Bilingualism and working memory capacity: A comprehensive meta-analysis

John G Grundy and Kalinka Timmer
York University, Toronto, ON, Canada

Abstract
Bilinguals often outperform monolinguals on executive function tasks, including tasks that tap cognitive flexibility, conflict monitoring, and task-switching abilities. Some have suggested that bilinguals also have greater working memory capacity than comparable monolinguals, but evidence for this suggestion is mixed. We therefore conducted a comprehensive meta-analysis on the effects of bilingualism on working memory capacity. Results from 88 effect sizes, 27 independent studies, and 2,901 participants revealed a significant small to medium population effect size of 0.20 in favor of greater working memory capacity for bilinguals than monolinguals. This suggests that experience managing two languages that compete for selection results in greater working memory capacity over time. Moderator analyses revealed that largest effects were observed in children than other age groups. Furthermore, whether the task was performed in the first (L1) or second (L2) language for bilinguals moderated the effect size of the bilingual advantage; this factor is often overlooked and our results point to the importance of defining language variables that influence critical cognitive outcomes.

I Introduction

Bilingualism has been associated with the enhancement of multiple executive functions, including cognitive flexibility (Adi-Japha et al., 2010), efficiency (Blumenfeld and Marian, 2014), task-switching (Gold et al., 2013; Prior and MacWhinney, 2010), and conflict resolution (Donnelly, Brooks, and Homer, 2015). This is believed to be the result

Corresponding author:
John G. Grundy, Department of Psychology, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada
Email: jggrundy@gmail.com

of lifelong experience managing multiple languages that compete for selection. Resolution of this competition requires higher-order executive control processes that may extend to enhance generalized executive functioning (Bialystok et al., 2012). Given that language processing is largely dependent on working memory and that there is a strong positive relationship between many higher-order executive functions and working memory (WM) capacity (Engle, 2002), it is logical to expect that bilinguals would also have greater WM capacity than monolinguals. Some research supports this prediction (e.g. Morales et al., 2013; Soliman, 2014) but other studies have found no effect (e.g. Namazi and Thordardottir, 2010; Ratiu, and Azuma, 2015). Recently, such failures to replicate bilingual cognitive benefits have led some to argue that there is no bilingual advantage (Gasquoine, 2016; Paap et al., 2015), sparking widespread responses to these claims (Bak, 2015; Gold, 2015; Kousaie and Taler, 2015; Woumans and Duyck, 2015). It is therefore especially important to discover the reality of these effects in light of the contentious nature of the field. A meta-analysis on the current state of the literature would be revealing in uncovering the strength and reality of these findings. Here, we present a meta-analysis of performance differences between monolinguals and bilinguals on WM span tasks.

There are reasons to expect that bilinguals might have greater working memory capacity than monolinguals. Substantial evidence exists that both languages are activated in the bilingual brain even when only one language is being used (Martin et al., 2009; Spivey and Marian, 1999; Thierry and Wu, 2007; Timmer et al., 2014; Wu and Thierry, 2010, 2012; for review, see Kroll et al., 2012). Managing languages that compete for selection requires resources from WM (Thorn and Gathercole, 1999), and continual use of WM resources might lead to enhanced WM capacity over time in order to ensure efficiency of processing in the future. Some researchers claim that larger WM capacity actually represents greater ability to control attention within the WM store, rather than having a greater store per se (Engle, 2002). From this perspective as well, bilinguals should have greater WM capacity than monolinguals because numerous studies have shown that bilinguals are better at controlling attention (Bialystok, 1999; Bialystok et al., 2004; Costa et al., 2008; Hernandez et al., 2010; Soveri et al., 2011).

Recent models of WM share the view that WM is made up of multiple component processes that rely heavily on selective attention (Eriksson et al., 2015). Eriksson et al. (2015) propose that selective and sustained attention processes that operate on perceptual and long-term memory (LTM) information are integral parts of WM. Specifically, they propose a model in which attentional and other components (e.g. inhibition, prospective planning, updating), all contribute to 'temporarily enhanced accessibility'. Learning a second language across the lifespan necessarily involves selective and sustained attention that operate on perceptual and LTM information. For example, preverbal infants that grow up in a bilingual environment already show enhanced attention to the mouth of talking faces at an earlier age than infants growing up in a monolingual context; focus on the mouth is a strategy used by more mature individuals (e.g. adults) in order to make use of the redundancy of audiovisual speech cues (Pons et al., 2015). This shows that bilingual infants have more mature attentional control than monolinguals. Furthermore, in order for infants and children to make an association between an item currently being processed and one of the two languages being learned, the item needs to be rehearsed, while previously stored information from the two languages (that compete for selection) is retrieved

from LTM. Lifelong practice with attentional control devoted to these rehearsal and retrieval processes should contribute to better WM for bilinguals than monolinguals.

A 2010 meta-analysis examining the relationship between bilingual status and executive function (EF) tasks revealed that bilinguals outperformed monolinguals on a range of EF tasks, including tasks that tapped WM processes (Adesope et al., 2010). However, only seven effect sizes from five independent studies were available for this meta-analysis and, since then, an upsurge of studies has published on the relationship between bilingual status and WM capacity. In fact, 19 of the 27 studies included in the present meta-analysis were from 2010 onward. Furthermore, some of the studies included in Adesope et al. (2010) meta-analysis used non-span tasks that only indirectly measured WM capacity (e.g. Bialystok, 2006). Given these limitations and the current conflicting results in the literature, a re-evaluation of the population effect size of WM capacity advantages for bilinguals over monolinguals (if present) is necessary.

The present study performed a comprehensive meta-analysis on the effects of bilingualism on working memory capacity to reveal an estimate of the population effect size. Furthermore, given the upsurge of new studies, we were able to examine age, the linguistic nature of the task, and the language in which the task was performed as potential moderating variables.

II Methods

1 Literature search

We first searched the PsycINFO database with the keywords 'bilingual' and 'working memory' or 'short-term memory'. Following this, we performed a Google Scholar search for 'bilingual working memory span' and 'bilingual short-term memory span' for any articles that PsycINFO might have missed. Finally, we combed through the reference sections of the articles obtained to find any additional studies that we might have missed. The search included any relevant articles (see below) found up to 13 July 2015.

2 Inclusion criteria

Studies included in the meta-analysis conformed to the following set of criteria:

1. The study had a clear monolingual and a clear bilingual group. Studies that examined working memory performance across different levels of bilingualism (e.g. low proficiency vs. high proficiency bilinguals) were not included.
2. All participants were typically developing individuals without learning disabilities or cognitive impairments of any sort.
3. The study included a working memory span task (e.g. forward digit span, backward digit span, sentence span, operation span, etc.). Studies that indirectly manipulated working memory demands within a non-span task were not included.
4. The study had enough information to calculate an effect size. In cases where an effect size could not be calculated based on descriptive statistics (i.e. means and standard deviations), *F*- and *t*-values were converted to effect sizes using the conversion equations reported in Appendix 1.

Following these guidelines, a total of 88 effect sizes from 27 independent studies with a total sample involving 2,901 participants were obtained. All studies included in the meta-analysis are shown in Table 1 and identified by an asterisk in the references section.

3   Main meta-analysis

Following the suggestion of Field and Gillett (2010; see also Rosenthal and DiMatteo, 2001), we chose the Pearson *r* correlation coefficient as our effect size measure for the analysis. This coefficient is understood by most psychologists and is a versatile measure of the strength of an experimental effect. Furthermore, the Pearson *r* correlation coefficient is easy to calculate from means and standards deviations, as well as from other statistics (e.g. *F*- and *t*- values; Field, 2005a, 2005b; Rosenthal, 1991; for conversion equations used in the present experiment, see Appendix 1).

For the present meta-analysis, we employed the Hunter–Schmidt random-effects model (Hunter and Schmidt, 2004) on these effect sizes. Most researchers recommend using a random-effects model when one wishes to generalize findings beyond studies included in the analysis (i.e. to the general population; Field and Gillett, 2010; Hedges and Vevea, 1998; Overton, 1998). Fixed-effects models should only be used when there are reasons to expect that the population variance is homogeneous, which rarely occurs in reality, unless one is only interested in a subset of the population (e.g. using the same task with the same researcher). Furthermore, applying a fixed-effects model to the population can inflate type-I error rates from the usual 5% to anywhere from 11%–80% (Field, 2003; Hunter and Schmidt, 2000). Given that we wished to estimate the mean effect size of the general population, we opted for a random-effects model. To employ the Hunter–Schmidt random-effects analysis, we used the SPSS syntax files described by Field and Gillett (2010) and supplied by Field and Gillett (2009). We report the estimated population effect size (ρ), along with the 95% confidence intervals, sample correlation variance, sampling error variance, and the estimated variance in population correlations.

4   Sensitivity analysis

We conducted a sensitivity analysis to combat the potential issue of non-independence of effect sizes in our sample given that many of the studies have multiple effect sizes (see Table 1). Following previous research (Orlitzky et al., 2003), mean effect sizes were created, resulting in one independent effect size per study. If this analysis turns out similar to the comprehensive one, we can be more confident in our findings.

5   Moderator analyses

We performed moderator analyses on the following variables:

- Age group (children, young adults, older adults)
- Task type (verbal vs. non-verbal)
- Language in which the verbal tasks were performed

Table 1. All studies included in the meta-analysis.

Study	Age group	Mean effect size (r)	Number of effect sizes
Namazi and Thordardottir (2010)[1]	Children	0.01	3
Soliman (2014)	Children	0.32	11
Ratiu and Azuma (2015)	Young Adults	−0.21	3
Ransdell et al. (2006)	Young Adults	0.11	1
Bonifacci et al. (2011)[2]	Children and Young Adults	−0.01	8
Kaushanskaya et al. (2014)	Children	0.26	2
Engel de Abreu (2011)[3]	Children	−0.14	4
Luo et al. (2013)	Young Adults and Old Adults	−0.05	8
Biedroń and Szczepaniak (2012)	Young Adults	0.52	3
Białystok et al. (2008)	Young Adults and Old Adults	0.05	6
Sasisekaran and Weisberg (2013)	Young Adults	0.14	2
Mumtaz and Humphreys (2001)	Children	−0.15	5
Danahy et al. (2007)	Children	−0.10	1
Ljungberg et al. (2013)[4]	Older adults	0.45	1
Thorn and Gathercole (1999)[5]	Children	0.26	2
Nguyen and Astington (2014)[6]	Children	0.32	1
Cole and Pickering (2010)	Young Adults	−0.15	3
Da Fontoura and Siegel (1995)	Children	0.11	1
Abu-Rabia and Siegel (2002)	Children	0.09	1
Smithson and Nicoladis (2013)	Young Adults	−0.09	3
D'Angiulli et al. (2001)[7]	Children	0.25	2
Engel de Abreu et al. (2012)	Children	−0.09	2
Leikin and Tovli (2014)	Children	0.91	1
Blom et al. (2014)[8]	Children	0.25	4
Kaushanskaya et al. (2011)	Young Adults	−0.02	2
Kaushanskaya (2012)[9]	Young Adults	0.32	4
Morales et al. (2013)[10]	Children	0.11	4

a   Age group (children, young adults, older adults).   Previous work has shown that bilinguals and monolinguals are more likely to show executive function differences in children and older adults than in younger adults because the latter individuals are at peak cognitive performance (Bialystok et al., 2005). Thus, larger working memory capacity might not be as evident in young adults as it is in the other two groups. On the other hand, it is possible that a linear increase in working memory capacity will be revealed across the lifespan, given that experience managing two languages increases over time. Finally, it is possible that working memory demands are largest when people are first learning a new language and the greatest gains in working memory capacity might be expected in children, after which point they plateau.

b   Task type (verbal vs. non-verbal).   Working memory capacity might vary as a function of task type. Bilingual performance disadvantages compared to monolinguals have been reported on numerous tasks that require language processing (Bialystok and Luk, 2012; Gollan et al., 2007; Ivanova and Costa, 2008). Thus, it is possible that verbal working memory tasks are masking the true working memory capacity of bilinguals (Bialystok et al., 2014).

c   Language in which the verbal tasks were performed.   Language in which the verbal tasks were performed: first language (L1) vs. second language (L2). All monolinguals necessarily have to perform WM span tasks in their L1, but bilinguals sometimes perform the tasks in their L1, sometimes in their L2, and sometimes in both. Performing WM span tasks in their L2 would disadvantage bilinguals on these tasks given that they are necessarily not as proficient in this language (compared to L1) and retrieval processes are slower. Thus, we explored whether these considerations moderated the results.

6   Publication bias

An issue with any meta-analysis is the fact that published studies are more likely to report significant findings than non-significant findings, and this can over-represent the estimated effect size of the population. Fortunately, there are some analyses that one can perform to instill confidence in the meta-analytic results. One of the most popular is Rosenthal's (1979) fail-safe N, which is an estimate of the number of missing or unpublished non-significant studies that are required in order to render the estimated effect size non-significant. A general rule of thumb is that the estimated effect size is likely safe from publication bias if the fail-safe N exceeds $5k + 10$, where k is the number of effect sizes included in the analysis. In the present study, the fail-safe N would need to exceed 470 to be safe from publication bias.

III   Results

1   Population estimate

Table 1 provides descriptive statistics and mean effect sizes for all studies included in the analysis. Following Rosenthal's (1995) advice, Table 2 shows stem-and-leaf plots

Table 2. Stem-and-leaf plot of all effect sizes.

Stem	Leaf												
−0.7	1												
−0.6													
−0.5	4												
−0.4	7												
−0.3	0	1	3										
−0.2	1	2	6	6									
−0.1	0	0	1	2	4	5	5	5	8	9			
−0.0	1	2	2	2	2	4	4	4	6	6	7	8	8
0.0	0	0	0	0	4	7	7	7	8	8	9		
0.1	0	1	1	2	2	5	5	7	8	8	8	9	9
0.2	0	1	2	2	2	2	4	4	5				
0.3	2	2	2	3	3	4	5	5	7	7			
0.4	2	3	3	4	5	5	5	7	7				
0.5													
0.6	1	3											
0.7													
0.8													
0.9	1												

for all effect sizes. A chi-squared test of homogeneity revealed considerable variation in effect sizes overall, $X^2(1, N = 88) = 892.92$, $p < 0.001$, which gives justification for use of the random-effects model. The weighted mean population estimate was highly significant, $\rho = 0.20$, $X^2(1, N = 88) = 799.65$, $p < 0.001$; the 95% confidence interval was −.253 (lower) to 0.653 (upper). The sample correlation variance was found to be 0.06 and the sampling error 0.01. Thus, the estimated population correlation variance was 0.05.

2 Sensitivity analysis

A chi-squared test of homogeneity revealed considerable variation in effect sizes overall, $X^2(1, N = 27) = 168.97$, $p < 0.001$. The weighted mean population estimate was highly significant, $\rho = 0.158$, $X^2(1, N = 27) = 132.04$, $p < 0.001$; the 95% confidence interval was −.205 (lower) to 0.521 (upper). The sample correlation variance was found to be 0.043 and the sampling error 0.009. Thus, the estimated population correlation variance was 0.034. These results are similar to the comprehensive meta-analysis, thus, we use the comprehensive list of effect sizes for the remaining analyses.

3 Publication bias

The fail-safe N was found to be 6496, which is much larger than the suggested minimum of 470 (i.e. $5k + 10$). This suggests that our population effect size estimate is likely safe from publication bias.

4   Moderator analyses

Table 3 shows stem and leaf plots of the effect sizes for each of the age (children, young adults, older adults) and task type (verbal, non-verbal) combinations. Using the moderator syntax script for meta-analyses provided by Field and Gillett (2009, 2010), age did not significantly moderate the overall population effect size, $X^2(2, N = 88) = 2.94$, $p = 0.23$, nor did type of task, $X^2(1, N = 88) = 2.48$, $p = 0.115$. However, because of their theoretical significance and our a priori predictions, we followed these analyses up with conventional ANOVAs on the weighted Pearson r values for the independent variables. The one-way ANOVA on age revealed a significant effect, $F(2, 85) = 5.01$, $p = 0.01$, but the ANOVA on type of task did not, $F(1, 86) = 1.97$, $p = 0.16$. The effect of age is explained by the finding that children ($\rho = 0.25$) showed larger effect sizes than young adults ($\rho = 0.03$; $t(54^{11}) = 3.56$, $p = 0.001$) and older adults ($\rho = 0.08$; $t(27^1) = 2.32$, $p = 0.03$). Younger and older adult effect sizes did not differ from each other ($p > 0.4$). However, caution should be taken when interpreting the older adult results because only 8 of the 88 effect sizes were from the older adult population and are therefore less reliable.

The next analysis involved whether studies that used verbal tasks were performed in the L1 or the L2 for bilinguals. This analysis involved 46 effect sizes (some of which were previously collapsed into one effect size; see Table 1 footnotes) from 18 independent studies. The analysis revealed a significant moderating effect on the population effect size, $X^2(1, N = 46) = 12.00$, $p = 0.001$. This is explained by the finding that when bilinguals did the WM task in their L1 ($\rho = 0.28$; N = 35) the effect size was much larger than when they did the task in their L2 ($\rho = -0.11$; N = 11), which was in the opposite direction.

IV   Discussion

We performed a comprehensive meta-analysis to examine the possibility that bilinguals have greater WM capacity than monolinguals because of their extensive experience managing two languages that compete for selection. Results revealed a robust small to medium population effect size of 0.20 showing greater working memory capacity for bilinguals than monolinguals. Our interpretation is that competition between languages places demands on WM capacity and that this results in greater WM capacity over time.

A recent study performed a meta-analysis on the influence of working memory capacity on second language comprehension and production scores amongst bilinguals (Linck et al., 2014). They found a population effect size estimate of 0.255 amongst 79 independent studies. They interpreted these findings as suggesting that WM capacity leads to greater proficiency in a second language amongst bilinguals. However, correlation does not equal causation, and it is equally likely that their data reflect that amongst bilinguals, greater second language proficiency leads to enhanced WM capacity. The latter possibility is in line with our interpretation of increased WM capacity for bilinguals over monolinguals. The present study can shed light on this potential bidirectional relationship given that we examined the difference between monolinguals and bilinguals, whereas Linck and colleagues only examined bilinguals. Most bilinguals do not choose to be

Table 3. Stem-and-leaf plots for effect sizes broken down by age group and task type.

Stem	Leaf							
Children: Non-verbal:								
−0.7								
−0.6								
−0.5	4							
−0.4	7							
−0.3								
−0.2								
−0.1	0	5						
−0.0	2	4						
0.0								
0.1	0	2	2	8				
0.2	1	2	2	2				
0.3	2							
0.4								
0.5								
0.6								
0.7								
0.8								
0.9								

Stem	Leaf							
Children: Verbal:								
−0.7	1							
−0.6								
−0.5								
−0.4								
−0.3	0							
−0.2	1							
−0.1	0							
−0.0	4	7						
0.0	0	0	0	7	9			
0.1	1	5	7	8	9	9		
0.2	4	5						
0.3	2	2	3	3	5	5	7	7
0.4	2	3	4	5				
0.5								
0.6	1							
0.7								
0.8								
0.9	1							

Stem	Leaf				
Young adults: Non-verbal:					
−0.7					
−0.6					
−0.5					
−0.4					
−0.3					
−0.2					
−0.1					
−0.0	1	4	6	6	8
0.0	4	7	7		
0.1					
0.2	3				
0.3	4				
0.4	5				
0.5					
0.6					
0.7					
0.8					
0.9					

Stem	Leaf			
Young adults: Verbal:				
−0.7				
−0.6				
−0.5				
−0.4				
−0.3	1	3		
−0.2	2	6	6	
−0.1	1	2	5	8
−0.0	2	2	8	
0.0	8			
0.1	1	5		
0.2	0	4		
0.3				
0.4	3	7	7	
0.5				
0.6	3			
0.7				
0.8				
0.9				

Stem	Leaf
Old adults: Non-verbal:	
−0.7	
−0.6	

Stem	Leaf
Old adults: Verbal:	
−0.7	
−0.6	

(Continued)

Table 3. (Continued)

Old adults: Non-verbal:				Old adults: Verbal:		
−0.5				−0.5		
−0.4				−0.4		
−0.3				−0.3		
−0.2				−0.2		
−0.1	4			−0.1	5	9
−0.0	2			−0.0		
0.0	0	8		0.0		
0.1	8			0.1		
0.2				0.2		
0.3				0.3		
0.4				0.4	5	
0.5				0.5		
0.6				0.6		
0.7				0.7		
0.8				0.8		
0.9				0.9		

bilingual; rather, they are raised in bilingual environments or become bilingual out of life necessity (Bialystok and Poarch, 2014). Therefore, it is not likely that bilinguals in the present study became bilingual because of greater WM capacity, nor is it likely that monolinguals in the present study became so because of smaller WM capacity. Differences in executive functions between bilinguals and monolinguals who are well matched on other environmental factors, such as socioeconomic status and intelligence, are therefore not likely due to genetic or biological factors, but rather to their linguistic experiences. We propose that second language experience has a positive effect on WM capacity, both between language groups (bilingual vs. monolingual; present study) and within language groups (low proficiency vs. high proficiency bilinguals; Linck et al., 2014).

The hypothesis that age would have an influence on the bilingual advantage for WM performance was partly supported. Specifically, the bilingual advantage appeared to be strongest in children. We hypothesized that age might moderate the effect of bilingual status on WM performance in three possible ways:

1. Largest effects would be seen in children and older adults because of ceiling performance amongst young adults (Bialystok et al., 2005).
2. A linear increase from youngest to oldest would be seen due to increasing practice managing two languages across the lifespan; or
3. Children would show the largest effects because of more cognitive demand when first learning a new language.

The present results support the third interpretation, with greatest group differences amongst children. Learning a new language is very cognitively demanding and, because of how plastic the brain is in children relative to other age groups (Johnston, 2009),

group differences in neural reorganization are more likely to appear, leading to observations of greater WM capacity for bilinguals.

It is interesting to note that type of task did not moderate the relationship between bilingual status and WM performance. It was possible that the relationship between bilingualism status and WM performance would be moderated by whether or not a verbal task was used due to subpar performance by bilinguals on tasks that required language processing. The present results did not support this conclusion, and instead support the idea that bilinguals generally outperform monolinguals on WM span tasks, regardless of the linguistic nature of the task. However, it is also possible that the WM tasks that we included in our analyses are not clearly defined as 'verbal' or 'non-verbal', and that this masks any group differences that one might expect. For example, the digit span task requires holding digits in mind and is considered to be a verbal task, but the actual linguistic demands are very minimal and may not be enough to differentiate these tasks from non-verbal span tasks. The reading span task on the other hand requires that participants read sentences and remember the last word of each sentence. This requires much more linguistic processing and it is possible that bilinguals are disadvantaged on such tasks because of their poorer language processing. The digit span and the reading span tasks are clearly quite different with respect to how much they tax language processing centers, yet both are classified as 'verbal'. It is possible that task type only moderates the bilingual advantage effect when more extreme language processing is required. The present study did not have the required number of observations to perform such an analysis, but future studies are encouraged to examine this possibility.

Another linguistic variable that was examined was whether the verbal tasks were performed in the L1 or the L2. Results revealed that when bilinguals performed the task in their L1, they performed better than monolinguals, but when they performed the task in their L2, they performed worse than monolinguals. This is in line with psycholinguistic studies showing slower lexical access for L2 compared to L1 production (for a review, see Kroll and Gollan, 2014). The L2 linguistic disadvantage likely masks the advantage of enhanced attentional WM processing for bilinguals. This exemplifies the importance of ensuring that both groups are performing the task in their dominant language; collapsing across this information masks important group effects.

There are a number of additional linguistic factors to take into consideration in the future that were not possible to analyze here due to lack of available data. For example, given Linck et al.'s (2014) findings, proficiency in the L2 likely influences the effect size between monolinguals and bilinguals, but studies often treat bilingualism/monolingualism as a dichotomous variable rather than a continuous one, even though bilingualism is not a categorical variable (Luk and Bialystok, 2013). Furthermore, proficiency is not always defined in the same way when it is reported. For example, some studies use objective measures of proficiency, and others use subjective self-reports, making comparisons across studies problematic. Although we did not have the available data to perform the analyses, we would expect that WM performance for high proficiency bilinguals would be even more distinct from monolinguals than low proficiency bilinguals; collapsing across these groups likely masks important group effects. Another factor worth investigating is whether bilinguals are bilingual by necessity or by choice. This often coincides with whether someone is an early or a late second language learner. Again, we

did not have the available data for such an analysis, but future studies are encouraged to explore this further.

There is a large literature on how bilinguals outperform monolinguals on a range of executive function tasks (for a review, see Bialystok et al., 2012), but until now, the effects of bilingualism on WM capacity were unclear. Results from the present meta-analysis indicate that WM capacity can be added to the list of executive function advantages that bilinguals have over monolinguals.

Declaration of Conflicting Interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Effect sizes for the monolinguals are based on the average scores for the two monolingual groups. Effect size for bilinguals is based on the average span score across both of their languages because they are simultaneous bilinguals.
2. The 'young adult' group here consists of individuals who are 14–22 years of age.
3. Monolingual and bilingual children were tested longitudinally at 3 different ages. We used the average score of the 3 time periods.
4. Older adults consisted of participants between the ages of 35–70.
5. We took the mean of 2 bilingual groups to compare with monolinguals.
6. We took the mean of 2 monolingual groups to compare with bilinguals. Statistics used to calculate the effect size were the statistics performed after controlling for age and verbal ability.
7. We took the mean of normally developing skilled and non-skilled readers. Only English WM tasks were examined for bilinguals because comparing bilinguals who were not fully proficient in Italian on an Italian task to monolingual Italians is not a fair comparison.
8. Statistics used were the statistics performed after controlling for age, SES, and vocabulary.
9. We took the mean of 2 monolingual groups to compare with bilinguals.
10. Both span length and accuracy scores were reported. Thus, effect sizes represent the average of length and accuracy.
11. The degrees of freedom are adjusted here to reflect that equal variance is not assumed. Corresponding adjusted *t*- and *p*-values are also reported.

References

Note. All studies included in the meta-analysis are shown in Table 1 and identified by an asterisk in the references section.

* Abu-Rabia S and Siegel LS (2002) Reading, syntactic, orthographic, and working memory skills of bilingual Arabic–English speaking Canadian children. *Journal of Psycholinguistic Research* 31: 661–78.

Adesope OO, Lavin T, Thompson T, and Ungerleider C (2010) A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research* 80: 207–45.

Adi-Japha E, Berberich-Artzi J, and Libnawi A (2010) Cognitive flexibility in drawings of bilingual children. *Child Development* 81: 1356–66.

Bak TH (2015) Beyond a simple 'yes' and 'no'. *Cortex: A journal devoted to the study of the nervous system and behavior* 73: 332–33.

Bialystok E (1999) Cognitive complexity and attentional control in the bilingual mind. *Child Development* 70: 636–44.

Bialystok E (2006) Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology* 60: 68–79.

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

Bialystok E and Poarch GJ (2014) Language experience changes language and cognitive ability. *Zeitschrift für Erziehungswissenschaft* 17: 433–46.

Bialystok E, Craik FI, Klein R, and Viswanathan M (2004) Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and Aging* 19: 290–303.

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

Bialystok E, Craik FI, and Luk G (2012) Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences* 16: 240–50.

Bialystok E, Martin MM, and Viswanathan M (2005) Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism* 9: 103–19.

Bialystok E, Poarch G, Luo L, and Craik FI (2014) Effects of bilingualism and aging on executive function and working memory. *Psychology and Aging* 29: 696.

* Biedroń A and Szczepaniak A (2012) Working memory and short-term memory abilities in accomplished multilinguals. *The Modern Language Journal* 96: 290–306.

* Blom E, Küntay AC, Messer M, Verhagen J, and Leseman P (2014) The benefits of being bilingual: Working memory in bilingual Turkish–Dutch children. *Journal of Experimental Child Psychology* 128: 105–19.

Blumenfeld HK and Marian V (2014) Cognitive control in bilinguals: Advantages in stimulus–stimulus inhibition. *Bilingualism: Language and Cognition* 17: 610–29.

* Bonifacci P, Giombini L, Bellocchi S, and Contento S (2011) Speed of processing, anticipation, inhibition and working memory in bilinguals. *Developmental Science* 14: 256–69.

* Cole RL and Pickering SJ (2010) Phonological and visual similarity effects in Chinese and English language users: Implications for the use of cognitive resources in short-term memory. *Bilingualism: Language and Cognition* 13: 499–512.

Costa A, Hernández M, and Sebastián-Gallés N (2008) Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition* 106: 59–86.

* D'Angiulli A, Siegel LS, and Serra E (2001) The development of reading in English and Italian in bilingual children. *Applied Psycholinguistics* 22: 479–507.

* Danahy K, Windsor J, and Kohnert K (2007) Counting span and the identification of primary language impairment. *International Journal of Language and Communication Disorders* 42: 349–65.

* Da Fontoura HA and Siegel LS (1995) Reading, syntactic, and working memory skills of bilingual Portuguese–English Canadian children. *Reading and Writing* 7: 139–53.

Donnelly S, Brooks PJ, and Homer BD (2015) *Examining the bilingual advantage on conflict resolution tasks: A meta-analysis*. In *Proceedings of the 37th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

* Engel de Abreu PM (2011) Working memory in multilingual children: Is there a bilingual effect? *Memory* 19: 529–37.

* Engel de Abreu PME, Cruz-Santos A, Tourinho CJ, Martin R, and Bialystok E (2012) Bilingualism enriches the poor: Enhanced cognitive control in low-income minority children. *Psychological Science* 23: 1364–71.

Engle RW (2002) Working memory capacity as executive attention. *Current Directions in Psychological Science* 11: 19–23.

Eriksson J, Vogel EK, Lansner A, Bergström F, and Nyberg L (2015) Neurocognitive architecture of working memory. *Neuron* 88: 33–46.

Field AP (2003) The problems in using fixed-effects models of meta-analysis on real-world data. *Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences* 2: 105–24.

Field AP (2005a) *Discovering statistics using SPSS*. 2nd edition. London: Sage.

Field AP (2005b) Is the meta-analysis of correlation coefficients accurate when population correlations vary? *Psychological methods* 10: 444–67.

Field AP and Gillett R (2009) *How to do a meta-analysis*. Retrieved from http://www.statistic-shell.com/meta_analysis (accessed October 2016).

Field AP and Gillett R (2010) How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology* 63: 665–94.

Gasquoine PG (2016) Effects of bilingualism on vocabulary, executive functions, age of dementia onset, and regional brain structure. *Neuropsychology*. Electronic publication ahead of print version. Published online: 12 May 2016 (doi: 10.1037/neu0000294).

Gold BT (2015) Executive control, brain aging and bilingualism. *Cortex* 73: 369.

Gold BT, Kim C, Johnson NF, Kryscio RJ, and Smith CD (2013) Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *The Journal of Neuroscience* 33: 387–96.

Gollan TH, Fennema-Nostestine C, Montoya RI, and Jernigan TL (2007) The bilingual effect on Boston naming test performance. *Journal of the International Neuropsychological Society* 13: 197–208.

Hedges LV and Vevea JL (1998) Fixed- and random-effects models in meta-analysis. *Psychological Methods* 3: 486–504.

Hernandez M, Costa A, Fuentes LJ, Vivas AB, and Sebastian-Galles N (2010) The impact of bilingualism on the executive control and orienting networks of attention. *Bilingualism: Language and Cognition* 13: 315–25.

Hunter JE and Schmidt FL (2000) Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment* 8: 275–92.

Hunter JE and Schmidt FL (2004) *Methods of meta-analysis: Correcting error and bias in research findings*. 2nd edition. Newbury Park, CA: Sage.

Ivanova I and Costa A (2008) Does the bilingualism hamper lexical access in speech production? *Acta Psychologica* 127: 277–88.

Johnston MV (2009) Plasticity in the developing brain: implications for rehabilitation. *Developmental Disabilities Research Reviews* 15: 94–101.

* Kaushanskaya M (2012) Cognitive mechanisms of word learning in bilingual and monolingual adults: The role of phonological memory. *Bilingualism: Language and Cognition* 3: 470–89.

* Kaushanskaya M, Blumenfeld HK, and Marian V (2011) The relationship between vocabulary and short-term memory measures in monolingual and bilingual speakers. *International Journal of Bilingualism* 15: 408–425.

* Kaushanskaya M, Gross M, and Buac M (2014) Effects of classroom bilingualism on task-shifting, verbal memory, and word learning in children. *Developmental Science* 17: 564–83.

Kousaie S and Taler V (2015) The bilingual advantage: Elusive but worth the effort? *Cortex* 73: 338–39.

Kroll JF and Gollan TH (2014) Speech planning in two languages: What bilinguals tell us about language production. In: Ferreira V, Goldrick M, and Miozzo M (eds) *The Oxford handbook of language production*. Oxford: Oxford University Press, pp. 165–81.

Kroll JF, Dussias PE, Bogulski CA, and Valdes Kroff JR (2012) Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In: Ross B (ed.) *The psychology of learning and motivation: Volume 56*. San Diego, CA: Academic Press, pp. 229–62.

* Leikin M and Tovli E (2014) Bilingualism and creativity in early childhood. *Creativity Research Journal* 26: 411–17.

Linck JA, Osthus P, Koeth JT, and Bunting MF (2014) Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin and Review* 21: 861–83.

* Ljungberg JK, Hansson P, Andrés P, Josefsson M, and Nilsson LG (2013) A longitudinal study of memory advantages in bilinguals. *PloS ONE* 8: e73029.

Luk G and Bialystok E (2013) Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology* 25: 605–21.

* Luo L, Craik FI, Moreno S, and Bialystok E (2013) Bilingualism interacts with domain in a working memory task: Evidence from aging. *Psychology and Aging* 28: 28–34.

Martin CD, Dering B, Thomas EM, and Thierry G (2009) Brain potentials reveal semantic priming in both the 'active' and the 'non-attended' language of early bilinguals. *NeuroImage* 47: 326–33.

* Morales J, Calvo A, and Bialystok E (2013) Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology* 114: 187–202.

* Mumtaz S and Humphreys G (2001) The effects of bilingualism on learning to read English: evidence from the contrast between Urdu–English bilingual and English monolingual children. *Journal of Research in Reading* 24: 113–34.

* Namazi M and Thordardottir E (2010) A working memory, not bilingual advantage, in controlled attention. *International Journal of Bilingual Education and Bilingualism* 13: 597–616.

* Nguyen TK and Astington JW (2014) Reassessing the bilingual advantage in theory of mind and its cognitive underpinnings. *Bilingualism: Language and Cognition* 17: 396–409.

Orlitzky M, Schmidt FL, and Rynes SL (2003) Corporate social and financial performance: A meta-analysis. *Organization Studies* 24: 403–41.

Overton RC (1998) A comparison of fixed-effects and mixed (random-effects) models for meta-analysis tests of moderator variable effects. *Psychological Methods* 3: 354–379.

Paap KR, Johnson HA, and Sawi O (2015) Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69: 265–78.

Pons F, Bosch L, and Lewkowicz DJ (2015) Bilingualism modulates infants' selective attention to the mouth of a talking face. *Psychological Science* 26: 490–98.

Prior A and MacWhinney B (2010) A bilingual advantage in task switching. *Bilingualism: Language and Cognition* 13: 253–62.

* Ransdell S, Barbier ML, and Niit T (2006) Metacognitions about language skill and working memory among monolingual and bilingual college students: When does multilingualism matter? *International Journal of Bilingual Education and Bilingualism* 9: 728–41.

* Ratiu I and Azuma T (2015) Working memory capacity: Is there a bilingual advantage? *Journal of Cognitive Psychology* 27: 1–11.

Rosenthal R (1979) The file drawer problem and tolerance for null results. *Psychological Bulletin* 86: 638–41.

Rosenthal R (1991) *Meta-analytic procedures for social research*. 2nd edition. Newbury Park, CA: Sage.

Rosenthal R (1995) Writing meta-analytic reviews. *Psychological Bulletin* 118: 183–92.

Rosenthal R and DiMatteo MR (2001) Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology* 52: 59–82.

* Sasisekaran J and Weisberg S (2013) The effects of cognitive: Linguistic variables and language experience on behavioural and kinematic performances in nonword learning. *Journal of Psycholinguistic Research* 42: 175–90.

* Smithson L and Nicoladis E (2013) Verbal memory resources predict iconic gesture use among monolinguals and bilinguals. *Bilingualism: Language and Cognition* 16: 934–44.

* Soliman AM (2014) Bilingual advantages of working memory revisited: A latent variable examination. *Learning and Individual Differences* 32: 168–77.

Soveri A, Laine M, Hämäläinen H, and Hugdahl K (2011) Bilingual advantage in attentional control: Evidence from the forced-attention dichotic listening paradigm. *Bilingualism: Language and Cognition* 14: 371–78.

Spivey MJ and Marian V (1999) Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science* 10: 281–84.

Thierry G and Wu YJ (2007) Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences* 104: 12530–35.

* Thorn AS and Gathercole SE (1999) Language-specific knowledge and short-term memory in bilingual and non-bilingual children. *The Quarterly Journal of Experimental Psychology: Section A* 52: 303–24.

Timmer K, Ganushchak LY, Ceusters I, and Schiller NO (2014) Second language phonology influences first language word naming. *Brain and Language* 133: 14–25.

Woumans E and Duyck W (2015) The bilingual advantage debate: moving toward different methods for verifying its existence. *Cortex* 73: 356–57.

Wu YJ and Thierry G (2010) Chinese–English bilinguals reading English hear Chinese. *The Journal of Neuroscience* 30: 7646–51.

Wu YJ and Thierry G (2012) Unconscious translation during incidental foreign language processing. *NeuroImage* 59: 3468–73.

Appendix 1. Equations used to convert F and t values to Pearson r correlation coefficients.

Conversion	Formula
t to r	$r = \sqrt{(t^2 / (t^2 + n - 2))}$, where n = sample size
F to r	$r = \sqrt{(F / (F + df_{error}))}$, where df = degrees of freedom in the error term

Note. Conversion to r is not possible when there are more than two groups being compared.