7-1-2021

The Bilingual Advantage in Children's Executive Functioning Is Not Related to Language Status: A Meta-Analytic Review

Cassandra J. Lowe
Western University

Isu Cho
Western University

Samantha F. Goldsmith
Western University

J. Bruce Morton
Western University, bmorton3@uwo.ca

Research Article

Psychological Science
2021, Vol. 32(7) 1115–1146

DOI: 10.1177/0956797621993108
www.psychologicalscience.org/PS
SAGE

The Bilingual Advantage in Children's Executive Functioning Is Not Related to Language Status: A Meta-Analytic Review

Cassandra J. Lowe[1,2], Isu Cho[1,3], Samantha F. Goldsmith[1,2], and J. Bruce Morton[1,2]

[1]Department of Psychology, The University of Western Ontario; [2]The Brain and Mind Institute, The University of Western Ontario; and [3]Department of Psychology, Brandeis University

Abstract

There is considerable debate about whether bilingual children have an advantage in executive functioning relative to monolingual children. In the current meta-analysis, we addressed this debate by comprehensively reviewing the available evidence. We synthesized data from published studies and unpublished data sets, which equated to 1,194 effect sizes from 10,937 bilingual and 12,477 monolingual participants between the ages of 3 and 17 years. Bilingual language status had a small overall effect on children's executive functioning (g = .08, 95% confidence interval = [.01, .14]). However, the effect of language status on children's executive functioning was indistinguishable from zero (g = −.04) after we adjusted for publication bias. Further, no significant effects were apparent within the executive-attention domain, in which the effects of language status have been hypothesized to be most pronounced (g = .06, 95% confidence interval = [−.02, .14]). Taken together, available evidence suggests that the bilingual advantage in children's executive functioning is small, variable, and potentially not attributable to the effect of language status.

Keywords

bilingual advantage, executive function, language status, meta-analysis, childhood development, open data, open materials

Received 7/6/20; Revision accepted 12/4/20

Questions concerning the bilingual advantage in children have become a critical focus in the broader debate about bilingual language status and its relation to executive functioning. According to the prevailing *bilingual-advantage hypothesis*, bilinguals become highly practiced at selecting and controlling attention owing to years of experience managing conflicts between competing phonological and lexical representations. Over the course of time, these practice effects generalize to problems outside the domain of language and contribute to a bilingual advantage in executive functioning (Bialystok, 2011, 2017; Bialystok et al., 2012; Kroll & Bialystok, 2013). A mounting number of null findings from large-scale comparisons of bilingual and monolingual adults (Nichols et al., 2020; Paap et al., 2015, 2017, 2018), however, have cast doubt on the bilingual-advantage account and shifted attention to studies involving children. Unlike adults, children do not perform at ceiling in executive-functioning tasks, which according to some researchers, leaves "more room for experience to push performance in a particular direction" (see Grundy et al., 2017, p. 43). Thus, although language-status effects might be small and difficult to detect in adults, they should be large and comparatively easy to detect in children (for a discussion, see Grundy et al., 2017).

In light of these claims, we conducted an exhaustive and comprehensive review of studies of the relationship between language status and executive functioning in children. In all, we included data from 136 peer-reviewed articles, 11 doctoral theses, and two unpublished data

Corresponding Author:
J. Bruce Morton, The University of Western Ontario, The Brain and Mind Institute, Department of Psychology
E-mail: bmorton3@uwo.ca

sets spanning the period from 1987 to November 2020 and that together reported findings from the study of 23,414 children (10,937 bilinguals and 12,477 monolinguals) between the ages of 3 and 17 years. We chose 3 years as the lower bound because it is around this age that children can complete measures of executive functioning that are comparable with tasks completed by older children. We chose 17 years as the upper bound because although age-related changes in executive functioning continue into early adulthood, children are furthest from a putative performance ceiling prior to the age of 18 years (Davidson et al., 2006).

Language-status effects were assessed on an exhaustive set of executive-functioning measures including operationalizations considered central to the bilingual-advantage hypothesis (e.g., Bialystok, 2017). In all, we included 1,194 separate effect sizes based on task-based measures of selective attention, flexibility, working memory, response inhibition, automatic attention (such as alerting and orienting), and planning, as well as global survey measures of executive functioning. We tested for an overall effect of language status on all measures of children's executive functioning aggregated together. We also tested for effects of language status within specific domains of executive functioning given that executive functioning is generally considered a multidimensional construct, and language-status effects have been hypothesized to be stronger in some domains of executive functioning than others (Bialystok, 2017; Bialystok et al., 2009; Carlson & Meltzoff, 2008). Specific effects of language status were therefore tested within nine different domains of executive functioning, each defined according to gold-standard definitions in the literature, and that included three domains of executive attention thought to be particularly germane to detecting the bilingual advantage (Bialystok, 2017).

In view of concerns surrounding the methodological rigor of studies examining the bilingual advantage in children, we examined the relationship between the magnitude of reported effects and the methodological quality of reporting studies (Morton, 2015). We applied an objective measure of study quality called the Appraisal Tool for Cross-Sectional Studies (AXIS), which evaluates studies according to their reported objective measurement of independent and dependent variables, use of representative samples, and transparent discussion of study limitations. Additionally, we examined specific indices of study quality that have been discussed in the literature, including the measured equivalence of groups and the control of socioeconomic status (SES).

Additional moderation analyses examined whether language-status measurement has implications for the assessment of language-status effects on children's executive functioning (DeLuca et al., 2019). We tested whether reported effect sizes varied depending on whether children's language status was measured by means of receptive vocabulary measures in both languages, language-use surveys, or an adult's nomination. We also compared effect sizes in bilingual children who showed full mastery of two languages with effect sizes in bilingual children who showed emerging proficiency but not mastery of a second language. These analyses were undertaken in response to calls for more nuanced characterizations of bilingualism and a recognition that bilingual language status is not all or nothing (Luk & Bialystok, 2013)

Finally, we tested for bias in the reporting of research findings by examining the relationship between the size and the precision of reported effects and testing whether there is a disproportionate number of large positive effects among studies reporting imprecise effect-size estimates. We then corrected for distortions in the literature by recalculating estimates of language-status effects on children's executive functioning while adjusting for bias.

Statement of Relevance

According to some accounts, bilingual language experience leads to a measurable advantage in executive functioning in children, a view that has gained substantial traction within the psychological sciences and the popular media. Critics, however, charge that empirical support for the bilingual advantage is weak because important confounding variables have not been consistently measured and controlled. The present meta-analysis synthesized data from 136 peer-reviewed articles, 11 doctoral theses, and two unpublished data sets, which equated to 1,194 effect sizes, and found a small effect of language status on children's executive functioning that was largely explained by moderating factors and bias. Therefore, the safest conclusion to be drawn from the current review is that the bilingual advantage in children's executive functioning is small, variable, and potentially not attributable to the effect of language status.

Primary Research Questions

There were four primary research questions. The first was, "Do bilingual children show an advantage in executive functioning relative to monolingual children?" The second question was, "Is the bilingual advantage in children's executive functioning more pronounced in some domains than others?" The third question was, "What additional variables moderate the relationship between language status and children's executive functioning?" And the fourth question was, "Is the literature on the bilingual advantage in children biased in favor of confirmatory over disconfirmatory evidence?"

Method

Literature search and study selection

A comprehensive search of PsycINFO, Scopus, and Web of Science databases was conducted using the search term *bilingual** combined with *executive function*, *executive control*, *cognition*, *cognitive*, *inhibitory control*, *inhibition*, *set shifting*, *task shifting*, *task switching*, *mental flexibility*, *working memory*, *updating*, *decision making*, *attentional control*, *attention*, *verbal fluency*, *temporal discounting*, or *delay discounting* (see Fig. 1 for a flowchart depicting the screening and inclusion process). To ensure transparency and reproducibility, we outlined the complete search documentation and Population, Intervention, Comparison, and Outcome (PICO) method in Table S1 in this study's OSF project (https://osf.io/jv7wt/). Additionally, reference lists of relevant articles and pertinent reviews were manually searched for additional articles. A search of the gray literature was conducted using Web of Science, PsycINFO, PsyArXiv Preprints, Google Scholar, and ProQuest Dissertations Thesis databases. The first search was conducted in July 2018 and then updated in November 2020. No limits were placed on publication date or language. Decisions to include or exclude studies were based on reviews of the abstract and full text of each article. For details regarding the inclusion and exclusion criteria adopted for the current review, see "Supplemental Methods" at https://osf.io/jv7wt/.

Coding procedure

Executive-function domains. To guide the classification of individual measures into distinct executive-function domains, we defined executive functioning as a set of higher order cognitive processes that support children's goal-directed behavior (Zelazo et al., 1997, 2003). These processes include planning, flexibility, decision-making, working memory, and selection. Domain boundaries were refined to ensure that tasks hypothesized to be the locus of language-status effects were aggregated together in the same domain and labeled as such (Bialystok, 2017). The result was nine different executive-function domains, including three "executive-attention" domains (i.e., selection, nonverbal working memory, flexibility) hypothesized to be the locus of language-status effects (Bialystok, 2017). A full list of domains and associated measures appears in Table 1; definitions appear at https://osf.io/jv7wt/.

Meta-analytic procedure and analyses. The data, R code for computing all analyses, and additional details on all aspects of the analysis are available at https://osf.io/jv7wt/.

Effect-size calculation. For studies that reported means and standard deviations, effect sizes were transformed to Hedges's g. For studies that did not report the means and standard deviations, effect sizes were calculated using F, t, or p values and converted to Hedges's g. Effect sizes were coded such that positive effect-size values reflect a bilingual advantage and negative effect-size values reflect a bilingual disadvantage. Unusually high effect-size estimates were observed ($g = 34.92$) for the data obtained from Laloi et al. (2017), and therefore, this article was excluded from all analyses.

Multilevel model. Individual effect sizes cannot be treated as statistically independent because individual effects can originate from different comparisons within experiments, different experiments within articles, or different articles from the same research group. Dependencies of this kind can produce artificially narrow confidence intervals (CIs) and artificially small estimates of the standard error of the effect (Van den Noortgate et al., 2013, 2015). Therefore, following the removal of outliers (six effects—0.005% of the data—whose absolute g value was greater than 3), we estimated the influence of several different dependencies on effect-size variance using a multilevel model containing separate levels for comparisons within experiments, experiments within studies, and studies within research groups. Akaike information criterion (AIC) values and likelihood-ratio tests (see Table 2) indicated that the addition of each level significantly improved model fit (for profile likelihood plots, see Fig. S1 in the Supplemental Material). The final model accounted for approximately 42% of variance (intraclass correlation coefficient = .42) in reported effect sizes and provided a better fit than any of the reduced models. Additional levels did not significantly improve model fit.

Moderation analysis. Residual effect sizes from the multilevel model were statistically heterogeneous with respect to both the overall effect of language status on children's executive functioning and the effect of language status within specific executive-function domains (see the Results section). Moderation analysis therefore tested whether the effect of language status on children's executive functioning was moderated by other variables, including (a) executive-function domain; (b) participant

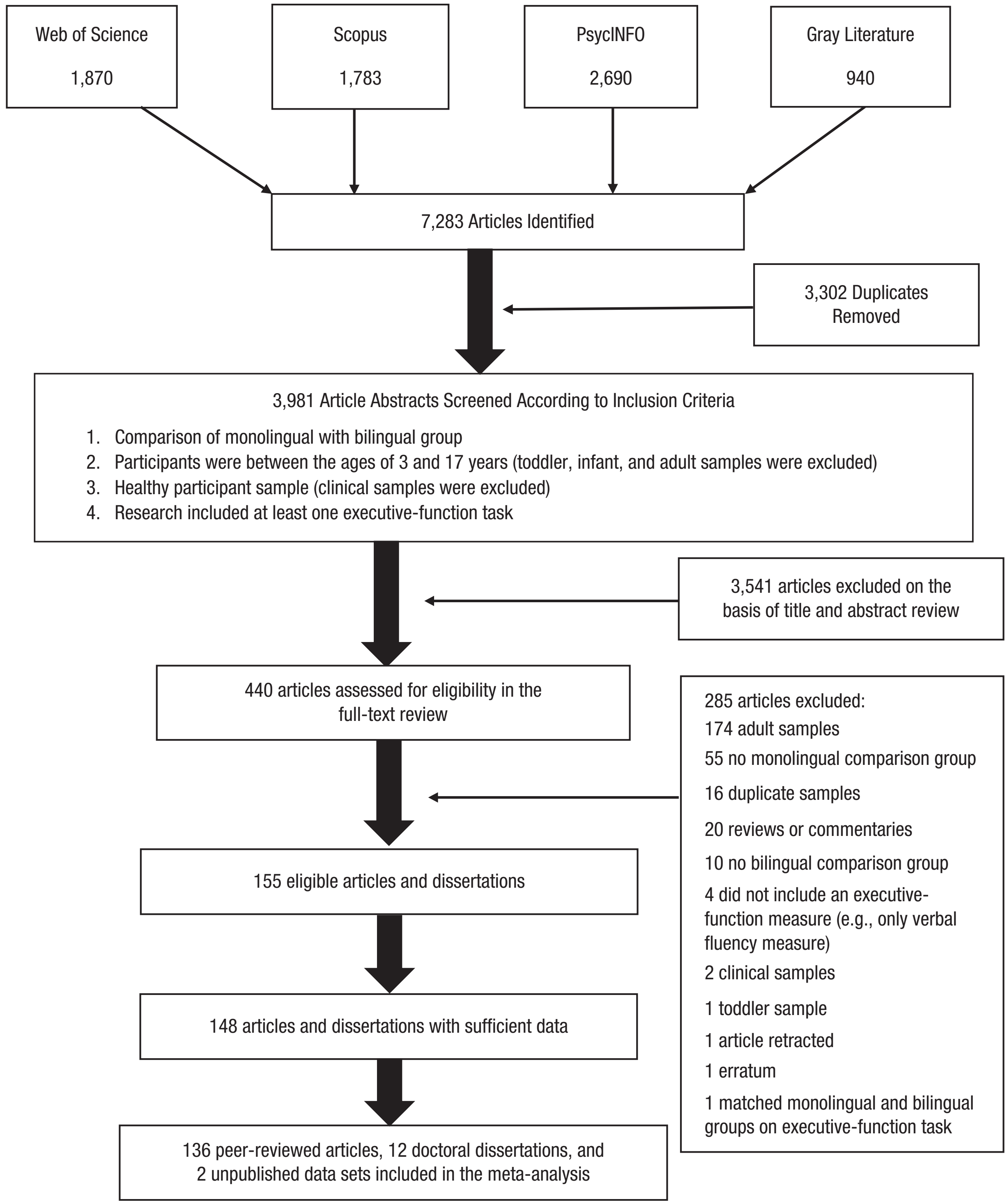

Fig. 1. Flowchart depicting the article screening and inclusion process.

characteristics, including age and degree of bilingualism (balanced, emergent, or unclassifiable); (c) study quality, including an overall assessment of study quality using the AXIS (Downes et al., 2016), measured equivalence of groups (yes or no), and reported objective measurement of SES (yes or no); (d) measure of language status (nomination, survey instrument, or receptive vocabulary test); (e) geographic origin of the sample (North America,

Table 1. Overview of Executive-Function Domains and Tasks Included in Each Domain

Domain and category	Example
Executive attention	
Selection	Stroop (sun/moon, grass/snow, happy/sad, day/night task, red/blue), Attention Network Task, Simon task, soccer task, flanker, flanker Attention Network Task (executive-functioning condition), bivalent shape task, opposite world
Flexibility	Trail Making Test-B, color-shape task, global local task, opposite world same word, dual-modality switching task (visual & auditory), faces task (switching condition), choice response time, pirate task, reverse-categorization task, creature-counting task, teddy bear test, tapping task (switch conditions), Wisconsin Card Sorting Task, Something's the Same, Dimensional Change Card Sorting Task
Working memory–nonverbal	Picture working memory task, Corsi block (forward, backward), maze memory, hand-position imitation, dot-matrix task, visually cued recall, odd one out, Mr. X task, frog matrices task, symbol search, block recall, visual pattern span, anticipation task (nonverbal)
Other executive function domains	
Working memory–verbal	Wechsler memory scale (memory story), reading span, listening span, counting recall, sentence recall, *n*-back, digit span, tapping task (match condition), word span, choice (auditory, visual conditions), Wechsler Intelligence Scale for Children (block design, digit span, arithmetic), Kaufman Assessment Battery for Children, Behaviour Rating Inventory of Executive Function (working memory subtest), houses, pick the picture, spy training
Response inhibition	Go/no go; Luria tapping (or pencil tapping); faces task (suppression, inhibitory control condition); continuous performance task (auditory, visual condition); statue task; stop-signal task; walk, don't run task; candy test, head-toes-shoulder task
Automatic attention	Moving word task, sky search task, pair-cancellation subtest, cancelation subtest, Weschler Intelligence Scale (verbal visual attention), Attention Network Task (alerting, orienting, overall; central/double cue), NEPSY (attention)
Reward-based learning/decision-making	Gift delay
Planning	Tower of Hanoi, Tower of London, NEPSY (Tower subtest)
Global executive functions	Global Behaviour Rating Inventory of Executive Function and NEPSY scores

Europe, East Asia, Middle East, or mixed); and (f) year of study publication. Details concerning the definition and measurement of moderator variables appear at https://osf.io/jv7wt/.

Analysis of publication bias. Publication bias was assessed by means of funnel plots that display effect-size estimates against the standard error of effect-size estimates (see Figs. 2 and 3). In the absence of publication bias, funnel plots should be symmetrical around the mean effect, and effect sizes should be more closely distributed around the mean effect as precision increases. Funnel-plot asymmetry suggests selective reporting of evidence and was evaluated by means of Egger's regression test.

Reestimate of language-status effects adjusting for publication bias. To correct for distortions introduced by the selective reporting of evidence, we estimated bias-adjusted estimates of language-status effects using the precision-effect test (PET) and PET with standard errors (PEESE; Stanley & Doucouliagos, 2014). Effect sizes were regressed onto their standard errors in a weighted

Table 2. Model-Fit Indices, Comparison Statistics, Estimated Effect Sizes (*g*s), and Variance Components for the Four-Level Multilevel Model

Model level	Added higher level	Model-fit index		Model-comparison index				Variance		
		AIC	Log likelihood	Model	LRT	*p*	*g*	σ^2_1	σ^2_2	σ^2_3
One		3,570.93	−1,784.47				.10			
Two	Participants	1,769.51	−882.76	1 vs. 2	1,803.43	< .001	.12	.095		
Three	Study	1,757.66	−875.83	2 vs. 3	13.85	< .001	.11	.082	.018	
Four	Research group	1,742.86	−867.43	3 vs. 4	16.80	< .001	.08	.041	.040	.017

Note: AIC = Akaike information criterion; LRT = likelihood-ratio test.

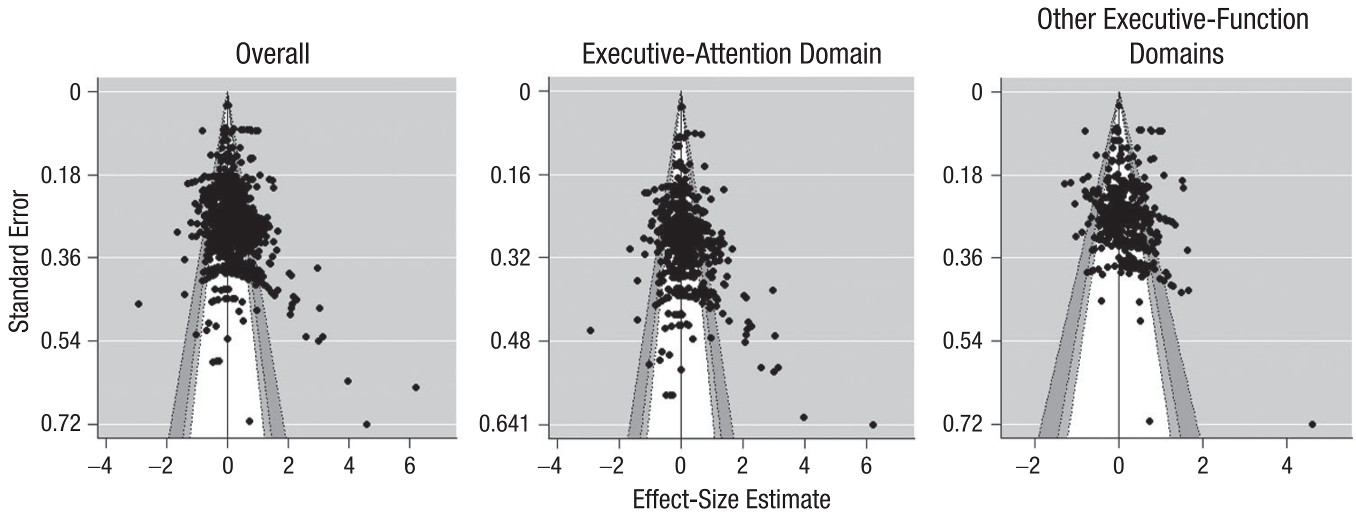

Fig. 2. Contour-enhanced funnel plots for the overall effect, executive-attention domain, and other executive-function domains. Effect-size estimates are plotted against the standard error of effect-size estimates. Dots represent individual studies. Shading in the triangular regions indicates significance (white area: p = .10, light gray area: p = .05, dark-gray area: p = .01).

least-squares regression model (i.e., PET) to test whether the bias-adjusted average effect size was distinct from zero. A significant and positive association between effect sizes and their standard errors is taken to suggest that studies with low precision report larger effects, and therefore, the overall effect may be potentially biased. The intercept of this model reflects the estimate of the true and unbiased effect in a hypothetical study with no bias or error (Stanley & Doucouliagos, 2014). Next, as recommended by Stanley and Doucouliagos, if the PET revealed a significant and positive association between effect sizes and their standard errors (i.e., the average bias-adjusted effect size, or intercept, was distinct from zero), then the PET was followed up by a PEESE to determine whether the average bias-adjusted effect size was statistically distinct from zero. The PEESE involves using variance as a predictor in the weighted least-squares regression model (Stanley & Doucouliagos, 2014).

Results

The final data set consisted of 1,194 effect sizes (1,105 following removal of outliers) drawn from 136 peer-reviewed publications, 11 doctoral dissertations, and two unpublished data sets (Cho et al., 2021; Goldsmith, 2021). Descriptive statistics for all included studies are presented in Table 3, and individual effect-size estimates are presented in Figure S2 in the Supplemental Material. Additional details can be found at https://osf.io/jv7wt/.

Results of the multilevel model revealed a small effect of language status across all domains of executive functioning that favored bilingual children (k = 1,188, g = .08, 95% CI = [.01, .14], p = .017). The effect was unchanged by the inclusion of outliers. Variability in reported effects was linked to dependencies in the data: Effects varied as a function of research group (σ^2 = .04) and studies within research group (σ^2 = .04). The prediction interval of the true effect size indicated that in 95% of populations, the true effect size would fall between an approximate range of −.54 and .70. However, even after we controlled for these dependencies, there was substantial variability in effect-size estimates between individual studies (Q = 4,539.44, p < .001, τ^2 = .10, I^2 = 67.27), and 67.27% of the total between-studies variability was attributable to true heterogeneity rather than sampling error alone. Subsequent moderation analyses therefore examined sources of unexplained effect-size variability.

Given a concern that the bilingual advantage may not be apparent in children who learned their second language through immersion schools or other educational programs, we evaluated the overall effect with these studies removed. Results indicated that the overall effect size was unchanged when these samples were removed from analyses (k = 1,053, g = .08, 95% CI = [.01, .14], p = .016).

Moderator analyses

Executive-function domain. Executive-function domain moderated the effect of language status on children's executive functioning, as reflected by a test for whether the moderator explained heterogeneity in the data, $Q_M(9)$ = 31.27, p < .001. Similar to the overall effect,

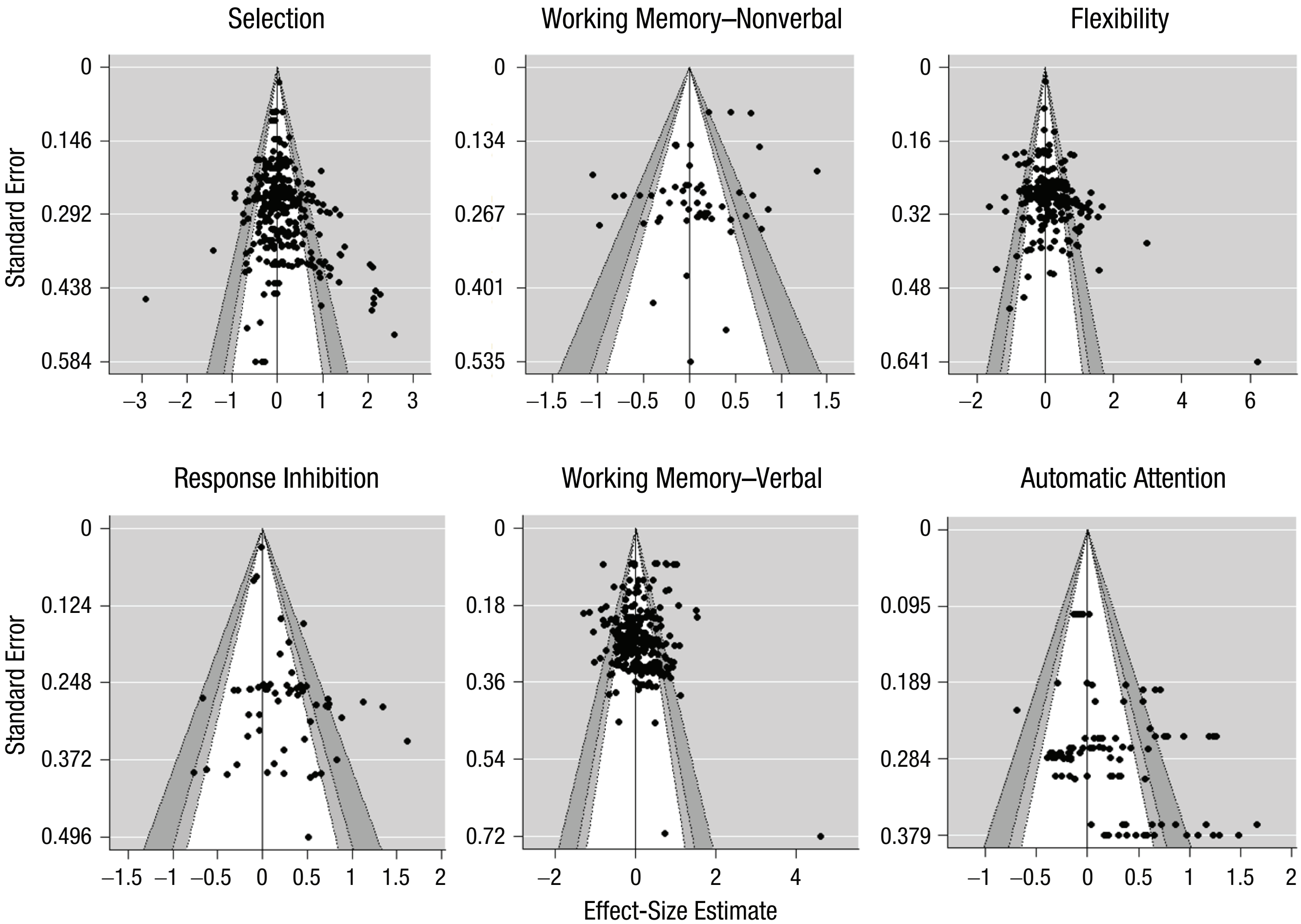

Fig. 3. Contour-enhanced funnel plots for each executive-function domain with sufficient data. Effect-size estimates are plotted against the standard error of effect-size estimates. Dots represent individual studies. Shading in the triangular regions indicates significance (white area: $p = .10$, light gray area: $p = .05$, dark-gray area: $p = .01$).

variability in reported effects after analyses accounted for executive-function domain was largest within research group ($\sigma^2 = .04$) and studies within research group ($\sigma^2 = .04$). Language-status effects were evident in the domain of response inhibition ($k = 57$, $g = .17$, 95% CI = [.05, .30], $p = .008$; see Fig. 4) but indistinguishable from zero in all other domains, including all three domains of executive attention (see Table 3). Effect-size estimates remained indistinguishable from zero when the multilevel model was run only on effects from the three executive-attention domains ($k = 694$, $g = .06$, 95% CI = [−.02, .14], $p = .118$, $\tau^2 = .12$, $I^2 = 70.39$). Effect sizes remained heterogeneous even after analyses accounted for the moderating influence of executive-function domain, as revealed by a test for residual heterogeneity, $Q_E(1178) = 4{,}495.15$, $p < .001$, $\tau^2 = .10$, $I^2 = 67.18$. Substantial heterogeneity was apparent within the domains of response inhibition, verbal working memory, automatic attention, and domains encompassing executive attention (see Table 4); therefore, these domains were included in subsequent moderator analyses.

Verbal versus nonverbal tasks. Use of verbal versus nonverbal tasks moderated the overall language-status effect on executive functioning, $Q_M(3) = 54.22$, $p < .001$. Specifically, a bilingual advantage was evident in studies using verbal tasks ($k = 331$, $g = .12$, 95% CI = [.05, .19], $p = .001$) but not nonverbal tasks ($k = 790$, $g = .06$, 95% CI = [−.002, .13], $p = .057$). A bilingual disadvantage was observed in studies that used tasks with both verbal and nonverbal stimuli or output ($k = 56$, $g = −.24$, 95% CI = [−.35, −.12], $p < .001$). There were, however, a limited number of effect sizes under this category, so these results should be interpreted with some caution. Because verbal tasks were domain specific (verbal working memory and selection), we did not conduct domain-level analyses for this moderator.

Table 3. Descriptive Statistics for Studies Included in the Meta-Analysis

Study and independent subgroup/comparison group	Mono-linguls (*n*)	Bi-linguls (*n*)	Mean age (years)	First language	Second language	Geographic location	Bilingual proficiency	Method used to assess language status[c]	Equivalence testing or matched samples[d]	Measure of SES[e]	Study-quality score
Abdelgafar & Moawad, 2015	25	25	8.8	Arabic	English	Saudi Arabia	Bilingual	RV	None	Y	10
Abu-Rabia & Siegel, 2002	45	18	11.5	Arabic	English	Canada	Bilingual	SR	Matched samples	N	10
Antón et al., 2014	180	180	9.8	Spanish	Basque	Spain	Bilingual	LU	Equivalence testing	Y	11
Antoniou et al., 2016[a]	25	44	7.5	Standard Modern Greek	Cypriot Greek	Cyprus	Bilingual	RV	Equivalence testing	Y	12
Arizmendi et al., 2018	167	80	7.8	Spanish	English	U.S.	Bilingual	RV	None	Y	14
Arredondo, 2017[a]	26	26	8.1	Spanish	English	U.S.	Bilingual	RV	Matched samples	Y	16
Arredondo et al., 2017[a]	14	13	9.9	Spanish	English	U.S.	Bilingual	RV	Equivalence testing	Y	12
Asadollahpour, 2015	70	70	7.6	Persian	Baluchi	Iran	Unclear	LU	None	N	9
Barac & Bialystok, 2012[a]											
Chinese-English bilinguals	26	30	5.9	Chinese	English	Canada	Bilingual	LU	Equivalence testing	Y	15
French-English bilinguals	26	28	5.9	French	English	Canada	Bilingual	LU	Equivalence testing	Y	15
Spanish-English bilinguals	26	20	5.9	Spanish	English	Canada	Bilingual	LU	Equivalence testing	Y	15
Barac et al., 2016[a]	37	25	5.3	Mix	English	Canada	Bilingual	LU	Equivalence testing	Y	15
Barbosa et al., 2019											
English monolinguals	40	40	5.2	Mandarin	English	Canada	Bilingual	RV	Equivalence testing	N	15
Mandarin monolinguals	38	40	5.3	Mandarin	English	Canada (bilinguals), China (monolinguals)	Bilingual	RV	Equivalence testing	N	15
Bastian et al., 2018[a]	40	23	5.1	German	English	Germany	Bilingual	RV	Equivalence testing	Y	15
Bialystok, 1999											
Younger participants	15	15	4.2	Mandarin	English	Canada	Bilingual	SR	None	N	13
Older participants	15	15	5.5	Mandarin	English	Canada	Bilingual	SR	None	N	13
Bialystok, 2010											
Study 1	25	26	6.0	Mix	English	Canada	Bilingual	LU	None	N	8
Study 2	25	25	5.8	Mix	English	Canada	Bilingual	LU	None	N	8
Study 3	25	25	6.1	Mix	English	Canada	Bilingual	LU	None	N	8
Bialystok, 2011[a]	32	31	8.6	Mix	English	Canada	Bilingual	LU	Equivalence testing	Y	12
Bialystok et al., 2010											
Younger Children/ English monolinguals	40	27	3.6	Mix	English	Canada	Bilingual	LU	None	N	8

(continued)

Table 3. *(continued)*

Study and independent subgroup/comparison group	Mono-linguls (*n*)	Bi-linguls (*n*)	Mean age (years)	First language	Second language	Geographic location	Bilingual proficiency	Method used to assess language status[c]	Equivalence testing or matched samples[d]	Measure of SES[e]	Study-quality score
Older Children/English monolinguals	29	29	4.6	Mix	English	Canada	Bilingual	LU	None	N	8
Younger Children/ French monolinguals	20	27	3.6	Mix	French	France (monolinguals), Canada (bilinguals)	Bilingual	LU	None	N	8
Older Children/ French monolinguals	17	29	4.6	Mix	French	France (monolinguals), Canada (bilinguals)	Bilingual	LU	None	N	8
Bialystok et al., 2009											
Bilinguals in Canada	30	30	8.5	Mix	English	Canada	Bilingual	LU	None	N	8
Bilinguals in India	30	30	8.6	Tamil or Telugu	English	India (B) Canada M	Bilingual	LU	None	N	8
Bialystok & Martin, 2004											
Study 1	36	31	4.9	Cantonese	English	Canada	Bilingual	SR	None	N	12
Study 2	15	15	4.8	French	English	Canada	Bilingual	RV	None	N	12
Study 3	27	26	4.3	Cantonese or Mandarin	English	Canada	Bilingual	SR	None	N	12
Bialystok & Senman, 2004											
4-year-old children	33	22	4.3	Mix	English	Canada	Bilingual	SR	None	N	12
5-year-old children	19	21	5.6	Mix	English	Canada	Bilingual	SR	None	N	12
Bialystok & Shapero, 2005											
Study 1	24	24	5.9	Mix	English	Canada	Bilingual	SR	None	N	12
Study 2	27	26	5.6	Mix	English	Canada	Bilingual	SR	None	N	12
Blom & Boerma, 2017[a]	30	30	5.96	Mix	Dutch	Netherlands	Bilingual	LU	Matched samples	Y	12
Blom et al., 2017[a]											
Frisian bilinguals	44	44	6.8	Frisian	Dutch	Netherlands	Bilingual	LU	Matched samples	Y	12
Limburgish bilinguals	44	44	6.8	Limburgish	Dutch	Netherlands	Bilingual	LU	Matched samples	Y	12
Polish bilinguals	44	44	6.8	Polish	Dutch	Netherlands	Bilingual	RV	Matched samples	Y	12
Boerma et al., 2017[a]	32	32	5.9	Mix	Dutch	Netherlands	Bilingual	LU	Matched samples	N	13
Bonifacci et al., 2011	18	18	9.4	Italian	Mix	Italy	Bilingual	SR	Matched samples	N	13
Bosman & Janssen, 2017	48	38	7.3	Turkish	Dutch	Netherlands	Bilingual	RV	None	Y	13
Brito & Noble, 2018[a]	281	281	13.5	Mix	English	U.S.	Emergent bilingual	SR	Matched samples	Y	15
Buac et al., 2016[a]	36	46	6.3	Spanish	English	U.S.	Bilingual	RV	Equivalence testing	Y	14
Buac & Kaushanskaya, 2014	46	39	8.4	Spanish	English	U.S.	Bilingual	RV	None	Y	14
Burch, 1987											
English monolinguals	26	59	9.0	English	Spanish	U.S.	Unclear	SR	Matched samples	N	16
Spanish monolinguals	29	59	9.0	Spanish	English	U.S.	Unclear	SR	Matched samples	N	16

(continued)

Table 3. *(continued)*

Study and independent subgroup/comparison group	Mono-linguls (n)	Bi-linguls (n)	Mean age (years)	First language	Second language	Geographic location	Bilingual proficiency	Method used to assess language status[c]	Equivalence testing or matched samples[d]	Measure of SES[e]	Study-quality score
Calvo & Bialystok, 2014											
Working class	20	44	6.7	Mix	English	Canada	Bilingual	LU	Equivalence testing	Y	10
Middle class	46	65	6.7	Mix	English	Canada	Bilingual	LU	Equivalence testing	Y	10
Cape et al., 2018[a]											
Bilingual	17	15	9.6	English	Gaelic	Scotland	Bilingual	LU	Equivalence testing	Y	15
Emergent bilingual	13	13	9.4	English	Gaelic	Scotland	Emergent bilingual	LU	Equivalence testing	Y	15
Carlson & Meltzoff, 2008[a]											
Bilingual	17	12	6.1	Spanish	English	U.S.	Bilingual	LU	Equivalence testing	Y	14
Emergent bilingual	17	21	6.0	English	Spanish or Japanese	U.S.	Emergent bilingual	LU	Equivalence testing	Y	14
Chan, 2004	29	31	4.4	English	Chinese	Canada	Bilingual	RV	Equivalence testing	N	13
Cho et al., 2021[a]											
Canadian monolinguals	34	32	4.6	Korean	English	Canada	Bilingual	RV	Equivalence testing	Y	17
Korean monolinguals	33	32	4.5	Korean	English	Canada (bilinguals), Korea (monolinguals)	Bilingual	RV	Equivalence testing	Y	17
Choi et al., 2018	475	210	4.5	English	Spanish	U.S.	Bilingual	SR	None	Y	14
Christoffels et al., 2015	29	30	17.2	Dutch	English	Netherlands	Emergent bilingual	SR	Equivalence testing	N	15
Chung-Fat-Yim, 2019[a]	33	32	16.1	English	Mix	Canada	Bilingual	LU	Equivalence testing	Y	12
Climie, 2008	29	39	NR	French	English	Canada	Bilingual	SR	None	N	13
Cockcroft, 2016[a]	67	53	6.7	isiZulu or isiXhosa	English	South Africa	Bilingual	SR	Equivalence testing	Y	14
Cockcroft & Alloway, 2012[a]											
South Africa monolinguals	42	37	7.2	Nguni or Sotho	English	South Africa	Bilingual	SR	Equivalence testing	Y	12
UK monolinguals	40	37	7.9	Nguni or Sotho	English	UK (monolinguals), South Africa (bilinguals)	Bilingual	SR	Equivalence testing	Y	12
Cottini et al., 2015											
Grade 3	25	28	8.2	Italian	German	Italy	Bilingual	LU	None	N	14
Grade 5	24	27	10.3	Italian	German	Italy	Bilingual	LU	None	N	14

(continued)

Table 3. *(continued)*

Study and independent subgroup/comparison group	Monolinguls (n)	Bilinguls (n)	Mean age (years)	First language	Second language	Geographic location	Bilingual proficiency	Method used to assess language status[c]	Equivalence testing or matched samples[d]	Measure of SES[e]	Study-quality score
Da Fontoura & Siegel, 1995	57	37	10.5	Portuguese	English	Canada	Unclear	RV	Matched samples	Y	9
Dahlgren et al., 2017	14	14	4.5	Serbo-Croatian	Swedish	Sweden	Unclear	SR	Matched samples	N	10
Danahy et al., 2007	50	22	10.3	Spanish	English	U.S.	Bilingual	RV	None	N	13
D'Angiulli et al., 2001											
9- to 10-year-old skilled readers/English monolinguals	37	23	NR	English	Italian	Canada	Bilingual	SR	Matched samples	N	10
11- to 13-year-old skilled readers English/English Canadian monolinguals	64	39	NR	English	Italian	Canada	Bilingual	SR	Matched samples	N	10
9- to 10-year-old skilled readers Italian/Italian monolinguals	25	23	NR	English	Italian	Canada	Bilingual	SR	Matched samples	N	10
11- to 13-year-old skilled readers/Italian monolinguals	42	39	NR	English	Italian	Canada	Bilingual	SR	Matched samples	N	10
Dell'Armi, 2015	19	30	8.1	Spanish	French	France and Spain	Bilingual	RV	Matched samples	Y	11
De Sousa, 2012	30	30	9.32	Afrikaans	English	South Africa	Unclear	SR	Equivalence testing	N	13
De Sousa et al., 2010	30	30	9.85	Afrikaans	English	South Africa	Emergent bilingual	SR	Equivalence testing	N	13
Diaz & Farrar (2018a)[a]	33	32	4.17	Spanish	English	U.S.	Bilingual	RV	Equivalence testing	Y	17
Diaz & Farrar (2018b)	38	40	3.98	Spanish	English	U.S.	Bilingual	LU	Equivalence testing	N	16
Dick et al., 2019	2784	1740	10.0	English	Mix	U.S.	Bilingual	SR	None	Y	11
Duñabeitia et al., 2014	252	252	10.5	Spanish	Basque	Spain	Bilingual	LU	Matched samples	N	10
Ebert et al., 2019[a]	27	27	4.5	Spanish	English	U.S.	Bilingual	RV	Equivalence testing	Y	15
Engel de Abreu, 2011[a]	22	22	6.3	Luxembourgish	Mix	Luxembourg	Bilingual	LU	Matched samples	Y	13
Engel de Abreu et al., 2013[a]											
Luxembourgish monolinguals	20	20	7.1	Portuguese	Luxembourgish	Luxembourg	Bilingual	RV	Matched samples	Y	14
Brazilian Portuguese monolinguals	20	20	7.1	Portuguese	Luxembourgish	Luxembourg and Brazil	Bilingual	RV	Matched samples	Y	14
Engel de Abreu et al., 2014[a]	33	33	8.2	Portuguese	Luxembourgish	Luxembourg	Bilingual	RV	Matched samples	Y	16

(continued)

Table 3. *(continued)*

Study and independent subgroup/comparison group	Monolinguls (n)	Bilinguls (n)	Mean age (years)	First language	Second language	Geographic location	Bilingual proficiency	Method used to assess language status[c]	Equivalence testing or matched samples[d]	Measure of SES[e]	Study-quality score
Engel de Abreu et al., 2012[a]	40	40	8.2	Luxembourgish	Portuguese	Portugal and Luxembourg	Bilingual	RV	Matched samples	Y	12
Esposito & Baker-Ward, 2013											
Kindergarten	16	18	6.0	English	Spanish	U.S.	Emergent bilingual	SR	None	N	16
Grade 2	22	17	8.3	English	Spanish	U.S.	Emergent bilingual	SR	None	N	16
Grade 4	17	23	10.2	English	Spanish	U.S.	Emergent bilingual	SR	None	N	16
Esposito et al., 2013	25	26	4.2	Spanish	English	U.S.	Bilingual	LU	None	N	16
Foy & Mann, 2014[a]	30	30	5.3	Spanish	English	U.S.	Bilingual	LU	Matched samples	Y	13
Gangopadhyay et al., 2016[a]	42	42	9.3	Spanish	English	U.S.	Bilingual	RV	Matched samples	Y	15
Gangopadhyay et al., 2019[a]	38	38	9.4	Spanish	English	U.S.	Bilingual	RV	Matched samples	Y	15
Gangopadhyay et al., 2018[a]	44	44	119	Spanish	English	U.S.	Bilingual	RV	Matched samples	Y	12
Garratt & Kelly, 2008	27	27	7.2	Mix	English	UK	Emergent bilingual	SR	Equivalence testing	Y	11
Goldman et al., 2014											
Younger children/ English monolinguals	32	40	4.8	English	Mix	U.S.	Unclear	LU	None	Y	17
Younger children/non-English monolinguals	20	40	4.9	Mix	Mix	U.S.	Unclear	LU	None	Y	17
Older children/English Monolinguals	32	40	4.8	English	Mix	U.S.	Unclear	LU	None	Y	17
Older children/non-English monolinguals	20	40	4.9	Mix	Mix	U.S.	Unclear	LU	None	Y	17
Goldsmith, 2021[b]											
Canada	14	135	10.3	English	French	Canada	Bilingual	LU		Y	
China	104	162	11.8	Mandarin	English	China	Bilingual	LU		Y	
Lebanon	3	190	10.8	Arabic	English	Lebanon	Bilingual	LU		Y	
Gonzalez, 2017[a]	49	40	8.5	English	Spanish	U.S.	Bilingual	LU	Matched samples	Y	15
Gonzalez-Barrero & Nadig, 2019[a]	13	13	8.3	French	English	Canada	Bilingual	RV	Equivalence testing	Y	16
Goriot et al., 2018											
4- to 5-year-olds	38	40	4.9	Dutch	English	Netherlands	Emergent bilingual	LU	None	N	13
8- to 9-year-olds	34	38	9.0	Dutch	English	Netherlands	Emergent bilingual	LU	None	N	13
11- to 12-year-olds	26	28	12.	Dutch	English	Netherlands	Emergent bilingual	LU	None	N	13
Goriot et al., 2016[a]											
Dutch-German bilinguals	23	25	9.2	Dutch	German	Netherlands	Bilingual	RV	Equivalence testing	Y	12
Dutch-Turkish bilinguals	23	23	9.0	Dutch	Turkish	Netherlands	Bilingual	RV	Equivalence testing	Y	12

(continued)

Table 3. *(continued)*

Study and independent subgroup/comparison group	Mono-linguls (n)	Bi-linguls (n)	Mean age (years)	First language	Second language	Geographic location	Bilingual proficiency	Method used to assess language status[c]	Equivalence testing or matched samples[d]	Measure of SES[e]	Study-quality score
8- to 9-year-olds	34	38	9.0	Dutch	English	Netherlands	Bilingual	RV	Equivalence testing	Y	12
11- to 12-year-olds	26	28	12.	Dutch	English	Netherlands	Bilingual	RV	Equivalence testing	Y	12
Grundy & Keyvani Chahi, 2017[a]	40	40	7.3	English	Mix	Canada	Bilingual	LU	Equivalence testing	Y	10
Haft et al., 2019[a]	16	24	5.6	English	Mix	U.S.	Bilingual	RV	None	Y	15
Hansen et al., 2016[a]											
Grade 2	19	19	NR	Spanish	English	Spain	Emergent bilingual	SR	Matched samples	Y	12
Grade 3	21	21	NR	Spanish	English	Spain	Emergent bilingual	SR	Matched samples	Y	12
Grade 5	21	21	NR	Spanish	English	Spain	Emergent bilingual	SR	Matched samples	Y	12
Grade 8	15	15	NR	Spanish	English	Spain	Emergent bilingual	SR	Matched samples	Y	12
Harvey, 2012	56	42	4.2	Mix	English	U.S.	Emergent bilingual	RV	None	N	14
Hutchison, 2012[a]											
Bilinguals	33	17	6.4	English	Mix	U.S.	Bilingual	RV	Equivalence testing	Y	12
Emergent bilinguals	33	29	6.1	English	Mix	U.S.	Emergent bilingual	RV	Equivalence testing	Y	12
Iarocci et al., 2017[a]	24	59	9.5	English	Mix	Canada	Unclear	LU	Equivalence testing	Y	13
Jaekel et al., 2019	95	242	9.5	Turkish	German	Germany	Bilingual	SR	None	Y	16
Jalali-Moghadam & Kormi-Nouri, 2015	59	45	10.5	Farsi	Swedish	Sweden	Emergent bilingual	SR	Equivalence testing	N	15
Janus & Bialystok, 2018	48	45	9.4	Mix	English	Canada	Unclear	LU	Equivalence testing	Y	10
Kalashnikova & Mattock, 2014[a]	33	33	4.6	English	Welsh	UK	Bilingual	SR	Matched samples	Y	13
Kalia et al., 2019	54	61	7.5	English	Spanish	U.S.	Emergent bilingual	RV	None	Y	15
Kalia et al., 2018											
Native English speakers	54	36	7.4	English	Spanish	U.S.	Emergent bilingual	RV	None	Y	15
Native Spanish speakers	54	26	7.6	Spanish	English	U.S.	Emergent bilingual	RV	None	Y	15
Kapa & Colombo, 2013[a]											
Early bilingual	22	21	9.5	English	Spanish	U.S.	Bilingual	RV	Equivalence testing	Y	12
Late bilingual	22	36	9.8	Spanish	English	U.S.	Emergent bilingual	RV	Equivalence testing	Y	12
Karlsson et al., 2015[a]											
Younger	25	24	7.4	Swedish	Mix	Sweden	Bilingual	LU	Equivalence testing	Y	14
Older	23	27	11.0	Swedish	Mix	Sweden	Bilingual	LU	Equivalence testing	Y	14

(continued)

Table 3. *(continued)*

Study and independent subgroup/comparison group	Mono-linguls (*n*)	Bi-linguls (*n*)	Mean age (years)	First language	Second language	Geographic location	Bilingual proficiency	Method used to assess language status[c]	Equivalence testing or matched samples[d]	Measure of SES[e]	Study-quality score
Kaushanskaya et al., 2014[a]	19	19	6.5	English	Spanish	U.S.	Emergent bilingual	RV	Matched samples	Y	15
Kempert & Hardy, 2015	29	28	10.2	Italian or Greek	German	Germany	Emergent bilingual	SR	Equivalence testing	Y	11
Kohnert et al., 2004	50	22	10.3	Spanish	English	U.S.	Bilingual	RV	None	Y	14
Krizman et al., 2012	25	23	14.7	Spanish	English	U.S.	Bilingual	LU	Matched samples	Y	7
Krizman et al., 2016											
Low socioeconomic status	15	17	14.6	Spanish	English	U.S.	Bilingual	LU	None	Y	11
High socioeconomic status	16	12	14.6	Spanish	English	U.S.	Bilingual	LU	None	Y	11
Krizman et al., 2014	27	27	14.6	Spanish	English	U.S.	Bilingual	LU	Matched samples	Y	8
Ladas et al., 2015[a]											
Study 1	24	26	9.4	Greek	Albanian	Greece	Bilingual	RV	Matched samples	Y	15
Study 2	32	28	6.6	Greek	Albanian	Greece	Bilingual	RV	Matched samples	Y	15
Lauchlan et al., 2013							Unclear	SR	None	N	15
Scotland	30	30	9.8	English	Gaelic	Scotland					
Sardinia	29	32	9.1	Italian	Sardinia	Italy					
Leikin & Tovli, 2014	16	15	5.9	Russian	Hebrew	Israel	Unclear	LU	None	Y	13
Lesaux & Siegel, 2003	757	181	7.8	Mix	English	Canada	Emergent bilingual	SR	None	N	13
Li et al., 2017[a]	16	11	8.6	Japanese	English	Japan	Bilingual	RV	Equivalence testing	N	16
Loe & Feldman, 2016[a]	53	26	4.3	English	Mix	U.S.	Unclear	LU	Equivalence testing	Y	14
Marini et al., 2019[a]	31	31	4.6	Italian	English	Italy	Emergent bilingual	RV	Matched samples	Y	16
Martin-Rhee & Bialystok, 2008											
Study 1	17	17	4.8	English	French	Canada	Bilingual	RV	None	N	10
Study 2	20	21	4.6	English	Mix	Canada	Bilingual	SR	None	N	9
Study 3	19	13	8.0	English	Hebrew or Russian	Canada	Bilingual	SR	None	N	9
McVeigh et al., 2019											
7-year-olds	34	32	7.1	English	Irish	Ireland	Emergent bilingual	LU	None	Y	15
9-year-olds	32	23	9.0	English	Irish	Ireland	Emergent bilingual	LU	None	Y	15
Mehrani & Zabihi, 2017	31	36	4.5	Persian	Turkish	Iran	Bilingual	RV	None	Y	15
Meir & Armon-Lotem, 2017[a]											
Low socioeconomic status	16	44	6.0	Russian	Hebrew	Israel	Bilingual	RV	Matched samples	Y	14
High socioeconomic status	16	44	6.1	Russian	Hebrew	Israel	Bilingual	RV	Matched samples	Y	14
Messer et al., 2010	67	60	4.4	Turkish	Dutch	Netherlands	Emergent bilingual	RV	None	Y	15

(continued)

Table 3. *(continued)*

Study and independent subgroup/comparison group	Monolinguls (n)	Bilinguls (n)	Mean age (years)	First language	Second language	Geographic location	Bilingual proficiency	Method used to assess language status[c]	Equivalence testing or matched samples[d]	Measure of SES[e]	Study-quality score
Mohades et al., 2014											
Bilingual	14	19	9.5	Dutch	Mix	Belgium	Bilingual	SR	None	N	15
Emergent bilingual	14	18	9.5	Dutch	Mix	Belgium	Emergent bilingual	SR	None	N	15
Mok et al., 2008											
Chinese monolinguals	34	25	10.0	Chinese	English	Hong Kong	Bilingual	SR	Equivalence testing	N	15
English monolinguals	20	25	10.0	Chinese	English	Hong Kong	Bilingual	SR	Equivalence testing	N	15
Morales et al., 2013											
Study 1	29	27	5.5	Mix	English	Canada	Bilingual	LU	Equivalence testing	N	12
Study 2	34	35	6.9	Mix	English	Canada	Bilingual	LU	Equivalence testing	N	12
Morton & Harper, 2007[a]	17	17	6.9	French	English	Canada	Bilingual	RV	Equivalence testing	Y	13
Mueller Gathercole et al., 2010											
7- to 8-year-olds/only English spoken at home	22	22	8.1	English	Welsh	UK	Unclear	LU	None	N	12
7- to 8-year-olds / Welsh and English spoken at home	22	23	8.1	English	Welsh	UK	Unclear	LU	None	N	12
7- to 8-year-olds/only Welsh spoken at home	22	23	8.1	Welsh	English	UK	Unclear	LU	None	N	12
13- to 15-year-olds/ Welsh and English spoken at home	20	25	14.5	English	Welsh	UK	Unclear	LU	None	N	12
13- to 15-year-olds/ only Welsh spoken at home	20	24	14.5	English	Welsh	UK	Unclear	LU	None	N	12
13- to 15-year-olds/ only English spoken at home	20	34	14.5	Welsh	English	UK	Unclear	LU	None	N	12
Mumtaz & Humphreys, 2001	60	60	7.8	Urdu	English	UK	Bilingual	LU	None	N	10
Namazi & Thordardottir, 2010[a]											
English monolinguals	15	15	4.9	French	English	Canada	Bilingual	RV	Matched samples	Y	13
French monolinguals	15		4.9	French	English	Canada	Bilingual	RV	Matched samples	Y	13
Nayak, 2018	66	56	4.2	English	Mix	U.S.	Bilingual	LU	None	Y	15
Nayak et al., 2020[a]	61	57	6.9	English	Mix	U.S.	Bilingual	LU	Equivalence testing	Y	13

(continued)

Table 3. *(continued)*

Study and independent subgroup/comparison group	Mono-linguls (*n*)	Bi-linguls (*n*)	Mean age (years)	First language	Second language	Geographic location	Bilingual proficiency	Method used to assess language status[c]	Equivalence testing or matched samples[d]	Measure of SES[e]	Study-quality score
Nayak & Tarullo, 2020[a]	62	53	4.2	English	Mix	U.S.	Bilingual	LU	Equivalence testing	Y	15
Nguyen & Astington, 2014[a]											
English monolinguals	24	24	3.9	English	French	Canada	Bilingual	RV	Equivalence testing	Y	14
French monolinguals	24	24	4.0	English	French	Canada	Bilingual	RV	Equivalence testing	Y	14
Nicolay & Poncelet, 2015	50	51	8.8	French	English	Belgium	Emergent Bilingual	SR	None	N	13
Niolaki & Masterson, 2012											
English monolingual/ weak Greek bilingual group	33	23	7.8	English	Greek	UK	Unclear	SR	Equivalence testing	N	13
English monolingual/ strong Greek bilingual group	33	23	7.8	English	Greek	UK	Unclear	SR	Equivalence testing	N	13
Greek monolingual/ weak Greek bilingual group	38	23	7.9	English	Greek	UK	Unclear	SR	Equivalence testing	N	13
Greek monolingual/ strong Greek bilingual group	38	23	7.9	English	Greek	UK	Unclear	SR	Equivalence testing	N	13
Okanda et al., 2010											
Monolinguals matched on age and verbal age	18	18	4.2	Japanese	French	Japan	Bilingual	LU	Matched samples	N	10
Monolinguals matched on age with higher-verbal-age monolinguals	18	18	4.2	Japanese	French	Japan	Bilingual	LU	Matched samples	N	10
Park et al., 2018[a]	41	41	9.4	Spanish	English	U.S.	Bilingual	RV	Matched samples	Y	17
Park, 2014[a]	22	20	10.2	Mix	English	U.S. and Canada	Unclear	LU	Matched samples	Y	14
Pawlicka et al., 2015	42	35	7.1	Polish	English	Poland	Emergent bilingual	SR	None	N	12
Pearson, 1988	18	18	NR	Spanish	English	U.S.	Bilingual	RV	None	N	11
Pino Escobar et al., 2018[a]	17	17	7.1	English	Mix	Australia	Bilingual	LU	Matched samples	Y	12
Poarch & Bialystok, 2015											
Bilingual	60	60	9.4	English	Mix	Canada	Bilingual	LU	Equivalence testing	Y	9
Poarch & van Hell, 2012											
Bilingual	20	18	6.9	German	English	Germany	Bilingual	LU	None	Y	12

(continued)

Table 3. *(continued)*

Study and independent subgroup/comparison group	Monolinguls (*n*)	Bilinguls (*n*)	Mean age (years)	First language	Second language	Geographic location	Bilingual proficiency	Method used to assess language status[c]	Equivalence testing or matched samples[d]	Measure of SES[e]	Study-quality score
Purić et al., 2017											
High-exposure bilingual group	22	19	7.8	Serbian	English or German	Serbia	Emergent bilingual	SR	Equivalence testing	Y	11
Low-exposure bilingual group	22	17	8.0	Serbian	English or German	Serbia	Emergent bilingual	SR	Equivalence testing	Y	11
Rainey et al., 2016											
Nonbrokers	26	30	9.5	English	Spanish	U.S.	Bilingual	RV	Equivalence testing	Y	12
Brokers	26	36	9.5	English	Spanish	U.S.	Bilingual	RV	Equivalence testing	Y	12
Raudszus et al., 2018	76	102	9.9	Dutch	Mix	Netherlands	Bilingual	RV	Equivalence testing	N	15
Riggs et al., 2014	53	129	10.7	English	Spanish	U.S.	Unclear	LU	None	N	12
Robinson & Sorace, 2019	36	26	5.3	English	Mix	UK	Emergent bilingual	LU	None	N	16
Ross & Melinger, 2017											
Study 1	45	54	7.7	English	Gaelic	Scotland	Bilingual	LU	Equivalence testing	N	15
Study 2	21	49	7.5	English	Gaelic	Scotland	Bilingual	LU	Equivalence testing	N	15
Rothou & Tsimpli, 2020											
Bilingual	78	24	NR	Greek	Albanian	Greece	Bilingual	SR	None	N	14
Emergent bilingual	78	66	NR	Greek	Albanian	Greece	Emergent bilingual	SR	None	N	14
Santillán & Khurana, 2018[a]	733	216	4.4	Spanish	English	U.S.	Bilingual	RV	Equivalence testing	N	15
Sawan (2015)	100	107	13.5	Arabic	English	Saudi Arabia	Emergent bilingual	SR	None	N	15
Serratrice & De Cat, 2020	87	87	5.9	Mix	English	UK	Bilingual	RV	None	N	14
Shoghi Javan & Ghonsooly, 2018	60	60	16.4	NR	English	Iran	Emergent bilingual	LU	None	N	11
Soliman, 2014[a]	306	306	NR	Arabic	English	Egypt	Bilingual	LU	Matched samples	Y	13
Stephens, 2013[a]	49	62	9.6	English	Irish	Ireland	Bilingual	RV	Matched samples	Y	13
Struys et al., 2018[a]											
6-year-old children	29	29	6.6	Dutch	Mix	Belgium	Bilingual	LU	Matched samples	Y	13
11-year-old children	29	29	11.6	Dutch	Mix	Belgium	Bilingual	LU	Matched samples	Y	13
Thorn & Gathercole, 1999											
Bilingual	17	14	5.8	French	English	UK	Bilingual	RV	None	N	13
Emergent bilingual	17	14	6.9	English	French	UK	Emergent Bilingual	RV	None	N	13
Timmermeister et al., 2020[a]	27	27	7.5	Dutch	Turkish	Netherlands	Bilingual	RV	Matched samples	Y	18
Tran et al., 2019											
Spanish-English bilingual	13	13	3.2	Spanish	English	U.S.	Bilingual	SR	None	Y	13
Vietnamese-English bilingual	13	15	3.3	Vietnamese	English	U.S.	Bilingual	SR	None	Y	13

(continued)

Table 3. *(continued)*

Study and independent subgroup/comparison group	Mono-linguls (*n*)	Bi-linguls (*n*)	Mean age (years)	First language	Second language	Geographic location	Bilingual proficiency	Method used to assess language status[c]	Equivalence testing or matched samples[d]	Measure of SES[e]	Study-quality score
Vietnamese-Cantonese bilingual	20	16	3.2	Vietnamese	Cantonese	Vietnam	Bilingual	SR	None	Y	13
Vega, 2009[a]	15	25	9.0	Spanish	English	U.S.	Bilingual	RV	Matched samples	Y	14
Weber, 2011[a]	48	19	6.3	English	Spanish	U.S.	Bilingual	RV	Equivalence testing	Y	17
White & Greenfield, 2017											
Bilingual	83	148	4.4	English	Spanish	U.S.	Bilingual	RV	None	N	12
Emergent bilingual	83	72	4.2	Spanish	English	U.S.	Emergent bilingual	RV	None	N	12
White, 2019	7	27	5.6	Mix	English	South Africa	Emergent bilingual	RV	None	Y	13
Yang & Yang, 2016	31	32	5.1	Korean	English	Korea	Bilingual	LU	None	Y	13
Yang et al., 2011											
English monolinguals	15	15	4.7	English	Korean	U.S.	Bilingual	LU	Equivalence testing	N	13
Korean monolinguals	13	15	4.6	Korean	English	U.S.	Bilingual	LU	Equivalence testing	N	13
Korean monolinguals	13	15	4.5	Korean	English	Korea (monolinguals), U.S. (bilinguals)	Bilingual	LU	Equivalence testing	N	13
Yao, 2014	19	41	NR	Spanish	English	U.S.	Bilingual	RV	None	Y	18
Yu et al., 2019[a]	63	59	9.9	Mongolian	Mandarin	China	Bilingual		Equivalence testing	Y	12
Zeng et al., 2019	17	20	8.3	English	Mix	Australia	Bilingual	LU	Equivalence testing	N	9

Note: NR = not reported.
[a]These studies were included in study-quality subgroup analyses (i.e., study-quality score > 12, matched samples, measured socioeconomic status [SES]). [b]This is an unpublished data set, and there was not enough information to calculate a study-quality score. [c]For the method used to assess language status, questionnaires that asked parents to indicate whether another language was spoken in the home or whether the child spoke another language were classified as self-report (SR) questionnaires. Language-use (LU) questionnaires asked parents to indicate the child's proficiency in the second language, the amount of time children spoke or were exposed to the second language in the home, and other questions designed to assess proficiency and exposure. Studies that indicated that parents were asked only if the child spoke another language at home were classified as having SR language status (by participant, parent, or school official). Studies that determined language status by enrollment in immersion programs were included in the SR category. RV = measured receptive vocabulary in both the first and second language. [d]Details pertaining to the classification of equivalence testing or matched samples are reported in Table S2 at https://osf.io/jv7wt/. [e]For measure of SES (Y = yes, N = no), authors had to report that SES was objectively measured (income, parental occupation, or parental education). Studies that recruited from low- or high-income neighborhoods or schools without additional measures to confirm that participants in the sample were indeed in that SES bracket received a "no" classification for this measure.

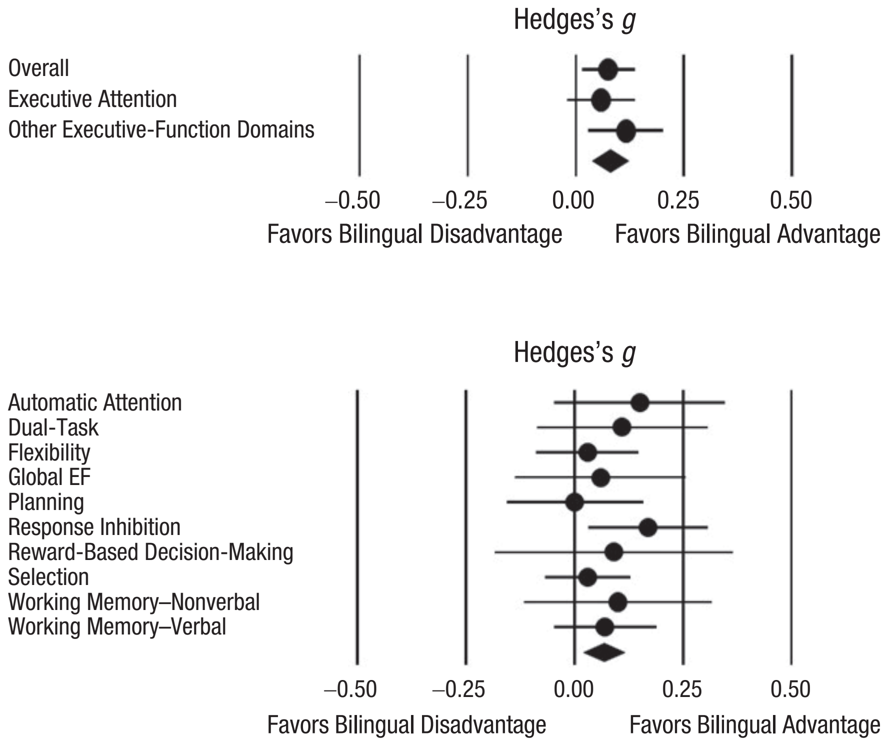

Fig. 4. Forest plots showing the mean effect-size estimate for (a) the overall effect of language status on executive functions (EFs), executive attention, and other EF domains and (b) each executive-function domain. Diamonds indicate overall effect sizes. Error bars represent 95% confidence intervals.

Participant characteristics.

Age. Mean age did not moderate the overall effect of language status on children's executive functioning, $Q_M(1) = 0.05$, β = 0.002, 95% CI = [–0.01, 0.02], $p = .824$, nor did it moderate the effect of language status on automatic attention (β = –0.02, 95% CI = [–0.06, 0.01], $p = .208$); selection (β = –0.01, 95% CI = [–0.03, 0.004], $p = .138$); nonverbal working memory (β = 0.03, 95% CI = [–0.007, 0.03], $p = .112$); flexibility (β = 0.01, 95% CI = [–0.006, 0.004], $p = .183$); or verbal working memory (β = 0.01, 95% CI = [–0.01, 0.03], $p = .308$). Age did, however, moderate the effect of language status on response inhibition (β = 0.03, 95% CI = [0.001, 0.06], $p = .044$).

Degree of bilingualism. Degree of bilingualism moderated the overall effect of language status on executive functioning, $Q_M(3) = 7.81$, $p = .050$; the test for residual heterogeneity was significant, $Q_E(1185) = 4{,}466.38$, $p < .001$. Across all executive-function domains, balanced bilinguals showed a small but significant advantage in executive functioning relative to monolinguals ($g = .08$, 95% CI = [.009, .15], $p = .027$), but emergent bilinguals ($g = .03$, 95% CI = [–.06, .13], $p = .489$) and unclassifiable bilinguals ($g = .18$, 95% CI = [–.01, .37], $p = .064$) did not. Within the individual executive-function domains, $Q_M(16) = 65.62$, $p < .001$, degree of bilingualism moderated the effect of language status within the flexibility domain; unclassifiable bilinguals showed an advantage relative to monolinguals ($g = .44$, 95% CI = [.15, .74], $p = .003$), but balanced bilinguals ($g = .01$, 95% CI = [–.21, .23], $p = .923$) and emergent bilinguals ($g = –.09$, 95% CI = [–.33, .15], $p = .466$) did not. Conversely, balanced bilinguals ($g = –.25$, 95% CI = [–.48, –.02], $p = .034$) showed a significant disadvantage relative to monolinguals on nonverbal working memory tasks, but emergent bilinguals ($g = .09$, 95% CI = [–.20, .38], $p = .545$) and unclassifiable bilinguals ($g = .24$, 95% CI = [–.32, .16], $p = .166$) did not. Language-status effects were indistinguishable from zero in all other domains.

Geographic origin of the sample. Geographic origin of the sample moderated the overall effect of language status on children's executive functioning, $Q_M(6) = 14.82$, $p = .022$; the test for residual heterogeneity remained significant, $Q_E(1161) = 3{,}999.91$, $p < .001$. An effect of language status favoring bilingual children was evident in European samples ($g = .11$, 95% CI = [.01, .21], $p = .028$) but not in samples from North America ($g = .07$, 95% CI = [–.02, .16], $p = .115$); East Asia ($g = –.13$, 95% CI = [–.30, .04], $p = .126$); the Middle East ($g = .07$,

Table 4. Effect Size (*g*), Heterogeneity, and Variance Components for the Overall Effect and Results for Processes Associated With Executive-Attention and Other Executive-Function Domains Separately

	Effect-size estimates and significance tests							Heterogeneity				Variance components		
Test and measure	*k*	*g*	*SE*	95% CI	*z*	*p*	Prediction interval	*Q*	*p*	I^2	τ^2	σ^2_1	σ^2_2	σ^3_2
Overall	1,188	.08	.03	[.01, .14]	2.36	.017	[−0.54, 0.70]	4,539.44	< .001	67.27	.10	.041	.040	.017
Executive-attention domains														
Executive attention: overall	694	.06	.04	[−.02, .14]	1.56	.118	[−0.62, 0.74]	2,388.05	< .001	70.39	.12	.037	.046	.040
Working memory–nonverbal	53	.10	.11	[−.11, .32]	0.93	.350	[−0.85, 1.05]	281.46	< .001	84.65	.21	.12	.10	.00
Selection	371	.03	.05	[−.07, .13]	0.58	.565	[−0.71, 0.77]	996.89	< .001	73.57	.14	.053	.00	.091
Stroop														
Congruent trials: accuracy	4	.06	.12	[−.17, .28]	0.50	.620	[−1.02, 1.14]	0.16	.984	0	.05	.00	.00	.00
Congruent trials: reaction time	11	.23	.20	[−.16, .92]	1.13	.258	[−1.19, 1.65]	43.66	< .001	82.01	.27	.09	.09	.09
Incongruent trials: accuracy	9	.09	.09	[−.09, .26]	0.96	.337	[−0.12, 0.30]	8.47	.387	0	.00	.00	.00	.00
Incongruent trials: reaction time	10	.10	.25	[−.40, .60]	0.40	.689	[−0.97, 1.17]	40.64	< .001	87.65	.45	.15	.15	.15
Simon														
Congruent trials: accuracy	12	.19	.17	[−.16, .53]	1.07	.284	[−0.76, 1.14]	24.26	.012	64.47	.15	.073	.073	.007
Congruent trials: reaction time	16	.30	.34	[−.37, .96]	0.88	.381	[−1.82, 2.42]	107.75	< .001	88.99	.86	.29	.57	.00
Incongruent trials: accuracy	12	−.02	.16	[−.32, .30]	−0.11	.915	[−1.06, 1.02]	28.64	< .001	69.06	.19	.00	.00	.19
Incongruent trials: reaction time	15	−.04	.08	[−.21, .12]	−0.51	.611	[−0.04, −0.27]	11.23	.668	5.05	.005	.00	.005	.00
Flanker														
Congruent trials: accuracy	18	.01	.18	[−.35, .36]	0.04	.966	[−1.07, 1.09]	68.95	< .001	77.14	.23	.23	.00	.00
Congruent trials: reaction time	22	.09	.13	[−.17, .35]	0.68	.495	[−0.71, 0.89]	47.98	< .001	67.05	.13	.13	.00	.00
Incongruent trials: accuracy	23	.10	.18	[−.26, .46]	0.54	.592	[−1.08, 1.28]	88.13	< .001	80.99	.29	.23	.056	.00
Incongruent trials: reaction time	21	.08	.13	[−.16, .33]	0.67	.499	[−0.66, 0.82]	42.55	.002	63.95	.11	.11	.00	.00
Flexibility	270	.05	.06	[−.06, .16]	0.84	.400	[−0.69, 0.79]	1,097.49	< .001	70.54	.14	.033	.091	.015
DCCS	61	.08	.10	[−.11, .28]	0.824	.410	[−0.77, 0.93]	156.00	< .001	80.23	.17	.17	.004	.00
TMT	21	.18	.22	[−.24, .60]	0.84	.402	[−1.03, 1.39]	92.29	< .001	79.13	.29	.156	.00	.132
WCST	25	−.40	.16	[−.72, .09]	−0.25	.012	[−1.21, 0.41]	132.89	< .001	60.82	.09	.030	.030	.030
Other executive-function domains														
Other executive-function domains: overall	494	.11	.04	[.03, .20]	2.43	.015	[−0.60, 0.82]	2,138.60	< .001	75.11	.13	.091	.039	.003
Automatic attention	105	.15	.10	[−.05, .35]	1.49	.137	[−0.62, 0.92]	284.94	< .001	73.91	.14	.029	.11	.00
Attention Network Task	86	.08	.15	[−.20, .37]	0.57	.569	[−0.72, 0.88]	225.49	< .001	73.35	.14	.10	.018	.018
Response inhibition	57	.17	.07	[.05, .30]	2.67	.008	[−0.34, 0.68]	162.52	< .001	61.42	.06	.060	.00	.00
Go/no go	21	.20	.11	[−.01, .41]	1.82	.068	[−0.36, 0.76]	30.32	.065	45.40	.06	.019	.019	.019
Working memory–verbal	278	.06	.06	[−.05, .18]	1.09	.275	[−0.76, 0.88]	1,600.59	< .001	78.69	.17	.12	.03	.01
N-back	33	.07	.21	[−.34, .48]	0.32	.746	[−0.85, 0.99]	123.91	< .001	75.47	.16	.079	.079	.00
Forward digit span	47	.07	.13	[−.18, .32]	0.57	.571	[−1.21, 1.35]	381.42	< .001	88.00	.39	.39	.00	.00
Backward digit span	44	.03	.09	[−.14, .20]	0.35	.726	[−0.77, 0.83]	203.49	< .001	77.13	.15	.14	.012	.00
Reward-based decision-making	9	.09	.14	[−.18, .37]	0.67	.500	[−0.49, 0.67]	10.49	.233	26.30	.04	.00	.00	.043
Planning	14	.0004	.08	[−.16, .16]	0.005	.996	[−0.27, 0.27]	20.17	.091	14.55	.009	.003	.003	.003
Global executive functioning	7	.06	.11	[−.15, .26]	0.54	.587	[−0.36, 0.48]	6.25	.396	21.62	.016	.005	.005	.005
Dual task	24	.11	.10	[−.08, .30]	1.14	.255	[−0.25, 0.47]	33.97	.066	27.57	.02	.006	.006	.006

Note: We provide statistics for key tasks within each domain. These are for descriptive purposes only, and because of small samples within each task, results should be interpreted with caution. For variance components, σ^2_1 represents variance in the effect-size estimate due to variability between research groups (highest level), σ^2_2 represents variance in the effect-size estimate between studies clustered within research groups, and σ^2_3 represents within-sample variance in the effect-size estimate. CI = confidence interval; DCCS = Dimensional Change Card Sorting Task; TMT = Trail Making Test; WCST = Wisconsin Card Sorting Task.

95% CI = [−.18, .31], p = .587); or Africa (g = −.12, 95% CI = [−.35, .12], p = .330). Likewise, no significant effects were observed for mixed samples (g = .11, 95% CI = [−.06, .29], p = .203). Because of variability in the number of effect sizes, we were unable to test for language-status effects within Australian samples and within executive-functioning domains.

Study quality.

The AXIS measure of study quality. Study quality as measured by the AXIS (see Table 5) moderated the overall effect of language status on children's executive functioning, $Q_M(1) = 6.82$, $p = .009$; β = −0.03, 95% CI = [−0.05, −0.008]; effect-size magnitude decreased as study quality increased. The test for residual heterogeneity remained significant, $Q_E(1164) = 4{,}432.04$, $p < .001$. AXIS scores similarly moderated the language-status effect, $Q_M(6) = 15.15$, $p = .02$, in the domains of selection (β = −0.03, 95% CI = [−0.06, −0.006], p = .018) and flexibility (β = −0.05, 95% CI = [−0.08, −0.02], p = .001).

Measured equivalence of groups. The equivalence of monolingual and bilingual groups needs to be established through measurement to ensure that between-groups differences reflect an effect of independent variables rather than unmeasured confounds. Thus, measured equivalence of groups, through either matching or statistical testing, on confounding factors including age, nonverbal IQ, gender, or SES is an important measure of study quality. In all, 41 of 159 studies reported matching monolingual and bilingual samples on at least a single variable, and an additional 32 of 159 studies reported using equivalence testing to ensure that groups were comparable on at least one demographic variable. Ensuring that monolingual and bilingual samples were comparable on any demographic variables by using either matched samples or equivalence testing was not a significant moderator of the language-status effect on overall executive functioning, $Q_M(2) = 5.38$, $p = .068$; the test of residual heterogeneity remained significant, $Q_E(1164) = 4{,}437.51$, $p < .001$.

The use of matched samples or equivalence testing was, however, a significant moderator within specific executive-function domains, $Q_M(12) = 52.35$, $p < .001$. Specifically, studies that did not ensure group equivalence by measuring confounding variables showed language-status effects favoring bilingual children in the domains of automatic attention (g = .28, 95% CI = [.03, .52], p = .027); response inhibition (g = .17, 95% CI = [.06, .29], p = .004); flexibility (g = .17, 95% CI = [.07, .28], p = .001); and selection (g = .17, 95% CI = [.07, .28], p = .001; see Table 6). By contrast, studies that ensured group equivalence through measurement showed language-status effects favoring bilingual children only in the domains of response inhibition (g = .22, 95% CI = [.09, .35], p = .001) and verbal working memory (g = .13, 95% CI = [.04, .22], p = .005).

Measurement of SES. Study quality as assessed by reported objective measurement and control of SES moderated the effect of language status on children's executive functioning. In all, 94 of 158 studies reported objectively measuring SES. Measurement of SES moderated the language-status effect on overall executive functioning, $Q_M(2) = 8.11$, $p = .017$; the test of residual heterogeneity remained significant, $Q_E(1164) = 4{,}429.53$, $p < .001$. The effect of language status was evident in studies for which an objective measure of SES was not reported (g = .13, 95% CI = [.04, .22], p = .005) but indistinguishable from zero among studies for which an objective measure of SES was reported (g = .04, 95% CI = [−.03, .12], p = .284). Reported objective measurement of SES similarly moderated the effect of language status within specific domains of children's executive functioning, $Q_M(12) = 47.22$, $p < .001$. Among studies that did not measure SES, language-status effects favoring bilinguals were evident in the domains of response inhibition (g = .31, 95% CI = [.13, .48], p < .001); selection (g = .22, 95% CI = [.12, .33], p < .001); and flexibility (g = .17, 95% CI = [.06, .28], p = .002). Among studies that did measure and control for SES, effect sizes in these three domains were indistinguishable from zero.

Language-status measure. Choice of language-status measure moderated the overall effect of language status on children's executive functioning, $Q_M(3) = 13.06$, $p = .004$; the test of residual heterogeneity remained significant, $Q_E(1185) = 4{,}384.66$, $p < .001$. Studies that measured language status via nomination (self, parental, or teacher; k = 216, g = .17, 95% CI = [.06, .27], p = .002) or through the use of language-use surveys (k = 593, g = .10, 95% CI = [.01, .19], p = .024) reported an effect of language status favoring bilingual children, whereas studies that used receptive vocabulary tests (k = 379, g = −.007, 95% CI = [−.09, .08], p = .876) reported language-status effects that were indistinguishable from zero.

Year of publication. Year of publication was not a significant moderator of the overall effect of language status on children's executive functioning, $Q_M(1) = 0.15$, $p = .699$. Publication year did, however, significantly moderate, $Q_M(6) = 14.25$, $p = .027$, the effect of language status within the domains of automatic attention (β = 0.04, 95% CI = [0.0001, 0.08], p = .050) and nonverbal working memory (β = 0.03, 95% CI = [0.005, 0.049], p = .017), indicating that publication year was associated with an

Table 5. Percentage of Studies Meeting the Yes, No, and Unclear Criteria for All Study-Quality Measurements

Question	Yes (*k*)	No (*k*)	Unclear (*k*)
AXIS questions			
1. Were the aims/objectives of the study clear?	100 (158)	0 (0)	0 (0)
2. Was the study design appropriate for the stated aims?	96.84 (153)	2.53 (4)	0.63 (1)
3. Was sample size justified?	17.09 (27)	82.91 (131)	0 (0)
4. Was the target/reference population clearly defined?	99.37 (157)	0.63 (1)	(0)
5. Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	79.11 (125)	8.23 (13)	12.66 (20)
6. Was the selection process likely to select subjects/participants who were representative of the target/reference population under investigation?	76.58 (121)	10.13 (16)	13.29 (21)
7. Were the risk factor and outcome variables measured appropriate to the aims of the study?	99.37 (157)	0 (0)	0.63 (1)
8. Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialed, piloted, or published previously?	96.20 (152)	0 (0)	3.80 (6)
9. Is it clear what was used to determined statistical significance and/or precision estimates (e.g., *p* values, confidence intervals)?	99.37 (157)	0.63 (1)	0 (0)
10. Were the methods (including statistical methods) sufficiently described to enable them to be repeated?	91.77 (145)	8.23 (13)	0 (0)
11. Were the basic data adequately described?	94.94 (150)	3.79 (6)	1.27 (2)
12. Does the response rate raise concerns about nonresponse bias?	95.57 (151)	1.90 (3)	2.53 (4)
13. If appropriate, was information about nonresponders described?	91.77 (145)	7.59 (12)	0.63 (1)
14. Were the results internally consistent?	98.73 (156)	0.63 (1)	0.63 (1)
15. Were the results presented for all the analyses described in the methods?	100 (158)	0 (0)	0 (0)
16. Were the author's discussion and conclusions justified by the results?	90.51 (143)	7.59 (12)	1.90 (3)
17. Were the limitations of the study discussed?	53.80 (85)	43.04 (68)	3.16 (5)
18. Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?	6.33 (10)	34.18 (54)	61.49 (94)
19. Was ethical approval or consent of participants attained?	32.91 (52)	0 (0)	67.09 (106)
Additional measures of methodological rigor			
Use of matched samples or equivalence testing	61.39 (97)	38.61 (61)	0 (0)
Objective measurement of socioeconomic status	58.49 (93)	41.51 (77)	0 (0)

Note: One unpublished data set did not include enough information to rate the study on any of the Appraisal Tool for Cross-Sectional Studies (AXIS) dimensions or to use matched samples or equivalence testing. AXIS scores ranged from 7 to 18 (out of a maximum score of 20; $M = 12.54$, $SD = 2.34$).

increase in effect-size estimates within these domains. Publication year did not significantly moderate the effects of language status on flexibility (β = −0.003, 95% CI = [−0.02, 0.01], $p = .630$); selection (β = 0.003, 95% CI = [−0.01, 0.02], $p = .644$); verbal working memory (β = −0.006, 95% CI = [−0.02, 0.005], $p = .294$); or response inhibition (β = 0.004, 95% CI = [−0.01, 0.02], $p = .673$).

Multiple meta-regression. To consider all moderator variables in tandem, we conducted a multiple meta-regression analysis that predicted residualized effect sizes from participant characteristics, AXIS study-quality scores, use of matched samples or equivalence testing, measurement of SES, language-status measure, and year of publication, $Q_M(10) = 19.07$, $p = .039$. AXIS study-quality scores (β = −0.06, 95% CI = [−0.10, −0.03], $p < .001$) emerged as the only significant moderator. Still, the overall largest source of variability in reported effect sizes was the effect of research group ($\sigma^2 = .04$), as revealed by the multilevel model.

Publication bias

Funnel plots, Egger's test of asymmetry. Contour-enhanced funnel plots for the overall effect and by executive-function domain are presented in Figures 2 and 3. Asymmetry of effect sizes was clearly observed for the overall effect and for many of the included domains. This asymmetry was confirmed using the modified Egger's regression test for funnel-plot asymmetry (Pustejovsky & Rodgers, 2019; see Table 7).

PET-PEESE correction for publication bias. PET-PEESE analysis also revealed evidence of publication bias in both the estimate of the overall effect of language status on children's executive functioning and the effect of

Table 6. Effect-Size Estimates and Confidence Intervals (CIs) for Studies That Measured Group Equivalence Using Either Matched Samples or Equivalence Testing and Studies That Did Not Measure Group Equivalence

	Equivalence measured		Equivalence not measured	
Domain and measure	*g*	95% CI	*g*	95% CI
Overall	.07	[−.002, .15]	.08	[−.01, .17]
Executive attention				
Selection	.05	[−.03, .14]	.17	[.07, .28]
Flexibility	.03	[−.06, .13]	.17	[.07, .28]
Nonverbal working memory	−.07	[−.18, .04]	−.02	[−.20, .15]
Other executive-function domains				
Automatic attention	.05	[−.05, .16]	.28	[.03, .52]
Response inhibition	.22	[.09, .35]	.17	[.06, .29]
Verbal working memory	.13	[.04, .22]	−.06	[−.16, .05]

language status within specific domains. Overall effect sizes and effect sizes within each domain were significantly associated with both their standard error ($p < .001$) and their variance ($p < .001$).

PET-PEESE analysis was then used to adjust for the influence of publication bias. Results indicated that after we adjusted for publication bias, the overall effect of language status on children's executive functioning was indistinguishable from zero ($g = -.04$, 95% CI = [−.13, .05], $p = .414$). Within domains, bias-corrected estimates revealed a statistically significant bilingual disadvantage for nonverbal working memory ($g = -.19$) and language-status effects that were indistinguishable from 0 for all remaining executive-function domains (see Table 7).

Discussion

A systematic review of available literature revealed no coherent evidence that bilingual children are advantaged in executive functioning relative to monolingual children. A multilevel model of 1,194 effect-size estimates revealed a small ($g = .08$) but statistically significant overall effect of language status on children's executive functioning after we controlled for unique samples, individual studies, and different research groups. The overall effect of language status was confined to studies using verbal ($g = .12$) task stimuli and outputs and was strongest in studies using European samples.

Language-status effects were, however, evident in only one of nine theoretically defined domains of executive function—response inhibition—and were indistinguishable from zero in all three domains of executive attention hypothesized to be the locus of language-status effects in children (Bialystok, 2017). Further, effect-size heterogeneity was elevated in almost every domain of executive functioning. Variability in the magnitude of reported effects derived primarily from the influence of different research groups and studies, suggesting that selected studies and research groups exert an inordinate influence on estimates of language-status effects.

Moderation analyses identified two additional factors that contribute to variability in reported effect sizes, including study quality and measurement of SES. Reported effects were larger in low-quality studies and those that did not measure SES and were statistically indistinguishable from zero in high-quality studies and those that measured SES. To be sure, a priori criteria for both moderator variables were not that stringent. To achieve a high score on the study quality AXIS, a study needed to objectively measure independent and dependent variables, provide evidence of the representativeness of experimental and control samples, and the authors had to be transparent in reporting conflicts of interest and study limitations. And to be classified as measuring SES, a study merely had to measure family income, parental education, or an objective proxy thereof.

The analysis also revealed evidence of a confirmatory bias in the reporting of research evidence. Funnel plots of the magnitude versus the standard error of effect-size estimates revealed asymmetries that were driven by a disproportionate number of large positive effects among studies with low precision estimates. Such asymmetries are considered a reflection of publication or small sample bias because they suggest that confirmatory findings are more likely to survive peer review than are disconfirmatory findings. After adjusting for the influence of publication and small sample bias using the PET-PEESE procedure (Stanley & Doucouliagos, 2014), we found that effect-size estimates for both executive functioning overall and almost all included executive-function domains were statistically indistinguishable from zero; the PET-PEESE

Table 7. PET-PEESE-Corrected Estimates and Results From the Modified Egger's Regression Test for the Overall Language-Status Effect

Domain and measure	PET–PEESE-corrected estimate		Egger's regression test (*p*)	
	g	*p*	Raw data	Trimmed data
Overall	−.04 [−.13, .05]	.414	< .001	< .001
Executive attention	.02 [−.09, .14]	.664	< .001	< .001
Working memory–nonverbal	−.19 [−.30, −.08]	.001	.042	.042
Selection	.006 [−.10, .09]	.909	.008	.010
Flexibility	−.01 [−.11, .08]	.789	< .001	< .001
Other executive-function domains	−.07 [−.19, .06]	.307	< .001	< .001
Automatic attention	−.01 [−.12, .10]	.815	.052	.052
Response inhibition	.04 [−.07, .14]	.495	< .001	< .001
Working memory–verbal	−.05 [−.15, .04]	.292	< .001	< .001
Reward-based decision-making	−.06 [−.32, .21]	.677	.015	.015
Planning	−.10 [−.29, .08]	.272	.012	.012
Global executive functioning	.04 [−.18, .26]	.720	.503	.503
Dual task	−.03 [−.21, .14]	.703	.098	.098

Note: Values in brackets are 95% confidence intervals. PET = precision-effect test; PEESE = PET with standard errors.

corrected estimate for nonverbal working memory indicated a statistically significant effect in favor of a bilingual disadvantage.

Taken together, the current findings parallel those of Gunnerud et al. (2020), who found little evidence of a bilingual advantage among children ages 2 through 15 years, considerable heterogeneity in the magnitude of reported effects, a moderating effect of SES, and evidence of publication bias in a substantially smaller survey of the pediatric literature (583 vs. the current 1,194 effect sizes). The current findings do, however, extend the findings of Gunnerud et al. (2020) in several important ways. First, we very specifically tested for—and found no evidence of—language-status effects in three domains of executive attention that Bialystok (2017) highlighted as particularly relevant for identifying the bilingual advantage. Thus, our findings show that null effects reported by Gunnerud and colleagues cannot be explained away by arguing that executive-functioning domains were not properly defined to reveal a bilingual advantage. Critical executive-attention domains used in the current analysis were defined according to recent theory (Bialystok, 2017) to maximize the likelihood of detecting language-status effects. Despite this, we found no evidence of any language-status effects. Second, we tested for and found evidence of the importance of study quality in explaining heterogeneity in reported effects. Gunnerud et al. also found substantial heterogeneity in reported effects but identified only two moderating variables: SES and research group. Our findings therefore provide additional insight into methodological considerations that contribute to variance in the magnitude of reported effects, as has been suggested by various critics (for a discussion, see Morton, 2015).

The findings challenge the view that bilingual language status favorably impacts children's executive functioning. In the face of null findings from the study of adults, proponents of the bilingual-advantage hypothesis have argued that language-status effects are more difficult to detect in adults than in children because adults perform at ceiling on executive-function tasks, whereas children do not. The implication is that if language-status effects are to be detected at all, they are more likely to be detected earlier rather than later in development (see Grundy et al., 2017). The results of the current meta-analysis challenge this argument by suggesting that language-status effects on executive functioning in children, should they exist at all, are diminishingly small and very difficult to detect. Based on the current review, the overall effect (*g*) of language status on children's executive functioning, uncorrected for the influence of study quality and publication bias, was .07. One would require two equal groups of more than 2,800 participants to detect this effect with a conservative level of power of .8. Detecting the effect of language status on children's inhibition (estimated as *g* = .17, uncorrected for the influence of study quality and publication bias) would require two equal groups of more than 550 participants. To date, only one study has had samples this size, and the authors of that study reported no differences between monolingual and bilingual children on measures of executive functioning (Dick et al., 2019).

The current findings have important implications for future research on the bilingual advantage in children.

First, there is a need to move away from the use of small samples. Given current estimates, language-status effects are far too small to be detected by comparisons of 20 or 30 children, which is the current standard. Samples need to be scaled up considerably if language-status effects are to be reliably detected, perhaps through the coordinated efforts of a consortium (for a discussion, see Morton, 2015). Second, there is a need to raise basic methodological standards on a number of fronts. This would include a more exhaustive cataloguing of, and matching of groups on, potentially confounding variables such as SES and immigration status. Although language status may influence children's executive functioning, to date, reported effects are highly variable from study to study and likely reflect the influence of factors other than language status. Finally, to properly appreciate the complex relationship between language status and children's executive functioning, it may be necessary to move away from simple binary characterizations of language status such as that utilized in the present review. However, to achieve this, we see no way forward other than to abandon the practice of measuring language status through basic self-nomination or paper-and-pencil measures and commit to more thorough measurements that yield continuous, standardized, and reliable measures of language proficiency. Only in this way will it be possible to examine the relation between levels of bilingualism and children's executive functioning across different studies.

Transparency

Action Editor: Sachiko Kinoshita
Editor: Patricia J. Bauer
Author Contributions
C. J. Lowe and J. B. Morton developed the study concept. C. J. Lowe conducted the literature search. C. J. Lowe, S. F. Goldsmith, and I. Cho reviewed the abstracts and articles and coded study quality. C. J. Lowe extracted data from individual articles, coded all moderators, and analyzed and interpreted the data. C. J. Lowe and J. B. Morton drafted the manuscript, and S. F. Goldsmith and I. Cho provided critical revisions. All the authors approved the final manuscript for submission.
Declaration of Conflicting Interests
The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.
Funding
This research was supported by a Social Sciences and Humanities Research Council Insight Grant awarded to J. B. Morton. C. J. Lowe received funding from the Canada First Research Excellence Fund 516 initiative BrainsCAN.
Open Practices
All data and analysis code have been made publicly available via OSF and can be accessed at https://osf.io/jv7wt. The design and analysis plans for the study were not preregistered. This article has received the badges for Open Data and Open Materials. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

ORCID iD

Cassandra J. Lowe https://orcid.org/0000-0003-3830-5283

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797621993108

References

*Asterisks indicate studies that were included in the meta-analysis.

*Abdelgafar, G. M., & Moawad, R. A. M. (2015). Executive function differences between bilingual Arabic–English and monolingual Arabic children. *Journal of Psycholinguistic Research*, *44*(5), 651–667. https://doi.org/10.1007/s10936-014-9309-3

*Abu-Rabia, S., & Siegel, L. S. (2002). Reading, syntactic, orthographic, and working memory skills of bilingual Arabic-English speaking Canadian children. *Journal of Psycholinguistic Research*, *31*(6), 661–678. https://doi.org/10.1023/A:1021221206119

*Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., Davidson, D. J., & Carreiras, M. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Frontiers in Psychology*, *5*, Article 398. https://doi.org/10.3389/fpsyg.2014.00398

*Antoniou, K., Grohmann, K. K., Kambanaros, M., & Katsos, N. (2016). The effect of childhood bilectalism and multilingualism on executive control. *Cognition*, *149*, 18–30. https://doi.org/10.1016/j.cognition.2015.12.002

*Arizmendi, G. D., Alt, M., Gray, S., Hogan, T. P., Green, S., & Cowan, N. (2018). Do bilingual children have an executive function advantage? Results from inhibition, shifting, and updating tasks. *Language Speech and Hearing Services in Schools*, *49*(3), 356–378. https://doi.org/10.1044/2018_LSHSS-17-0107

*Arredondo, M. M. (2017). *A bilingual advantage? The functional organization of linguistic competition and attentional networks in the bilingual developing brain* [Doctoral dissertation, University of Michigan]. Deep Blue. https://deepblue.lib.umich.edu/handle/2027.42/136988

*Arredondo, M. M., Hu, X. S., Satterfield, T., & Kovelman, I. (2017). Bilingualism alters children's frontal lobe functioning for attentional control. *Developmental Science*, *20*(3). https://doi.org/10.1111/desc.12377

*Asadollahpour, F. (2015). The performance of bilingual and monolingual children on working memory tasks. *Iranian Rehabilitation Journal*, *13*(3), 53–57.

*Barac, R., & Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child Development*, *83*(2), 413–422. https://doi.org/10.1111/j.1467-8624.2011.01707.x

*Barac, R., Moreno, S., & Bialystok, E. (2016). Behavioral and electrophysiological differences in executive control between monolingual and bilingual children. *Child Development*, *87*(4), 1277–1290. https://doi.org/10.1111/cdev.12538

*Barbosa, P. G., Jiang, Z., & Nicoladis, E. (2019). The role of working and short-term memory in predicting receptive vocabulary in monolingual and sequential bilingual children. *International Journal of Bilingual Education and Bilingualism*, *22*, 801–817. https://doi.org/10.1080/13670050.2017.1314445

*Bastian, L., Reichenbach, K., Helbig, L., Rohrbach, S., Pollex-Fischer, D., Schäuble, M., Gross, M., Lenz, K., & Sarrar, L. (2018). Effects of bilingualism on cognitive functions in early childhood. *Praxis Der Kinderpsychologie Und Kinderpsychiatrie*, *67*(1), 2–17. https://doi.org/10.13109/prkk.2018.67.1.2

*Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development*, *70*(3), 636–644. https://doi.org/10.1111/1467-8624.00046

*Bialystok, E. (2010). Global-local and trail-making tasks by monolingual and bilingual children: Beyond inhibition. *Developmental Psychology*, *46*(1), 93–105. https://doi.org/10.1037/a0015466

*Bialystok, E. (2011). Coordination of executive functions in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *110*(3), 461–468. https://doi.org/10.1016/j.jecp.2011.05.005

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

*Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010). Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognition and Development*, *11*(4), 485–508. https://doi.org/10.1080/15248372.2010.516420

*Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, *10*(3), 89–129. https://doi.org/10.1177/1529100610387084

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

*Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, *7*(3), 325–339. https://doi.org/10.1111/j.1467-7687.2004.00351.x

*Bialystok, E., & Senman, L. (2004). Executive processes in appearance-reality tasks: The role of inhibition of attention and symbolic representation. *Child Development*, *75*(2), 562–579. https://doi.org/10.1111/j.1467-8624.2004.00693.x

*Bialystok, E., & Shapero, D. (2005). Ambiguous benefits: The effect of bilingualism on reversing ambiguous figures. *Developmental Science*, *8*(6), 595–604. https://doi.org/10.1111/j.1467-7687.2005.00451.x

*Blom, E., & Boerma, T. (2017). Effects of language impairment and bilingualism across domains. *Linguistic Approaches to Bilingualism*, *7*(3–4), 277–300. https://doi.org/10.1075/lab.15018.blo

*Blom, E., Boerma, T., Bosma, E., Cornips, L., & Everaert, E. (2017). Cognitive advantages of bilingual children in different sociolinguistic contexts. *Frontiers in Psychology*, *8*, Article 552. https://doi.org/10.3389/fpsyg.2017.00552

*Boerma, T., Leseman, P., Wijnen, F., & Blom, E. (2017). Language proficiency and sustained attention in monolingual and bilingual children with and without language impairment. *Frontiers in Psychology*, *8*, Article 1241. https://doi.org/10.3389/fpsyg.2017.01241

*Bonifacci, P., Giombini, L., Bellocchi, S., & Contento, S. (2011). Speed of processing, anticipation, inhibition and working memory in bilinguals. *Developmental Science*, *14*(2), 256–269. https://doi.org/10.1111/j.1467-7687.2010.00974.x

*Bosman, A. M. T., & Janssen, M. (2017). Differential relationships between language skills and working memory in Turkish–Dutch and native-Dutch first-graders from low-income families. *Reading and Writing*, *30*(9), 1945–1964. https://doi.org/10.1007/s11145-017-9760-2

*Brito, N. H., & Noble, K. G. (2018). The independent and interacting effects of socioeconomic status and dual-language use on brain structure and cognition. *Developmental Science*, *21*(6), 1–10. https://doi.org/10.1111/desc.12688

*Buac, M., Gross, M., & Kaushanskaya, M. (2016). Predictors of processing-based task performance in bilingual and monolingual children. *Journal of Communication Disorders*, *62*, 12–29. https://doi.org/10.1016/j.jcomdis.2016.04.001

*Buac, M., & Kaushanskaya, M. (2014). The relationship between linguistic and non-linguistic cognitive control skills in bilingual children from low socio-economic backgrounds. *Frontiers in Psychology*, *5*, Article 1098. https://doi.org/10.3389/fpsyg.2014.01098

*Burch, R. K. (1987). *The relationship between English proficiency and digit span performance in Mexican-American children* [Doctoral dissertation, The University of Arizona]. UA Campus Repository. https://repository.arizona.edu/handle/10150/184047

*Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, *130*(3), 278–288. https://doi.org/10.1016/j.cognition.2013.11.015

*Cape, R., Vega-Mendoza, M., Bak, T. H., & Sorace, A. (2018). Cognitive effects of Gaelic medium education on primary school children in Scotland. *International Journal of Bilingual Education and Bilingualism*. Advance online publication. https://doi.org/10.1080/13670050.2018.1543648

*Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, *11*(2), 282–298. https://doi.org/10.1111/j.1467-7687.2008.00675.x

*Chan, K. T. (2004). *Chinese-English bilinguals' theory-of-mind development* [Doctoral dissertation]. University of Toronto.

*Cho, I., Park, J., Song, H., & Morton, J. B. (2021). *Disentangling language status and country-of-origin explanations of the* *bilingual advantage in preschoolers* [Manuscript submitted for publication].

*Choi, J. Y., Jeon, S., & Lippard, C. (2018). Dual language learning, inhibitory control, and math achievement in Head Start and kindergarten. *Early Childhood Research Quarterly*, *42*, 66–78. https://doi.org/10.1016/j.ecresq.2017.09.001

*Christoffels, I. K., de Haan, A. M., Steenbergen, L., van den Wildenberg, W. P. M., & Colzato, L. S. (2015). Two is better than one: Bilingual education promotes the flexible mind. *Psychological Research*, *79*(3), 371–379. https://doi.org/10.1007/s00426-014-0575-3

*Chung-Fat-Yim, A., Himel, C., & Bialystok, E. (2019). The impact of bilingualism on executive function in adolescents. *International Journal of Bilingualism*, *23*, 1278–1290. https://doi.org/10.1177/1367006918781059

*Climie, E. (2008). *Cognitive flexibility in bilingual preschool children* [Doctoral dissertation, University of Calgary]. https://central.bac-lac.gc.ca/.item?id=MR44254&op=pdf&app=Library&oclc_number=694268301

*Cockcroft, K. (2016). A comparison between verbal working memory and vocabulary in bilingual and monolingual South African school beginners: Implications for bilingual language assessment. *International Journal of Bilingual Education and Bilingualism*, *19*(1), 74–88. https://doi.org/10.1080/13670050.2014.964172

*Cockcroft, K., & Alloway, T. (2012). Phonological awareness and working memory: Comparisons between South African and British children. *Southern African Linguistics and Applied Language Studies*, *30*(1), 13–24. https://doi.org/10.2989/16073614.2012.693706

*Cottini, M., Pieroni, L., Spataro, P., Devescovi, A., Longobardi, E., & Rossi-Arnaud, C. (2015). Feature binding and the processing of global–local shapes in bilingual and monolingual children. *Memory and Cognition*, *43*(3), 441–452. https://doi.org/10.3758/s13421-014-0467-1

*Da Fontoura, H. A., & Siegel, L. S. (1995). Reading, syntactic, and working memory skills of bilingual Portuguese-English Canadian children. *Reading and Writing*, *7*(1), 139–153. https://doi.org/10.1007/BF01026951

*Dahlgren, S. O., Almén, H., & Dahlgren Sandberg, A. (2017). Theory of mind and executive functions in young bilingual children. *Journal of Genetic Psychology*, *178*(5), 303–307. https://doi.org/10.1080/00221325.2017.1361376

*Danahy, K., Windsor, J., & Kohnert, K. (2007). Counting span and the identification of primary language impairment. *International Journal of Language and Communication Disorders*, *42*(3), 349–365. https://doi.org/10.1080/13682820600940083

*D'Angiulli, A., Siegel, L. S., & Serra, E. (2001). The development of reading in English and Italian in bilingual children. *Applied Psycholinguistics*, *22*(4), 479–507. https://doi.org/10.1017/S0142716401004015

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

*Dell'Armi, M. (2015). L'impact du bilinguisme sur les capacités inhibitrices des enfants: Une étude-pilote [The impact of bilingualism on children's inhibitory abilities: A pilot study]. *Enfance*, *2015*(2), 245–259. https://doi.org/10.4074/S0013754515002050

DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences, USA*, *116*(15), 7565–7574. https://doi.org/10.1073/pnas.1811513116

*De Sousa, D. (2012). Cognitive processing skills in monolingual and bilingual South African children: Implications for assessment in linguistically diverse societies. *Language Matters*, *43*(1), 97–112. https://doi.org/10.1080/10228195.2011.590516

*De Sousa, D. S., Greenop, K., & Fry, J. (2010). Simultaneous and sequential cognitive processing in monolingual and bilingual children in South Africa. *South African Journal of Psychology*, *40*(2), 165–173. https://doi.org/10.1177/008124631004000206

*Diaz, V., & Farrar, M. J. (2018a). Do bilingual and monolingual preschoolers acquire false belief understanding similarly? The role of executive functioning and language. *First Language*, *38*(4), 382–398. https://doi.org/10.1177/0142723717752741

*Diaz, V., & Farrar, M. J. (2018b). The missing explanation of the false-belief advantage in bilingual children: A longitudinal study. *Developmental Science*, *21*(4), Article e12594. https://doi.org/10.1111/desc.12594

*Dick, A. S., Garcia, N. L., Pruden, S. M., Thompson, W. K., Hawes, S. W., Sutherland, M. T., Riedel, M. C., Laird, A. R., & Gonzalez, R. (2019). No evidence for a bilingual executive function advantage in the nationally representative ABCD study. *Nature Human Behaviour*, *3*(7), 692–701. https://doi.org/10.1038/s41562-019-0609-3

Downes, M. J., Brennan, M. L., Williams, H. C., & Dean, R. S. (2016). Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open*, *6*(12), Article 011458. https://doi.org/10.1136/bmjopen-2016-011458

*Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., & Carreiras, M. (2014). The inhibitory advantage in bilingual children revisited: Myth or reality? *Experimental Psychology*, *61*(3), 234–251. https://doi.org/10.1027/1618-3169/a000243

*Ebert, K. D., Rak, D., Slawny, C. M., & Fogg, L. (2019). Attention in bilingual children with developmental language disorder. *Journal of Speech, Language, and Hearing Research*, *62*(4), 979–992. https://doi.org/10.1044/2018_JSLHR-L-18-0221

*Engel de Abreu, P. M. J. (2011). Working memory in multilingual children: Is there a bilingual effect? *Memory*, *19*(5), 529–537. https://doi.org/10.1080/09658211.2011.590504

*Engel de Abreu, P. M. J., Baldassi, M., Puglisi, M. L., & Befi-Lopes, D. M. (2013). Cross-linguistic and cross-cultural effects on verbal working memory and vocabulary: Testing language-minority children with an immigrant background. *Journal of Speech, Language, and Hearing*

Research, *56*(2), 630–642. https://doi.org/10.1044/1092-4388(2012/12-0079)
*Engel de Abreu, P. M. J., Cruz-Santos, A., & Puglisi, M. L. (2014). Specific language impairment in language-minority children from low-income families. *International Journal of Language and Communication Disorders*, *49*(6), 736–747. https://doi.org/10.1111/1460-6984.12107
*Engel de Abreu, P. M. J., Cruz-Santos, A., Tourinho, C. J., Martin, R., & Bialystok, E. (2012). Bilingualism enriches the poor: Enhanced cognitive control in low-income minority children. *Psychological Science*, *23*(11), 1364–1371. https://doi.org/10.1177/0956797612443836
*Esposito, A. G., & Baker-Ward, L. (2013). Dual-language education for low-income children: Preliminary evidence of benefits for executive function. *Bilingual Research Journal*, *36*(3), 295–310. https://doi.org/10.1080/15235882.2013.837848
*Esposito, A. G., Baker-Ward, L., & Mueller, S. T. (2013). Interference suppression vs. response inhibition: An explanation for the absence of a bilingual advantage in preschoolers' Stroop task performance. *Cognitive Development*, *28*(4), 354–363. https://doi.org/10.1016/j.cogdev.2013.09.002
*Foy, J. G., & Mann, V. A. (2014). Bilingual children show advantages in nonverbal auditory executive function task. *International Journal of Bilingualism*, *18*(6), 717–729. https://doi.org/10.1177/1367006912472263
*Gangopadhyay, I., Davidson, M. M., Weismer, S. E., & Kaushanskaya, M. (2016). The role of nonverbal working memory in morphosyntactic processing by school-aged monolingual and bilingual children. *Journal of Experimental Child Psychology*, *142*, 171–194. https://doi.org/10.1016/j.jecp.2015.09.025
*Gangopadhyay, I., Ellis Weismer, S., & Kaushanskaya, M. (2019). Domain-general inhibition and lexical processing in monolingual and bilingual children: A longitudinal approach. *Cognitive Development*, *49*, 68–80. https://doi.org/10.1016/J.COGDEV.2018.11.008
*Gangopadhyay, I., McDonald, M., Weismer, S. E., & Kaushanskaya, M. (2018). Planning abilities in bilingual and monolingual children: Role of verbal mediation. *Frontiers in Psychology*, *9*, Article 323. https://doi.org/10.3389/fpsyg.2018.00323
*Garratt, L. C., & Kelly, T. P. (2008). To what extent does bilingualism affect children's performance on the NEPSY? *Child Neuropsychology*, *14*(1), 71–81. https://doi.org/10.1080/09297040701218405
*Goldman, M. C., Negen, J., & Sarnecka, B. W. (2014). Are bilingual children better at ignoring perceptually misleading information? A novel test. *Developmental Science*, *17*(6), 956–964. https://doi.org/10.1111/desc.12175
*Goldsmith, S. F., El-Baba, M., He, X., Lewis, D. J., Dirani, L. A., Liu, J., & Morton, J. B. (2021). *No bilingual advantage in children's executive attention: Congruency and sequential congruency effects in a large sample of monolingual and bilingual children* [Manuscript submitted for publication].
*Gonzalez, M. R. (2017). *Language experience and socioeconomic status (SES): Implications for language, cognitive, and brain development in bilingual children* [Doctoral dissertation, University of California, San Diego]. Electronic Theses and Dissertations. https://escholarship.org/uc/item/8p1840nr
*Gonzalez-Barrero, A. M., & Nadig, A. S. (2019). Can bilingualism mitigate set-shifting difficulties in children with autism spectrum disorders? *Child Development*, *90*(4), 1043–1060. https://doi.org/10.1111/cdev.12979
*Goriot, C., Broersma, M., McQueen, J. M., Unsworth, S., & van Hout, R. (2018). Language balance and switching ability in children acquiring English as a second language. *Journal of Experimental Child Psychology*, *173*, 168–186. https://doi.org/10.1016/j.jecp.2018.03.019
*Goriot, C., Denessen, E., Bakker, J., & Droop, M. (2016). Benefits of being bilingual? The relationship between pupils' perceptions of teachers' appreciation of their home language and executive functioning. *International Journal of Bilingualism*, *20*(6), 700–713. https://doi.org/10.1177/1367006915586470
Grundy, J. G., Chung-Fat-Yim, A., Friesen, D. C., Mak, L., & Bialystok, E. (2017). Sequential congruency effects reveal differences in disengagement of attention for monolingual and bilingual young adults. *Cognition*, *163*, 42–55. https://doi.org/10.1016/J.COGNITION.2017.02.010
*Grundy, J. G., & Keyvani Chahi, A. (2017). Post-conflict slowing effects in monolingual and bilingual children. *Developmental Science*, *20*(1), Article e12488. https://doi.org/10.1111/desc.12488
Gunnerud, H. L., ten Braak, D., Reikerås, E. K. L., Donolato, E., & Melby-Lervåg, M. (2020). Is bilingualism related to a cognitive advantage in children? A systematic review and meta-analysis. *Psychological Bulletin*, *146*(12), 1059–1083. https://doi.org/10.1037/bul0000301
*Haft, S. L., Kepinska, O., Caballero, J. N., Carreiras, M., & Hoeft, F. (2019). Attentional fluctuations, cognitive flexibility, and bilingualism in kindergarteners. *Behavioral Sciences*, *9*(5), Article 58. https://doi.org/10.3390/bs9050058
*Hansen, L. B., Macizo, P., Duñabeitia, J. A., Saldaña, D., Carreiras, M., Fuentes, L. J., & Bajo, M. T. (2016). Emergent bilingualism and working memory development in school aged children. *Language Learning*, *66*, 51–75. https://doi.org/10.1111/lang.12170
*Harvey, H. A. (2012). *Executive function development and early mathematics: Examination of dual language learners* [Doctoral dissertation, University of Denver]. Electronic Theses and Dissertations. https://digitalcommons.du.edu/etd/276
*Hutchison, L. A. (2012). *Relations between executive functioning, second language fluency, and externalizing behavior problems in early childhood* [Doctoral dissertation, George Mason University]. https://hdl.handle.net/1920/7955
*Iarocci, G., Hutchison, S. M., & O'Toole, G. (2017). Second language exposure, functional communication, and executive function in children with and without Autism Spectrum Disorder (ASD). *Journal of Autism and Developmental Disorders*, *47*(6), 1818–1829. https://doi.org/10.1007/s10803-017-3103-7
*Jaekel, N., Jaekel, J., Willard, J., & Leyendecker, B. (2019). No evidence for effects of Turkish immigrant children's

bilingualism on executive functions. *PLOS ONE*, *14*(1), Article e0209981. https://doi.org/10.1371/journal.pone.0209981

*Jalali-Moghadam, N., & Kormi-Nouri, R. (2015). The role of executive functions in bilingual children with reading difficulties. *Scandinavian Journal of Psychology*, *56*(3), 297–305. https://doi.org/10.1111/sjop.12198

*Janus, M., & Bialystok, E. (2018). Working memory with emotional distraction in monolingual and bilingual children. *Frontiers in Psychology*, *9*, Article 1582. https://doi.org/10.3389/fpsyg.2018.01582

*Kalashnikova, M., & Mattock, K. (2014). Maturation of executive functioning skills in early sequential bilingualism. *International Journal of Bilingual Education and Bilingualism*, *17*(1), 111–123. https://doi.org/10.1080/13670050.2012.746284

*Kalia, V., Daneri, M. P., & Wilbourn, M. P. (2019). Relations between vocabulary and executive functions in Spanish–English dual language learners. *Bilingualism*, *22*(1), 1–14. https://doi.org/10.1017/S1366728917000463

*Kalia, V., Lane, P. D., & Wilbourn, M. P. (2018). Cognitive control and phonological awareness in the acquisition of second language vocabulary within the Spanish-English dual immersion context. *Cognitive Development*, *48*, 176–189. https://doi.org/10.1016/J.COGDEV.2018.08.010

*Kapa, L. L., & Colombo, J. (2013). Attentional control in early and later bilingual children. *Cognitive Development*, *28*(3), 233–246. https://doi.org/10.1016/j.cogdev.2013.01.011

*Karlsson, L. C., Soveri, A., Räsänen, P., Kärnä, A., Delatte, S., Lagerström, E., Mård, L., Steffansson, M., Lehtonen, M., & Laine, M. (2015). Bilingualism and performance on two widely used developmental neuropsychological test batteries. *PLOS ONE*, *10*(4), Article e0125867. https://doi.org/10.1371/journal.pone.0125867

*Kaushanskaya, M., Gross, M., & Buac, M. (2014). Effects of classroom bilingualism on task-shifting, verbal memory, and word learning in children. *Developmental Science*, *17*(4), 564–583. https://doi.org/10.1111/desc.12142

*Kempert, S., & Hardy, I. (2015). Children's scientific reasoning in the context of bilingualism. *International Journal of Bilingualism*, *19*(6), 646–664. https://doi.org/10.1177/1367006914527803

*Kohnert, K., Windsor, J., & Kohnert, K. (2004). The search for common ground: Part II. Nonlinguistic performance by linguistically diverse learners. *Journal of Speech, Language, and Hearing Research*, *47*, 877–890.

*Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences, USA*, *109*(20), 7877–7881. https://doi.org/10.1073/pnas.1201575109

*Krizman, J., Skoe, E., & Kraus, N. (2016). Bilingual enhancements have no socioeconomic boundaries. *Developmental Science*, *19*(6), 881–891. https://doi.org/10.1111/desc.12347

*Krizman, J., Skoe, E., Marian, V., & Kraus, N. (2014). Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. *Brain and Language*, *128*(1), 34–40. https://doi.org/10.1016/j.bandl.2013.11.006

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497–514. https://doi.org/10.1080/20445911.2013.799170

*Ladas, A. I., Carroll, D. J., & Vivas, A. B. (2015). Attentional processes in low-socioeconomic status bilingual children: Are they modulated by the amount of bilingual experience? *Child Development*, *86*(2), 557–578. https://doi.org/10.1111/cdev.12332

Laloi, A., de Jong, J., & Baker, A. (2017). Can executive functioning contribute to the diagnosis of SLI in bilingual children?: A study on response inhibition. *Linguistic Approaches to Bilingualism*, *7*(3–4), 431–459. https://doi.org/10.1075/lab.15020.lal

*Lauchlan, F., Parisi, M., & Fadda, R. (2013). Bilingualism in Sardinia and Scotland: Exploring the cognitive benefits of speaking a "minority" language. *International Journal of Bilingualism*, *17*(1), 43–56. https://doi.org/10.1177/1367006911429622

*Leikin, M., & Tovli, E. (2014). Bilingualism and creativity in early childhood. *Creativity Research Journal*, *26*(4), 411–417. https://doi.org/10.1080/10400419.2014.961779

*Lesaux, N. K., & Siegel, L. S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology*, *39*(6), 1005–1019. https://doi.org/10.1037/0012-1649.39.6.1005

*Li, H. K., Oi, M., Gondo, K., & Matsui, T. (2017). How does being bilingual influence children with autism in the aspect of executive functions and social and communication competence? *Journal of Brain Science*, *47*, 21–49. https://doi.org/10.20821/jbs.47.0_21

*Loe, I. M., & Feldman, H. M. (2016). The effect of bilingual exposure on executive function skills in preterm and full-term preschoolers. *Journal of Developmental and Behavioral Pediatrics*, *37*(7), 548–556. https://doi.org/10.1097/DBP.0000000000000318

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

*Marini, A., Eliseeva, N., & Fabbro, F. (2019). Impact of early second-language acquisition on the development of first language and verbal short-term and working memory. *International Journal of Bilingual Education and Bilingualism*, *22*(2), 165–176. https://doi.org/10.1080/13670050.2016.1238865

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

*McVeigh, C., Wylie, J., & Mulhern, G. (2019). Verbal and visuospatial working memory in immersion-educated bilingual children. *International Journal of Bilingual Education and Bilingualism*, *22*(4), 505–517. https://doi.org/10.1080/13670050.2016.1271769

*Mehrani, M. B., & Zabihi, R. (2017). A comparative study of shifting ability, inhibitory control and working memory in monolingual and bilingual children. *Psychological Studies*, *62*(4), 421–427. https://doi.org/10.1007/s12646-017-0432-8

*Meir, N., & Armon-Lotem, S. (2017). Independent and combined effects of socioeconomic status (SES) and bilingualism on children's vocabulary and verbal short-term memory. *Frontiers in Psychology*, *8*, Article 1442. https://doi.org/10.3389/fpsyg.2017.01442

*Messer, M. H., Leseman, P. P. M., Boom, J., & Mayo, A. Y. (2010). Phonotactic probability effect in nonword recall and its relationship with vocabulary in monolingual and bilingual preschoolers. *Journal of Experimental Child Psychology*, *105*(4), 306–323. https://doi.org/10.1016/j.jecp.2009.12.006

*Mohades, S. G., Struys, E., Van Schuerbeek, P., Baeken, C., Van De Craen, P., & Luypaert, R. (2014). Age of second language acquisition affects nonverbal conflict processing in children: An fMRI study. *Brain and Behavior*, *4*(5), 626–642. https://doi.org/10.1002/brb3.246

*Mok, N., Tsang, L., Lee, T. M. C., & Llorente, A. M. (2008). The impact of language on the equivalence of trail making tests: Findings from three pediatric cohorts with different language dominance. *Applied Neuropsychology*, *15*(2), 123–130. https://doi.org/10.1080/09084280802083962

*Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *114*(2), 187–202. https://doi.org/10.1016/j.jecp.2012.09.002

Morton, J. B. (2015). Still waiting for real answers. *Cortex*, *73*, 352–353. https://doi.org/10.1016/j.cortex.2015.07.010

*Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science*, *10*(6), 719–726. https://doi.org/10.1111/j.1467-7687.2007.00623.x

*Mueller Gathercole, V. C., Thomas, E. M., Jones, L., Guasch, N. V., Young, N., & Hughes, E. K. (2010). Cognitive effects of bilingualism: Digging deeper for the contributions of language dominance, linguistic knowledge, socioeconomic status and cognitive abilities. *International Journal of Bilingual Education and Bilingualism*, *13*(5), 617–664. https://doi.org/10.1080/13670050.2010.488289

*Mumtaz, S., & 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*(2), 113–134. https://doi.org/10.1111/1467-9817.t01-1-00136

*Namazi, M., & Thordardottir, E. (2010). A working memory, not bilingual advantage, in controlled attention. *International Journal of Bilingual Education and Bilingualism*, *13*(5), 597–616. https://doi.org/10.1080/13670050.2010.488288

*Nayak, S. (2018). *Beyond bilingual advantages: Contexts, mechanisms, and correlates of executive function in bilingual and monolingual children* [Doctoral dissertation, Boston University]. https://hdl.handle.net/2144/27179

*Nayak, S., Salem, H. Z., & Tarullo, A. R. (2020). Neural mechanisms of response-preparation and inhibition in bilingual and monolingual children: Lateralized readiness potentials (LRPs) during a nonverbal Stroop task. *Developmental Cognitive Neuroscience*, *41*, Article 100740. https://doi.org/10.1016/j.dcn.2019.100740

*Nayak, S., & Tarullo, A. R. (2020). Error-related negativity (ERN) and 'hot' executive function in bilingual and monolingual preschoolers. *Bilingualism: Language and Cognition*, *23*(4), 897–908. https://doi.org/10.1017/S1366728919000725

*Nguyen, T. K., & Astington, J. W. (2014). Reassessing the bilingual advantage in theory of mind and its cognitive underpinnings. *Bilingualism*, *17*(2), 396–409. https://doi.org/10.1017/S1366728913000394

Nichols, E. S., Wild, C. J., Stojanoski, B., Battista, M. E., & Owen, A. M. (2020). Bilingualism affords no general cognitive advantages: A population study of executive function in 11,000 people. *Psychological Science*, *31*(5), 548–567. https://doi.org/10.1177/0956797620903113

*Nicolay, A. C., & Poncelet, M. (2015). Cognitive benefits in children enrolled in an early bilingual immersion school: A follow up study. *Bilingualism*, *18*(4), 789–795. https://doi.org/10.1017/S1366728914000868

*Niolaki, G. Z., & Masterson, J. (2012). Transfer effects in spelling from transparent Greek to opaque English in seven-to-ten-year-old children. *Bilingualism*, *15*(4), 757–770. https://doi.org/10.1017/S1366728911000721

*Okanda, M., Moriguchi, Y., & Itakura, S. (2010). Language and cognitive shifting: Evidence from young monolingual and bilingual children. *Psychological Reports*, *107*(1), 68–78. https://doi.org/10.2466/03.10.28.PR0.107.4.68-78

Paap, K. R., Anders-Jefferson, R., Mason, L., Alvarado, K., & Zimiga, B. (2018). Bilingual advantages in inhibition or selective attention: More challenges. *Frontiers in Psychology*, *9*, Article 1409. https://doi.org/10.3389/fpsyg.2018.01409

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265–278. https://doi.org/10.1016/j.cortex.2015.04.014

Paap, K. R., Myuz, H. A., Anders, R. T., Bockelman, M. F., Mikulinsky, R., & Sawi, O. M. (2017). No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *Journal of Cognitive Psychology*, *29*(2), 89–112. https://doi.org/10.1080/20445911.2016.1248436

*Park, J., Weismer, S. E., & Kaushanskaya, M. (2018). Changes in executive function over time in bilingual and monolingual school-aged children. *Developmental Psychology*, *54*(10), 1842–1853. https://doi.org/10.1037/dev0000562

*Park, J. S. (2014). *Bilingual influence on cognitive processes related to linguistic abilities in school-age children* [Doctoral dissertation, Pennsylvania State University]. Electronic Theses and Dissertations. https://etda.libraries.psu.edu/catalog/23658

*Pawlicka, P., Lipowska, M., & Gajdzińska, H. (2015). Linguistic stimulation impact on verbal working memory in the early stages of school education. *Acta Neuropsychologica*, *13*(1), 49–62. https://doi.org/10.5604/17307503.1148332

*Pearson, C. A. (1988). Cognitive differences between bilingual and monolingual children on the Kaufman Assessment Battery for Children. *Journal of Psychoeducational Assessment*, *6*(3), 271–279. https://doi.org/10.1177/073428298800600310

*Pino Escobar, G., Kalashnikova, M., & Escudero, P. (2018). Vocabulary matters! The relationship between verbal fluency and measures of inhibitory control in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *170*, 177–189. https://doi.org/10.1016/j.jecp.2018.01.012

*Poarch, G. J., & Bialystok, E. (2015). Bilingualism as a model for multitasking. *Developmental Review*, *35*, 113–124. https://doi.org/10.1016/j.dr.2014.12.003

*Poarch, G. J., & van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: Evidence from second-language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, *113*(4), 535–551. https://doi.org/10.1016/j.jecp.2012.06.013

*Purić, D., Vuksanović, J., & Chondrogianni, V. (2017). Cognitive advantages of immersion education after 1 year: Effects of amount of exposure. *Journal of Experimental Child Psychology*, *159*, 296–309. https://doi.org/10.1016/j.jecp.2017.02.011

Pustejovsky, J. E., & Rodgers, M. A. (2019). Testing for funnel plot asymmetry of standardized mean differences. *Research Synthesis Methods*, *10*(1), 57–71. https://doi.org/10.1002/jrsm.1332

*Rainey, V. R., Davidson, D., & Li-Grining, C. (2016). Executive functions as predictors of syntactic awareness in English monolingual and English-Spanish bilingual language brokers and nonbrokers. *Applied Psycholinguistics*, *37*(4), 963–995. https://doi.org/10.1017/S0142716415000326

*Raudszus, H., Segers, E., & Verhoeven, L. (2018). Lexical quality and executive control predict children's first and second language reading comprehension. *Reading and Writing*, *31*(2), 405–424. https://doi.org/10.1007/s11145-017-9791-8

*Riggs, N. R., Shin, H.-S. S., Unger, J. B., Spruijt-Metz, D., & Pentz, M. A. (2014). Prospective associations between bilingualism and executive function in Latino children: Sustained effects while controlling for biculturalism. *Journal of Immigrant and Minority Health*, *16*(5), 914–921. https://doi.org/10.1007/s10903-013-9838-0

*Robinson, M. G., & Sorace, A. (2019). The influence of collaborative language learning on cognitive control in unbalanced multilingual migrant children. *European Journal of Psychology of Education*, *34*(1), 255–272. https://doi.org/10.1007/s10212-018-0377-x

*Ross, J., & Melinger, A. (2017). Bilingual advantage, bidialectal advantage or neither? Comparing performance across three tests of executive function in middle childhood. *Developmental Science*, *20*(4), Article e12405. https://doi.org/10.1111/desc.12405

*Rothou, K. M., & Tsimpli, I. (2020). Biliteracy and reading ability in children who learn Greek as a second language. *International Journal of Bilingual Education and Bilingualism*, *23*(8), 1036–1050. https://doi.org/10.1080/13670050.2017.1386614

*Santillán, J., & Khurana, A. (2018). Developmental associations between bilingual experience and inhibitory control trajectories in Head Start children. *Developmental Science*, *21*(4), Article e12624. https://doi.org/10.1111/desc.12624

*Sawan, M. (2015). *The relationship between bilingualism and cognitive development among Saudi Arabian school-aged children* [Doctoral dissertation, Howard University].

*Serratrice, L., & De Cat, C. (2020). Individual differences in the production of referential expressions: The effect of language proficiency, language exposure and executive function in bilingual and monolingual children. *Bilingualism*, *23*(2), 371–386. https://doi.org/10.1017/S1366728918000962

*Shoghi Javan, S., & Ghonsooly, B. (2018). Learning a foreign language: A new path to enhancement of cognitive functions. *Journal of Psycholinguistic Research*, *47*(1), 125–138. https://doi.org/10.1007/s10936-017-9518-7

*Soliman, A. M. (2014). Bilingual advantages of working memory revisited: A latent variable examination. *Learning and Individual Differences*, *32*, 168–177. https://doi.org/10.1016/j.lindif.2014.02.005

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*(1), 60–78. https://doi.org/10.1002/jrsm.1095

*Stephens, C. (2013). *Executive function development: A comparison of monolingual and bilingual children in Ireland* [Doctoral dissertation, Queen's University Belfast].

*Struys, E., Duyck, W., & Woumans, E. (2018). The role of cognitive development and strategic task tendencies in the bilingual advantage controversy. *Frontiers in Psychology*, *9*, Article 1790. https://doi.org/10.3389/fpsyg.2018.01790

*Thorn, A. S. C., & Gathercole, S. E. (1999). Language-specific knowledge and short-term memory in bilingual and nonbilingual children. *The Quarterly Journal of Experimental Psychology Section A*, *52*(2), 303–324. https://doi.org/10.1080/713755823

*Timmermeister, M., Leseman, P., Wijnen, F., & Blom, E. (2020). No bilingual benefits despite relations between language switching and task switching. *Frontiers in Psychology*, *11*, Article 1832. https://doi.org/10.3389/fpsyg.2020.01832

*Tran, C. D., Arredondo, M. M., & Yoshida, H. (2019). Early executive function: The influence of culture and bilingualism. *Bilingualism*, *22*(4), 714–732. https://doi.org/10.1017/S1366728918000160

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods*, *45*(2), 576–594. https://doi.org/10.3758/s13428-012-0261-6

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2015). Meta-analysis of multiple outcomes: A multilevel approach. *Behavior Research Methods*, *47*(4), 1274–1294. https://doi.org/10.3758/s13428-014-0527-2

*Vega, C. (2009). *Cognitive flexibility advantages in bilingual children* [Doctoral dissertation, Carlos Albizu University].

*Weber, R. C. (2011). *How hot or cool is it to speak two languages: Executive function advantages in bilingual children* [Doctoral dissertation, Texas A&M University]. https://hdl.handle.net/1969.1/ETD-TAMU-2011-08-10028

*White, L. J., & Greenfield, D. B. (2017). Executive functioning in Spanish- and English-speaking Head Start preschoolers. *Developmental Science*, *20*(1), Article e12502. https://doi.org/10.1111/desc.12502

*White, M. J. (2019). The development of English proficiency and working memory in 5–6 year old ELLs in their first year of formal education. *International Journal of Bilingual Education and Bilingualism*. Advance online publication. https://doi.org/10.1080/13670050.2019.1571009

*Yang, S., & Yang, H. (2016). Bilingual effects on deployment of the attention system in linguistically and culturally homogeneous children and adults. *Journal of Experimental Child Psychology*, *146*, 121–136. https://doi.org/10.1016/j.jecp.2016.01.011

*Yang, S., Yang, H., & Lust, B. (2011). Early childhood bilingualism leads to advances in executive attention: Dissociating culture and language. *Bilingualism*, *14*(3), 412–422. https://doi.org/10.1017/S1366728910000611

*Yao, L. (2014). *Monolingualism, bilingualism, and executive function of children from low-income families* [Doctoral dissertation, University of Nebraska]. https://digitalcommons.unl.edu/dissertations/AAI3632757/

*Yu, H., Yao, N., Zhang, J., & Gao, B. (2019). Time-course of attentional bias for culture-related cues in Mongolian-Chinese bilingual children. *International Journal of Bilingualism*, *23*(6), 1483–1501. https://doi.org/10.1177/1367006918808047

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology*, *1*(2), 198–226. https://doi.org/10.1037/1089-2680.1.2.198

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., Chiang, J. K., Hongwanishkul, D., Schuster, B. V., & Sutherland, A. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, *68*(3), vii–137.

*Zeng, Z., Kalashnikova, M., & Antoniou, M. (2019). Integrating bilingualism, verbal fluency, and executive functioning across the lifespan. *Journal of Cognition and Development*, *20*(5), 656–679. https://doi.org/10.1080/15248372.2019.1648267