolve interference comes from the recent probe task (Jonides & Nee, 2006). Participants must make yes/no decisions to a single probe as to whether it appeared in the previous memory set slide. The key manipulations are trials designed to elicit facilitation, i.e., the probe was present in the memory set and trial n-1, or interference effects, i.e., the probe was not present in the memory set but was present for trial n-1. This proactive interference condition was of interest for the current study for two reasons. First, it is well documented that older adults are more susceptible to proactive interference effects (e.g., Jonides, Marshuetz, Smith, Reuter-Lorenz, & Koeppe, 2000; May, Hasher, & Kane, 1999). Second, the presence of proactive interference introduces a situation in which previous information that is highly familiar creates a response bias to respond 'yes' when in reality, a negative response is required (Jonides et al., 2000); creating the need for conflict resolution. This presents a scenario in WM that is similar to the nature of the bilingual experience where conflict resolution/attention is required to deal with the interference from two separate languages. With respect to the PDP previously described, the probe in the interference condition elicits a sense of familiarity in memory, however recollection/controlled processing is required to establish the appropriate contextual details as to whether the probe was indeed in the current memory set (See Wodniecka et al., 2010). In a study by Bialystok, Poarch, Luo, & Craik (2014), younger and older bilingual and monolingual adults performed a letter version and a nonverbal stick figure version of the recent probe task. Younger adults performed better than older adults for both versions, but there were no effects of language group for the letter task. The key findings were that for the nonverbal figure task, bilinguals were faster than monolinguals on negative interference trials and showed smaller costs with greatest effects in older adults.

The current study was designed to determine the conditions under which bilingual processing differences could be found in WM tasks. The hypothesis was that better bilingual performance is tied to the need for EC in performing the task, with the largest EC demands recruited by interference. Older and younger adult participants of monolingual or bilingual language backgrounds were included to assess whether bilingualism would also act as a protective factor against typical age-related declines in memory performance. Domain-specific effects were assessed by including verbal and nonverbal materials.

In the verbal domain, participants performed a star counting task. The EC demands consisted of manipulating the required number of switches from forward to backward counting of stars that were presented in rows on a card. WM demands were increased in a condition in which a more effortful counting rule had to be remembered, namely, count forward by 2s and backward by 1s in contrast to the standard condition in which all counting was by 1s. The star counting task was developed to assess attention regulation and shows a stronger correlation with the processing aspects of numerical span than storage (Das-Smaal, de Jong, & Koopmans, 1993). Bilingual participants were allowed to use their preferred language of counting to limit possible slowing due to lexical retrieval.

In the nonverbal domain, a flanker task was developed that required responses to the correct or opposite central arrow direction depending on the color. This manipulation added WM demands to a well-established EC task, as opposed to the other two tasks (i.e., star counting and recent probe) where EC conditions were present within primarily WM tasks. Finally, a complex nonverbal recent probe memory task was included that contained a proactive interference manipulation and used different nonverbal stimuli from Bialystok et al. (2014). The first hypothesis was that complex task conditions that require more EC, particularly interference resolution, will be performed better by younger participants and bilinguals, with larger language group effects in old age. The second hypothesis was that performance will depend on whether the task is verbal or nonverbal. Better performance by bilinguals should be tied to nonverbal tasks, and once vocabulary knowledge is controlled for, bilinguals' ability to handle EC requirements in the verbal task should outweigh problems due to lexical retrieval.

2. Method

2.1 Participants

There were 115 participants, consisting of older and younger adults with monolingual or bilingual language backgrounds. Thirteen participants were excluded in total for either: having unclear language backgrounds ($n = 6$), low English vocabulary or nonverbal intelligence (standardized scores < 70; $n = 3$), history of a lobotomy/cerebral palsy ($n = 1$), full vision in only one eye ($n = 1$), less than high school education ($n = 1$), and one older monolingual for having very low performance across all three experimental tasks (star counting accuracy = 6.3%, modified flanker accuracy = 33%, recent probe accuracy = 53%).The final younger adult sample included 53 participants between the ages of 18 and 38 ($M = 21.1$, $SD = 4.1$), of whom 26 were monolingual English speakers and 27 were bilinguals who reported being fluent in English and at least one other language.[1] The final older adult sample consisted of 49 participants between the ages of 63 and 80 ($M = 71.0$, $SD = 4.9$), of whom 23 were English monolinguals and 26 were bilinguals.[2]

2.2 Background measures

Participants completed the Language and Social Background Questionnaire (LSBQ; Luk & Bialystok, 2013) to obtain information about language experience.

1. The non-English language of the younger adult bilinguals included Cantonese (3), Portuguese (1), Ilocano (2), Armenian (1), Bisaya (1), French (3), Hindi (2), Farsi (1), Vietnamese (1), Ukrainian (1), Punjabi (1), Bangla (1), Amharic (1), Spanish (2), Korean (1), Albanian (1), Urdu (2), Pashto (1) and Creole (1).

2. The non-English languages of the older adult bilinguals included Bengali (1), French (4), Spanish (1), Yiddish (3), Swiss German (1), Turkish (1), Filipino (2), Estonian (1), Marathi (1), Russian (1), Dutch (1), German (2), Mandarin Chinese (Fookien Dialect) (1), Hindi (1), Tamil (1), Italian (1), Hebrew (1), Ukrainian (1), and Urdu (1).

Participants answered questions about their language use and proficiency for all known languages and rated their level of bilingualism on a global self-assessment scale. Additionally, they answered questions regarding age, gender, handedness, vision/hearing problems, neurological impairments, psychoactive medication use, education level, and country of birth. Older adults also reported their occupation, and younger adults answered questions about their parents' education levels, occupations, and known languages.

Vocabulary and nonverbal fluid intelligence were assessed by the paper-based versions of the Shipley-2 Institute of Living Scale Verbal and Blocks (Shipley, Gruber, Martin, & Klein, 2009). Responses were scored and standardized according to the published instructions. Each test has a population mean of 100 and a standard deviation of 15.

2.3 Tasks

Paper-and-pencil and computer-based tasks were used to assess verbal (star counting) and nonverbal (flanker, recent probe) working memory.

2.3.1 *Star counting task*

The star counting task required participants to follow specific rules and count out loud the number of stars that appeared on a page (adapted from Das-Smaal, de Jong, & Koopmans, 1993). Laminated 8-1/2 x 11 inch sheets with arrangements of black stars and interspersed plus and minus signs at unpredictable locations were presented. The signs indicated the counting direction, with plus signalling count forward and minus signalling count backward. Each sheet had a number in the upper left corner beside the first row of stars indicating the number from which counting was to begin. Participants were to move across the rows from left to right, and to proceed down the rows from the top to bottom of the card. In the standard condition, both forward and backward counting proceeded by intervals of one; in the working memory condition, forward counting proceeded by twos but backward counting by ones. This counting rule (2 Forward, 1 Backward) was printed in red on the left side of each card in the working memory condition (See Appendix A for a sample card). Both the standard and working memory conditions were further divided into low switch and high switch conditions. Each low switch card had four signs indicating a change in counting direction, and each high switch card contained ten signs. There were sixteen cards, with four in each of the four conditions. The order of conditions and the order of cards within each condition were randomized across participants.

2.3.2 *Modified flanker task*

To increase working memory demands in a simple executive control task, a modified flanker task was developed. Stimuli were presented in the center of the screen with a response key on either side of the screen. Each trial began with a central fixation cross for 250 ms, followed by a response screen (see Appendix B). Trials timed out after 2000 ms. For baseline trials, a single blue or pink arrow appeared and participants indicated the direction it was pointing if the arrow was blue (same condition) but the opposite direction if it was pink (opposite condition). Conflict block trials consisted of five arrows presented in a horizontal line across the centre of the screen consisting of two flanking black arrows on either side of a central blue or pink arrow. The flanking arrows pointed in the same direction as the centre arrow for congruent trials, but in the opposite direction for incongruent trials. For both congruent and incongruent trials, central blue and pink arrows indicated same and opposite conditions respectively. Participants completed two blocks of each trial type (single and conflict) in alternating order. There were a total of 48 single arrow trials, consisting of 24 same trials and 24 opposite trials, and 96 conflict block trials, consisting of 48 congruent trials (24 same and 24 opposite) and 48 incongruent trials (24 same and 24 opposite). Colors assigned for the same and opposite trials were counterbalanced across participants.

2.3.3 *Recent probe task*

This task examines the effect of proactive interference on the ability to perform a simple memory task. The stimuli were 26 Microsoft Word Wingdings symbols (e.g., ♈, ♉, ⌘). Trials began with a central fixation cross presented for 1000 ms, followed by a memory set for 2500 ms (see Appendix C). The memory set then disappeared and the fixation cross remained on the screen for 1500 ms until the probe appeared. The probe slide timed out after 3000 ms. Each memory set contained four symbols arranged in a square surrounding the fixation cross. Probe screens contained a single symbol in the centre of the screen. There were four trial types created by two factors. The first was whether or not the probe appeared in the memory set, creating positive ('yes') and negative ('no') trials. Second, the probe may also have appeared in the previous (n-1) set, creating facilitation for positive trials but interference for negative trials. Positive and negative baseline trials were those in which the probe did not appear in the previous memory set. The task consisted of a pure block of 32 trials (16 positive baseline and 16 negative baseline), two mixed blocks with 64 trials each (16 of each of the four trial types), and another pure block of 32 trials (16 positive baseline and 16 negative baseline). The task was programmed using a pseudorandomized order, such that no more than three of the same trial type would be presented in sequence.

2.4 Procedure

Participants completed all tasks within a single 2-hour session. Upon arrival, participants completed the consent form and the LSBQ. The experimenter then administered the Shipley-2 Verbal and Shipley-2 Blocks according to standardized instructions.

For the star counting task, the experimenter sat at a table across from the participant and presented each card on the table individually. Bilingual participants were told they were allowed to count in their preferred language. A practice card was presented before the first experimental card to familiarize participants with the counting rules. For each card, the experimenter recorded the participant's final time in seconds and answer.

The modified flanker and recent probe tasks were completed on a Dell Dimension 8400 desktop using E-Prime (Version 2.0, Psychology Software Tools) software. The order of the two tasks was counterbalanced across participants. For the modified flanker task, the 'Q' key on the left side of the keyboard was used as the left response key and the 'P' key on the right side of the keyboard was the right response key. Participants were instructed to press the button on the same side as the arrow was pointing in for blue arrows (same) and on the opposite side for pink arrows (opposite). Participants completed eight practice trials with feedback before each block, and were told to respond as quickly as possible while avoiding errors. For the recent probe task, participants completed eight practice trials with verbal feedback, and were told to respond as quickly as possible while avoiding errors.

3. Results

3.1 Background measures

Background measures are reported in Table 1 by age group and language group. Two-way ANOVAs for age group and language group were run on the variables age in years, years of education, English vocabulary scores, and nonverbal intelligence scores. For age in years, there was an expected main effect of age group, $F\,(1, 98) = 3163.31, p < .0001, \eta^2_p = .97$, but importantly, no main effect or interactions with language group, *F*s < 1.5. For years of education, there was a main effect of age group, $F\,(1, 98) = 51.23, p < .0001, \eta^2_p = .34$, with older adults having more years of education than younger adults, but no main effect or interactions with language group, *F*s < 2.3. English vocabulary scores were significantly higher in older adults than younger adults, $F\,(1, 98) = 31.43, p < .0001, \eta^2_p = .24$, and higher in monolinguals than in bilinguals, $F(1, 98) = 9.36$, $p = .003$, $\eta^2_p = .09$, with no significant interaction, $F < 1$. There were no significant main effects or interactions for nonverbal intelligence scores *F*s < 3.9. Language profiles are shown in Table 2.

Table 1. Means (and *SDs*) of background variables by age group and language group

Age	Younger Adults		Older Adults	
Group	**Monolinguals**	**Bilinguals**	**Monolinguals**	**Bilinguals**
N	26	27	23	26
Age (years)	21.2 (4.3)	21.0 (4.0)	70.0 (4.4)	71.9 (5.1)
Education (years)	13.0 (1.1)	13.3 (1.5)	15.1 (1.8)	15.8 (2.1)
Vocabulary	101.5 (7.9)	95.1 (11.3)	111.2 (6.8)	106.3 (10.3)
Nonverbal Intelligence	100.7 (9.6)	97.6 (13.9)	101.7 (14.1)	95.3 (10.7)

Note. Vocabulary and nonverbal intelligence were measured using the Shipley-2 Verbal and Shipley-2 Blocks respectively, which are standardized by age group.

Table 2. Language profile means (and *SDs*) by language group and age group

	Bilingual		Monolingual	
	Younger	**Older**	**Younger**	**Older**
Age Learned English	6.2 (5.3)	8.8 (6.0)	1.1 (1.2)	0.9 (0.9)
Age Learned Non-English Language	2.0 (3.0)	1.9 (3.7)	8.1 (5.3) n = 11	11.6 (5.5) n = 7
N Indicated English as L1[a]	13/27	15/26	26/26	23/23
English Proficiency	91.6 (10.2)	97.3 (4.4)	99.9 (0.5)	99.2 (2.5)
Non-English Proficiency	91.9 (10.2)	90.7 (12.5)	8.0 (12.0)	6.0 (10.6)
English Usage	60.4 (19.8)	70.8 (22.8)	98.6 (4.2)	99.3 (2.2)

a *L1*. Language listed first when asked to list known languages in order of fluency.
Note. Self-report ratings of proficiency range from 0 = "no proficiency" and 100 = "fully fluent" and usage from 0 = "All L2" and 100 = "All English".

3.2 Star counting task

Mean counting time and accuracy rates for the star counting task are presented in Table 3. RT analyses were conducted on correct trials only. Any trials noted by the experimenter where participants restarted counting, made a counting infringement (i.e., did not count each individual star), made a substantially long pause to correct themselves,[3] or stopped before completing the entire card were classified

3. This occurrence was noted by the experimenter for a single card for two older adult bilinguals; with the trial length containing the pause to be substantially longer than the mean of

Table 3. Mean counting times in seconds (and *SDs*), LS Means (and *SDs*) using English vocabulary as a covariate, and mean accuracy (and *SDs*) for the star counting task by age group and language group

	Younger Adults		Older Adults	
Condition	**Monolinguals**	**Bilinguals**	**Monolinguals**	**Bilinguals**
		Counting Time		
Standard				
Low Switch	28.5 (7.6)	31.3 (8.3)	31.4 (7.2)	36.6 (8.8)
High Switch	39.1 (9.9)	45.9 (12.9)	41.6 (9.5)	45.4 (11.3)
WM				
Low Switch	44.0 (12.0)	49.7 (11.0)	43.8 (10.0)	51.2 (15.5)
High Switch	50.8 (13.5)	56.3 (15.8)	53.1 (13.7)	57.1 (15.4)
		LS Means		
Standard				
Low Switch	28.0 (7.7)	29.1 (8.3)	33.5 (8.2)	37.4 (7.7)
High Switch	38.6 (10.6)	43.0 (11.5)	44.3 (11.3)	46.4 (10.7)
WM				
Low Switch	43.5 (12.0)	46.9 (13.1)	46.4 (12.8)	52.2 (12.1)
High Switch	50.0 (13.8)	51.8 (15.0)	57.3 (14.8)	58.6 (13.9)
		Accuracy[a]		
Standard				
Low Switch	3.4 (0.9)	3.3 (0.8)	3.5 (0.8)	3.5 (0.8)
High Switch	3.0 (1.0)	2.8 (0.9)	2.7 (1.1)	3.0 (0.9)
WM				
Low Switch	3.0 (0.8)	2.8 (1.1)	3.0 (0.9)	3.0 (1.1)
High Switch	2.6 (1.1)	2.7 (1.1)	2.5 (0.8)	2.7 (1.1)

a Accuracy rates are based on the number of cards performed correctly out of 4.

as errors. One younger bilingual did not perform the task. Initially, a four-way mixed ANOVA with age group and language group as the between-subjects variables, and switch condition (low or high) and counting rule (standard or WM) as within-subjects variables was used to analyze the counting time data. A significant main effect of language group was found, $F\ (1, 91) = 5.43$, $p = .02$, $\eta^2_p = .06$, such that bilinguals had slower counting times than monolinguals. A main effect of

the remaining three cards of that condition (12.2 seconds longer for one participant and 26.8 seconds longer for the other participant).

switch, showed that participants were faster in low switch than high switch conditions, $F(1, 91) = 253.48$, $p < .0001$, $\eta^2_p = .74$, and a main effect of counting rule indicated that participants were faster in standard conditions than in WM conditions, $F(1, 91) = 428.80$, $p < .0001$, $\eta^2_p = .82$. There was also, a significant three-way interaction of switch by counting by age, $F(1, 91) = 4.29$, $p = .04$, $\eta^2_p = .05$. Separate univariate analyses by condition, revealed that this interaction was driven by older adults being significantly slower than younger adults on the low switch/standard counting condition ($p = .01$), but not the remaining conditions (*F*s < 1). There were no other significant main effects or interactions, *F*s < 3.5.

Because the groups were not equivalent in vocabulary and due to the verbal nature of the task, a correlational analysis was conducted to determine whether there was a relationship between vocabulary and counting times across the four conditions. There was a significant negative correlation for the full sample,[4] such that higher vocabulary scores were associated with faster production, $r(93) = -.31$, $p = .003$. A four-way mixed ANCOVA with vocabulary as a covariate was used to re-examine the data. Based on these adjusted scores, there was a significant main effect of age, $F(1, 90) = 5.73$, $p = .02$, $\eta^2_p = .06$, with slower counting time by older adults than younger adults, but not language group, $F < 1.9$, and no interaction, $F < 1$. Thus, the slower performance of bilinguals disappeared when vocabulary levels were taken into account. There was again a main effect of switch, $F(1, 90) = 14.76$, $p = .0002$, $\eta^2_p = .14$ and a main effect of counting rule, $F(1, 90) = 15.21$, $p = .0002$, $\eta^2_p = .14$. The three-way interaction of switch by counting by age remained significant, $F(1, 90) = 6.23$, $p = .01$, $\eta^2_p = .06$, however now older adults were significantly slower for both the low switch/standard condition ($p = .0002$) and the high switch/WM condition ($p = .03$), and not the remaining two conditions (*F*s < 3.5). There were no other significant interactions, *F*s < 3.9.

Analyses of the star counting task accuracy data (See Table 3) using a four-way mixed ANOVA revealed significant task effects, with higher accuracy on low switch than high switch conditions, $F(1, 96) = 30.68$, $p < .0001$, $\eta^2_p = .24$ and on standard than WM conditions, $F(1, 96) = 15.08$, $p = .0002$, $\eta^2_p = .14$. There were no other significant main effects or interactions, *F*s < 1.8.

4. To justify use of the ANCOVA, the assumption of homogeneity of the regression slope was examined by also looking at the correlations between vocabulary and production time separately by age and language group. The assumption was met as similar slopes were shown across all four subgroups, with all correlations ranging between −.27 to −.42.

3.3 Modified flanker task

Six participants (1 younger monolingual, 1 younger bilingual, and 4 older bilinguals) who had an overall accuracy rate below 60% (−2.5 *SDs* from the full study sample mean) on this task were excluded from the analyses. This resulted in all accuracies being around 90% and was therefore not analyzed further. One additional older monolingual did not complete the task due to difficulty. RT trimming procedures consisted of removing trials with RTs below 200 ms, eliminating 0.03% of trials for younger adults and 0.03% of trials for older adult participants. RT analyses were conducted on correct trials only.

Results for the flanker task are presented in Table 4. Reaction times for the single arrow condition was analyzed using a three-way mixed ANOVA using age, language group, and direction (same or opposite). There was a significant main effect of age, $F(1, 91) = 50.87$, $p < .0001$, $\eta^2_p = .36$, with younger adults responding faster than older adults, and a significant main effect of direction, $F(1, 91) = 16.25$, $p = .0001$, $\eta^2_p = .15$, with same trials faster than opposite trials. There were no other significant main effects or interactions, *F*s < 2.8.

Table 4. Mean RTs (and *SDs*) for the modified flanker task by age group and language group

	Younger Adults		Older Adults	
Condition	**Monolinguals**	**Bilinguals**	**Monolinguals**	**Bilinguals**
Single Arrow Blocks RT				
Same Direction	680 (145)	745 (156)	890 (132)	949 (151)
Opposite Direction	710 (151)	763 (157)	950 (154)	976 (177)
Conflict Blocks RT				
Congruent				
Same Direction	664 (137)	736 (118)	839 (118)	903 (149)
Opposite Direction	681 (118)	762 (160)	905 (132)	970 (181)
Incongruent				
Same Direction	664 (135)	725 (129)	867 (126)	927 (160)
Opposite Direction	709 (155)	762 (157)	929 (124)	970 (182)

For conflict block reaction times, a four-way mixed ANOVA, with age, language group, congruency (congruent or incongruent) and direction (same or opposite) revealed a main effect of age, $F(1, 91) = 51.70$, $p < .0001$, $\eta^2_p = .36$, with faster responding by younger adults, and language group, $F(1, 91) = 4.96$, $p = .03$, $\eta^2_p = .05$, with faster responding by monolinguals. There was a main effect of congruency, $F(1, 91) = 6.08$, $p = .02$, $\eta^2_p = .06$, with the average response to congruent trials being faster than incongruent. There was also a main effect of direction, $F(1, 91) = 47.72$, $p < .0001$, $\eta^2_p = .34$, with longer RTs for opposite trials, and a significant direction by age interaction, $F(1, 91) = 4.57$, $p = .04$, $\eta^2_p = .05$, with older adults showing substantially longer RTs to opposite direction trials compared to younger adults. There were no other significant interactions, all *F*s < 3.3.

A conflict direction cost score was calculated to investigate the RT costs due to keeping the opposite direction responding rule in mind (WM cost score). This was operationalized as ((congruent opposite + incongruent opposite)/2 – (congruent same + incongruent same)/2). For this cost score, there was a significant main effect of age group, $F(1, 91) = 4.57$, $p = .04$, $\eta^2_p = .05$, but not language group, $F < 1$, and no interaction, $F < 1$ ($M_{YML} = 30.7$, $SD_{YML} = 58.5$, $M_{YBL} = 31.9$, $SD_{YBL} = 55.6$, $M_{OML} = 63.8$, $SD_{OML} = 67.7$, $M_{OBL} = 55.0$, $SD_{OBL} = 73.9$).

3.4 Recent probe task

Results from the recent probe task are reported in Table 5. Two younger bilingual participants were excluded because one had RTs greater than 3 *SDs* above the group mean for 3 of the 6 task conditions, and the other due to a technical error. Two further participants were removed because of an overall accuracy rate below 54% (1 younger bilingual, and 1 older bilingual; −2.5 *SDs* from the full study sample mean). RT trimming procedures consisted of removing trials with RTs below 300 ms or above 2500 ms. This eliminated 0.7% of trials for younger adults and 1.2% of trials for older adults. All RT analyses were conducted on correct trials only.

For the pure blocks, a three-way mixed ANOVA for age, language group, and response type ('no' or 'yes') on RT revealed a main effect of age, $F(1, 94) = 53.53$, $p < .0001$, $\eta^2_p = .36$, with faster responding by younger adults. There were no other significant main effects or interactions, all *F*s < 3.2. Pure block accuracy analysis revealed a main effect of age, $F(1, 94) = 21.54$, $p < .0001$, $\eta^2_p = .19$, with higher accuracy for younger adults, but no effect of or interaction with language group, *Fs* < 2. There was also an effect of response type, $F(1, 94) = 39.07$, $p < .0001$, $\eta^2_p = .29$, with higher accuracy for negative responses. There were no significant interactions, all *F*s < 2.4.

Mixed block trials for RT were analyzed using a four-way mixed ANOVA for age, language group, response type, and trial type (baseline or interference/facilitation) and revealed a main effect of age, $F(1, 94) = 42.65$, $p < .0001$, $\eta^2_p = .31$, with faster responding by younger adults. There was a main effect of trial type, $F(1, 94) = 26.59$, $p < .0001$, $\eta^2_p = .22$, with faster responding on baseline than experimental (facilitation/interference) trials. There was also a response by trial interaction, $F(1, 94) = 33.23$, $p < .0001$, $\eta^2_p = .26$, in which participants overall

Table 5. Mean RTs (and *SDs*) and percent accuracy (and *SDs*) for the recent probe task by age group and language group

	Younger Adults		Older Adults	
Condition	**Monolinguals**	**Bilinguals**	**Monolinguals**	**Bilinguals**
		RT		
Pure blocks				
Negative baseline	971 (159)	873 (141)	1192 (181)	1153 (209)
Positive baseline	986 (196)	904 (132)	1192 (218)	1173 (195)
Mixed blocks				
Negative baseline	969 (199)	880 (127)	1164 (197)	1150 (240)
Interference	1069 (199)	932 (141)	1244 (226)	1210 (249)
Positive baseline	963 (171)	871 (141)	1206 (265)	1160 (213)
Facilitation	967 (204)	897 (132)	1184 (234)	1149 (195)
		Accuracy		
Pure blocks				
Negative baseline	87.6 (8.2)	91.0 (7.8)	84.9 (12.2)	82.5 (13.3)
Positive baseline	81.7 (11.0)	78.5 (13.9)	73.4 (13.1)	66.4 (17.2)
Mixed blocks				
Negative baseline	87.5 (10.8)	88.5 (8.6)	87.9 (9.1)	84.9 (12.2)
Interference	82.6 (12.5)	84.8 (9.1)	79.9 (11.4)	79.4 (12.9)
Positive baseline	81.7 (10.9)	72.8 (15.9)	73.2 (16.7)	67.3 (15.8)
Facilitation	83.9 (7.6)	79.7 (11.8)	74.6 (16.7)	71.1 (14.4)

responded similarly to facilitation trials ($M = 1047$, $SD = 226$) and positive baseline trials ($M = 1048$, $SD = 241$), but were slower to interference trials ($M = 1112$, $SD = 239$) than negative baseline trials ($M = 1039$, $SD = 227$), indicating that only interference affected RT. Importantly, there was also a significant three-way interaction of response by trial by language, $F\,(1, 94) = 4.03$, $p = .0475$, $\eta^2_p = .04$. Univariate analyses by condition indicated that bilinguals were faster on interference trials ($p = .04$), but not on the other three types (*F*s < 2.9). No other interactions were significant, *F*s < 3.9.

An analysis of mixed block accuracy showed a significant effect of age, $F\,(1, 94) = 12.81$, $p = .0005$, $\eta^2_p = .12$, such that younger adults were more accurate than older adults. There was a significant effect of response type, $F\,(1, 94) = 25.77$, $p < .0001$, $\eta^2_p = .22$, with higher accuracy for negative responses, as well as a significant response by trial interaction, $F\,(1, 94) = 38.56$, $p < .0001$, $\eta^2_p = .29$. This interaction represents the expected experimental manipulation, with overall higher accuracy for facilitation ($M = 77.4$, $SD = 13.7$) than positive baseline trials ($M = 73.9$, $SD = 15.6$), and lower accuracy for interference ($M = 81.7$, $SD = 11.6$) than negative baseline trials ($M = 87.2$, $SD = 10.3$).There were no other significant main effects or interactions, all *F*s < 3.7.

The same data were analyzed as RT cost scores that were calculated following the procedures in Bialystok et al. (2014). Facilitation effects were represented as mixed block positive baseline - facilitation trials, and interference effects were represented as mixed block negative baseline - interference trials. These data are presented in Figure 1. A three-way mixed ANOVA with age group, language group, and experimental trial type (facilitation effect, interference effect) was run on these cost scores. One younger bilingual outlier was removed for having a large negative cost score close to 3 *SDs* minus the group mean. There was a significant main effect of trial type, $F(1, 93) = 31.53$, $p < .0001$, $\eta^2_p = .25$, indicating again the presence of a large cost on RTs from interference ($M = -70.7$, $SD = 92.6$) and no effect of facilitation ($M = 0.8$, $SD = 93.0$).There was a significant trial by language interaction, $F(1, 93) = 5.05$, $p = .027$, $\eta^2_p = .05$. One-way ANOVAs examining the effect of language on each experimental trial type indicated that bilinguals had smaller interference costs than monolinguals ($p = .03$), but the groups did not differ on facilitation ($F < 1$). There were no other significant main effects or interactions, all *F*s < 1.5.

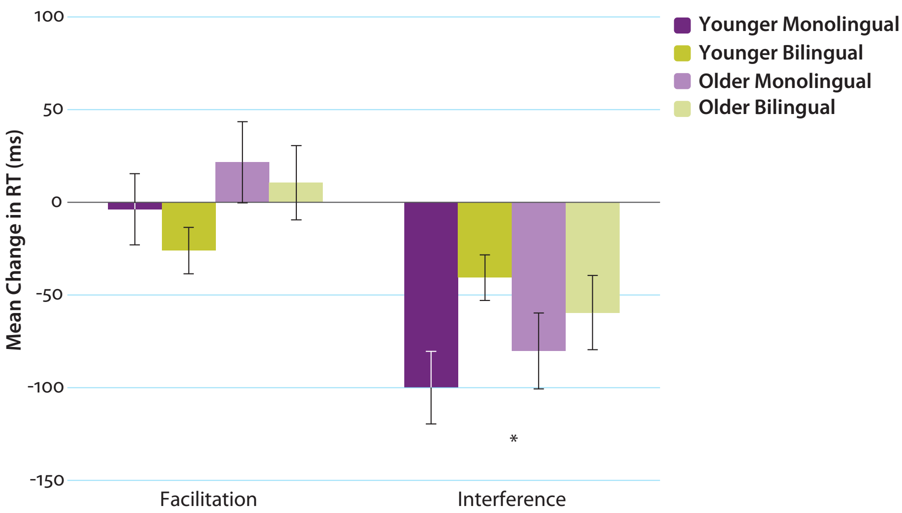


Figure 1. RT cost scores for facilitation and interference on the recent probe task by age and language group. Error bars represent standard errors.

A summary of the main findings for the three experimental tasks is shown in Table 6.

Table 6. Summary of main results by experimental task

	Age Group		Language Group	
Task	**Simple Condition**	**Complex Condition**	**Simple Condition**	**Complex Condition**
Star Counting	Y < O	Y < O	ML = BL	ML = BL
Modified Flanker	Y < O	Y < O	ML = BL	ML < BL
Recent Probe	Y < O	Y = O	ML = BL	ML > BL

4. Discussion

Both aging and bilingualism influenced performance on the three tasks used in the current study but the nature of the effects depended on the task (See Table 6). The participant groups were matched on relevant background measures with the exception of English vocabulary knowledge. This pattern of lower vocabulary scores in bilinguals and higher vocabulary scores in older adults, which is typical for the literature (e.g., Luo et al., 2013), had a significant impact on performance in the star counting task. As predicted, prior to controlling for vocabulary knowledge bilinguals were slower than monolinguals to count stars, even though participants were informed they could use the language in which they were most comfortable counting. The groups were equivalent on counting accuracy for standard and WM conditions, but the negative correlations between vocabulary and counting time reflect the verbal nature of this production task. Indeed, previous work has provided evidence that verbal memory and numerical working memory are part of the same construct (Oberauer, Süss, Schulze, Wilhelm, & Wittman, 2000). Once vocabulary was controlled, the main effect of bilingualism disappeared, but aging effects remained such that older adults were slower than younger adults. Thus, vocabulary was a factor in the slower task performance by bilinguals at both age groups but was not a factor in the slower performance by older adults in both language groups. Put another way, individuals became slower on this task with age but there was no evidence that performance was further moderated by language experience. The interaction of age with task conditions, such that older adults were slower than younger adults on conditions requiring the minimum level of EC and the maximal level of EC suggests that in this largely processing based task that younger adults surpass older adults when processing demands are few, but are also are better able to handle the dual demands of high switching and remembering the complex counting rule.

For the modified flanker task, typical age-related slowing effects were seen for both the single arrow and conflict blocks, and older adults were particularly affected by the demands of having to remember to respond to the opposite arrow direction during conflict blocks. Single arrows blocks did not elicit any language group differences indicating no differences in simple speed of processing. However, monolinguals performed significantly faster on the conflict blocks, contrary to our expected results. Although we have no explanation for the difference shown on this version of the task other than the possibility that the bilinguals in the current sample attended more to the overall context of the conflicting cues (see explanation below), the finding minimally demonstrates that the bilinguals were not inherently faster responders even on a simple conflict task.

WM during conflict trials on the modified flanker task was taken into account by calculating RT cost scores from having to remember the opposite response rule. There were no language group differences, reflecting equivalent effects of WM on performance. The cues from the flanking arrows in this version of the task are more complex than in the traditional flanker task. For example, in the opposite direction condition congruent flankers are really incongruent and therefore do not facilitate performance. This suggests that a greater amount of inhibition of flanking arrows is required for this version of the task, and undeniably, some of the older adults found this task to be particularly difficult and is in line with research showing that older adults have worse inhibitory control compared to younger adults (Hasher, Zacks, & May, 1999). Therefore, successful performance on this particular task compared to the other tasks used in the current study may be largely based on rapid inhibitory control processes more so than WM ability.

The recent probe task was the same as in Bialystok et al. (2014), but now used stimuli that are more distinctive and easier to encode. Task accuracy declined with aging, but there were no effects or interactions with language. Similar accuracy rates across language groups allows for an interpretation of response processing without any speed-accuracy trade-offs. Additionally, our sample consisted of healthy older and younger adults, without diagnosed memory impairment. For analyses of RTs, older adults were slower than younger adults, a result that was obviously expected. Crucially, however, bilinguals were faster than monolinguals on the proactive interference condition, replicating the finding in Bialystok et al. (2014). This indicates that when the need for conflict resolution in memory is high, (i.e., to overcome a strong sense of familiarity and engage in accurate recollection); bilinguals are able to engage in these processes at a faster rate than monolinguals. The analyses of cost scores that represented the difference in time to respond to baseline and experimental trials for each of the positive and negative responses additionally showed that bilinguals had smaller interference costs than monolinguals, although the groups did not differ in the amount of facilitation effects from repeated trials. In fact, the facilitation effects were small altogether (seen mainly for accuracy and not RT) so it is not surprising that there was no difference between groups on this measure. Aging and bilingualism did not interact, contrary to our initial prediction, indicating independent effects on WM performance for this sample of healthy adults.

Together, these results indicate that the impact of bilingualism on working memory depends on task demands such as the task domain, and presence of interference. Bilingual processing tends to be worse for verbal tasks, due to less efficient lexical access and retrieval, unless vocabulary level is controlled, and only then is bilingual performance similar to monolingual. This was seen both in the current study on the star counting task and previous work with verbal materials (e.g., Fernandes, Craik, Bialystok, and Kreuger, 2007). Bilinguals showed similar WM costs as monolinguals on performance on the nonverbal modified flanker task, whereas better bilingual performance in WM was tied to overcoming the influence of proactive interference. This was supported by the results seen on the nonverbal recent probe task. This replication strengthens the findings from Bialystok et al. (2014) using a new set of stimuli.

This work has important implications for research involving bilingualism as a contributing factor to cognitive reserve in aging, and suggests that the underlying mechanism may be the ability to utilize the executive control system to deal with the detrimental effects of interference. The current sample was comprised of healthy, well-educated adults. Future studies will need to assess if enhanced nonverbal interference resolution in bilinguals remain stable through aging and investigate possible changes that accrue with the onset of neuropathology. In addition, the contributing role of WM capacity versus processing in language group differences needs further exploration. The current study examined the role of WM processing in the form of varying levels of EC, however evidence from a recent meta-analysis also suggests that bilinguals may outperform monolinguals in terms of WM capacity (Grundy & Timmer, submitted). There is also an existing literature on the role of WM and L2 proficiency development and use. In a recent meta-analysis by Linck, Osthus, Koeth, and Bunting (2014), a positive relationship was shown between WM and L2 outcomes, and stronger relationships were shown for complex WM span tasks (i.e., greater need for EC) compared to simple WM span tasks and for verbal compared to nonverbal measures, providing evidence that the relationship between WM and language development/use may be bidirectional. We conclude however, with stating that bilingualism research is complex and accurate assessments of how bilingualism and aging influence WM will depend on the level of bilingual experience, pre-existing abilities, as well as the task domain.

Acknowledgements

This research was supported by grant A2559 from the Natural Sciences and Engineering Research Council, Canada (NSERC) to Ellen Bialystok. We thank Cari Bogulski, Lorinda Mak, Karen Lau, Angelina Pinhasov, Kara Hounsell, Zehra Kamani, Ashley Etherington, and Aram Keyvani for their assistance with the study, as well as all of our participants.

References

Adesope, O.O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, 80 (2), 207–245. doi: 10.3102/0034654310368803

Bialystok, E. (2015). Bilingualism and the development of executive function: The role of attention. *Child Development Perspectives*, 9 (2), 117–121. doi: 10.1111/cdep.12116

Bialystok, E., Craik, F.I.M., Green, D.W., & Gollan, T.H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, 10, 89–129. doi: 10.1177/1529100610387084

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

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, 19(2), 290–303. doi: 10.1037/0882-7974.19.2.290

Bialystok, E., & Feng, X. (2009). Language proficiency and executive control in proactive interference: Evidence from monolingual and bilingual children and adults. *Brain & Language*, 109, 93–100. doi: 10.1016/j.bandl.2008.09.001

Bialystok, E., & Luk, G. (2012). Receptive vocabulary differences in monolingual and bilingual adults. *Bilingualism: Language and Cognition*, 15, 397–401. doi: 10.1017/S136672891100040X

Bialystok, E., Luk, G., Peets, K.F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 13 (4), 525–531. doi: 10.1017/S1366728909990423

Bialystok, E., Poarch, G. J., Luo, L., & Craik, F. I. M. (2014). Effects of bilingualism and aging on executive function and working memory. *Psychology and Aging*, 29 (3), 696–705. doi: 10.1037/a0037254.

Das-Smaal, E., de Jong, P. F., & Koopmans, J. R. (1993). Working memory, attentional regulation and the star counting test. *Personality and Individual Differences*, 14(6), 815–824. doi: 10.1016/0191-8869(93)90094-J

Eriksson, J., Vogel, E.K., Lansner, A., Bergström, F., & Nyberg, L. (2015). Neurocognitive architecture of working memory. *Neuron*, 88 (1), 33–46. doi: 10.1016/j.neuron.2015.09.020

Fernandes, M.A., Craik, F.I.M., Bialystok, E., & Kreuger, S. (2007). Effects of bilingualism, aging, and semantic relatedness on memory under divided attention. *Canadian Journal of Experimental Psychology*, 6 (2), 128–141. doi: 10.1037/cjep2007014

Gollan, T.H., Montoya, R.I., Cera, C.M., & & Sandoval, T.C., (2008). More use almost always means smaller a frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, 58, 787–814. doi: 10.1016/j.jml.2007.07.001

Gollan, T.H., & Silverberg, N.B. (2001). Tip-of-the-tongue states in Hebrew-English bilinguals. *Bilingualism: Language and Cognition*, 4, 63–83. doi: 10.1017/S136672890100013X

Grundy, J.G., & Timmer, K. (submitted). Bilingualism and working memory capacity: A comprehensive meta-analysis.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher and A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application. Attention and performance* (pp. 653–675). Cambridge, MA: The MIT Press.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, 127, 277–288. doi: 10.1016/j.actpsy.2007.06.003

Jacoby. L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30 (5), 513–541. doi: 10.1016/0749-596X(91)90025-F

Jonides, J., Marshuetz, C., Smith, E.E., Reuter-Lorenz, P.A., Koeppe, R.A. (2000). Age differences in behavior and PET activation reveal differences in interference resolution in verbal working memory. *Journal of Cognitive Neuroscience*, 12 (1), 188–196. doi: 10.1162/089892900561823

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience*, 139, 181–193. doi: 10.1016/j.neuroscience.2005.06.042

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Valdes-Kroff, J. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In B. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 56, pp. 229–262). San Diego: Academic Press.

Linck, J.A., Osthus, P., Koeth, J.T., & Bunting, M.F. (2014). Working memory and second language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 21, 861–883. doi: 10.3758/s13423-013-0565-2

Lu, C.H, & Proctor, R.W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review*, 2, 174–207. doi: 10.3758/BF03210959

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

Luo, L., Craik, F. I. M., Moreno, S., & Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: Evidence from aging. *Psychology and Aging*, 28(1), 28–34. doi: 10.1037/a0030875

May, C., Hasher, L., & Kane, M.J. (1999). The role of interference in memory span. *Memory & Cognition*, 27, 759–767. doi: 10.3758/BF03198529

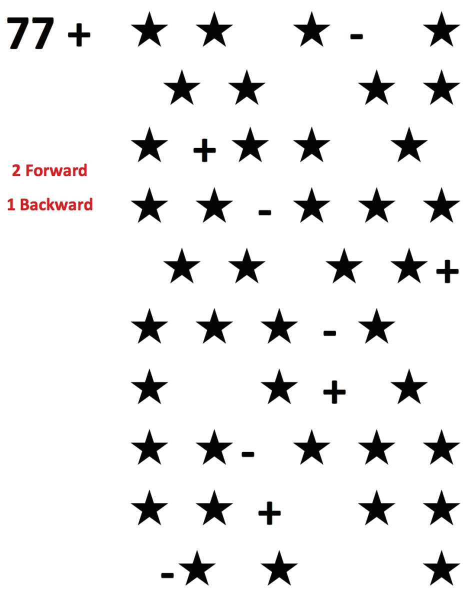
Oberauer, K., Süss, H.M., Schulze, R., Wilhelm, O., & Wittman, W.W. (2000). Working memory capacity- facets of a cognitive ability construct. *Personality and Individual Differences*, 29, 1017–1045. doi: 10.1016/S0191-8869(99)00251-2

Shipley, W. C., Gruber, C. P., Martin, T. A., & Klein, A. M., (2009). *Shipley-2 Manual.* Los Angeles, CA: Western Psychological Services.

van Merrienboer, J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17, 147–177. doi: 10.1007/s10648-005-3951-0

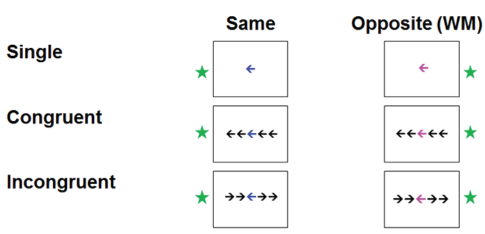
Wodniecka, Z., Craik, F.I.M., Luo, L., & Bialystok, E. (2010). Does bilingualism help memory? Competing effects of verbal ability and executive control. *International Journal of Bilingual Education and Bilingualism*, 13 (5), 575–595. doi: 10.1080/13670050.2010.488287

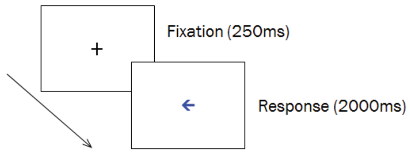
Appendix A



A sample card in the working memory condition of the star counting task

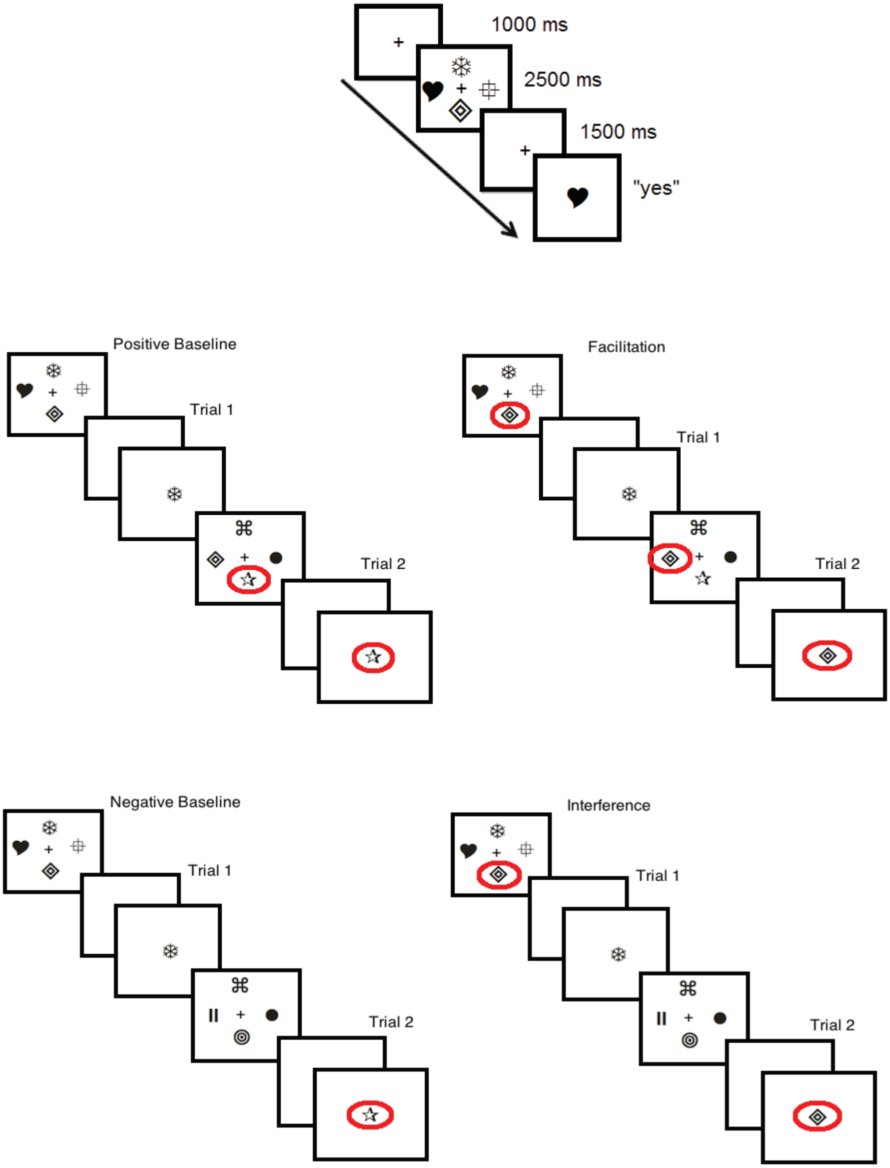
Appendix B





Procedure and conditions in the modified flanker task

Appendix C



Procedure and trial types in the recent probe task

Authors' addresses

Margot D. Sullivan
Yolanda Prescott
Devora Goldberg

Ellen Bialystok
York University
Department of Psychology
4700 Keele Street
TORONTO, ON M3J 1P3
CANADA

ellenb@yorku.ca

Publication history

Date received: 3 December, 2015
Date accepted: 11 January, 2016
Published online: 19 February 2016