Review of Educational Research
June 2010, Vol. 80, No. 2, pp. 207–245
DOI: 10.3102/0034654310368803

A Systematic Review and Meta-Analysis of the Cognitive Correlates of Bilingualism

Olusola O. Adesope
Washington State University
Tracy Lavin and Terri Thompson
Directions Evidence and Policy Research Group
Charles Ungerleider
University of British Columbia

A number of studies have documented the cognitive outcomes associated with bilingualism. To gain a clear understanding of the extent and diversity of these cognitive outcomes, the authors conducted a meta-analysis of studies that examined the cognitive correlates of bilingualism. Data from 63 studies (involving 6,022 participants) were extracted and analyzed following established protocols and procedures for conducting systematic reviews and guidelines for meta-analysis. Results indicate that bilingualism is reliably associated with several cognitive outcomes, including increased attentional control, working memory, metalinguistic awareness, and abstract and symbolic representation skills. Overall mean effect sizes varied from small to large, depending on the cognitive outcomes measured, and were moderated by methodological features of the studies.

KEYWORDS: bilingual, cognitive correlates, biliteracy, meta-analysis, systematic review.

Early research on bilingualism warned that bilingualism could be deleterious to learning. These early studies concluded that monolingual students outperformed bilingual students on a range of cognitive tasks (for reviews, see Bhatia & Ritchie, 2006; Hakuta, 1986; Macnamara, 1966).

In a seminal article, Peal and Lambert (1962) introduced the concept of "balanced bilinguals" and demonstrated the methodological weaknesses of previous bilingual studies, providing a new approach for research on bilingualism. Peal and Lambert noted that early studies on the effects of bilingualism did not properly match bilingual and monolingual participants along several dimensions, including socioeconomic status (SES), second language proficiency (pseudobilingualism), language of assessment, gender, age, and urban–rural contexts. They noted that these and other factors may have confounded earlier results showing bilingual disadvantages on cognitive measures. Controlling for these extraneous factors, Peal and Lambert found that bilingual participants significantly outperformed monolinguals on several measures of verbal and nonverbal intelligence.

Since Peal and Lambert's (1962) original studies, a considerable body of evidence has accumulated suggesting that bilingualism confers a number of cognitive benefits. For example, researchers have observed that bilinguals may have greater metalinguistic awareness (Bialystok, 1987, 1988, 2001b; Diaz, 1985; Diaz & Klinger, 1991; Ferdman & Hakuta, 1985; Goetz, 2000; Hakuta, 1990; Huber & Lasagabaster, 2000; Ricciardelli, 1993; Titone, 1997) and enhanced metacognitive skills (Duncan, 2005). Bilinguals may have stronger symbolic representation and abstract reasoning skills (Bamford & Mizokawa, 1990, 1992; Berguno & Bowler, 2004; Chan, 2005; Diaz, 1985; Goncz, 1988; Johnson, 1991; McLeay, 2003), as well as better learning strategies (Bochner, 1996; Ponomarev, 1992). Bilinguals may also have enhanced problem-solving skills because of their ability to selectively attend to relevant information and disregard misleading information (Bamford & Mizokawa, 1991; Bialystok, 1999, 2001a, 2005; Bialystok & Majumber, 1998; Duncan, 2005; Stephens, 1997) and may be able to use this selectivity to succeed at theory-of-mind tasks, which require the ability to attribute the behavior of others to their own distinct beliefs, desires, and intentions (Chan, 2005; Goetz, 2000). Bilinguals may have enhanced creative and divergent thinking skills (Braccini & Cianchi, 1993; Ho, 1987; Konaka, 1997; Ricciardelli, 1993; Srivastava, 1991) and greater cognitive flexibility (Hakuta, 1990; Iannaccone, Fraternali, & Vaccia, 1992; Kovacs & Teglas, 2002; Kozulin, 1999).

Although many studies have documented advantages for bilinguals on cognitive tasks, other studies have reported negative, null, or mixed effects of bilingualism (Macnamara, 1966; Rosenblum & Pinker, 1983). To make sense of these conflicting findings, the current work synthesizes the available research on the cognitive correlates of bilingualism. This review does not investigate the effectiveness of bilingual education programs because previous reviews, meta-analyses, and best-evidence syntheses have addressed the question of program effectiveness, albeit with inconclusive results (Baker & de Kanter, 1981; Rossell & Baker, 1996; Slavin & Cheung, 2005; Willig, 1985, 1987). Specifically, the current review focuses on examining the cognitive correlates of bilingualism and the associated effect sizes. The following section discusses potential relationships between bilingualism and various cognitive skills.

Attentional Control

There is considerable evidence that bilingual speakers are more readily able to control their attention while engaged in linguistic and nonverbal tasks compared to monolingual learners (Bialystok, 2001a; Bialystok, Craik, Klein, & Viswanathan, 2004; Bialystok, Craik, & Ryan, 2006; Emmorey, Luk, Pyers, & Bialystok, 2008). Several explanations have been advanced for this cognitive advantage. A dominant perspective suggests that the regular use of two languages requires that bilinguals control their attention and select the target language. Some researchers have claimed that the ability to selectively attend to different representations may be responsible for the greater attentional control exhibited by bilingual participants in many studies (Bialystok, 2001a; Bialystok, Martin, & Viswanathan, 2005; Yoshida, 2008). In other words, these researchers speculate that bilingual learners' ability to concurrently hold two languages in the mind, resisting intrusions of words and grammar from one language into the

other, might explain the greater control reflected by improved performance on tasks with conflicting or distracting information.

More recently, researchers have also shown that the cognitive control of attention found in studies with bilingual children appears to be sustained into adulthood. For example, Bialystok et al. (2004) found that adults who have been bilingual since childhood are more capable than comparable monolingual adults of managing their attention when presented with tasks requiring cognitive control.

In addition, there is a growing body of evidence that bilingualism may help offset some age-related cognitive declines by building cognitive reserves that slow the aging process for adults (Bialystok, Craik, & Freedman, 2007; Bialystok et al., 2004). In a recent study on the effect of lifelong bilingualism on age-related cognitive decline, Bialystok et al. (2007) found that bilingual adults showed symptoms of dementia 4 years later than comparable monolinguals, even when other factors remained constant. The preliminary findings in the literature suggest that "the lifelong experience of managing two languages attenuates the age-related decline in the efficiency of inhibitory processing" (Bialystok et al., 2004, p. 301).

Working Memory

There are at least two contrasting hypotheses about the relationship between bilingualism and working memory. First, the need to manage two languages concurrently could place greater demands on working memory. This hypothesis suggests that bilingualism may impede efficient processing of information in working memory because of the cognitive load imposed on working memory (Lee, Plass, & Homer, 2006; Sweller & Chandler, 1994; van Merrienboer & Sweller, 2005). Conversely, bilinguals' well-developed ability to inhibit one language while using the other may increase the efficiency of their working memory capacity because working memory resources are properly managed through such inhibitory processing (Bialystok et al., 2004; Bialystok, Craik, & Luk 2008; Fernandes, Craik, Bialystok, & Kreuger, 2007; Just & Carpenter, 1992; Michael & Gollan, 2005; Rosen & Engle, 1997).

Research on these competing hypotheses has yielded inconclusive findings with results depending on the nature of the task (Bialystok et al., 2008). On tasks that require greater attentional control, bilinguals appear to have greater working memory capacity than monolinguals (Engle, 2002; Kane, Bleckley, Conway, & Engle, 2001). In attention-aided tasks, however, the bilingual advantage disappears (Yang, Yang, Ceci, & Wang, 2005).

Metalinguistic Awareness

Metalinguistic awareness is the ability to think about language. It is the explicit awareness of linguistic forms and structures and an understanding of how these relate to and produce meaning (Cazden, 1974). It is hypothesized that the experience of acquiring and maintaining two different languages—with different forms and structures—allows bilingual speakers to develop an explicit and articulated understanding of how language works. For example, bilingual speakers have two different words for most concepts. Reflecting on this can point to the insight that words are only arbitrarily and symbolically related to their underlying concepts (e.g., knowing that *dog* and *chien* are concepts for dog makes it obvious that the

word *dog* is only an arbitrary symbol). Similarly, when syntactic rules differ across languages, bilingual speakers of those languages may notice the differences and become explicitly aware of the syntactic rules—which most monolingual speakers will know only implicitly. For example, English–Japanese bilinguals may note that objects follow verbs in English sentences (e.g., "I like chocolate") but objects precede verbs in Japanese sentences (e.g., "I chocolate like"). Noting this distinction provides insight into the specific grammatical rules in each language as well as into the universal properties of human language.

Over the past decades, researchers have investigated the effects of bilingualism on children's metalinguistic development. The majority of studies have found that bilingual speakers, particularly those highly proficient in both languages, demonstrate greater metalinguistic awareness than their monolingual counterparts (Bialystok, Majumder & Martin, 2003; Campbell & Sais, 1995; Galambos & Hakuta, 1988).

Metacognitive Awareness

Metacognitive awareness refers to knowledge about one's own cognitive processes. It is an awareness of one's own learning strategies and the mental activities required to self-regulate the learning process (Flavell, 1978). The process of learning the vocabulary, syntax, phonology, and morphology of more than one language, as well as learning how to use this body of knowledge in contextually appropriate fashion, may provide bilingual speakers special insight into their own cognitive processes and learning strategies (Kemp, 2007). Research comparing the metacognitive awareness of bilinguals and monolinguals is scant but has generally found that bilinguals show greater metacognitive awareness than monolinguals (Ransdell, Barbier, & Niit, 2006; Vorstman, De Swart, Ceginskas, & Van Den Bergh, 2009).

Abstract or Symbolic Reasoning and Creative and Divergent Thinking

Across a number of studies, bilinguals have shown enhanced skills with respect to creative and divergent thinking and to abstract and symbolic reasoning. In an investigation on creativity and bilingualism, Ricciardelli (1992) found that bilinguals outperformed monolinguals in 20 of the 24 studies reviewed, showing a clear positive relationship between bilingualism and creativity or divergent thinking. Peal and Lambert (1962) suggested that bilingual children develop greater cognitive flexibility and creativity as a result of switching between two languages and two different perspectives. As well, Cummins (1976) has proposed that bilingualism spurs the development of abstract and symbolic reasoning through the experience of having two different words for most concepts. This helps bilingual children understand that the relationship between words and their referents is entirely arbitrary and represents an abstract symbolic relationship.

Problem Solving

Bilinguals also show evidence of enhanced problem-solving skills, particularly on tasks requiring executive control (i.e., planning, cognitive flexibility, abstract thinking, rule acquisition, initiating appropriate actions and inhibiting inappropriate actions, and selecting relevant sensory information; Baddeley, 1996). A bilingual advantage has been demonstrated using the Simon task, the

dimensional change card sort task, and other similar tasks used to assess executive control for problem-solving tasks (Bialystok, 1999, 2006). Simon tasks refer to a family of tasks typically used to investigate interference effects. In the Simon task, stimuli are presented with different target features and in different positions. For example, participants may be asked to indicate the color of either a red or a green square presented on one side of the screen by pressing a left or a right key. The general finding in the Simon task is that reaction times are slowed when the spatial location of the target and its response coding do not correspond (incongruent condition) versus when spatial location and response coding correspond (congruent condition). An *incongruent* trial occurs when a signal is presented to the right but its color requires a left-hand button press. Conversely, signals that require a left-hand response and are also presented on the left side are referred to as *congruent trials*. Typically, reaction times are slower to incongruent compared to congruent trials, a finding referred to as the *congruency effect* or *interference effect*.

The enhanced problem-solving ability may be because of the cognitive flexibility associated with bilingualism. Because bilinguals have the capacity to choose between two languages, they may develop more flexibility with respect to thinking that can be applied to solve problems.

Purpose of the Study

Since Peal and Lambert's (1962) seminal work, a number of studies have documented the positive cognitive correlates of bilingualism; however, the magnitude of these effects remains unclear. The majority of recent studies have shown positive effects of bilingualism, but some studies have shown that bilingual performance is relatively impaired on some cognitive tasks. A few studies have also demonstrated mixed effects of bilingualism on performance on cognitive tasks. The current analysis is an attempt to synthesize these results. The goal of the present study is to meta-analyze research on the cognitive outcomes of bilingualism by estimating the effects of bilingualism on specific cognitive measures. Specifically, the meta-analysis addresses the following research questions:

1. What are the cognitive correlates associated with bilingualism?
2. How do the effects of bilingualism vary when cognitive outcomes are measured in different geographical locations, in different settings, and at different educational levels?
3. How are effect sizes influenced by different combinations of languages spoken by bilinguals?
4. Are the effect sizes influenced by methodological features of the research?

Method

Selection Criteria

To capture all relevant studies on the cognitive benefits of bilingualism, specific criteria for inclusion were developed. Studies were deemed eligible if:

a. Bilingual participants were reported to be equally (or almost equally) proficient in two languages. Thus, participants who were learning second

languages were not regarded as bilinguals but rather as second language learners. Studies with such second language learners were excluded from this meta-analysis. Bilinguals with learning disabilities or other cognitive disabilities were excluded.

b. They had an experimental group of bilingual participants and a control group of monolingual participants.
c. Measured outcomes (cognitive benefits) were clearly reported. These include attentional control, problem-solving skills, creative and divergent thinking, cognitive flexibility, learning strategies, symbolic representation and abstract reasoning skills, metalinguistic awareness, metacognitive skills, and working memory. We excluded studies that measured only psychosocial outcomes such as employability or social problem solving and other activities such as code switching, cross-language priming, and social identity.
d. They reported sufficient data to allow for effect size calculations. When basic descriptive statistics were not included in a study, other statistics were used (e.g., *t* and *F* statistics), but we coded for reviewer's confidence in effect size derivation. Studies with insufficient data for effect size calculations were excluded.
e. They were publicly available, either online or in library archives.

For multiple reports of the same study (e.g., dissertation, conference proceedings, and journal article), the version published as a journal article was coded, but in some cases other versions of the published article (e.g., conference proceedings) were used to make the coding features more extensive and accurate.

Location and Selection of Studies

A comprehensive and systematic search was conducted in the following electronic databases to locate appropriate studies: Academic Search Premier, Education Full Text, ERIC (including British and Australian ERIC), Linguistic and Language Behavior Abstracts, PsycINFO, and Web of Science. A primary search was conducted utilizing Boolean combinations of the controlled vocabulary within each database for the terms *immigrant students*, *bilingualism*, and *cognition*. A manual search of the reference lists of earlier reviews of the literature on bilingualism (e.g., Bialystok, 2002; Cenoz, 2003) was subsequently conducted.

A total of 5,185 articles were obtained from the search procedure. Two researchers reviewed the titles, abstracts, and keywords of these articles for possible inclusion by applying the selection criteria stated above. When abstracts did not contain sufficient information to determine inclusion or exclusion, the full text of the article was obtained and read. Duplicate studies were removed, and articles that did not meet the selection criteria were excluded. Interrater agreement was computed to determine the reliability of including or excluding articles based on reading only the abstracts. This yielded a Cohen's kappa of .88. Researchers discussed all disagreements until they were fully resolved. A total of 157 articles were retained for secondary screening.

Two researchers independently read the full texts of each of the 157 articles retained after first inclusion to further determine their suitability based on the specified criteria for inclusion. Only 39 articles met the second inclusion criterion, and data from these articles were extracted using EPPI-Reviewer, an online application for managing and conducting systematic reviews (Thomas & Brunton,

2006). Coded variables were organized into 11 major categories in the database. These include (a) study identification, (b) study characteristics and measured outcomes, (c) research questions, (d) research design, (e) groups and randomization strategy, (f) sampling strategy, (g) characteristics of samples in the study, (h) recruitment and consent, (i) data collection, (j) data analysis, and (k) results and conclusion. In cases where some variables were not explicitly stated in the study, reviewers made reasonable inferences and noted the absence of explicit information. The appendix shows the coding book containing a description of all variables coded under each category.

Although there are many variants of bilingualism (early bilinguals, late bilinguals, balanced bilinguals, etc.), sufficient information was not provided in many of the studies to code this variable. Nevertheless, as highlighted earlier in the article, all of the studies that met our inclusion criteria had participants who were proficient in two languages before the start of each study.

Some articles reported multiple studies. Hence, 63 studies with an overall sample size of 6,022 participants were reported in all 39 articles and were included for meta-analysis. Another interrater reliability analysis was conducted to determine agreement among researchers on inclusion or exclusion judgments based on full-text review of all 63 studies, yielding a Cohen's kappa of .92. Again, researchers discussed all disagreements until they were fully resolved.

Throughout the design and implementation of this review, guidelines for meta-analysis provided by Lipsey and Wilson (2001) were followed. The coding scheme prevented inappropriately combining statistically dependent comparisons in calculating mean effect sizes. To generate a single distribution of effect sizes, a mean effect size was obtained for each set of statistically dependent effect sizes by averaging over different cognitive outcomes and study characteristics.

Extraction and Calculation of Effect Sizes

Effect size is a standardized metric obtained by calculating the difference between the means of the experimental (bilingual) and control (monolingual) groups divided by the pooled standard deviation of the two groups. Hedges (1981) observed that estimates may yield inflated effect sizes when samples are small. To correct for such bias in effect size estimation, especially with small sample sizes (Lipsey & Wilson, 2001, p. 48), the obtained Cohen's d values were converted to Hedges's g, an unbiased estimate (Hedges & Olkin, 1985, p. 81) of the standardized mean difference effect size. When other statistics such as F or t were provided, these were also used to derive effect sizes or, in some cases, to verify the obtained d (Cooper & Hedges, 1994).

Data Analysis

Standard meta-analytic guidelines and equations were followed in all data analyses (Cooper & Hedges, 1994; Hedges & Olkin, 1985; Lipsey & Wilson, 2001). All data analyses were conducted using Comprehensive Meta-Analysis Version 2.2.048 (Borenstein, Hedges, Higgins, & Rothstein, 2008) and SPSS Version 16.0 for Windows.

Aggregating Effect Sizes

To aggregate effect sizes, the inverse variance weight was computed for each finding. An aggregate effect size was then obtained from the weighted effect sizes

to derive an overall weighted mean estimate of the effect of the treatment. This allowed more weight to be assigned to studies with larger sample sizes. The standard error of Hedges's unbiased estimate of the mean effect size was then computed.

A 95% confidence interval was computed around each weighted mean effect size to determine statistical significance. Confidence intervals spanning a range above zero were interpreted as indicating a statistically detectable result favoring bilinguals with respect to the associated cognitive outcome.

An important aspect of meta-analysis involves the determination of whether the various effect sizes that are averaged into a mean value all estimate the same population effect size. This assumption of homogeneity of variance was tested by the Q statistic. When all findings share the same population value, Q has an approximate chi-square distribution with $k - 1$ degrees of freedom, where k is the number of effect sizes or studies for a particular subset. When Q exceeded the critical value of the chi-square distribution (i.e., $p < .05$), the mean effect size was reported to be significantly heterogeneous, meaning that there was more variability in the effect sizes than would be expected from sampling error and suggesting that each effect size did not estimate a common population mean (Lipsey & Wilson, 2001). The I^2 statistic is reported as a complement to interpret the result of the homogeneity test (Higgins & Thompson, 2002; Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006).[1]

Results

After resolving statistical independence, 63 studies (from 39 articles) were analyzed.

Table 1 shows a summary of the variables coded for each of the 63 studies. This includes the study, grade level of participants, total number of participants involved in each study, languages spoken by the bilingual participants, cognitive benefits measured, and unbiased effect size, Hedges's g. In this and subsequent tables, positive effect sizes show bilingual advantages whereas negative effect sizes show an advantage for monolinguals over bilinguals on cognitive measures.

For outlier analysis, we used the Comprehensive Meta-Analysis software to determine the effect of removing a number of effect sizes from the distribution of effect. The forest plot of the 63 standardized mean difference effect sizes for the cognitive benefits of bilingualism was examined, and 3 potential outlying studies were removed. The recalculated results did not increase the fit of the remaining effect sizes to a simple model of homogeneity ($g = 0.39$; $Q_{\text{total}(59)} = 287.61$, $p < .001$; $I^2 = 79.50\%$). Because the removal of potential outliers did not produce a homogeneous model, a decision was made not to remove any effect sizes from the original distribution.

Overall Relationship Between Cognitive Outcomes and Bilingualism

Table 2 shows the overall weighted mean and homogeneity statistics of all statistically independent effect sizes. Table 2 and subsequent tables include the number of participants (N) in each category, the number of findings (k), the weighted mean effect size (g) and its standard error (SE), the 95% confidence interval around the mean, a test of the null hypothesis (z), a test of heterogeneity (Q) with its

TABLE 1
Summary of coded studies and associated effect sizes

Study	*Grade range*	*N*	*Languages (bilingual)*[a]	*Dominant cognitive benefits measured*	*Effect size (g)*
Abu-Rabia and Siegel (2002)	4–7	63	Arabic–English	Working memory for normally developed	0.20
Abu-Rabia and Siegel (2002)	4–7	31	Arabic–English	Working memory for reading disabled	–0.23
Bialystok (1997)	PreK	81	French–English	ASR: Symbolic representation of print[b]	1.04*
Bialystok (1997)	PreK	87	Chinese–English	ASR: Symbolic representation of print	1.32*
Bialystok (1997)	PreK	81	French–English	ASR: Symbolic representation[c]	0.07
Bialystok (1997)	PreK	87	Chinese–English	ASR: Symbolic representation	–0.26
Bialystok (2006)	PSec	97	Mixed–English	Working memory using Simon tasks	0.51*
Bialystok (1999)	PreK	30	Chinese–English	Attentional Control[d]	0.82*
Bialystok (1999)	PreK	30	Chinese–English	Attentional Control[e]	0.69
Bialystok, Craik, et al. (2005)	PSec	29	French–English	Attentional Control	0.63
Bialystok, Craik, and Ruocco (2006)[f], Exp. 1	PSec	48	Mixed–English	Working memory[g]	0.48
Bialystok, Craik, and Ruocco (2006), Exp. 2	PSec	48	Mixed–English	Working memory[h]	–0.30
Bialystok et al. (2004), Exp. 1	PSec	20	Tamil–English	Attentional control: MRT for younger adults[i]	2.25*
Bialystok et al. (2004), Exp. 1	PSec	20	Tamil–English	Attentional control: MRT for older adults	1.03*
Bialystok et al. (2004), Exp. 2	PSec	64	Tamil–English	Attentional control: MRT for younger	2.63*
Bialystok et al. (2004), Exp. 2	PSec	30	Mixed–English	Attentional control: MRT for older adults	1.35*
Bialystok et al. (2004), Exp. 2	PSec	20	French–English	Attentional control: MRT for older adults	2.45*
Bialystok, Luk, et al. (2005b)	1	132	Mixed–English	Metalinguistic awareness[j]	0.95*
Bialystok and Majumder (1998)	K–3	71	Mixed–English	Problem solving[k]	0.35
Bialystok et al. (2003), Exp. 1	K–3	72	French–English	Metalinguistic awareness	–0.02
Bialystok et al. (2003), Exp. 2	K–3	75	French–English	Metalinguistic awareness	–0.45
Bialystok et al. (2003), Exp. 3	K–3	64	Chinese–English	Metalinguistic awareness[l]	–0.92*
Bialystok et al. (2003), Exp. 3	K–3	58	Spanish–English	Metalinguistic awareness	0.51
Bialystok and Martin (2004), Exp. 1	PreK	67	Cantonese–English	Attentional control using the DCCS task[m]	0.35
Bialystok and Martin (2004), Exp. 2	PreK	30	French–English	Attentional control using manual DCCS task	0.83*
Bialystok and Martin (2004), Exp. 3	PreK	53	Chinese–English	Attentional control: color shape and color object	0.63*
Bialystok and Martin (2004), Exp. 3	PreK	53	Chinese–English	Attentional control: functional location and place	0.42
Bialystok, McBride-Chang, et al. (2005), Exp. 1	K–1	67	Chinese–English	Metalinguistic awareness	–0.04

(continued)

TABLE 1 (continued)

Study	*Grade range*	*N*	*Languages (bilingual)*[a]	*Dominant cognitive benefits measured*	*Effect size (g)*
Bialystok, McBride-Chang, et al. (2005), Exp. 2	K–1	67	Chinese–English	Metalinguistic awareness	–0.30
Bialystok and Senman (2004), Exp. 2	PreK	95	Mixed–English	Attentional control[n]	1.59*
Bialystok and Shapero (2005), Exp. 1	K–1	48	Mixed–English	ASR: Reversibility of ambiguous figures[o]	1.14*
Bialystok and Shapero (2005), Exp. 2	K	53	Mixed–English	Attentional control using different tasks[p]	0.18
Bruck and Genesee (1995)	K	137	French–English	Metalinguistic awareness	–0.44*
Campbell and Sais (1995)	PreK	30	Italian–English	Metalinguistic awareness[q]	0.64
Chen et al. (2004), Exp. 1	2	125	Canton-Mandarin	Metalinguistic awareness: tone, onset and rime tasks	0.20
Chen et al. (2004), Exp. 1	4	150	Canton-Mandarin	Metalinguistic awareness: tone, onset and rime tasks	–0.30
Chen et al. (2004), Exp. 2	1	94	Canton-Mandarin	Metalinguistic awareness tasks[r]	0.36
Chen et al. (2004), Exp. 2	2	89	Canton-Mandarin	Metalinguistic awareness: tone, onset and rime tasks	0.49*
Chen et al. (2004), Exp. 2	4	106	Canton-Mandarin	Metalinguistic awareness: tone, onset and rime tasks	0.06
Clarkson and Galbraith (1992)	6	96	Pidgin–English	Problem solving with mathematical tests	0.40
Cromdal (1999)	K	54	Swedish–English	Metalinguistic awareness tasks[s]	0.71*
Davidson and Tell (2005), Exp. 1	PreK	40	Urdu–English	ASR: Mutual exclusivity in naming whole objects[t]	0.18
Davidson and Tell (2005), Exp. 1	K–1	40	Urdu–English	Mutual exclusivity in naming whole objects	1.22*
Davidson and Tell (2005), Exp. 2	K–1	20	Urdu–English	Mutual exclusivity in naming whole objects	0.34
Demie and Strand (2006)	10	1026	Mixed–English	Problem solving with a standardized test	0.23*
Demont (2001)	1	43	German-French	Metalinguistic awareness: several tasks[u]	0.56
Eviatar and Ibrahim (2000)	K–1	76	Russian-Hebrew	Metalinguistic awareness using several tasks[v]	0.96*
Galambos and Goldin-Meadow (1990)	PreK–1	64	Spanish–English	Metalinguistic awareness—correcting errors	0.39
Galambos and Goldin-Meadow (1990)	PreK–1	64	Spanish–English	Metalinguistic awareness—explaining errors	0.09
Gutierrez-Clellen et al. (2004)	3	44	Spanish–English	Working memory using two tasks[w]	–0.09
Humphreys and Mumtaz (2001)	2–3	120	Urdu–English	Working memory	1.52*
Jordà-Maria (2003)	PSec	160	Catalan-Castilian	Metalinguistic awareness: Metapragmatic	0.66*
Keshavarz and Astaneh (2004)	PSec	60	Mixed–Persian[x]	Metalinguistic awareness	0.56*
Loizou and Stuart (2003)	1	32	Greek–English	Metalinguistic awareness using several tasks[y]	0.96*
Loizou and Stuart (2003)	1	36	English-Greek	Metalinguistic awareness using several tasks	–0.34
Love et al. (2003)	PSec	71	Mixed–English	Metalinguistic awareness: syntactic processing	0.20
Merriman and Kutlesic (1993)	K–2	78	Serbian–English	Metalinguistic awareness	0.48*
Oller et al. (2007)	2	288	Spanish–English	Metalinguistic awareness[z]	1.07*

(continued)

TABLE 1 (continued)

Study	*Grade range*	*N*	*Languages (bilingual)*[a]	*Dominant cognitive benefits measured*	*Effect size (g)*
Oller et al. (2007)	4	296	Spanish–English	Metalinguistic awareness with several measures	0.49*
Ransdell et al. (2006)	PSec	137	Mixed–English	Metacognitive awareness[aa]	0.16
Reichard and Mokhtari (2004)	PSec	350	Mixed–English	Metacognitive awareness and strategies	0.39*
Sanz (2000)	8–12	201	Catalan-Spanish	Metalinguistic awareness with several measures	0.68*
Sheng et al. (2006)	K–3	24	Mandarin–English	Metalinguistic awareness with several measures[bb]	0.22

Note. PreK = prekindergarten; PSec = postsecondary; *N* = number of participants; ASR = abstract and symbolic representation; MRT = mean reaction time.

a. The language spoken by the monolingual group is presented last in the bilingual coupling (e.g., Arabic–English bilingual means that English was the language of the monolingual group).
b. Moving word problem was used to determine abstract and symbolic representation (ASR) for literacy.
c. Consistent and inconsistent word size problems were used to assess participants' understanding of the principles of representing spoken language in printed text.
d. Postswitch items in the dimensional change card sort (DCCS) task were used for this sample.
e. Knowledge items in the DCCS task were used for this sample.
f. Results from 24 monolingual and 24 balanced full bilingual participants were used. We discarded results from the 24 unbalanced bilinguals because they reported minimal use of the second language. For example, "a participant reporting moderate knowledge of Italian but the opportunity to use it only several times per month when visiting relatives was considered an unbalanced bilingual" (Bialystok et al., 2006, p. 1971). Experiment 1 was conducted with younger participants and Experiment 2 with older participants.
g. A dual-task classification paradigm was used in both experiments.
h. Although the overall result here shows a bilingual disadvantage for adults, the breakdown of outcome measures actually shows bilingual superiority in letters or numbers (LN) conditions, but overall that advantage was subdued by a much larger bilingual disadvantage in the animals or musical instruments (A-M) task.
i. Mean reaction times (MRT) for both congruent and incongruent trials on Simon task. Note that MRT was reversed to reflect the inverse relationship between RT and benefit. MRT was categorized as a cognitive benefit closely associated with attentional control.
j. Chinese–English, Hebrew–English, and Spanish–English bilinguals were compared to English monolinguals on phoneme counting measure.
k. French–English and Bengali–English bilinguals were compared to English monolinguals on block design and water-level tasks to examine whether balanced bilinguals carry over their linguistic advantage into nonlinguistic tasks.
l. This includes the phoneme segmentation task with first and second graders.
m. This was measured by a computerized DCCS task measuring inhibition of attention—categorized as a cognitive benefit closely associated with attentional control.
n. Appearance–reality tasks were used.
o. Outcome measures derived by using children's embedded figures task and ambiguous figures task, categorized as abstract and symbolic representation.
p. Outcome measures for inhibition of attention include opposite worlds task and computerized DCCS.
q. MLA (metalinguistic awareness) outcomes include letter detection task.
r. Tone, onset, and rime awareness tasks were each assessed with two pholonogical tasks: the oddity task and the same–different task.
s. Outcome measures include symbol substitution, sentence judgment, and correction tasks.
t. Familiar and unfamiliar objects with or without spare parts were used as outcome measures in all experiments. These measures were categorized as abstract and symbolic representation.
u. This article was written in French. Outcomes used to measure linguistic awareness include the grammatical judgment and correction tasks.
v. Outcome measures include the arbitrariness test, initial and final phoneme tests, and phoneme–syllable deletion test.
w. Outcome measures are the competing language processing task and dual processing comprehension task.
x. Turkish–Persian and Armenian–Persian bilinguals were compared to Persian monolinguals using controlled productive ability test at 2,000- and 3,000-word levels.
y. Measures include the rhyme oddity, syllable completion, cluster onset oddity, single phoneme onset oddity, phoneme elision, and initial phoneme identification tasks.
z. Several metalinguistic awareness measures used including passage comprehension, proofreading, verbal analogies, and oral vocabulary.
aa. Metacognitive awareness was measured by self-ratings of reading, writing, speaking, and listening skills.
bb. Paradigmatic and syntagmatic Trials 1, 2, and 3.

*$p < .05$.

TABLE 2
Overall weighted mean effect size

			Effect size		*95% confidence interval*		*Test of null*	*Test of heterogeneity*			
	N	*k*	*g*	*SE*	*Lower*	*Upper*	*z*	*Q*	*df*	*p*	I^2 (%)
All	6,022	63	0.41	0.03	0.36	0.46	15.00*	362.62	62	.00	82.90
Not conducted by Bialystok	4,245	33	0.38	0.03	0.31	0.44	11.54*	141.26	32	.00	77.35
Conducted by Bialystok	1,777	30	0.48	0.05	0.39	0.58	9.74*	218.17	29	.00	86.71

*$p < .05$.

associated degrees of freedom (*df*), and the percentage of variability that is attributable to true heterogeneity, that is, over and above the sampling error (I^2).

Table 2 shows that the overall weighted mean effect size is moderate ($g = 0.41$) but with substantial variability among studies ($Q_{total} = 362.62$, $p < .001$). Heterogeneity among the full set of studies was anticipated as different studies measured different cognitive outcomes and there was no reason to expect similar effect sizes for different outcomes. Following up on this heterogeneity, separate analyses were conducted for each category of cognitive outcomes.

It was observed that 30 of the 63 independent effect sizes were obtained from studies conducted by Bialystok and colleagues. A sensitivity analysis was performed so as to investigate the potential bias of including about 47% of the entire studies conducted by a single principal investigator. Table 2 shows statistically detectable mean effect sizes irrespective of whether studies were authored by Bialystok or not. There was an overlap in the confidence intervals across the two categories. Hence, a decision was made to combine the data from both categories in subsequent analyses.

Cognitive Correlates of Bilingualism

Table 3 shows the weighted mean effect sizes for different cognitive outcomes associated with bilingualism. All the outcome measures produced statistically detectable mean effect sizes in favor of bilingualism. Attentional control produced the largest effect with a weighted mean effect size of .96 across 14 studies. Although all the cognitive outcomes in Table 3 had statistically detectable mean effect sizes, most of the effect size distributions were highly heterogeneous, indicating that the variability among effect sizes was greater than that expected from sampling error.

Moderator analyses were conducted to investigate this heterogeneity, but given the small number of studies within some of the categories of cognitive outcomes, a decision was made to collapse across categories with similar outcomes. We collapsed studies that investigated metalinguistic and metacognitive awareness into one group, and abstract and symbolic representation, attentional control, and problem solving were collapsed into another group. We did not include working memory studies with either of the recategorized groups because the dependent variables used to measure working memory were markedly different from those used in studies subsumed under the other two groups. Given the small number of studies

TABLE 3
Weighted mean effect sizes for bilingual studies by outcome constructs

			Effect size		*95% confidence interval*		*Test of null*	*Test of heterogeneity*			
	N	*k*	*g*	*SE*	*Lower*	*Upper*	*z*	*Q*	*df*	*p*	I^2 (%)
Outcome constructs (cognitive outcomes)											
Metalinguistic awareness	2,813	29	0.33	0.04	0.26	0.41	8.44*	151.56	28	.00	81.53
Metacognitive awareness	487	2	0.32	0.09	0.14	0.50	3.47*	1.26	1	.26	20.80
Working memory	451	7	0.48	0.10	0.29	0.67	4.90*	41.36	6	.00	85.49
Abstract and symbolic representation	484	8	0.57	0.10	0.39	0.76	6.04*	42.10	7	.00	83.37
Attentional control	594	14	0.96	0.09	0.79	1.13	10.86*	73.30	13	.00	82.26
Problem solving	1,193	3	0.26	0.07	0.13	0.38	3.96*	0.65	2	.72	0.00
Collapsed outcome constructs											
Metalinguistic and metacognitive awareness	3,300	31	0.33	0.04	0.26	0.41	9.14*	152.61	30	.00	80.34
Abstract and symbolic representation, attentional control, and problem solving	2,271	25	0.52	0.05	0.43	0.61	11.33*	157.69	24	.000	84.78

*$p < .05$.

in the working memory category, moderator analyses were not conducted on this category. However, we note that bilingualism was associated with greater working memory, resulting in a moderate effect size of .48. Henceforth, the results section deals with the two new categories (metalinguistic or metacognitive awareness and attention and representation).

Table 3 shows weighted mean effect sizes for the recategorized outcome constructs. The new categories produced statistically detectable mean effect sizes, and significant heterogeneity was observed in both categories. In subsequent analyses, we examined the different moderators that may help explain the variability within each of these two categories.

Moderator Variable Analyses

A mixed-effects model was used for all moderator variable analyses. A mixed-effects model uses a random-effects model to combine studies within subgroups and a fixed-effect model to combine studies across subgroups to yield an overall effect (Borenstein, Hedges, Higgins, & Rothstein, 2009). By using the random-effects model to combine studies within subgroups in moderator analyses, a mixed-effects model typically allows for population parameters to vary across studies, reducing the probability of committing a Type I error, and is usually regarded as a more rigorous meta-analytical model than a fixed-effects model only (Borenstein et al., 2009; Denson, 2009; Hedges & Vevea, 1998; National Research Council, 1992).

Relationship Between Bilingualism and Metalinguistic or Metacognitive Awareness Across Different Locations, Educational Levels, Settings, Language Groups, SES, and Region

Table 4 shows weighted mean effect sizes for metalinguistic or metacognitive awareness outcomes under various conditions. When disaggregated by the geographical location of the research, the effect of bilingualism was statistically detectable across studies conducted in the United States, Europe, and the Middle East.

Because the total between-levels variance was statistically detectable, $Q_B(5) = 14.47$, $p = .01$, further analyses showed that studies conducted in Europe were significantly different from those conducted in other geographical locations. Although studies conducted in the United States and the Middle East produced a statistically detectable effect, they were not significantly different from those conducted in China or Canada and/or those that are part of the various or mixed category.[2]

In Table 4, a majority of the studies were conducted with primary school students in kindergarten through third grade. A mean effect size was statistically detectable only for these early primary-level students, possibly because of the small number of studies with students at other levels.

Classroom studies in which learning activities contributed toward performance assessment in a program produced statistically detectable effect sizes along with studies that did not specify the setting. The classroom studies were significantly different from studies conducted in the laboratory. However, the certainty of this interpretation is limited by the high number of studies that did not specify the setting under which those studies were undertaken.

Bilingualism was reliably associated with greater metalinguistic or metacognitive awareness when bilinguals spoke Spanish and English or a mix of language

TABLE 4

Weighted mean effect sizes for studies investigating bilinguals' metalinguistic and metacognitive awareness under various conditions

			Effect size		95% confidence interval		Test of null	Test of heterogeneity		
	N	*k*	*g*	*SE*	*Lower*	*Upper*	*z*	Q_B	*df*	*p*
Geographical locations								14.47	5	.01
Canada	672	8	–0.08	0.15	–0.37	0.20	–0.57			
United States	757	5	0.55	0.18	0.19	0.90	2.99*			
Europe	556	7	0.57	0.17	0.24	0.89	3.41*			
Middle East	136	2	0.77	0.30	0.18	1.35	2.56*			
China	564	5	0.15	0.18	–0.19	0.49	0.87			
Mixed	615	4	0.26	0.20	–0.12	0.65	1.34			
Educational level								1.99	4	.74
Preschool	30	1	0.64							
Primary (K–3)	1,739	21	0.26	0.11	0.04	0.48	2.33*			
Intermediate (4–7)	552	3	0.09	0.27	–0.45	0.63	0.33			
Secondary (8–12)	261	2	0.63	0.35	–0.05	1.31	1.80			
Postsecondary	718	4	0.36	0.24	–0.11	0.83	1.50			
Setting								10.54	2	.01
Classroom	1,232	6	0.65	0.16	0.34	0.96	4.11*			
Laboratory	1,359	16	0.07	0.10	–0.13	0.27	0.72			
Not specified	709	9	0.42	0.14	0.15	0.70	2.99*			
Language of the monolingual group								2.99	3	.56
English	2,042	18	0.22	0.11	0.01	0.44	1.97*			
Romance (Castilian, French, and Spanish)	468	4	0.58	0.23	0.13	1.04	2.51*			

(continued)

TABLE 4 (continued)

			Effect size		*95% confidence interval*		*Test of null*	*Test of heterogeneity*		
	N	*k*	*g*	*SE*	*Lower*	*Upper*	*z*	Q_B	*df*	*p*
Chinese	564	5	0.15	0.20	–0.24	0.55	0.76			
Other[a]	226	4	0.50	0.25	0.02	0.99	2.03*			
Languages of the bilingual group								16.52	3	.00
French–English	284	3	–0.32	0.22	–0.74	0.11	–1.44			
Spanish–English	770	5	0.54	0.17	0.21	0.87	3.24*			
Chinese–English	918	10	0.08	0.12	–0.16	0.32	0.69			
Mixed[b]	1,328	13	0.51	0.11	0.29	0.72	4.57*			
Socioeconomic status								5.99	3	.11
Medium	485	4	–0.04	0.22	–0.47	0.39	–0.19			
High	313	4	0.28	0.23	–0.17	0.74	1.23			
Mixed	681	4	0.72	0.22	0.28	1.16	3.21*			
Not reported	1,821	19	0.27	0.11	0.07	0.48	2.60*			
Region								3.05	3	.39
Urban or inner city	552	8	0.28	0.18	–0.07	0.63	1.59			
Suburban	225	3	0.00	0.27	–0.54	0.54	0.01			
Rural	76	1	0.96							
Not reported	2,447	19	0.30	0.11	0.09	0.51	2.83*			

Note. $Q_B = Q_{\text{Between}}$.

a. The "other" subset refers to studies in which participants spoke languages other than English, Spanish, French, and Mandarin or Chinese. Specifically, monolingual control participants in the "other" subset spoke Swedish (Cromdal, 1999), Hebrew (Eviatar & Ibrahim, 2000), Persian (Keshavarz & Astaneh, 2004), or Greek (Loizou & Stuart, 2003).

b. The "mixed" subset refers to studies in which participants spoke various language couplings apart from French–English, Spanish–English, and Chinese–English.

*$p < .05$.

not including French–English or Chinese–English. Similarly, bilinguals showed a significant metalinguistic or metacognitive advantage over monolinguals for most monolingual groups except Chinese-speaking monolinguals.

Many of the studies did not report the SES of participants involved in the studies in Table 4. Statistically detectable mean effect sizes were found among participants with mixed SES and those studies that did not report the SES of participants. However, mean effect sizes did not differ statistically among the four SES categories (i.e., middle, high, mixed, and not reported).

Relationship Between Bilingualism and Attention and Representation Across Different Locations, Educational Levels, Settings, Language Groups, SES, and Region

Table 5 shows weighted mean effect sizes for studies investigating bilinguals' abstract and symbolic representation, attentional control, and problem-solving outcomes under various conditions. Studies conducted in Canada and various geographical locations (mixed) produced statistically detectable mean effect sizes. The mean effect size for studies conducted in different locations was much higher than those conducted specifically in the United States, Canada, and Europe.

Studies conducted with preschool, primary, and postsecondary students produced statistically detectable mean effect sizes, with bilingual postsecondary students showing the largest mean effect size ($g = 1.76$). Post hoc comparisons were conducted because the total between-levels variance for educational level was significant, $Q_B(4) = 12.92$, $p = .01$. Results showed that studies conducted with participants from preschool up to 12th grade were not significantly different from one another but that they were all significantly different and produced lower mean effect sizes than studies conducted with postsecondary students.

There were no classroom studies in which participants engaged in learning activities that contributed toward performance assessment. Most of the studies did not report the setting in which they were conducted. Nevertheless, studies conducted in the laboratory and those that did not report the setting both produced statistically detectable effect sizes.

Except Urdu–English bilinguals (who were included in only three studies), all categories of language pairs spoken by bilinguals showed statistically detectable mean effect sizes, indicating that bilingualism is reliably associated with better attention and representation skills. Because the total between-levels variance was statistically detectable, $Q_B(4) = 11.95$, $p = .02$, further analyses showed that studies conducted with bilinguals who spoke Tamil–English were significantly different from those conducted with Chinese–English bilinguals and those in the "mixed" category.

In Table 5, statistically detectable mean effect sizes were found among participants with medium SES and those studies that did not report the SES of participants. Mean effect sizes did not differ statistically among the four SES categories (i.e., low, middle, mixed, and not reported). These findings indicate that bilingualism contributes to cognitive benefits irrespective of the SES of participants. Across all regional categories, bilinguals showed more positive and statistically detectable mean effect sizes than monolinguals.

TABLE 5

Weighted mean effect sizes for studies investigating bilinguals' abstract and symbolic representation, attentional control, and problem solving under various conditions

			Effect size		95% confidence interval		Test of null	Test of heterogeneity		
	N	*k*	*g*	*SE*	*Lower*	*Upper*	*z*	Q_B	*df*	*p*
Geographical locations								12.06	4	.02
Canada	915	16	0.72	0.16	0.40	1.03	4.52*			
United States	100	3	0.58	0.38	–0.17	1.33	1.52			
Europe	1,026	1	0.23							
Papua New Guinea	96	1	0.40							
Mixed	134	4	1.92	0.36	1.23	2.62	5.42*			
Educational level								12.92	4	.01
Preschool	734	12	0.63	0.18	0.28	0.99	3.47*			
Primary (K–3)	232	5	0.64	0.29	0.07	1.21	2.20*			
Intermediate (4–7)	96	1	0.40							
Secondary (8–12)	1,026	1	0.23							
Postsecondary	183	6	1.76	0.30	1.18	2.34	5.93*			
Setting								0.30	1	.59
Laboratory	421	8	0.72	0.24	0.25	1.19	2.98*			
Not specified	1,850	17	0.88	0.17	0.55	1.20	5.28*			
Language of the monolingual group								0.44	1	.51
English	2,175	24	0.85	0.14	0.57	1.12	6.01*			
Pidgin English	96	1	0.40							
Languages of the bilingual group								11.95	4	.02

(continued)

TABLE 5 (continued)

			Effect size		95% confidence interval		Test of null	Test of heterogeneity		
	N	*k*	*g*	*SE*	*Lower*	*Upper*	*z*	Q_B	*df*	*p*
French–English	342	7	0.80	0.24	0.33	1.28	3.30*			
Chinese–English	407	7	0.55	0.23	0.09	1.01	2.36*			
Tamil–English	134	4	1.92	0.35	1.23	2.62	5.47*			
Urdu–English	100	3	0.58	0.38	–0.16	1.33	1.53			
Mixed[a]	1,288	4	0.63	0.29	0.05	1.20	2.14*			
Socioeconomic status								4.40	3	.22
Low	1,026	1	0.23							
Medium	716	13	1.12	0.20	0.73	1.52	5.52*			
Mixed	100	3	0.58	0.43	–0.26	1.42	1.36			
Not reported	429	8	0.56	0.25	0.07	1.06	2.23*			
Region								4.99	2	.08
Urban or inner city	356	8	0.57	0.27	0.05	1.09	2.16*			
Suburban	472	7	0.57	0.27	0.04	1.10	2.09*			
Not reported	1,443	10	1.25	0.24	0.78	1.73	5.23*			

a. The "mixed" subset refers to studies in which bilinguals spoke languages other than French–English, Chinese–English, Tamil–English, and Urdu–English. Specifically, bilingual participants spoke Bengali/French–English (Bialystok & Majumder, 1998); Armenia, Hebrew, Russian, Spanish, and a host of other languages (Bialystok & Senman, 2004, p. 573); Pidgin–English (Clarkson & Galbraith, 1992); or various languages (Demie & Strand, 2006).
*$p < .05$.

Effects of Different Methodological Features

Tables 6 and 7 show how effect sizes varied with research design and implementation features. The studies were categorized according to the researchers' confidence in the calculated effect size (rated as medium or high depending on whether sufficient data were provided to calculate an effect size), the reliability and validity of instruments used in the studies, and a measure of the level of control for bias in the studies. An additional source of variance often reported in meta-analyses is the research designs of the studies under consideration (Abrami & Bernard, 2006). However, because none of the studies in the current meta-analysis used random assignment—almost all the studies included were matched quasi-experimental designs except two studies that did not clearly report the design—we could not meta-analyze the variations based on research designs. As well, all the analyzed studies were published, hence precluding further analyses based on whether studies were published or not but heightening the potential for publication bias.

Table 6 shows the weighted mean effect sizes for metalinguistic or metacognitive awareness studies by different methodological features. High coder confidence in the effect size calculation was associated with statistically detectable mean effect size, but medium coder confidence was not. Studies in which the reliability of measures was not reported produced a statistically detectable mean effect size, but studies in which reliability was reported did not. Studies produced a statistically detectable mean effect size whether they reported validity measures or not. Studies in which bias was tightly controlled or not controlled produced a statistically detectable mean effect size, but studies in which biases were loosely controlled did not.

Table 7 shows the weighted mean effect sizes for representation and attention studies by different methodological features. High coder confidence in the effect size calculation was associated with a statistically detectable mean effect size, but medium coder confidence was not. Mean effect sizes were statistically detectable whether reliability of the measures used were reported or not. Similarly, mean effect sizes were statistically significant whether bias was tightly or loosely controlled.

Publication Bias

Researchers have observed that published studies are a biased sample of studies in a particular domain because research reports are more likely to be published when significant results are reported (Orwin, 1983; Rosenthal, 1979). Hence, studies with nonsignificant findings are often either tucked away in file drawers or reported in the less accessible gray literature. This problem, referred to as the "file-drawer problem," becomes apparent in meta-analyses, where there is a tendency to exclude unpublished and gray literature studies, thereby potentially skewing meta-analytical findings toward a positive mean effect size. This poses a threat to the validity of results obtained from any meta-analyses. Researchers have proposed different methods to examine the validity of results obtained from meta-analyses. In the current work, publication bias seems to be a potential threat to the validity of results obtained because all the studies analyzed were published in peer-reviewed journals. Two statistical tests were performed using Comprehensive Meta-Analysis software to assess the potential for publication bias in this meta-analysis. First, a "classic fail-safe *N*" test was performed to determine the number

TABLE 6

Weighted mean effect sizes for studies investigating bilinguals' metalinguistic and metacognitive awareness by different methodological features

			Effect size		*95% confidence interval*		*Test of null*	*Test of heterogeneity*		
	N	*k*	*g*	*SE*	*Lower*	*Upper*	*z*	Q_B	*df*	*p*
Confidence in effect size derivation								0.84	1	.36
Medium	238	2	0.58	0.33	–0.06	1.21	1.77			
High	3,062	29	0.27	0.09	0.09	0.44	3.02*			
Reliability								0.44	1	.51
Not reported	2,028	19	0.33	0.11	0.12	0.55	3.06*			
Reported	1,272	12	0.22	0.14	–0.05	0.49	1.57			
Validity								0.54	1	.46
Not reported	2,644	26	0.26	0.09	0.08	0.45	2.76*			
Reported	656	5	0.44	0.22	0.01	0.86	2.00*			
Control for bias in studies								3.11	2	.21
Tightly controlled	1,556	17	0.34	0.12	0.12	0.57	2.99*			
Loosely controlled	1,023	11	0.12	0.14	–0.15	0.39	0.86			
Not controlled	721	3	0.58	0.25	0.10	1.06	2.35*			

*$p < .05$.

TABLE 7
Weighted mean effect sizes for studies investigating bilinguals' abstract and symbolic representation, attentional control, and problem solving by different methodological features

			Effect size		*95% confidence interval*		*Test of null*	*Test of heterogeneity*		
	N	*k*	*g*	*SE*	*Lower*	*Upper*	*z*	Q_B	*df*	*p*
Confidence in effect size derivation								0.65	1	.42
Medium	1,180	4	0.59	0.34	–0.07	1.24	1.75			
High	1,091	21	0.88	0.15	0.58	1.18	5.69*			
Reliability								1.43	1	.23
Not reported	1,779	18	0.94	0.17	0.61	1.27	5.58*			
Reported	492	7	0.57	0.26	0.07	1.08	2.21*			
Control for bias in studies								3.11	2	.21
Tightly controlled	1,862	18	0.89	0.16	0.57	1.21	5.43*			
Loosely controlled	409	7	0.67	0.26	0.17	1.17	2.61*			

*$p < .05$.

of null effect studies needed to raise the *p* value associated with the average effect above an arbitrary alpha level (set at $\alpha = .05$). This test revealed that 3,791 additional studies would be required to invalidate the overall effect found in this meta-analysis.

The second test, Orwin's fail-safe *N*, was used to estimate the number of file-drawer studies with null results required to nullify the effects found in this meta-analysis. Using a criterion trivial level of .05, the fail-safe *N* was found to be 453 studies, meaning that 453 missing null studies would be needed to bring the current mean effect size found in this meta-analysis to .05. Researchers have claimed that meta-analytical results could be interpreted as valid and thus resistant to the "file-drawer problem" if the fail-safe *N* reaches the $5k+10$ limit (Carson, Schriesheim, & Kinicki, 1990; Rosenthal, 1979). The results of the two computed statistical tests suggest that it is unlikely that publication bias poses a significant threat to the validity of findings reported in the current work because both fail-safe *N* values are larger than the $5k+10$ limit.

Discussion

The meta-analysis presented here synthesized data from 63 studies to examine the cognitive correlates of bilingualism and the magnitude of such benefits. We found a moderate positive overall effect of bilingualism on different cognitive measures. Nevertheless, significant variability existed among studies, with some yielding a positive cognitive effect of bilingualism and others a negative cognitive effect. To more appropriately explain the variability among findings, we examined the study features that may account for the variable effects. Specifically, this meta-analysis provided answers to the following research questions.

What Cognitive Correlates Is Bilingualism Associated With?

Results of this meta-analysis show that bilingualism is positively associated with a range of cognitive benefits. Specifically, bilinguals were found to outperform monolinguals on the combined measures of metalinguistic and metacognitive awareness ($g = 0.33$) and on measures of abstract and symbolic representation, attentional control, and problem solving ($g = 0.52$). There was, however, significant variability in these effect sizes.

These results indicate that the process of acquiring two languages and of simultaneously managing those languages—of inhibiting one so the second can be used without interference—allows bilinguals to develop skills that extend into other domains. These skills appear to give bilingual speakers insight into the abstract features of language and into their own learning processes. They also appear to give bilingual speakers an enhanced capacity to appropriately control and distribute their attentional resources, to develop abstract and symbolic representations, and to solve problems.

How Do the Effects of Bilingualism Vary When Cognitive Outcomes Are Measured in Different Geographical Locations and Settings, for Different Language Groups, at Different SESs, and at Different Educational Levels?

Sociocultural attitudes toward bilingualism and the use of particular languages vary across different countries and communities (Hamers & Blanc, 2000). Consistent effects associated with geography were not observed in the current work. For example, Canada and the United States provide rather different contexts for bilingualism: Canada maintains two official languages, and a number of policies supporting bilingualism in both official languages are in place (Government of Canada, 2009). As well, Canada's multicultural attitude toward immigration encourages immigrants to maintain their native language while acquiring at least one of Canada's official languages (Esses & Gardner, 1996). In contrast, the United States is officially unilingual and has adopted a "melting pot" rather than multicultural approach to immigration (Ravitch, 1990). Despite these differences, we observed no consistent differences regarding cognitive correlates of bilingualism in Canada and the United States. It may be that limiting our inclusion criteria to cover only studies in which participants were balanced bilinguals eliminated any potential differences across bilingual speakers in different countries.

Bilinguals who acquire their second language at an early age often master that second language more fully than those who acquire their second language later in life (Johnson & Newport, 1989). The evidence reviewed in the current analysis suggests that earlier, rather than later, acquisition of a second language is also more likely to be associated with greater metalinguistic and metacognitive awareness. Although bilingual speakers of all ages demonstrated significant advantages with respect to representation and attention, only the youngest bilinguals (who, by definition, must have acquired their second language early in life) showed significant advantages with respect to metalinguistic or metacognitive awareness.

Studies conducted in the classroom yielded a moderately high effect of bilingualism over monolingualism on measures of metalinguistic and metacognitive awareness, but this effect was not observed in laboratory studies conducted outside

of the classroom. Because of the high stakes associated with classroom studies in which testing contributes to students' performance assessments, it is possible that bilinguals in such studies exhibited their cognitive skills to the fullest relative to bilinguals in the laboratory studies in which outcome measures typically do not contribute to performance assessments.

How Are Effect Sizes Influenced by Different Combinations of Languages Spoken by Bilinguals?

For most language pairings—including French–English and Chinese–English pairings—bilinguals outperformed monolinguals on the set of representation and attention measures that includes abstract and symbolic representations, attentional control, and problem solving. However, when measuring metalinguistic or metacognitive outcomes, French–English and Chinese–English pairings did not show a significant advantage over monolinguals. This may be a result of the specific manners in which metalinguistic or metacognitive outcomes were measured in the French and Chinese speaker studies. Alternatively, this pattern of results may indicate that any language pairing can yield general cognitive benefits for bilinguals but specific pairs of languages are necessary for bilinguals to develop metalinguistic or metacognitive advantages. For example, it may be necessary for both languages to contain certain features for bilinguals to become metalinguistically aware of those features. This would be in line with the findings from research on cross-linguistic transfer of reading skills, where it has been observed that the amount of cross-linguistic transfer is maximized when both languages have alphabetic writing systems. As a result, cross-linguistic transfer occurs easily for Spanish–English bilinguals but not for Mandarin–English bilinguals (Bialystok, Luk, & Kwan, 2005; Koda & Zehler, 2008).

Are the Effect Sizes of Cognitive Benefits Influenced by Methodological Features of the Research?

Across different methodological features, the meta-analysis found cognitive advantages of bilingualism. Specifically, on measures of metalinguistic or metacognitive awareness, statistically detectable benefits of bilingualism were obtained when researchers' confidence in calculating effect sizes was high and when biases in studies were tightly controlled. However, the large number of studies that did not report any reliability or validity measures somewhat undermines the methodological quality of the studies we analyzed. On measures of representation and attention, a similar result was found with researchers' confidence in calculating effect sizes. A large, statistically detectable effect size was obtained showing that a high coder confidence in the effect size was associated with a high mean effect size. Also, statistically detectable effect sizes were obtained whether reliability of instruments used was reported or not and irrespective of how biases in studies were controlled.

Conclusion

Although monolingualism is often depicted as normative, the best available evidence indicates that, around the world, bilingual and multilingual speakers outnumber monolingual speakers (Tucker, 1998). The current work suggests that bilingualism (and, presumably, multilingualism) is associated with a number of

cognitive benefits. These findings point to the need for further work investigating the utility of these benefits in a variety of contexts. For example, cognitive benefits documented in the current work may be of use to bilingual speakers in classrooms where the language of instruction is not their native language. As the pace of immigration to developed countries increases, the incidence of bilingualism and multilingualism in these countries will also increase—as will the number of second language learners in public school classrooms. Although second language learners often present challenges within the classroom, the current analysis suggests they may also bring a number of advantages. It remains unclear how, in practice, second language learners and their instructors may capitalize on these advantages. Further work investigating the cognitive correlates of bilingualism within educational contexts is required to clarify this issue.

APPENDIX

Codebook on the cognitive benefits of bilingualism: Coded variables

Section A: Identification of studies/ reviewer	
A.1 Name of the reviewer	A.1.1 Details
A.2 Date of the review	A.2.1 Details
A.3 Please enter the details of the paper	A.3.1 Author (last name, first name)
	A.3.2 Date of publication
	A.3.3 Title of the article
	A.3.4 Journal, issue number, pages
	A.3.5 Credentials—institutional affiliation of the authors
Section B: General information	
B.1 What is the focus of the study?	B.1.1 ESL instructional practices
	B.1.2 Cognitive benefits of bilingualism
B.2 In what country was the study conducted?	B.2.1 USA
	B.2.2 Canada
	B.2.3 United Kingdom
	B.2.4 Australia/New Zealand
	B.2.5 Other (please specify)
B.3 What is the language of the monolingual group?	B.3.1 English
	B.3.2 Other (please specify)
B.4 What are the languages of the bilingual group?	B.4.1 *(Please list)*
	B.4.2 *(Please list)*
B.5 What is/are the cognitive variable(s) being measured?	B.5.1 Problem-solving skills
	B.5.2 Theory-of-mind tasks
	B.5.3 Creative and divergent thinking
	B.5.4 Cognitive flexibility
	B.5.5 Abstract representation and reasoning skills

(continued)

APPENDIX (continued)

	B.5.6 Learning strategies
	B.5.7 Metalinguistic awareness
	B.5.8 Metacognitive skills
	B.5.9 Working memory
	B.5.10 Others *(Please provide details)*
B.6 What is the age range of the participants?	B.6.1 Grades K–3 *(5–8 years old)* B.6.2 Grades 4–7 *(9–12 years old)* B.6.3 Grades 8–12 *(13–18 years old)*
	B.6.4 Longitudinal range *(Please provide details)*
B.7 What is the location of the study?	B.7.1 Classroom B.7.2 Pull-out room *(for instance, a resource room or another room within the school)* B.7.3 Laboratory
	B.7.4 Other (please specify)
Section C: Study research questions	
C.1 What is the overarching question the researcher is trying to address? *Please write in authors' description if there is one. Elaborate if necessary, but indicate which aspects are reviewers' interpretations. Other, more specific questions about the research questions and hypotheses are asked later.*	C.1.1 Explicitly stated (please specify) C.1.2 Implicit (please specify) C.1.3 Not stated/unclear (please specify)
C.2 What is the author's specific research question? *Research questions operationalize the overarching question. Please write in author's description if there is one. Elaborate if necessary, but indicate which aspects are reviewer's interpretations.*	C.2.1 Explicitly stated (please specify) C.2.2 Implicit (please specify) C.2.3 Not stated/unclear (please specify)
C.3 What is the author's hypothesis? *Research questions or hypotheses operationalize the aims of the study. Please write in authors' description if there is one. Elaborate if necessary, but indicate which aspects are reviewers' interpretations.*	C.3.1 Explicitly stated (please specify) C.3.2 Implicit (please specify) C.3.3 Not stated/unclear (please specify)
C.4 What is the theoretical/empirical basis for this study? *Please write in author's description if there is one. Elaborate if necessary, but indicate which aspects are reviewers' interpretations.*	C.4.1 Explicitly stated (please specify) C.4.2 Implicit (please specify) C.4.3 Not stated/unclear (please specify)

(continued)

APPENDIX (continued)

Section D: Methods–designs	
D.3 What variables are included?	D.3.1 Independent variables *List independent moderator variables* D.3.2 Dependent (outcome) variables *List outcome variables*
D.4 What measurement tool(s) is/are used?	D.4.1 Standardized test *Please provide the name of the test if listed* D.4.2 Classroom or teacher developed test *Please describe and give page number* D.4.3 Observation *Please give a description and the page number* D.4.4 Other *Please describe and give page number*
Section E: Methods–groups	
E.1 What is the design of the study	E.1.1 Nonrandomized with treatment and control groups. *How were the groups assigned/created?* E.1.2 Repeated measures design *(Where the same sample of individuals is measured in all of the conditions)* E.1.3 Other (please specify)
E.2 How do the groups differ?	E.2.1 Explicitly stated (please specify) E.2.2 Implicitly stated (please specify) E.2.3 Not applicable (not more than one group) E.2.4 Not stated/unclear (please specify)
E.3 Number of groups. *For instance, in studies in which comparisons are made between groups, this may be the number of groups into which the dataset is divided for analysis (e.g., social class, or form size).*	E.3.1 Not applicable E.3.2 One E.3.3 Two E.3.4 Three E.3.5 Four or more (please specify) E.3.6 Other/unclear (please specify)
Section F: Method–sampling strategy	
F.1 Are the authors trying to produce findings that are representative of a given population? *Please write in authors' description. If authors do not specify, please indicate reviewers' interpretation.*	F.1.1 Explicitly stated (please specify) F.1.2 Implicit (please specify) F.1.3 Not stated/unclear (please specify)
F.2 What is the sampling frame (if any) from which the participants are chosen? *e.g., telephone directory, electoral register, postcode, school listing, etc. There may be two stages—e.g., first sampling schools and then classes or pupils within them.*	F.2.1 Not applicable (please specify) F.2.2 Explicitly stated (please specify) F.2.3 Implicit (please specify) F.2.4 Not stated/unclear (please specify)

(continued)

APPENDIX (continued)

F.3 Which method does the study use to select people, or groups of people (from the sampling frame)? *e.g., selecting people at random, systematically—selecting for example every 5th person, etc.*	F.3.1 Not applicable (no sampling frame) F.3.2 Explicitly stated (please specify) F.3.3 Implicit (please specify) F.3.4 Not stated/unclear (please specify)
Section G: Sample information	
G.1 What was the total number of participants in the study (the actual sample)? *If more than one group is being compared, please give numbers for each group.*	G.1.1 Not applicable (e.g., study of policies, documents, etc.) G.1.2 Explicitly stated (please specify) G.1.3 Implicit (please specify) G.1.4 Not stated/unclear (please specify)
G.2 What is the sex of the individuals in the actual sample? *Please give the numbers of the sample that fall within each of the given categories. If necessary refer to a page number in the report (e.g., for a useful table). If more than one group is being compared, please describe for each group.*	G.2.1 Not applicable (e.g., study of policies, documents, etc.) G.2.2 Single sex (please specify) G.2.3 Mixed sex (please specify) G.2.4 Not stated/unclear (please specify) G.2.5 Coding is based on: Authors' description G.2.6 Coding is based on: Reviewers' inference
G.3 What is the socioeconomic status of the individuals within the actual sample? *If more than one group is being compared, please describe for each group.*	G.3.1 Not applicable (e.g., study of policies, documents, etc.) G.3.2 Explicitly stated (please specify) G.3.3 Implicit (please specify) G.3.4 Not stated/unclear (please specify)
G.4 What is the ethnicity of the individuals within the actual sample? *If more than one group is being compared, please describe for each group.*	G.4.1 Not applicable (e.g., study of policies, documents, etc.) G.4.2 Explicitly stated (please specify) G.4.3 Implicit (please specify) G.4.4 Not stated/unclear (please specify)
G.5 What is known about the special educational needs of individuals within the actual sample? (choose all that apply) *Please note whether it was explicitly stated or implicit.*	G.5.1 Normally developing children G.5.2 Language impaired G.5.3 Learning disabled G.5.4 Reading disabled G.5.5 Late talkers G.5.6 Intellectual difficulties G.5.7 Other (please specify) G.5.8 Not applicable (e.g., study of policies, documents, etc.) G.5.9 Not stated/unclear (please specify)
G.6 What are the regional characteristics of individuals/groups in sample?	G.6.1 Not applicable (please specify) G.6.2 Urban/inner city G.6.3 Suburban G.6.4 Rural G.6.5 Not stated/unclear (please specify)

(continued)

APPENDIX (continued)

	G.6.6 Other
G.7 Do the authors describe strategies used to control for bias from confounding variables and groups? *Please include information on: Were the groups similar at the start of the study?*	G.7.1 Age (please specify) G.7.2 Gender (please specify) G.7.3 Social class (please specify) G.7.4 Not stated/Unclear G.7.5 Other (please specify)
G.8 What are additional sample information/characteristics if any?	G.8.1 Not applicable G.8.2 Sample was obtained from another study (specify study) G.8.3 Other G.8.4 Not stated/unclear (please specify)
G.9 How many participants left before the end of the study? *If more than one group, please give numbers for each group.*	
G.10 If the study involves following samples prospectively over time, do authors provide baseline values of key variables such as those being used as outcomes and relevant socio-demographic variables?	G.10.1 Not applicable (e.g., study of policies, documents, etc.) G.10.2 Not applicable (not following samples prospectively over time) G.10.3 Yes (please specify) G.10.4 No
Section H: Recruitment and consent	
H.1 Which methods are used to recruit people into the study? *e.g., letters of invitation, telephone contact, face-to-face contact.*	H.1.1 Letter of invitation H.1.2 Telephone contact H.1.3 Other (please specify) H.1.4 Not applicable (please specify) H.1.5 Not stated/unclear (please specify)
H.2 Were any incentives provided to recruit people into the study?	H.2.1 Not applicable (please specify) H.2.2 Explicitly stated (please specify) H.2.3 Not stated/unclear (please specify)
H.3 Whose consent was sought? *Please comment on the quality of consent if relevant.*	H.3.1 Participants H.3.2 Parental consent H.3.3 Other (please specify) H.3.4 No consent was sought H.3.5 Not stated/unclear
Section I: Data collection	
I.1 Which methods were used to collect the data? *Please indicate all that apply and give further detail where possible.*	I.1.1 Experimental I.1.2 Curriculum-based assessment I.1.3 Focus group I.1.4 Group interview I.1.5 One to one interview (face to face or by phone) I.1.6 Observation I.1.7 Self-completion questionnaire

(continued)

	I.1.8 Self-completion report or diary
	I.1.9 Exams
	I.1.10 Clinical test
	I.1.11 Practical test
	I.1.12 Psychological test
	I.1.13 Hypothetical scenario including vignettes
	I.1.14 School/college records (e.g., attendance records, etc.)
	I.1.15 Secondary data such as publicly available statistics
	I.1.16 Other documentation
	I.1.17 Not stated/unclear (please specify)
	I.1.18 Coding is based on: authors' description
	I.1.19 Coding is based on: reviewers' interpretation
I.2 Who collected the data? *Please indicate all that apply and give further detail where possible.*	I.2.1 Researcher
	I.2.2 Head teacher/senior management
	I.2.3 Teaching or other staff
	I.2.4 Parents
	I.2.5 Pupils/students
	I.2.6 Governors
	I.2.7 LEA/government officials
	I.2.8 Other educational practitioner
	I.2.9 Other (please specify)
	I.2.10 Not stated/unclear
	I.2.11 Coding is based on: Authors' description
	I.2.12 Coding is based on: Reviewers' inference
I.3 Do the authors describe any ways they addressed the reliability of their data collection tools/methods? *e.g., test–retest methods (Where more than one tool was employed, please provide details for each.)*	I.3.1 Details
I.4 Do the authors describe any ways they have addressed the validity of their data collection tools/methods? *e.g., mention previous validation of tools, published version of tools, involvement of target population in development of tools. (Where more than one tool was employed, please provide details for each.)*	I.4.1 Details

(continued)

APPENDIX (continued)

Section J: Data analysis	
J.1 Which statistical methods, if any, were used in the analysis? (check all that apply)	J.1.1 Descriptive J.1.2 Correlation J.1.3 Group differences (e.g., *t* test, ANOVA) (please specify) J.1.4 Growth Curve analysis/multilevel modeling (HLM) J.1.5 Structural equation modeling (SEM) J.1.6 Path analysis J.1.7 Regression J.1.8 Latent growth curve J.1.9 Other (please specify)
Section K: Results and conclusion	
K.1 What are the results of the study as reported by authors? *Please give details and refer to page numbers in the report(s) of the study, where necessary (e.g., for key tables).*	Group means: *SD*: *N*: Estimated effect size: Appropriate *SD*: *F*, *t* statistic: Significance: Inverse variance weight:
K.2 Are there any shortcomings in the reporting of the data? *Please list all implicit and explicit shortcomings of the study.*	K.2.1 Yes (please specify) K.2.2 No
K.3 Do the authors report on all variables they aimed to study as specified in their aims/research questions? *This excludes variables just used to describe the sample.*	K.3.1 Yes K.3.2 No (please specify)
K.4 What do the author(s) conclude about the findings of the study? *Please give details and refer to page numbers in the report of the study, where necessary.*	K.4.1 Details

Note. ESL = English as a second language; LEA = local educational agencies; HLM = hierarchical linear modeling. Thanks to the EPPI-Reviewer database system team (Thomas & Brunton, 2006).

Notes

This research was supported by a grant from the Canadian Language & Literacy Research Network (CLLRNet) and funds from the Canadian Council on Learning (CCL).

[1]When the value of Q is less than or equal to the degree of freedom associated with a subset, I^2 is assigned a value of zero. Similarly, negative values of I^2 are assigned a value of zero so that I^2 lies between 0% and 100%. A value of 0% indicates no observed heterogeneity, and larger values show increasing heterogeneity. Higgins and Thompson (2002) recommend that percentages of around 25% ($I^2 = .25$), 50% ($I^2 = .50$), and 75% ($I^2 = .75$) should be interpreted to mean low, medium, and high heterogeneity, respectively.

[2]The "mixed" category consists of studies conducted in multiple locations (e.g., Ransdell, Barbier, & Niit, 2006, reported a study conducted in three disparate geographical locations—Estonia, France, and the United States).

References

References marked with an asterisk indicate articles included in the meta-analysis.

Abrami, P. C., & Bernard, R. M. (2006). Research on distance education: In defense of field experiments. *Distance Education, 27,* 5–26.

*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*, 661–678.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A***,** 5–28.

Baker, K. A., & de Kanter, A. A. (1981). *Effectiveness of bilingual education: A review of the literature.* Washington, DC: U.S. Department of Education, Office of Planning, Budget and Evaluation.

Bamford, K. W., & Mizokawa, D. T. (1990, April). *Cognitive development of children in an additive-bilingual program: The third report*. Paper presented at the meeting of the American Educational Research Association, Boston, MA.

Bamford, K. W., & Mizokawa, D. T. (1991, April). *Cognitive and language development in an additive-bilingual program: Report after four observations.* Paper presented at the meeting of the American Educational Research Association, Chicago, IL.

Bamford, K. W., & Mizokawa, D. T. (1992, April). *Spanish-immersion children in Washington State: Fourth year of a longitudinal study*. Paper presented at the meeting of the American Educational Research Association, Seattle, WA.

Berguno, G., & Bowler, D. M. (2004). Communicative interactions, knowledge of a second language, and theory of mind in young children. *Journal of Genetic Psychology, 165*, 293–309.

Bhatia, T., & Ritchie, W. (Eds.). (2006). *Handbook of bilingualism*. Cambridge, MA: Blackwell.

Bialystok, E. (1987). Words as things: Development of word concept by bilingual children. *Studies in Second Language Acquisition, 9*, 133–140.

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology, 24*, 560–567.

*Bialystok, E. (1997). Effects of bilingualism and biliteracy on children's emerging concepts of print. *Developmental Psychology, 33*, 429–440.

*Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70*, 636–644.

Bialystok, E. (2001a). Against isolationism: Cognitive perspectives on second language research. *Selected Proceedings of the Second Language Research Forum,* 97–103.

Bialystok, E. (2001b). *Bilingualism in development: Language, literacy, and cognition.* Cambridge, UK: Cambridge University Press.

Bialystok, E. (2002). Acquisition of literacy in bilingual children: A framework for research. *Language Learning, 52,* 159–199.

Bialystok, E. (2005). Consequences of bilingualism for cognitive development. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 417–432). New York, NY: Oxford University Press.

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

Bialystok, E., Craik, F. I. M., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia, 45,* 459–464.

*Bialystok, E., Craik, F. I. M., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effect of bilingualism on cognitive control in the Simon task: Evidence from MEG. *NeuroImage, 24*, 40–49.

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

Bialystok, E., Craik, F. I. M., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition, 34,* 859–873.

*Bialystok, E., Craik, F. I. M., & Ruocco, A. C. (2006). Dual-modality monitoring in a classification task: The effects of bilingualism and aging. *Quarterly Journal of Experimental Psychology, 59*, 1968–1983.

Bialystok, E., Craik, F. I. M., & Ryan, J. (2006). Executive control in a modified anti-saccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1341–1354.

*Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading, 9*, 43–61.

*Bialystok, E., & Majumder, S. (1998). The relationship between bilingualism and the development of cognitive processes in problem solving. *Applied Psycholinguistics, 19*, 69–85.

*Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics, 24*, 27–44.

*Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science, 7*, 325–339.

Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism, 9,* 103–119.

*Bialystok, E., McBride-Chang, C., & Luk, G. (2005). Bilingualism, language proficiency, and learning to read in two writing systems. *Journal of Educational Psychology, 97*, 580–590.

*Bialystok, E., & Senman, L. (2004). Executive processes in Appearance-reality tasks: The role of inhibition of attention and symbolic representation. *Child Development, 75*, 562–579.

*Bialystok, E., & Shapero, D. (2005). Ambiguous benefits: The effect of bilingualism on reversing ambiguous figures. *Developmental Science, 8*, 595–604.

Bochner, S. (1996). The learning strategies of bilingual versus monolingual students. *British Journal of Educational Psychology, 66*, 83–93.
Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2008). *Comprehensive meta-analysis* (Version 2.2.048). Englewood, NJ: Biostat.
Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta analysis*. Chichester, UK: Wiley.
Braccini, F., & Cianchi, R. (1993). The influence on some linguistic and cognitive skills of the early learning of a foreign language. *Rassegna Italiana Di Linguistica Applicata, 25*, 53–66.
*Bruck, M., & Genesee, F. (1995). Phonological awareness in young second language learners. *Journal of Child Language, 22*, 307–324.
*Campbell, R., & Sais, E. (1995). Accelerated metalinguistic (phonological) awareness in bilingual children. *British Journal of Developmental Psychology, 13*, 61–68.
Carson, K. P., Schriesheim, C. A., & Kinicki, A. J. (1990). The usefulness of the "fail-safe" statistic (N) in meta-analysis. *Educational and Psychological Measurement, 50,* 233–243.
Cazden, C. R. (1974). Play with language and metalinguistic awareness: One dimension of language experience. *Urban Review, 7*, 28–29.
Cenoz, J. 2003. The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism, 7,* 71–87.
Chan, K. T. (2005). Chinese-English bilinguals' theory-of-mind development. *Dissertation Abstracts International, A: The Humanities and Social Sciences, 65* (10-A).
*Chen, X., Anderson, R. C., Li, W., Hao, M., Wu, X., & Shu, H. (2004). Phonological awareness of bilingual and monolingual Chinese children. *Journal of Educational Psychology, 96*, 142–151.
*Clarkson, P. C., & Galbraith, P. (1992). Bilingualism and mathematics learning—another perspective. *Journal for Research in Mathematics Education, 23*, 34–44.
Cooper, H. M., & Hedges, L. V. (1994). *The handbook of research synthesis*. New York, NY: Russell Sage.
*Cromdal, J. (1999). Childhood bilingualism and metalinguistic skills: Analysis and control in young Swedish English bilinguals. *Applied Psycholinguistics, 20*, 1–20.
Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism*, *9,* 1–43.
*Davidson, D., & Tell, D. (2005). Monolingual and bilingual children's use of mutual exclusivity in the naming of whole objects. *Journal of Experimental Child Psychology, 92*, 25–45.
*Demie, F., & Strand, S. (2006). English language acquisition and educational attainment at the end of secondary school. *Educational Studies, 32*, 215–231.
*Demont, E. (2001). Contribution of early second language learning in the development of language awareness and learning to read. *International Journal of Psychology, 36,* 274–285.
Denson, N. (2009). Do curricular and co-curricular diversity activities influence racial bias? A meta-analysis. *Review of Educational Research, 79,* 805–838.
Diaz, R. M. (1985). The intellectual power of bilingualism. *Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 7*, 16–22.
Diaz, R. M., & Klinger, C. (1991). Towards an explanatory model of the interaction between bilingualism and cognitive development. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 167–192). Cambridge, UK: Cambridge University Press.
Duncan, S. E. (2005). Child bilingualism and cognitive functioning: A study of four Hispanic groups. *Dissertation Abstracts International, A: The Humanities and Social Sciences, 65*(8), 2895–A.

Emmorey, K., Luk, G., Pyers, J. E., & Bialystok, E. (2008). The source of enhanced cognitive control in bilinguals. *Psychological Science*, *19,* 1201–1206.

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

Esses, V. M., & Gardner, R. C. (1996). Multiculturalism in Canada: Context and current status. *Canadian Journal of Behavioural Science, 28,* 145–152.

*Eviatar, Z., & Ibrahim, R. (2000). Bilingual is as bilingual does: Metalinguistic abilities of Arabic-speaking children. *Applied Psycholinguistics, 21*, 451–471.

Ferdman, B. M., & Hakuta, K. (1985, April). *A population perspective on bilingualism in Puerto Rican children*. Paper presented at the meeting of the Society for Research in Child Development, Toronto, Canada.

Fernandes, M. A., Craik, F. I. M., Bialystok, E., & Kreuger, S. (2007). Effects of bilingualism, aging, and semantic relatedness on memory under divided attention. *Canadian Journal of Experimental Psychology, 61,* 128–141.

Flavell, J. H. (1978). Metacognitive development. In J. M. Scandura & C. J. Brainerd (Eds.), *Structural process theories of complex human behavior* (pp. 213–245). Alphen aan den Rijn, Netherlands: Sijthoff & Noordhoff.

*Galambos, S. J., & Goldin-Meadow, S. (1990). The effects of learning two languages on levels of metalinguistic awareness. *Cognition, 34*, 1–56.

Galambos, S., & Hakuta, K. (1988). Subject-specific and task-specific characteristics of metalinguistic awareness in bilingual children. *Applied Psycholinguistics, 9,* 141–162.

Goetz, P. J. (2000). A comparison of 3- and 4-year-old English monolinguals, mandarin monolinguals, and mandarin-English bilinguals on a series of theory of mind-related tasks. *Dissertation Abstracts International, B: Sciences and Engineering, 61*, 1110–B.

Goncz, L. (1988). A research study on the relation between early bilingualism and cognitive development. *Psychologische Beitrage, 30*, 75–91.

Government of Canada. (2009). *Roadmap for Canada's linguistic duality, 2008–2013: Acting for the future*. Retrieved from http://www.pch.gc.ca/pgm/slo-ols/strat-eng.cfm

*Gutierrez-Clellen, V. F., Calderon, J., & Weismer, S. E. (2004). Verbal working memory in bilingual children. *Journal of Speech, Language, and Hearing Research, 47*, 863–876.

Hakuta, K. (1986). *Mirror of language: The debate on bilingualism*. New York, NY: Basic Books.

Hakuta, K. (1990). *Bilingualism and bilingual education: A research perspective* (Occasional Papers in Bilingual Education No. 1). Washington, DC: National Clearinghouse for Bilingual Education.

Hamers, J. F., & Blanc, M. H. A. (2000). *Bilinguality and bilingualism* (2nd ed.). Cambridge, UK: Cambridge University Press.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107–128.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Hedges, L. V., & Vevea, J. L. (1998). Fixed and random-effects models in meta-analysis. *Psychological Methods, 3,* 486–504.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21,* 1539–1558.

Ho, D. Y. (1987). Bilingual effects on language and cognitive development: With special reference to Chinese-English bilinguals. *Bulletin of the Hong Kong Psychological Society, 18*, 61–69.

Huber, E., & Lasagabaster, D. (2000). The cognitive effects of bilingualism. *ITL, Review of Applied Linguistics, 129–130*, 191–224.

Huedo-Medina, T. B., Sánchez-Meca, F., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: I^2 or Q statistic? *Psychological Methods*, *11,* 193–206.

*Humphreys, G. W., & Mumtaz, S. (2001). The effects of bilingualism on learning to read English: Evidence from the contrast between Urdu-English bilingual and English monolingual children. *Journal of Research in Reading, 24*, 113–134.

Iannaccone, A., Fraternali, O. T., & Vaccia, R. (1992). Cognitive flexibility and early bilingual education. *Rassegna Italiana Di Linguistica Applicata, 24*, 149–165.

Johnson, J. (1991). Constructive processes in bilingualism and their cognitive growth effects. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 193–221). Cambridge, UK: Cambridge University Press.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21,* 60–99.

*Jordà-Maria, P. S. (2003). Metapragmatic awareness and pragmatic production of third language learners of English: A focus on request acts realizations. *International Journal of Bilingualism, 7*, 43–68.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 98,* 122–149.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130,* 169–183.

Kemp, C. (2007). Strategic processing in grammar learning: Do multilinguals use more strategies? *International Journal of Bilingualism, 4,* 241–261.

*Keshavarz, M. H., & Astaneh, H. (2004). The impact of bilinguality on the learning of English vocabulary as a foreign language (L3). *International Journal of Bilingual Education and Bilingualism, 7*, 295–302.

Koda, K., & Zehler, A. (Eds.). (2008*). Learning to read across languages: Cross-linguistic relationships in first- and second-language literacy development.* New York, NY: Routledge.

Konaka, K. (1997). The relationship between degree of bilingualism and gender to divergent thinking ability among native Japanese-speaking children in the New York Area. *Dissertation Abstracts International,* AAT 9718715.

Kovacs, A. M., & Teglas, E. (2002). Integrating two languages, theories of minds, and executive functions. *Odense Working Papers in Language and Communication, 3*, 1.

Kozulin, A. (1999). Reality monitoring, psychological tools, and cognitive flexibility in bilinguals: Theoretical synthesis and pilot experimental investigation. In P. Lloyd & C. Fernyhough (Eds.), *Lev Vygotsky: Critical assessments: Future directions* (Vol. 4, pp. 187–198). New York, NY: Routledge.

Lee, H., Plass, J. L., & Homer, B. D. (2006). Optimizing cognitive load for learning from computer-based science simulations. *Journal of Educational Psychology*, *98*, 902–913.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

*Loizou, M., & Stuart, M. (2003). Phonological awareness in monolingual and bilingual English and Greek five-year-olds. *Journal of Research in Reading, 26*, 3–18.

*Love, T., Maas, E., & Swinney, D. (2003). The influence of language exposure on lexical and syntactic language processing. *Experimental Psychology, 50*, 204–216.

Macnamara, J. (1966). *Bilingualism and primary education*. Edinburgh, UK: Edinburgh University Press.

McLeay, H. (2003). The relationship between bilingualism and the performance of spatial tasks. *International Journal of Bilingual Education and Bilingualism, 6*, 423–438.

*Merriman, W. E., & Kutlesic, V. (1993). Bilingual and monolingual children's use of two lexical acquisition heuristics. *Applied Psycholinguistics, 14*, 229–249.

Michael, E., & Gollan, T. H. (2005). Being and becoming bilingual: Individual differences and consequences for language production. In J. F. Kroll & A. M. B. de Groot (Eds.), *The handbook of bilingualism: Psycholinguistic approaches* (pp. 389–407). New York, NY: Oxford University Press.

National Research Council. (1992). *Combining Information: Statistical issues and opportunities for research*. Washington, DC: National Academy Press.

*Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics, 28*, 191–230.

Orwin, R. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics, 8,* 157–159.

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs, 76,* 1–23.

Ponomarev, S. Y. (1992). Bilingualism and its influence on personality development. *Vestnik Sankt-Peterburgskogo Universiteta, Seriya 2: Istoriya, Yazykoznanie, Literaturovedenie, 3*, 46–49.

*Ransdell, S., Barbier, M., & Niit, T. (2006). Metacognitions about language skill and working memory among monolingual and bilingual college students: When does multilingualism matter? *International Journal of Bilingual Education and Bilingualism, 9*, 728–741.

Ravitch, D. (1990). Multiculturalism: E pluribus plures. *American Scholar, 59*, 339–340.

*Reichard, C., & Mokhtari, K. (2004). Investigating the strategic reading processes of first and second language readers in two different cultural contexts. *System, 32*, 379–394.

Ricciardelli, L. A. (1992). Creativity and bilingualism. *Journal of Creative Behavior, 26*, 242–254.

Ricciardelli, L. A. (1993). An investigation of the cognitive development of Italian-English bilinguals and Italian monolinguals from Rome. *Journal of Multilingual and Multicultural Development, 14*, 345–346.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General, 126*, 211–227.

Rosenblum, T., & Pinker, S. A. (1983). Word magic revisited: Monolingual and bilingual children's understanding of the word-object relationship. *Child Development*, *54*, 773–780.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86,* 638–641.

Rossell, C. H., & Baker, K. (1996). The educational effectiveness of bilingual education. *Research in the Teaching of English, 30*, 7–69.

*Sanz, C. (2000). Bilingual education enhances third language acquisition: Evidence from Catalonia. *Applied Psycholinguistics, 21*, 23–44.

*Sheng, L., McGregor, K. K., & Marian, V. (2006). Lexical-semantic organization in bilingual children: Evidence from a repeated word association task. *Journal of Speech, Language, and Hearing Research, 49*, 572–587.

Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research*, *75*, 247–284.

Srivastava, B. (1991). Creativity and linguistic proficiency. *Psycho-Lingua, 21*, 105–109.

Stephens, M. A. (1997). Bilingualism, creativity, and social problem-solving. *Dissertation Abstracts International,* AAT 9729615.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition & Instruction, 12,* 185–233.

Thomas, J., & Brunton, J. (2006). *EPPI-Reviewer 3.0: Analysis and management of data for research synthesis.* London, UK: Social Science Research Unit, Institute of Education.

Titone, R. (1997). Early second-language learning: Bilingualism and metalinguistic development. *Europe Plurilingue, 12–13*, 138–149.

Tucker, G. R. (1998). A global perspective on multilingualism and multilingual education. In J. Cenoz & F. Genesee (Eds.), *Beyond bilingualism: Multilingualism and multilingual education* (pp. 3–15). Clevedon, UK: Multilingual Matters.

van Merrienboer, J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17*, 147–177.

Vorstman, E., De Swart, H., Ceginskas, V., & Van Den Bergh, H. (2009). Language learning experience in school context and metacognitive awareness of multilingual children. *International Journal of Multilingualism, 6,* 258–280.

Willig, A. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research, 55*, 269–317.

Willig, A. (1987). Examining bilingual education research through meta-analysis and narrative review: A response to Baker. *Review of Educational Research, 57*, 363–376.

Yang, H., Yang, S., Ceci, S., & Wang, Q. (2005). Effects of bilinguals' controlled-attention on working memory and recognition. In J. Cohen, K. McAlister, K. Rolstad, & J. MacSwan (Eds.), *Proceedings of the 4th international symposium on bilingualism.* Somerville, MA: Cascadilla Press.

Yoshida, H. (2008). The cognitive consequences of early bilingualism. *Journal of Zero to Three, 29,* 26–30.

Authors

OLUSOLA ADESOPE holds a Ph.D. in Educational Psychology from Simon Fraser University, Canada. He recently accepted a position as Assistant Professor of Educational Psychology in the Department of Educational Leadership & Counseling Psychology at Washington State University, Pullman, WA 99164-2136; email: *adesope@yahoo.com*. His research interests include the cognitive and psychological processes underlying multimedia learning, teaching and learning with interactive concept maps and diagrams, language pedagogy, technology-mediated instruction and research synthesis, especially meta-analysis.

TRACY LAVIN is a Researcher with the Directions Evidence and Policy Research Group, 1055 Dunsmuir–Suite 1254, Four Bentall Centre, Vancouver, BC V7X 1A2; email: *lavin.t@gmail.com*. She received a Ph.D. from the University of British Columbia and pursued postdoctoral studies at Northwestern University. She has studied and written

about a number of learning and educational issues from both academic and public policy perspectives. Her research interests include studying how children and adults learn language and acquire its underpinning cognitive concepts. She has written several research papers for the National Literacy Strategy initiative and also leads the team of analysts responsible for developing, analyzing, and reporting on Canadian Council on Learning's annual Survey of Canadian Attitudes toward Learning.

TERRI THOMPSON is a Researcher with the Directions Evidence and Policy Research Group, 1055 Dunsmuir–Suite1254, Four Bentall Centre, Vancouver, BC V7X 1A2; email: *territho@gmail.com*. Over the past five years, she has participated in, and managed the completion of, over 60 systematic reviews and question scans devoted to issues about education and learning. She holds a Master's degree in Education from the University of British Columbia where she studied issues of social justice and education policy.

CHARLES UNGERLEIDER is a professor of the sociology of education at the University of British Columbia, 2125 Main Mall, Vancouver, BC V6T 1Z4, Canada; e-mail: *charles.ungerleider@ubc*.ca. His research interests include educational policy, finance and governance, educational change, and the educational trajectories of students in elementary and secondary schools.