medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Is there a bilingual advantage in auditory attention among children? A systematic review and meta-analysis of standardized auditory attention tests

Wenfu Bao[1, 2*], Claude Alain[3, 4, 5, 6], Michael Thaut[2, 5, 6, 7], and Monika Molnar[1, 2]

[1] Department of Speech-Language Pathology, University of Toronto, Toronto, Ontario, Canada

[2] Rehabilitation Sciences Institute, University of Toronto, Toronto, Ontario, Canada

[3] Rotman Research Institute, Baycrest Health Centre, Toronto, Ontario, Canada

[4] Department of Psychology, University of Toronto, Toronto, Ontario, Canada

[5] Institute of Medical Sciences, University of Toronto, Toronto, Ontario, Canada

[6] Music and Health Science Research Collaboratory, University of Toronto, Toronto, Ontario, Canada

[7] Faculty of Music, University of Toronto, Toronto, Ontario, Canada

* Corresponding author

E-mail: wenfu.bao@mail.utoronto.ca (WB)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

A wealth of research has investigated the effects of bilingualism on cognition, especially on executive function. Developmental studies reveal different cognitive profiles between monolinguals and bilinguals in (audio-)visual attention tasks, which might stem from their attention allocation differences. Yet, whether such distinction exists in the auditory domain alone is unknown. In this study, we compared differences in auditory attention, measured by standardized tests, between monolingual and bilingual children. A comprehensive literature search was conducted in three electronic databases: OVID Medline, OVID PsycInfo, and EBSCO CINAHL. Twenty studies using standardized tests to assess auditory attention in monolingual and bilingual participants aged less than 18 years were identified. We assessed the quality of these studies using a scoring tool for evaluating primary research. For statistical analysis, we pooled the effect size in a random-effects meta-analytic model, where between-study heterogeneity was quantified using the I^2 statistic. No substantial publication bias was observed based on the funnel plot. Further, meta-regression modelling suggests that test measure (accuracy vs. response times) significantly affected the studies' effect sizes whereas other factors (e.g., participant age, stimulus type) did not. Specifically, studies reporting accuracy observed marginally greater accuracy in bilinguals (g = 0.10), whereas those reporting response times indicated faster latency in monolinguals (g = -0.34). There was little difference between monolingual and bilingual children's performance on standardized auditory attention tests. We also found that studies tend to include a wide variety of bilingual children but report limited language background information of the participants. This, unfortunately, limits the potential theoretical contributions of the reviewed studies. Recommendations to improve the quality of future research are discussed.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

Over the past decade, the growing body of research investigating the effects of bilingualism on cognition suggests that bilingual experience can shape the brain and cognitive systems [1, 2]. Some studies [3, 4] demonstrate that bilingual participants outperform their monolingual counterparts on a wide range of cognitive tasks while controlling for other factors, which is often interpreted as a bilingual advantage. For example, bilingual children develop conflict resolution ability earlier than monolinguals and perform better on memory tasks based on executive control [3] (but see [5] for equal performance in inhibitory tasks).

Cognition has many facets, among which executive function (EF) is most frequently measured in bilingual developmental studies [6]. Recent meta-analyses present mixed results about whether bilingualism confers an EF benefit in children. For instance, Gunnerud et al. [7] analyzed 143 studies examining different EF components in children aged 18 months through 14.5 years; no bilingual advantage was found in overall EF after adjusting for publication bias. Similarly, Lowe et al. [8] reported a negligible effect of bilingualism on overall EF (g = -0.04) in children aged between 3 and 17 years. It is worth noting that when addressing attention, nearly all studies included in these meta-analyses have focused on the (audio-)visual domain, and very few measure auditory attention through tools like behavioral tasks or standardized tests. Using the Bayesian statistical approach, Grundy [9] revealed that when group differences do appear on EF tasks, bilinguals outperform monolinguals far more likely than chance. Given these discrepant findings, the bilingual advantage appears small at best and might be subject to specific circumstances [10]. However, what these “circumstances” are remains unclear.

Addressing whether a general bilingual advantage exists is beyond the scope of this study. We do not intend to engage in a dichotomous discussion to this debate, as it oversimplifies

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

the effects of bilingualism on cognitive development. Nevertheless, we acknowledge that bilingualism can exert influence on cognition at least in some types of bilinguals, which has been supported by empirical studies [11, 12]. Further, attention has been hypothesized to be a plausible domain responsible for the bilingualism effects [2]. That is, the habitual use of two languages over years possibly enables bilinguals to become more practiced in managing conflicts and controlling attention. In particular, early bilingual exposure can affect how attention is allocated to the environment. For instance, infants growing up in bilingual homes pay more attention to subtle environmental differences [13, 14], which could improve their attentional processing [15].

However, a definition of attention remains vague in cognitive literature [16, 17]. Here, we conceptualize it as a system with three primary components: sustained attention, selective attention, and executive control [18, 19]. Specifically, based on Petersen and Posner [18], *sustained attention* (or alerting) refers to maintaining alertness over a long period of time, which is usually measured in the form of vigilance tasks that involve monitoring a target stimulus interspersed with non-targets. *Selective attention* (or orienting) refers to the ability to select certain input for enhanced processing while suppressing other irrelevant information, and thus is often measured by focusing on the target stimuli and ignoring the distractors. Lastly, *executive control* deals with resolving conflicts, shifting attention, and regulating thoughts and behavior. These components are also associated with different neural substrates in the human brain [18].

In bilingualism research, attention was not clearly conceptualized prior to a recent paper by Bialystok and Craik [16]. According to it, lifelong bilingual experience enhances attentional control, which is defined as a repertoire of processing operations that higher-level cognition utilizes to fulfill various goals [16]. However, visual or audio-visual attention has been primarily

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

studied in the past, whereas auditory attention has received little consideration. Though auditory attention is important to language processing and development: language is often processed through the auditory domain alone (e.g., speech perception), and infants start learning about language in utero without any visual support [20]. Likewise in the visual domain, bilingualism might also affect how auditory attention is allocated in bilingual children.

The current systematic review and meta-analysis assess whether there are reliable differences in auditory attention, measured by standardized tests, between monolingual and bilingual children. An initial search in our laboratory indicates that standardized tests are often used to assess children's auditory attention in research and clinical settings [21], and are an ideal comparison across contexts given consistent administration guidelines. These tests use different experimental paradigms to target different auditory attention components. For example, the Go/No-Go task is often employed to assess sustained attention, during which participants are asked to respond in some conditions but not to respond in others. Accordingly, depending on the task, different outcome measures are reported, such as response speed and accuracy. In addition, these tests use different types of auditory stimuli, which are either linguistic (e.g., syllables, words) or non-linguistic (e.g., tones, animal sounds; see Results for further information).

As attention might support monolingual and bilingual development differently [2], investigating auditory attention in monolingual and bilingual children illuminates different adaptations to their language environments. Additionally, examining the bilingualism effects on auditory attention addresses the current literature gap by focusing on the auditory domain, which has been overlooked in prior attention research. Considering evidence on bilingualism modulating audiovisual speech processing [22], our prediction is that if auditory attention development is shaped by bilingual experience, bilingual children might have more accurate

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

and/or faster responses than their monolingual counterparts in standardized tests. We are also interested to explore whether the difference would vary by the attention components assessed.

An additional goal of our study is to determine whether certain bilingual characteristics (e.g., age of acquisition, language proficiency) mitigate the potential differences between monolinguals and bilinguals. Since the bilingualism effects are more evident among those with higher language proficiency and greater exposure [11, 12], we hypothesize that simultaneous bilinguals (i.e., children who learn both languages before the age of three) would more likely show enhanced auditory attention than sequential bilinguals (i.e., children who learn additional languages after the age of three). Finally, given that it is unclear whether bilingualism exerts influence beyond linguistic domain and standardized tests vary by stimulus type (linguistic vs. non-linguistic), we assess if different auditory stimuli affect attention performance in monolingual and bilingual children.

Materials and methods

This study is part of a larger systematic review that investigates auditory attention development from infancy to adolescence, in which the population's language experience (i.e., monolingual vs. bilingual) and research methods are not controlled. The review protocol was registered a priori with OSF [21]. We followed the *Cochrane Handbook for Systematic Reviews of Interventions* [23] as the methods guidance, and developed the search strategy in consultation with a health sciences librarian. A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram illustrating the screening process is presented in Fig 1. The methods described below are specific to our current (more focused) literature search embedded in Bao and Molnar [21].

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Fig 1. PRISMA flow diagram presenting the study screening process.

Eligibility criteria

To include the most relevant studies on auditory attention measured by standardized tests, we adhered to the following criteria when determining a study's eligibility:

- Studies that used standardized tests to measure auditory attention were included. *Standardized tests* are norm-referenced tests administered and scored in a consistent manner. Studies using non-standardized measures or focusing on visual attention and other cognitive abilities were excluded.
- Studies that tested typically developing individuals below the age of 18 and had a monolingual group and a bilingual group were included. The bilingual group should be raised in bilingual families (i.e., at least one parent speaks another language than English to the child) or study in immersion schools where most of the curriculum is taught in an unfamiliar language. Studies only having monolingual or bilingual participants were excluded.
- Studies that controlled for participants' age and socio-economic status (SES) were included, as both factors can affect children's cognitive measures independently [24].
- Published, empirical studies that reported primary results in peer-reviewed articles were included.
- Only articles written in English were included.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Information sources and search

A comprehensive literature search was conducted in three electronic databases: OVID Medline, OVID PsycInfo and EBSCO CINAHL. Search terms were defined after consulting the librarian and tailored to each database (see S1-S3 Tables). The search contained articles from the databases' start date up to March 8, 2023. Based on the search result of the larger systematic review, we identified all studies that compared monolingual and bilingual performance on standardized auditory attention tests. In addition, a manual search was performed by checking the reference lists of relevant articles, and an updated manual search was undertaken in Google Scholar to identify more studies.

Study selection

Records and data were managed using the Covidence systematic review software [25]. Duplicates were identified and removed during reference importing and screening. Two independent reviewers screened studies in two phases: (1) title/abstract screening required to indicate "Yes," "No" or "Maybe" for relevance; (2) full-text screening required to indicate "Include" or "Exclude" for eligibility and specify the exclusion reasons. If discrepancies occurred, a third reviewer was called to resolve the conflicts.

Data collection

For eligible studies, two independent reviewers extracted the following data items on Covidence: (1) sample size (monolingual vs. bilingual); (2) participant age; (3) language background (first and second language), bilingual type (simultaneous vs. sequential), and bilingualism assessment; (4) standardized tests used to assess auditory attention; (5) test measure

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

(accuracy vs. response times or RTs); (6) stimulus type (linguistic vs. non-linguistic); (7) auditory attention components measured; (8) reporting of SES. In the case of unclear or missing items, we contacted the investigators to confirm and obtain additional information. Afterwards, a third reviewer compared the data extracted by the two reviewers and built consensus: for items where there was a conflict, a final decision was made by selecting or entering the most accurate response. Then the consensus data was exported for analysis.

Quality assessment

The methodological quality of the studies was assessed by two independent reviewers using a modified version of the *Standard Quality Assessment Criteria for Evaluating Primary Research Papers: Quality Scoring for Quantitative Studies* (i.e., the "QualSyst" tool; [26]). Specifically, each study was evaluated according to 11 items based on description or reporting of objective, study design, participant selection, participant characteristics, outcome measure, analytic methods, estimate of variance, results, and conclusions, as well as sample size determination and confounding factors control. Each item was scored depending on the degree to which the specific criteria were met ("yes" = 2, "partial" = 1, "no" = 0), and then a summary score was calculated for each study and averaged between reviewers as the final rating. All studies received a score above 0.8, indicating high quality. Further, we used this tool to identify limitations of the reviewed studies and provide recommendations for future research.

Data analysis

Apart from a narrative synthesis of included data, we performed a meta-analysis on the pooled effect size using the "meta" package (version 5.2-0) [27] in *R* [28]. Specifically, we

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

extracted raw effect size data in the form of means and standard deviations of the two groups (monolingual and bilingual) from the included studies. In view of insufficient details being reported, we contacted the original investigators to obtain them or used the WebPlotDigitizer tool [29] to extract data from the graphs (e.g., in Krizman et al. [30). As we anticipated considerable between-study heterogeneity, a random-effects model was built using the "metacont" function to pool effect sizes. Given that our effect size data was continuous, we used the restricted maximum-likelihood estimator [31] to calculate the heterogeneity variance τ^2, and the Hedges method to calculate the standardized mean difference (i.e., Hedges' *g*). To control for uncertainty in heterogeneity estimates, we used Knapp-Hartung adjustments [32] to calculate the confidence interval (CI) around the pooled effect. In terms of quantifying heterogeneity, we reported the I^2 statistic (i.e., percentage of variability not caused by sampling error) [33] along with its confidence intervals, as suggested in [34].

To further explore sources of statistical heterogeneity, we built mixed-effects meta-regression models using the "metafor" package (version 3.0-2) [35]. Specifically, the dependent variable was the unbiased estimate of the population effect size Hedges' *g*, calculated for each study. The independent variables were test measure (binary: accuracy vs. RTs), stimulus type (binary: linguistic vs. non-linguistic), participant age (continuous, in years) and attention components (categorical: selective attention, sustained attention, executive control, or auditory attention overall). To evaluate the effects of these predictors, we conducted model comparison in a forward stepwise manner: a likelihood ratio test was performed using the *anova* function to compare a reduced model and a full model, which had one additional component. Specifically, we inspected the estimated *p*-value and Akaike's Information Criterion (AIC) value to assess model performance (Bayesian Information Criterion or BIC is preferred over AIC when the

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

heterogeneity is large in the studies [36]). The full model was favored only when the difference was significant as indicated by the *p*-value (less than the conventional threshold of 0.05) and when it provided a better fit for the data as suggested by lower AIC value. For categorical variables that significantly predicted the effect size, subgroup analyses were conducted using the "meta" package and forest plots were generated to visualize the effects. Of note, a limitation of meta-regression is that it describes an observational association across studies rather than a causal relationship; thus its findings should be interpreted with caution [37, 38].

Furthermore, publication bias was analyzed for the included studies by means of the funnel plot, created using the "meta" package. The funnel plot displays the studies' effect size against its standard error. Usually, symmetry indicates the absence of publication bias, which is reflected by data points scattered around the mean effect size forming an upside-down funnel [39]. This symmetry was further quantified through the Egger's regression test.

Results

Synthesis

The database search yielded 634 records for title/abstract screening. Forty articles remained for full-text review, and three articles qualified for extraction. Our manual search contributed another nine articles, four of which did not appear in the database search as they were not indexed. In total, 12 articles were extracted, which included 20 studies (see Table 1). Note that studies from the same article represent independent effect sizes, because multiple factors, such as age, outcome measure, stimulus type, etc. were considered across these studies.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Population characteristics

With an age range between 5 to 14 years, participants can be categorized into preschoolers (5-6 years), primary school-aged children (7-11 years), and young adolescents (12-14 years). This categorization has considered different education systems across countries. For example, while primary school starts at the age of four in the Netherlands, the first two years are comparable to kindergarten. Therefore, the five-year-old children in Boerma et al. [40] and Kwakkel et al. [41] were treated as preschoolers. Table 1 presents a detailed description of all participants. Taking language background for an example, there are a great variety of first and second languages, with French-English and Spanish-English being the most common language combinations within bilinguals.

Table 1. Descriptive statistics of included studies.

Study	Country	Monolingual (*n*)	Bilingual (*n*)	Age (mean; years)	First language; second language	Bilingual type	Bilingualism assessment	Standardized test	Test measure	Stimulus type	Attention component	SES
Barbu et al. 2019	Belgium	57	59	6	French; English	Sequential* (French children enrolled in an immersion program since the age of 5)	Customized language background questionnaire	KiTAP: The Owls	RT	Non-linguistic: animal sounds	Selective attention	Matched for two groups, from diverse levels
Boerma et al. 2017	The Netherlands	32	32	5	Turkish, Tarifit-Berber, Arabic; Dutch	Sequential* (children learned Dutch as an L2 since the age of 4)	PaBiQ	IVA + Plus: Auditory task	Accuracy	Linguistic: numbers	Sustained attention	No group difference
Foy & Mann 2014 (1) Foy & Mann 2014 (2) Foy & Mann 2014 (3) Foy & Mann 2014 (4)	USA	30	30	5	Spanish; English	Simultaneous* (American children exposed to Spanish since at least 12 months of age)	Customized language background questionnaire; Language Dominance Survey from EOWPVT-SBE	ACPT-P (modified): Verbal Go/No-Go task ACPT-P (modified): Nonverbal Go/No-Go task	RT	Linguistic: speech syllables Non-linguistic: animal and nature sounds	Auditory attention (sustained attention, executive control)	Matched for two groups, all from low status
Garratt & Kelly 2007	UK	27	27	7	English; Bengali, Urdu/Punjabi, Malay, Arabic	Not reported	Customized language background questionnaire	NEPSY: Auditory Attention and Response Set	Accuracy	Linguistic: words	Auditory attention	Comparable between two groups, most from low status
Karlsson et al. 2015 (1) Karlsson et al. 2015 (2) Karlsson et al. 2015 (3) Karlsson et al. 2015 (4)	Finland	25 23 25 23	24 27 24 27	7 11 7 11	Swedish; Finish and other ("other" second languages were not specified in the study)	Simultaneous (almost all bilingual children learned both languages by the age of 3)	Parental report	NEPSY-II: Auditory Attention NEPSY-II: Response Set	Accuracy	Linguistic: words	Selective and sustained attention Executive control	Matched for two groups, except that bilinguals had a higher level in the younger sample
Krizman et al. 2012	USA	25	23	14	Spanish; English	Simultaneous (children had their first exposure to both languages about the age of 3)	LEAP-Q	IVA + Plus: Auditory task	Accuracy	Linguistic: numbers	Sustained and selective attention	Matched for two groups
Krizman et al. 2014	USA	27	27	14	Spanish; English	Simultaneous (children learned both languages before/about the age of 3)	LEAP-Q; parental report	IVA + Plus: Auditory task	Accuracy	Linguistic: numbers	Attentional control	44% of monolinguals and 59% of bilinguals from low status

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Kwakkel et al. 2021	The Netherlands	80	89	5	Dutch; English	Sequential* (Dutch children enrolled in a bilingual program since the age of 4)	Parental questionnaire	ACPT: Go/No-Go task	Accuracy	Non-linguistic: tones	Sustained attention	Matched for two groups, from high status overall
Nicolay & Poncelet 2013	Belgium	51	53	8	French; English	Sequential* (French children enrolled in an immersion program since the age of 5)	Not explicitly assessed	KiTAP: The Owls	RT	Non-linguistic: animal sounds	Selective attention	Matched for two groups, from medium and high levels
Nicolay & Poncelet 2015	Belgium	50	51	8	French; English	Sequential* (French children enrolled in an immersion program since the age of 5)	Not explicitly assessed	KiTAP: The Owls	RT	Non-linguistic: animal sounds	Selective attention	Matched for two groups, from diverse levels
Simonis et al. 2020 (1)	Belgium	129	156	12	French; Dutch, English	Sequential* (French children enrolled in an immersion program since the age of 5)	Background information questionnaire	TEA-Ch: Code Transmission (adapted)	Accuracy	Linguistic: numbers	Sustained attention	Bilinguals had higher SES
Simonis et al. 2020 (2) Simonis et al. 2020 (3)		153	173					TAP: Divided Attention (adapted)	Accuracy RT	Non-linguistic: tones	Selective attention	
Strydom et al. 2022	South Africa	20	20	7	Non-English (first languages were not specified in the study); English	Simultaneous* (children learned English as a second language before the age of 3)	Not explicitly assessed	SAAT: Words Intelligibility by Picture Identification in quiet	Accuracy	Linguistic: words	Selective attention	Same for two groups

PaBiQ: Parents of Bilingual Children Questionnaire. EOWPVT-SBE: Expressive One Word Picture Vocabulary Test – Spanish-English Bilingual Edition. LEAP-Q: Language Experience and Proficiency Questionnaire. KiTAP: Test for Attentional Performance for Children; IVA + Plus: Integrated Visual and Auditory Continuous Performance Test; ACPT: Auditory Continuous Performance Test; ACPT-P: Auditory Continuous Performance Test-Preschoolers; TEA-Ch: Test of Everyday Attention for Children; TAP: Test for Attentional Performance; NEPSY(-II): A Developmental Neuropsychological Assessment (Second Edition); SAAT: Selective Auditory Attention Test.

Numbers following the names signify individual studies from the same article, which considers participant age, test measure, stimulus type, and attention components.

*Asterisk indicates that the bilingual type was not directly reported by the investigators but inferred from the original articles.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

However, quality assessment suggests that there was not sufficient information reported about bilingual characteristics in the reviewed papers. For instance, bilingual type (simultaneous vs. sequential) was rarely reported, except that Karlsson et al. [42] and Krizman et al. [30, 43] recruited "early bilinguals." Usually, bilingual children were from immigrant families learning the societal language as their second language (L2) [40, 44], or enrolled in immersion schools acquiring a foreign language [45-48]. Note that for longitudinal studies, we considered participants' language condition when they were tested. For example, Nicoley and Poncelet [47] collected data at two time points; data from the second one (i.e., three years after immersion school enrollment) were included only, because children cannot be deemed bilingual yet during the initial testing when they just started the program. Generally, immersion school children were considered sequential bilinguals, as most of them were from a monolingual environment and did not learn an L2 until they started school. Furthermore, bilingual experience was often evaluated using parental questionnaires that focus on one aspect (e.g., language use, exposure, or proficiency), which led to inconsistent definitions of bilinguals and the difficulty of comparing them across studies.

SES was reported in all studies, and mostly assessed through education level of the mother or of both parents (parental occupation was also used occasionally, e.g., Garratt and Kelly [44]). Monolingual and bilingual participants had a comparable SES across articles, except in Simonis et al. [48] where the bilingual group had a higher SES than the monolingual group. Given no significant difference in test performance between the two groups, this article was included in the final analysis.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Test characteristics

Various standardized tests (see Table 1) were used to assess auditory attention. Considering how they were described in the included studies, we categorized them into the following three groups: (1) Test for Attentional Performance (TAP) [49], Test of Attentional Performance for Children (KiTAP) [50], and Selective Auditory Attention Test (SAAT) [51]: assess *selective attention*; (2) Integrated Visual and Auditory Continuous Performance Test (IVA + Plus) [52], Auditory Continuous Performance Test for Preschoolers (ACPT-P) [53], and Test of Everyday Attention for Children (TEA-Ch) [54]: assess *sustained attention*; (3) Developmental Neuropsychological Assessment (NEPSY) [55] and its second version (NEPSY-II) [56]: assess *auditory attention overall* (selective and sustained attention, and executive control). For tests evaluating audio-visual attention, we included data from subtests using only auditory stimuli. Further, we found that both linguistic and non-linguistic stimuli were utilized. For example, some tests used non-linguistic stimuli like auditory tones (TAP) and animal sounds (KiTAP; e.g., [45, 48]), whereas others employed linguistic stimuli such as numbers (IVA + Plus; TEA-Ch), speech syllables (ACPT-P), and words (NEPSY, NEPSY-II, SAAT; e.g., [30, 40, , 43, 57]). Though it remains unknown whether different types of auditory stimuli affect attentional performance.

Moreover, some researchers used translated versions (e.g., Swedish version of NEPSY-II in [42]) or adapted the tests [48], and reported the assessed attention components differently. For example, modelled after ACPT-P, Foy and Mann [58] used an auditory Go/No-Go task with two blocks. We infer that the first block assessed sustained attention and the second one evaluated executive control, as the authors did not explicitly report the attention components. Using the same task, however, Kwakkel et al. [41] stated that they measured “sustained attention.”

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Another source of variance we observed involved the test measure reported. All studies presented either accuracy or RTs, except that Simonis et al. [48] did both. Nevertheless, how accuracy was reported in different studies was not always consistent, even for the same test. For instance, Garratt and Kelly [44] reported standard scores for the Auditory Attention and Response Set subtests of NEPSY together, whereas Karlsson et al. [42] reported raw scores for each subtest of NEPSY-II individually. Simonis et al. [48] used decimal numbers to indicate accuracy (i.e., 1 = 100% accuracy), whereas Kwakkel et al. [41] used reversed omission scores. Despite these discrepancies, for all accuracy statistics in this meta-analysis, larger values indicate better auditory attention. However, we did not include data that were presented in other forms and could not be converted to accuracy scores or proportions, e.g., misses in Foy and Mann [58] and errors in Barbu et al. [45].

Meta-analysis

Effect size was pooled in the meta-analysis. According to the random-effects model (see S4 Table), the between-study heterogeneity variance was estimated at $\tau^2 = 0.09$ (95% CI: 0.03-0.29), with an I^2 value of 65.8% (95% CI: 45.1-78.6%) that indicates moderate to substantial heterogeneity. Meta-regression modelling was further implemented to address the heterogeneity effect between studies. We first built a mixed-effects model, with each study's effect size as the dependent variable and test measure as the moderator. Model summary suggested that test measure significantly influenced the studies' effect size ($p = 0.0067$; see S5 Table). Further, results of a subgroup analysis confirmed that there was a significant difference in effect size between studies reporting accuracy (i.e., accuracy studies) and studies reporting RTs (i.e., RT studies; $p = 0.0014$; see S6 Table). Fig 2 shows a forest plot stratified by test measure: bilingual

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

children presented more accurate responses in accuracy studies (g = 0.10), but slower latency in RT studies (g = -0.34) than their monolingual counterparts. However, results favored monolingual children (g = -0.09) in standardized auditory attention tests when taking both measures together.

Fig 2. Forest plot stratified by test measure, based on random-effects meta-analytic model analysis. The upper and lower panels display the results for accuracy studies (n = 12) and RT studies (n = 8), respectively.

To investigate the effects of other variables, we divided our data into two subsets (i.e., accuracy studies and RT studies) and analyzed them separately. For accuracy studies (n = 12), a meta-regression model with participant age as the predictor was built. Age did not influence the effect size significantly (p = 0.1282; see S7 Table). However, a positive trend favoring young bilingual adolescents was observed (see S1 Fig). Following the same approach, we tested the effects of stimulus type and attention components, but neither of them was significant (stimulus type: p = 0.9550; attention components: p = 0.4865; see S8 and S9 Tables). For RT studies (n = 8), none of the variables above was significant (participant age: p = 0.2056, stimulus type: p = 0.5494, attention components: p = 0.9396; see S10-S12 Tables). In addition, there was not substantial publication bias in our data according to the funnel plot (S2 Fig), which was further confirmed by the result of Egger's regression test (p = 0.414; see S13 Table). However, we did not include unpublished research, which is a potential limitation.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Discussion

The current systematic review and meta-analysis analyzed 20 studies that compared monolingual and bilingual children's performance on standardized auditory attention tests. Results suggest that test measure was significantly related to differences in effect sizes: accuracy studies (n = 12) indicated marginally greater accuracy in bilinguals, whereas RT studies (n = 8) indicated faster responses in monolinguals. However, no other factors (i.e., participant age, stimulus type, attention components) resulted in significant differences or interactions. Overall, there was little difference between monolingual and bilingual children's performance on standardized auditory attention tests. This finding is consistent with those of Gunnerud et al. [7] and Lowe et al. [8], though both of them analyzed tasks focusing on visual attention.

In answer to the question in the title, a bilingual advantage in auditory attention, when measured using standardized tests, likely does not exist or is limited to certain conditions (e.g., accuracy measure or more proficient bilinguals). To be fair in our interpretation of this finding, two points should be considered: (1) a very heterogeneous bilingual population was synthesized across studies, wherein bilingualism was poorly assessed; (2) most standardized tests are developed for monolingual English speakers, thus they might not be suitable to measure bilinguals. As far as we know, there is no clear theoretical framework where we could place our results directly or interpret possible bilingual advantages [59]. Our work, however, contributes to the field by uncovering the relation between bilingualism and auditory attention.

In the reviewed papers, standardized auditory attention tests were administered to participants as young as five years of age, and to children (5-11 years) more often than to young adolescents (12-14 years). Monolingual children seemingly had faster RTs, but bilingual children showed somewhat higher accuracy. Prior work from other domains than auditory attention has

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

presented mixed findings about the RT difference between groups. For instance, Bialystok et al. [60] used the Simon task—a non-linguistic interference task—and observed faster latency in bilingual children, although this effect was later only found when the demands for inhibitory control were high [61]. However, studies using language processing tasks show longer RTs in bilingual adults. For example, bilinguals have slower RTs to target words during a lexical decision task [62], and perform more slowly in a picture naming task than monolinguals [63]. These linguistic tasks specifically measure cross-language interference and switching, and bilinguals' slower RTs have been often associated with their need to navigate more than one language system.

In our meta-analysis, bilingual children seemed to have slightly more accurate responses, but at the cost of longer latency relative to their monolingual peers. However, these findings should be interpreted with caution, because most accuracy and RT measures were derived from different studies across the reviewed articles. Only Simonis et al. [48] reported both measures, but neither accuracy nor latency showed group differences between monolingual and bilingual children in their study. To draw a firm conclusion on the relationship between accuracy and RTs, one should only consider studies that report both measures within the same group.

With regards to the measured auditory attention components, selective attention and sustained attention are mostly assessed, followed by executive control. While one may argue that it is hard to disentangle these components and measure them independently, we distinguished them by considering the specific tasks used by investigators and the attention model developed by Posner and colleagues [18]. We also noticed inconsistency across publications when reporting the measured components with a given test. For example, using the auditory section of IVA + Plus, Krizman et al. [43] measured "attentional control," whereas Boerma et al. [40] assessed

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

"sustained attention." This observation is in line with Williams et al. [6] who also found variation in targeted cognitive abilities for the same task in bilingualism research.

Moreover, rarely were sufficient details about participants' bilingual background reported. This is problematic because differences in language experience (e.g., age of acquisition, proficiency; [4, 11, 64]) can affect cognitive performance. For example, more proficient bilinguals tend to exhibit a positive influence of bilingualism. Consistent with this, we observed a trend of the bilingual advantage toward adolescents relative to younger children, which can be explained by the fact that bilingual children become more proficient in their languages when growing older.

In addition, bilingualism was poorly assessed and hardly deemed a continuum. We identified two common themes: (1) most language background questionnaires fail to assess bilingualism comprehensively (i.e., including the use of, exposure to, and proficiency of each language; see [30]); (2) bilingual participants are usually recruited as a comparison group to monolinguals, without explaining inclusion criteria. This practice overlooks the nuances within bilinguals. For example, an English L2 learner from Spain is not comparable to a Spanish heritage speaker from the U.S., although both considered as Spanish-English bilinguals. Thus, variance in different bilingual communities should be considered, captured, and reported in future work.

In future studies, bilingualism should be considered as a continuum, and bilingual variables should be appropriately assessed and reported. This will facilitate our understanding of how various facets of bilingualism (e.g., age of acquisition, language proficiency, language exposure, language use, etc.) interact with auditory attention development. Moreover, further investigation should test monolinguals and bilinguals across age groups and report both accuracy

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

and RTs. Given that most tests are designed for monolingual English speakers, we also advocate for developing tests that are normed to assess both monolinguals and bilinguals. Lastly, there was little consistency regarding the targeted attention components even for the same test across publications, which may be due to the lack of a clear definition of attention in the field itself. Therefore, we need to consider further what attention along with its components refers to in (bilingualism) research and theoretical models.

References

1. Anderson JAE, Grundy JG, Grady CL, Craik FIM, Bialystok E. Bilingualism contributes to reserve and working memory efficiency: evidence from structural and functional neuroimaging. Neuropsychologia. 2021 Dec 10;163:108071. doi: 10.1016/j.neuropsychologia.2021.108071. Epub 2021 Oct 27. PMID: 34715120.
2. Bialystok E. The bilingual adaptation: how minds accommodate experience. Psychol Bull. 2017 Mar;143(3):233-62. doi: 10.1037/bul0000099. PMID: 28230411.
3. Bialystok E. Bilingualism: the good, the bad, and the indifferent. Biling (Camb Engl). 2009;12(1):3-11. doi: 10.1017/s1366728908003477.
4. Tse CS, Altarriba J. The relationship between language proficiency and attentional control in Cantonese-English bilingual children: evidence from Simon, Simon switching, and working memory tasks. Front Psychol. 2014 Sep 3;5:954. doi: 10.3389/fpsyg.2014.00954. PMID: 25232345.
5. Dunabeitia JA, Hernandez JA, Anton E, Macizo P, Estevez A, Fuentes LJ, Carreiras M. The inhibitory advantage in bilingual children revisited: myth or reality? Exp Psychol.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

2014;61(3):234-51. doi: 10.1027/1618-3169/a000243. PMID: 24217139.

6. Williams L, Parthasarathy P, Molnar M. Measures of bilingual cognition - from infancy to adolescence. J Cogn. 2021 Aug 26;4(1):45. doi: 10.5334/joc.184. PMID: 34514316.
7. Gunnerud HL, Ten Braak D, Reikeras EKL, Donolato E, Melby-Lervag M. Is bilingualism related to a cognitive advantage in children? A systematic review and meta-analysis. Psychol Bull. 2020 Dec;146(12):1059-83. doi: 10.1037/bul0000301. Epub 2020 Sep 10. PMID: 32914991.
8. Lowe CJ, Cho I, Goldsmith SF, Morton JB. The bilingual advantage in children's executive functioning is not related to language status: a meta-analytic review. Psychol Sci. 2021 Jul;32(7):1115-46. doi:10.1177/0956797621993108. Epub 2021 Jul 2. PMID: 34213379.
9. Grundy JG. The effects of bilingualism on executive functions: an updated quantitative analysis. J Cult Cogn Sci. 2020 Aug;4(2):177-99. doi: 10.1007/s41809-020-00062-5.
10. de Bruin A, Dick AS, Carreiras M. Clear theories are needed to interpret differences: perspectives on the bilingual advantage debate. Neurobiol Lang (Camb). 2021 Dec;2(4):433-51. doi: 10.1162/nol_a_00038.
11. Festman J, Czapka S, Winsler A. How many moderators does it take till we know … that too many bilingual advantage effects have died? In Kersten K, Winsler A, editors. Understanding variability in second language acquisition, bilingualism, and cognition. New York: Routledge; 2022. p. 80-127.
12. Crivello C, Kuzyk O, Rodrigues M, Friend M, Zesiger P, Poulin-Dubois D. The effects of bilingual growth on toddlers' executive function. J Exp Child Psychol. 2016 Jan;141:121-32. doi: 10.1016/j.jecp.2015.08.004. Epub 2015 Sep 22. PMID: 26402219.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

13. Kuipers JR, Thierry G. Bilingualism and increased attention to speech: evidence from event-related potentials. Brain Lang. 2015 Oct;149:27-32. doi:10.1016/j.bandl.2015.07.004. Epub 2015 Jul 13. PMID: 26185046.
14. Sebastian-Galles N, Albareda-Castellot B, Weikum WM, Werker JF. A bilingual advantage in visual language discrimination in infancy. Psychol Sci. 2012 Sep 1;23(9):994-9. doi: 10.1177/0956797612436817. Epub 2012 Jul 18. PMID: 22810164.
15. Bialystok E. Bilingualism and the development of executive function: the role of attention. Child Dev Perspect. 2015 Jun 1;9(2):117-21. doi: 10.1111/cdep.12116. PMID: 26019718.
16. Bialystok E, Craik FIM. How does bilingualism modify cognitive function? Attention to the mechanism. Psychon Bull Rev. 2022 Aug;29(4):1246-69. doi: 10.3758/s13423-022-02057-5. Epub 2022 Jan 28. PMID: 35091993.
17. Styles E. The psychology of attention. 2nd ed. London: Psychology Press; 2006.
18. Petersen SE, Posner MI. The attention system of the human brain: 20 years after. Annu Rev Neurosci. 2012;35(1):73-89. doi: 10.1146/annurev-neuro-062111-150525. Epub 2012 Apr 12. PMID: 22524787.
19. Alain C, Arnott SA, Dyson B. Varieties of auditory attention. In: Ochsner KN, Kosslin S, editors. The Oxford handbook of cognitive neuroscience, volume 1: core topics. New York: Oxford University Press; 2013. p. 215-36.
20. Moon C, Lagercrantz H, Kuhl PK. Language experienced in utero affects vowel perception after birth: a two-country study. Acta Paediatr. 2003 Feb;102(2):156-60. Doi: 10.1111/apa.12098. Epub 2013 Jan 9. PMID: 23173548.
21. Bao W, Molnar M. Auditory attention development from infancy to adolescence: a

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

systematic review protocol. OSF [Preprint]. 2021 [cited 2023 Mar 17]. Available from https://doi.org/10.17605/OSF.IO/GWKUQ.

22. Pons F, Bosch L, Lewkowicz DJ. Bilingualism modulates infants' selective attention to the mouth of a talking face. Psychol Sci. 2015 Apr;26(4):490-8. doi: 10.1177/0956797614568320. Epub 2015 Mar 12. PMID: 25767208; PMCID: PMC4398611.
23. McKenzie JE, Brennan SE. Cochrane handbook for systematic reviews of interventions. 2nd ed. Chichester (UK): John Wiley & Sons; 2019. Chapter 12, Synthesizing and presenting findings using other methods; p. 321-47.
24. Calvo A, Bialystok E. Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. Cognition. 2014 Mar;130(3):278-88. doi: 10.1016/j.cognition.2013.11.015. Epub 2013 Dec 25. PMID: 24374020.
25. Covidence. Covidence systematic review software, Veritas Health Innovation. Melbourne, Australia; 2021. Available from: http://www.covidence.org
26. Kmet LM, Lee RC, Cook LS. Standard quality assessment criteria for evaluating primary research papers from a variety of fields. Alberta Heritage Foundation for Medical Research. Alberta, Canada; 2004. Available from https://doi.org/10.7939/R37M04F16.
27. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health. 2019 Nov;22(4):153-60. Doi: 10.1136/ebmental-2019-300117. Epub 2019 Sep 28.
28. R Core Team. R: a language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2021. Available from: http://www.R-Project.org.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

29. Rohatgi A. WebPlotDigitizer. Version 4.6 [software]. 2022 Sep [cited 2023 Mar 17]. Available from: https://automeris.io/WebPlotDigitizer.

30. Krizman J, Marian V, Shook A, Skoe E, Kraus N. Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. Proc Natl Acad Sci USA. 2012 May 15;109(20):7877-81. doi:10.1073/pnas.1201575109. Epub 2012 Apr 30. PMID: 22547804.

31. Viechtbauer W. Bias and efficiency of meta-analytic variance estimators in the random-effects model. J Edu Behav Stat. 2005 Sep;30(3):261-93. doi: 10.3102/1076998603000326.

32. Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. Stat Med. 2003 Sep 15;22(17):2693-710. doi: 10.1002/sim.1482. PMID: 12939780.

33. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002 Jun 15;21(11):1539-58. doi: 10.1002/sim.1186. PMID: 12111919.

34. Ioannidis JP, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. BMJ. 2007 Nov 3;335(7626):914-6. doi: 10.1136/bmj.39343.408449.80. PMID: 17974687; PMCID: PMC2048840.

35. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw. 2010 Aug; 36(3):1-48. doi: 10.18637/jss.v036.i03.

36. Brewer MJ, Bulter A, Cooksley SL. The relative performance of AIC, AICc and BIC in the presence of unobserved heterogeneity. Methods Ecol Evol. 2016;7(6):679-92. doi: 10.1111/2041-210x.12541.

37. Baker WL, White CM, Cappelleri JC, Kluger J, Coleman CI; Health Outcomes, Policy,

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

and Economics (HOPE) Collaborative Group. Understanding heterogeneity in meta-analysis: the role of meta-regression. Int J Clin Pract. 2009 Oct;63(10):1426-34. doi: 10.1111/j.1742-1241.2009.02168.x. PMID: 19769699.

38. Spineli LM, Pandis N. Problems and pitfalls in subgroup analysis and meta-regression. Am J Orthod Dentofacial Orthop. 2020 Dec;158(6):901-904. doi: 10.1016/j.ajodo.2020.09.001. PMID: 33250104.

39. Balduzzi S, Rucker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health. 2019 Nov;22(4):153-60. doi: 10.1136/ebmental-2019-300117. Epub 2019 Sep 28. PMID: 31563865.

40. Boerma T, Leseman P, Wijnen F, Blom E. Language proficiency and sustained attention in monolingual and bilingual children with and without language impairment. Front Psychol. 2017 Jul 21;8:1241. doi: 10.3389/fpsyg.2017.01241. PMID: 28785235.

41. Kwakkel H, Droop M, Verhoeven L, Segers E. The impact of lexical skills and executive functioning on L1 and L2 phonological awareness in bilingual kindergarten. Learn Individ Differ. 2021;88:102009. doi:10.1016/j.lindif.2021.102009.

42. Karlsson LC, Soveri A, Rasanen P, Karna A, Delatte S, Lagerstrom E, Mard L, Steffansson M, Lehtonen M, Laine M. Bilingualism and performance on two widely used developmental neuropsychological test batteries. PLoS One. 2015 Apr 29;10(4):e0125867. doi:10.1371/journal.pone.0125867. PMID: 25922937.

43. Krizman J, Skoe E, Marian V, Kraus N. Bilingualism increases neural response consistency and attentional control: evidence for sensory and cognitive coupling. Brain Lang. 2014 Jan;128(1):34-40. doi:10.1016/j.bandl.2013.11.006. Epub 2014 Jan 9. PMID: 24413593.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

44. Garratt LC, Kelly TP. To what extent does bilingualism affect children's performance on the NEPSY? Child Neuropsychol. 2007;14(1):71-81. doi: 10.1080/09297040701218405. PMID: 17852119.

45. Barbu C, Gonzalez A, Gillet S, Poncelet M. Cognitive advantage in children enrolled in a second-language immersion elementary school program for one year. Psychol Belg. 2019 Oct 21;59(1):416-35. doi: 10.5334/pb.469. Erratum in: Psychol Belg. 2020 Jan 27;60(1):35-6. PMID: 31709070.

46. Nicolay AC, Poncelet M. Cognitive advantage in children enrolled in a second-language immersion elementary school program for three years. Biling (Camb Engl). 2013;16(3):597-607. doi: 10.1017/s1366728912000375.

47. Nicolay AC, Poncelet M. Cognitive benefits in children enrolled in an early bilingual immersion school: a follow-up study. Biling (Camb Engl). 2015;18(4):789-95. doi: 10.1017/s1366728914000868.

48. Simonis M, Galand B, Hiligsmann P, Szmalec A. Attentional abilities of children enrolled in immersion education in French-speaking Belgium. Int J Biling Educ Biling. 2020:1-18. doi: 10.1080/13670050.2020.1856033.

49. Zimmermann P, Fimm B. Test for attentional performance (TAP). Herzogenrath: Psytest; 1995.

50. Zimmermann P, Gondan M, Fimm B. Test of attentional performance for children (KiTAP). Herzogenrath: Psytest; 2005.

51. Cherry R. Selective auditory attention test. St. Louis: Auditec of St. Louis; 1980.

52. Sandford JA, Turner A. IVA + Plus: integrated visual and auditory continuous performance test administration manual. Richmond: Brain Train; 2004.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

53. Mahone EM, Pillion JP, Hiemenz JR. Initial development of an auditory continuous performance test for preschoolers. J Atten Disord. 2001;5(2):93-106. doi:10.1177/108705470100500203.

54. Manly T, Robertson IH, Anderson V, Nimmo-Smith I. TEA-Ch: the test of everyday attention for children manual. Bury St. Edmunds, UK: Thames Valley Test Company; 1999.

55. Korkman M, Kirk U, Kemp S. NEPSY: A developmental neuropsychological assessment. San Antonio: The Psychological Corporation; 1998.

56. Korkman M, Kirk U, Kemp S. NEPSY-II (2nd ed.). San Antonio: Harcourt Assessment; 2007.

57. Strydom L, Pottas L, Soer M, Graham MA. Effects of language experience on selective auditory attention and speech-in-noise perception among English second language learners: preliminary findings. Int J Pediatr Otorhinolaryngol. 2022 Mar;154:111061. doi: 10.1016/j.ijporl.2022.111061. Epub 2022 Feb 7. PMID: 35149369.

58. Foy JG, Mann VA. Bilingual children show advantages in nonverbal auditory executive function task. Int J Biling. 2014 Nov;18(6):717-29. doi: 10.1177/1367006912472263.

59. Blanco-Elorrieta E, Caramazza A. On the need for theoretically guided approaches to possible bilingual advantages: an evaluation of the potential loci in the language and executive control systems. Neurobiol Lang (Camb). 2021;2(4):452-63. doi: 10.1162/nol_a_00041.

60. Bialystok E, Martin MM, Viswanathan M. Bilingualism across the lifespan: the rise and fall of inhibitory control. Int J Biling. 2005 Mar;9(1):103-19. doi: 10.1177/13670069050090010701.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

61. Martin-Rhee MM, Bialystok E. The development of two types of inhibitory control in monolingual and bilingual children. Biling (Camb Engl). 2008;11(1):81-93. doi:10.1017/s1366728907003227.

62. Soares C, Grosjean F. Bilinguals in a monolingual and a bilingual speech mode: the effect on lexical access. Mem Cognit. 1984 Jul;12(4):380-6. doi: 10.3758/bf03198298. PMID: 6503701.

63. Gollan TH, Montoya RI, Fennema-Notestine C, Morris SK. Bilingualism affects picture naming but not picture classification. Mem Cognit. 2005 Oct;33(7):1220-34. doi: 10.3758/BF03193224. PMID: 16532855.

64. Sabourin L, Vinerte S. The bilingual advantage in the Stroop task: simultaneous vs. early bilinguals. Biling (Camb Engl). 2015;18(2):350-5. doi: 10.1017/s1366728914000704.

Supporting information

S1 Checklist. PRISMA 2020 checklist.

S1 Fig. Forest plot stratified by participant age groups, based on random-effects model analysis. Green lines represent accuracy studies, and blue lines represent RT studies.

S2 Fig. Contour-enhanced funnel plot. Effect size (i.e., standardized mean difference or SMD) is plotted against its standard error. Each dot represents an individual study. Grey-shaded areas indicate different *p*-value intervals.

S1 Table. Search terms used in the electronic databases under the concept of "auditory attention".

medRxiv preprint doi: https://doi.org/10.1101/2023.03.31.23286915; this version posted December 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

S2 Table. Search terms used in the electronic databases under the concept of "population".

S3 Table. Search terms used in the electronic databases under the concept of "study types and methods".

S4 Table. Random-effects meta-analytic model summary.

S5 Table. Mixed-effects meta-regression model summary, with test measure as the moderator.

S6 Table. Subgroup analysis result, stratified by test measure.

S7 Table. Mixed-effects meta-regression model summary for accuracy studies, with participant age as the moderator.

S8 Table. Mixed-effects meta-regression model summary for accuracy studies, with stimulus type as the moderator.

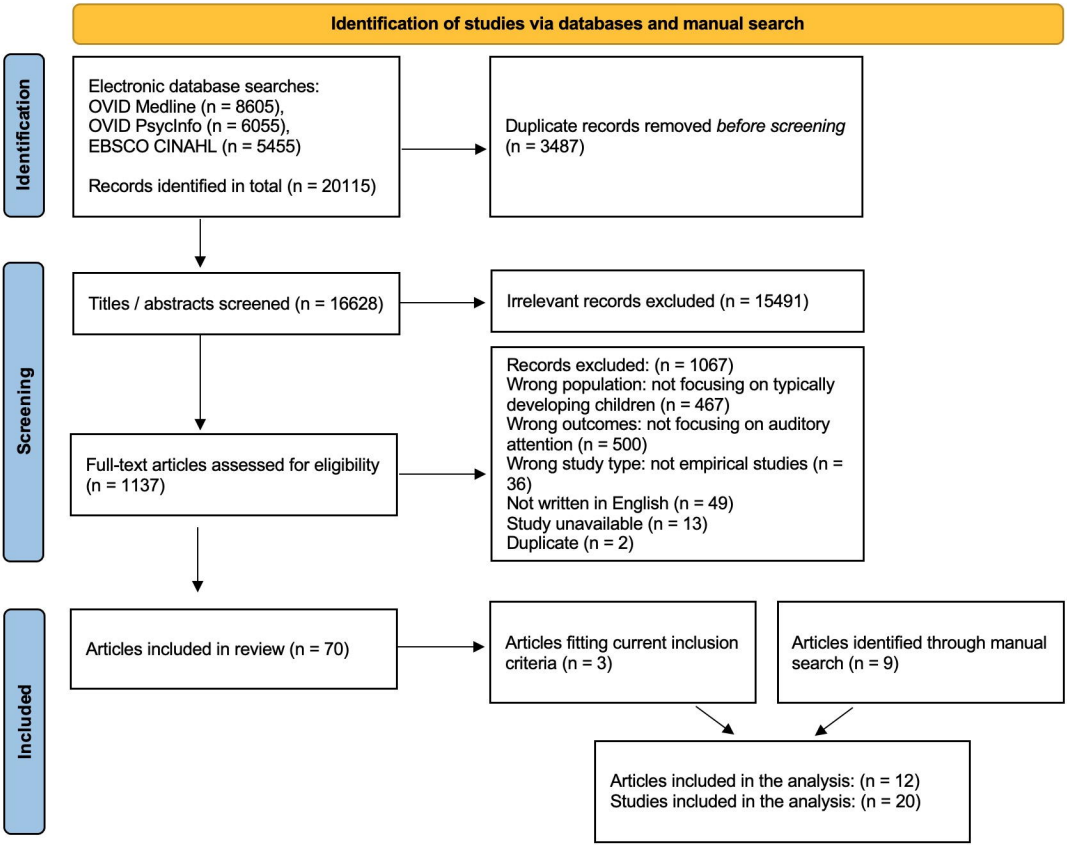
S9 Table. Mixed-effects meta-regression model summary for accuracy studies, with attention components as the moderator.

S10 Table. Mixed-effects meta-regression model summary for RT studies, with participant age as the moderator.

S11 Table. Mixed-effects meta-regression model summary for RT studies, with stimulus type as the moderator.

S12 Table. Mixed-effects meta-regression model summary for RT studies, with attention components as the moderator.

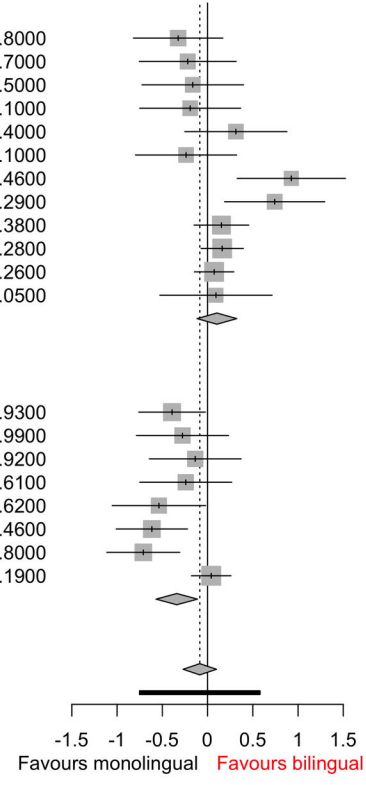
S13 Table. Egger's regression test result.


Identification of studies via databases and manual search
Identification
Electronic database searches:
OVID Medline (n = 8605),
OVID PsycInfo (n = 6055),
EBSCO CINAHL (n = 5455)
Records identified in total (n = 20115)
Duplicate records removed *before screening* (n = 3487)
Screening
Titles / abstracts screened (n = 16628)
Irrelevant records excluded (n = 15491)
Full-text articles assessed for eligibility (n = 1137)
Records excluded: (n = 1067)
Wrong population: not focusing on typically developing children (n = 467)
Wrong outcomes: not focusing on auditory attention (n = 500)
Wrong study type: not empirical studies (n = 36)
Not written in English (n = 49)
Study unavailable (n = 13)
Duplicate (n = 2)
Included
Articles included in review (n = 70)
Articles fitting current inclusion criteria (n = 3)
Articles identified through manual search (n = 9)
Articles included in the analysis: (n = 12)
Studies included in the analysis: (n = 20)

Study	Monolingual Total	Monolingual Mean	Monolingual SD	Bilingual Total	Bilingual Mean	Bilingual SD	SMD	95%-CI	Weight
Accuracy									
Boerma et al. 2017	32	2.33	0.9000	32	2.05	0.8000	-0.32	[-0.82; 0.17]	4.7%
Garratt & Kelly 2007	27	10.20	2.7000	27	9.60	2.7000	-0.22	[-0.75; 0.32]	4.4%
Karlsson et al. 2015 (1)	25	52.70	9.5000	24	51.20	8.5000	-0.16	[-0.72; 0.40]	4.2%
Karlsson et al. 2015 (2)	23	58.50	2.0000	27	58.10	2.1000	-0.19	[-0.75; 0.37]	4.2%
Karlsson et al. 2015 (3)	25	49.30	11.8000	24	52.70	9.4000	0.31	[-0.25; 0.88]	4.2%
Karlsson et al. 2015 (4)	23	66.00	4.4000	27	64.70	6.1000	-0.24	[-0.80; 0.32]	4.2%
Krizman et al. 2012	25	69.00	21.6500	23	88.00	18.4600	0.93	[0.33; 1.52]	3.9%
Krizman et al. 2014	27	69.19	21.8800	27	84.37	18.2900	0.74	[0.19; 1.29]	4.3%
Kwakkel et al. 2021	80	15.31	11.2300	89	17.22	13.3800	0.15	[-0.15; 0.46]	6.3%
Simonis et al. 2020 (1)	129	0.86	0.3400	156	0.91	0.2800	0.16	[-0.07; 0.40]	6.9%
Simonis et al. 2020 (2)	153	0.90	0.2900	173	0.92	0.2600	0.07	[-0.14; 0.29]	7.0%
Strydom et al. 2022	20	96.30	3.2600	20	96.60	3.0500	0.09	[-0.53; 0.71]	3.8%
Random effects model	589			649			0.10	[-0.12; 0.32]	58.0%
Heterogeneity: I^2 = 48%, τ^2 = 0.0404, p = 0.03									
RT									
Barbu et al. 2019	57	875.36	152.4900	59	817.20	141.9300	-0.39	[-0.76; -0.02]	5.7%
Foy & Mann 2014 (1)	30	1245.79	388.8200	30	1142.52	345.9900	-0.28	[-0.79; 0.23]	4.6%
Foy & Mann 2014 (2)	30	1400.64	344.1700	30	1344.98	455.9200	-0.14	[-0.64; 0.37]	4.6%
Foy & Mann 2014 (3)	30	833.73	381.7700	30	732.98	441.6100	-0.24	[-0.75; 0.27]	4.6%
Foy & Mann 2014 (4)	30	988.54	402.8000	30	788.11	329.6200	-0.54	[-1.05; -0.02]	4.5%
Nicolay & Poncelet 2013	51	820.63	127.2500	53	751.87	92.4600	-0.62	[-1.01; -0.22]	5.5%
Nicolay & Poncelet 2015	50	824.50	109.0000	51	752.80	90.8000	-0.71	[-1.11; -0.31]	5.4%
Simonis et al. 2020 (3)	153	671.88	193.2700	173	679.34	174.1900	0.04	[-0.18; 0.26]	7.0%
Random effects model	431			456			-0.34	[-0.57; -0.11]	42.0%
Heterogeneity: I^2 = 59%, τ^2 = 0.0476, p = 0.02									
Random effects model	**1020**			**1105**			**-0.09**	**[-0.27; 0.10]**	**100.0%**
Prediction interval								**[-0.75; 0.58]**	

Heterogeneity: I^2 = 66%, τ^2 = 0.0931, p < 0.01

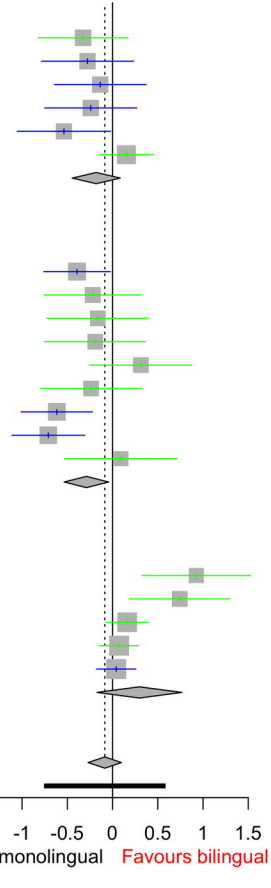
Test for subgroup differences: χ^2_1 = 10.20, df = 1 (p < 0.01)

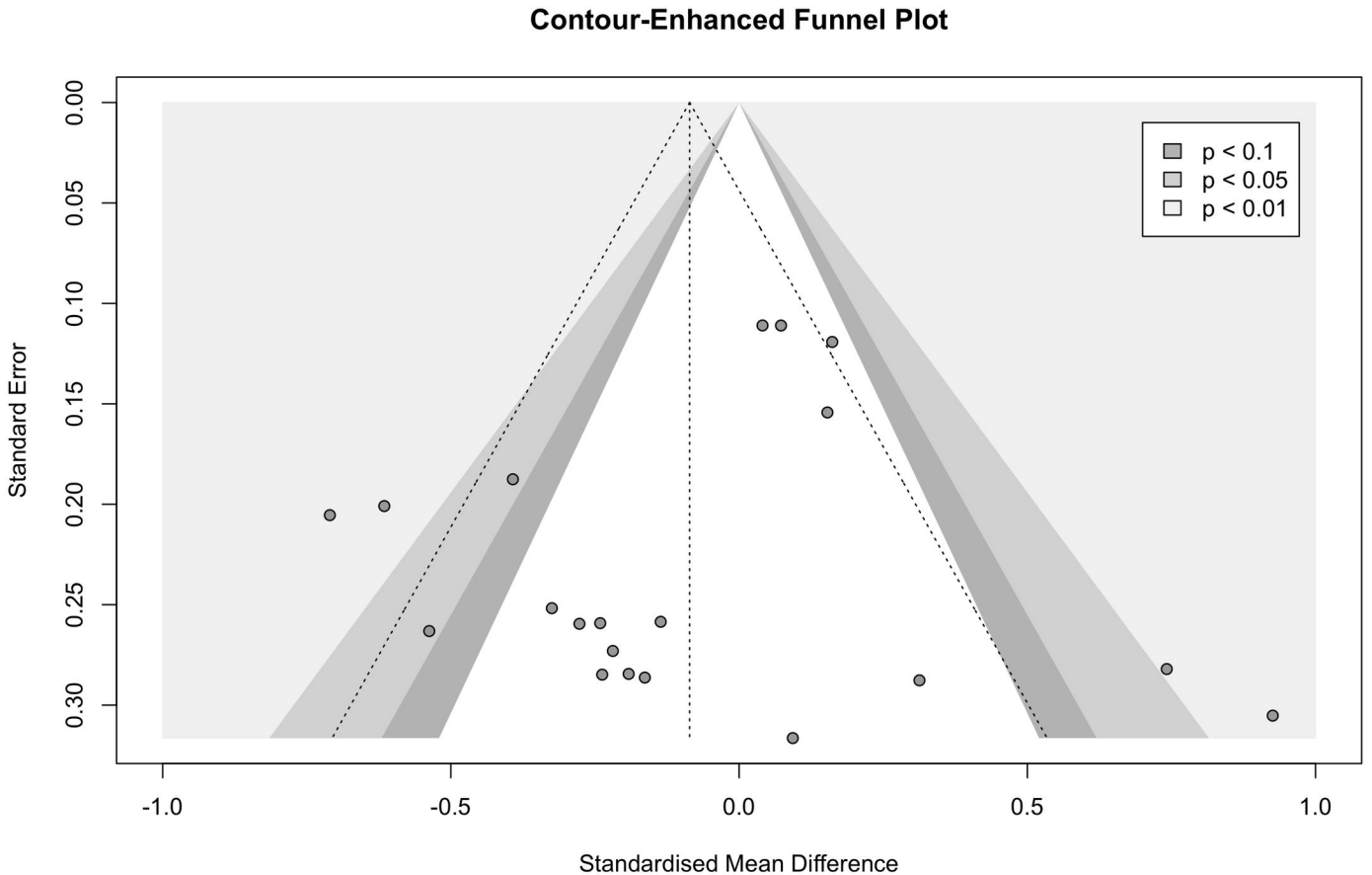


Study	Monolingual Total	Monolingual Mean	Monolingual SD	Bilingual Total	Bilingual Mean	Bilingual SD	SMD	95%-CI	Weight
Preschoolers									
Boerma et al. 2017	32	2.33	0.9000	32	2.05	0.8000	-0.32	[-0.82; 0.17]	4.7%
Foy & Mann 2014 (1)	30	1245.79	388.8200	30	1142.52	345.9900	-0.28	[-0.79; 0.23]	4.6%
Foy & Mann 2014 (2)	30	1400.64	344.1700	30	1344.98	455.9200	-0.14	[-0.64; 0.37]	4.6%
Foy & Mann 2014 (3)	30	833.73	381.7700	30	732.98	441.6100	-0.24	[-0.75; 0.27]	4.6%
Foy & Mann 2014 (4)	30	988.54	402.8000	30	788.11	329.6200	-0.54	[-1.05; -0.02]	4.5%
Kwakkel et al. 2021	80	15.31	11.2300	89	17.22	13.3800	0.15	[-0.15; 0.46]	6.3%
Random effects model	232			241			-0.18	[-0.45; 0.09]	29.3%
Heterogeneity: $I^2 = 27\%$, $\tau^2 = 0.0279$, $p = 0.23$									
Primary school-aged children									
Barbu et al. 2019	57	875.36	152.4900	59	817.20	141.9300	-0.39	[-0.76; -0.02]	5.7%
Garratt & Kelly 2007	27	10.20	2.7000	27	9.60	2.7000	-0.22	[-0.75; 0.32]	4.4%
Karlsson et al. 2015 (1)	25	52.70	9.5000	24	51.20	8.5000	-0.16	[-0.72; 0.40]	4.2%
Karlsson et al. 2015 (2)	23	58.50	2.0000	27	58.10	2.1000	-0.19	[-0.75; 0.37]	4.2%
Karlsson et al. 2015 (3)	25	49.30	11.8000	24	52.70	9.4000	0.31	[-0.25; 0.88]	4.2%
Karlsson et al. 2015 (4)	23	66.00	4.4000	27	64.70	6.1000	-0.24	[-0.80; 0.32]	4.2%
Nicolay & Poncelet 2013	51	820.63	127.2500	53	751.87	92.4600	-0.62	[-1.01; -0.22]	5.5%
Nicolay & Poncelet 2015	50	824.50	109.0000	51	752.80	90.8000	-0.71	[-1.11; -0.31]	5.4%
Strydom et al. 2022	20	96.30	3.2600	20	96.60	3.0500	0.09	[-0.53; 0.71]	3.8%
Random effects model	301			312			-0.29	[-0.53; -0.04]	41.7%
Heterogeneity: $I^2 = 39\%$, $\tau^2 = 0.0423$, $p = 0.11$									
Young adolescents									
Krizman et al. 2012	25	69.00	21.6500	23	88.00	18.4600	0.93	[0.33; 1.52]	3.9%
Krizman et al. 2014	27	69.19	21.8800	27	84.37	18.2900	0.74	[0.19; 1.29]	4.3%
Simonis et al. 2020 (1)	129	0.86	0.3400	156	0.91	0.2800	0.16	[-0.07; 0.40]	6.9%
Simonis et al. 2020 (2)	153	0.90	0.2900	173	0.92	0.2600	0.07	[-0.14; 0.29]	7.0%
Simonis et al. 2020 (3)	153	671.88	193.2700	173	679.34	174.1900	0.04	[-0.18; 0.26]	7.0%
Random effects model	487			552			0.30	[-0.17; 0.77]	29.0%
Heterogeneity: $I^2 = 68\%$, $\tau^2 = 0.0876$, $p = 0.02$									
Random effects model	**1020**			**1105**			**-0.09**	**[-0.27; 0.10]**	**100.0%**
Prediction interval								**[-0.75; 0.58]**	

Heterogeneity: $I^2 = 66\%$, $\tau^2 = 0.0931$, $p < 0.01$

Test for subgroup differences: $\chi^2_2 = 8.78$, df = 2 ($p = 0.01$)




Contour-Enhanced Funnel Plot
p < 0.1
p < 0.05
p < 0.01
Standard Error
0.00
0.05
0.10
0.15
0.20
0.25
0.30
-1.0
-0.5
0.0
0.5
1.0
Standardised Mean Difference