RAN and two languages: a meta-analysis of the RAN-reading relationship in bilingual children

Victoria Kishchak[1] · Anna Ewert[1] · Paulina Halczak[1] · Paweł Kleka[2] ·
Marcin Szczerbiński[3]

Abstract

RAN (Rapid Automatized Naming) is known to be a robust predictor of reading development in different languages. Much less is known about RAN predictive power in bilingual contexts. This is the first meta-analysis of research with bilingual children, assessing the strength of the RAN-reading relationship both within and across languages. It also explored the moderators that may affect this relationship. The search identified 38 published studies of bilingual children with 47 samples, 313 effect sizes and 5312 participants. Analyses of random-effects models with robust variance estimation revealed weak-to-moderate overall effect sizes of RAN and reading concurrently ($r=-.39$) and longitudinally ($r=-.38$). Moderator analyses of concurrent correlations revealed that RAN correlated more strongly with reading fluency ($r=-.56$) than accuracy ($r=-.38$). Alphanumeric RAN tasks (digits $r=-.39$, letters $r=-.42$) showed stronger associations with reading than non-alphanumeric RAN tasks (objects $r=-.38$, colors $r=-.25$). RAN-reading correlation was statistically significant both within and across languages. It was somewhat weaker when the two skills were measured in different languages (rL1RAN—L2 reading $=-.34$, rL2RAN—L1 reading $= -.36$) compared to when they were measured in the same language (rL1 $=-.40$, rL2 $= -.44$), though those differences failed to reach statistical significance. In addition, the type of bilingualism was found to be a potential moderator of the RAN-reading relationship longitudinally, with its magnitude being the strongest in simultaneous bilinguals. Our results suggest that, as a predictor, RAN taps into general, language-independent processes underlying reading.

Keywords RAN · Bilingualism · Reading · Literacy acquisition · Reading development

✉ Victoria Kishchak
   wikkis@amu.edu.pl

Extended author information available on the last page of the article

Introduction

Rapid automatized naming (RAN) has been considered an important predictor of reading abilities in children since Denckla and Rudel (1974, 1976). Recently, five meta-analyses of research on the RAN-reading relationship have been published (Araújo et al., 2015; Chen et al., 2021; McWeeny et al., 2022; Song et al., 2016; Swanson et al., 2003), reporting broadly consistent results (see Table 1). However, they all focused specifically on monolingual children, while much less is known on the RAN-reading relationship in bilingual children. A number of studies in bilingual reading development investigating the RAN-reading relationship concurrently and over time reported inconsistent results, finding moderate (e.g., Chung et al., 2021; Vataja et al., 2021) to weak RAN-reading correlations (e.g., Fleury & Avila, 2015). Likewise, empirical reports on cross-linguistic transfer (i.e., RAN measured in L1 as a predictor of reading success in L2) vary in findings, with some studies identifying it (McBride-Chang & Suk-Han Ho, 2005), yet others failing to do so (Li et al., 2011). With large numbers of second language speakers in early primary classrooms all over the world, it is essential to understand the nature of the RAN-reading relationship in bilingual children. A comprehensive meta-analysis of RAN and reading in bilingual children is in place to provide more insight into the nature of both within language and cross-language RAN-reading relationships. Since this is the first meta-analysis of such kind, our main aim is to examine the strength of concurrent as well as longitudinal RAN-reading relationships in bilingual participants, both within one language (L1 RAN–L1 reading and L2 RAN–L2 reading) and between languages (L1 RAN–L2 reading and L2 RAN–L1 reading).

Rapid Automatized Naming (RAN) is the speed of naming highly familiar stimuli, such as colour patches, line drawings of common objects, letters or digits. RAN assessment usually follows the setup adopted in the seminal work of Denckla and Rudel (1974, 1976): 5 different stimuli, each repeated 10 times in a pseudorandom order, are presented on an A4 sheet; the outcome is the time taken to name all 50 stimuli. However, alternative arrangements (e.g., the number of stimuli named in 30 s) are also used.

RAN has attracted a great deal of research and clinical interest, as this seemingly simple task correlates with ecologically valid outcomes, such as ADHD diagnoses (Ryan et al., 2017), arithmetic performance (Hornung et al., 2017) and, most notably, literacy skills across languages in monolingual children (e.g.,Caravolas et al., 2013; Georgiou et al., 2008; Landerl et al., 2019). Numerous studies have found slow RAN performance to be a core deficit in developmental dyslexia (Fernandes et al., 2017; Kirby et al., 2008; Wolf et al., 2002).

Detecting reading impairments in bilingual children is challenging due to the necessity of disentangling between problems caused by insufficient language proficiency and core reading impairments. Consequently, the validity of established predictors of reading impairment ought to be re-examined in the bilingual context. This applies to RAN too: are RAN measures a valid predictor for reading impairment in bilinguals? While the current study does not investigate reading

Table 1 The RAN-reading relationship in monolingual participants: A summary of the main findings

Scope of analysis	
Swanson et al. (2003)	Cross-sectional only. Alphabetic orthographies only (mostly English). Some L2 learners included. Mostly children (mean age 10 years) but some adults also included
Araújo et al. (2015)	Cross-sectional only. Only children (mean sample age below 18)
Song et al. (2016)	Cross-sectional only. Chinese native speakers only. Children only (from kindergarten till grade 6)
Chen et al. (2021)	Cross-sectional only. Alphabetic orthographies only. Children only
McWeeney et al. (2022)	Longitudinal only. English speakers (English monolinguals or English as primary language of education). Kindergarten or preschool children, where reading was subsequently measured at some point in Grades 1–5
The effect of RAN task	
Araújo et al. (2015)	RAN-reading correlations are stronger for RAN letters [.51] and numbers [.48] compared to objects [.35] and colours [.33]
Song et al. (2016)	Reading accuracy: RAN-reading correlations are stronger for alphanumeric RAN [.39] than non-alphanumeric RAN [.30]
	Reading fluency: correlations stronger for alphanumeric RAN [.53] than non-alphanumeric RAN [.37]
Chen et al. (2021)	RAN-reading correlations are stronger for RAN letters [.46] and numbers [.46] than colours [.35]. The difference with RAN objects [.36] not significant due to large confidence interval
McWeeney et al. (2022)	RAN-reading correlations are stronger for alphanumeric RAN (letters: .46; numbers: .45) than non-alphanumeric RAN (objects: .34; colours: .32)
The effect of writing system	
Araújo et al. (2015)	Reading accuracy: no difference in the strength of RAN-reading correlations between alphabetic [.42] and non-alphabetic [.42] writing systems
	Reading fluency: RAN-reading correlations are stronger for non-alphabetic [.54] than alphabetic [.48] writing systems
Consistency (alphabetic orthographies only)	
Araújo et al. (2015)	Reading accuracy: RAN-reading correlations are stronger for opaque alphabetic orthographies [.44] than transparent ones [.35]
	Reading fluency: RAN-reading correlations are stronger for opaque alphabetic orthographies [.57] than transparent ones [.48]
Chen et al. (2021)	No difference in the strength of RAN-reading correlations between opaque [.46] and transparent [.43] alphabetic orthographies

Table 1 (continued)

Reading domain	
Swanson et al. (2003)	Strength of RAN-reading correlations for word reading [.41], pseudoword reading [.38] and reading comprehension [.45] were not compared directly
Araújo et al. (2015)	Strength of RAN-reading correlations for word reading [.45], pseudoword reading [.40], text reading [.45] and reading comprehension [.39] were not compared directly, but examination of confidence intervals imply a significantly stronger correlation for word reading compared to pseudoword reading and reading comprehension
McWeeney et al. (2022)	RAN-reading correlations are higher for word [.41] than pseudoword [.33] reading. Confidence intervals imply that neither differ significantly from reading comprehension [.38]
Chen et al. (2021)	No difference in the strength of RAN-reading correlations between words [.45] and pseudowords [.45]
Reading outcome (accuracy vs. fluency)	
Araújo et al. (2015)	Transparent orthographies: RAN-reading correlations are higher for fluency [.45] than accuracy [.35]
	Opaque orthographies: RAN-reading correlations are higher for fluency [.55] than accuracy [.44]
Song et al. (2016)	RAN-reading correlations are higher for fluency [.52] than accuracy [.39]
McWeeney et al. (2022)	No difference in correlation strength between efficiency [.40] and accuracy [.37]
Chen et al. (2021)	RAN-reading correlation are higher for fluency [.46] than accuracy [.38]
Grade	
Swanson et al. (2003)	Word reading: The strength of RAN-reading correlations correlates significantly with grade [.46], i.e. correlations get stronger with grade
	Pseudoword reading: The strength of RAN-reading relationship does not correlate with grade [−.14]
	Reading comprehension: The strength of RAN-reading relationship does not correlate with grade [−.20]
Araújo et al. (2015)	Transparent orthographies: RAN-reading correlations are: [.38] for kindergarten, [.43]for grades 1–2, [.47] for grades 3–4, and [.37] for grades 5+. Those differences are not significant
	Opaque orthographies: RAN-reading correlations are:[.43] for kindergarten, [.50] for grades 1–2, [.46] for grades 3–4, and [.36] for grades 5+. Some of those differences are significant but there is no clear direction
	Reading accuracy: RAN-reading correlation for kindergarten [.41] and grades 1–2 [.47] are stronger than for grades 3–4 [.35] and 5+[.35]
	Reading fluency: RAN-reading correlations for kindergarten [.40], grades 1–2 [.45], grades 3–4 [.50], and grades 5+[.48] do not differ significantly

Table 1 (continued)

Song et al. (2016)	Reading accuracy: RAN-reading correlations tend to get stronger with grade, but this trend is falling short of significance ($\beta = 0.377$, $p = .058$)
	Reading fluency: RAN-reading correlations show no significant relationship with grade ($\beta = 0.474$, $p = .236$)
Chen et al. (2021)	RAN-reading correlations in kindergarten [.37], grades 1–2 [.41], grades 3–4 [.50] and grades 5 + [.51] do not differ significantly
Reading difficulties	
Swanson et al. (2003)	The strength of RAN-reading correlations does not differ between average/skilled readers [.43] and poor readers [.41]
Araújo et al. (2015)	The strength of RAN-reading correlations does not differ between average readers [.45] and impaired readers [.49]
McWeeney et al. (2022)	The strength of RAN-reading correlations does not differ as a function of dyslexia risk (no difference between low, medium, or high proportion of children at risk groups; correlation values not reported)

impairment, it can address this question indirectly, by examining whether the pattern of RAN-reading correlations in bilingual learners is similar to that observed in monolinguals. If it is similar (in terms of strength of correlations and their moderators)—especially in longitudinal studies—then the validity of RAN measures in the bilingual context is probably secure. The results of this meta-analysis are thus of practical importance for anyone concerned with the detection of reading impairments in bilingual children.

RAN-reading relationships: theoretical accounts

While there is broad agreement that differences in RAN performance predict differences in reading skills (both concurrently and longitudinally), the debate continues as to the mechanisms underlying their relationship. That debate is certainly warranted given that RAN tends to be a *non-redundant* predictor: it often accounts for a significant (if small) proportion of variance in reading even when other obvious predictors (such as prior reading skills or phonological awareness) are controlled for (for a review, see Landerl et al., 2022; for specific examples, see e.g. Caravolas et al., 2019; Kirby et al., 2008; Landerl et al., 2019).

Proponents of the *phonological processing account* argue that RAN tasks measure the rate of access and retrieval of phonological codes from long-term memory (Torgesen et al., 1994; Wagner et al., 1994, 1997). From that perspective, RAN performance and reading performance share variance insofar as they both rely on efficient phonological access and phonological retrieval processes (which can ultimately be explained by the quality of underlying phonological representations).

Wolf and colleagues (2000) proposed *the orthographic processing account,* which emphasises the importance of *precise temporal coordination* (i.e., timing) of

several cognitive subprocesses (visual, phonological, semantic, and articulatory). That domain general timing mechanism is crucial for fast naming (of alpha- and nonalphanumeric stimuli) as well as the efficient mapping of *letter strings* onto phonology.

The speed of processing account (Kail & Hall, 1994, 1999) suggests that skilled performance in reading as well as rapid naming rely on the rapid execution of underlying processes.

The working memory account (Amtmann et al., 2007) postulates that the RAN-reading relationship reflects a shared requirement for the maintenance of a set of names in working memory (specifically, its phonological loop component) that allows to integrate phonological and orthographic representations of words in a timely fashion.

The attentional processes account of the RAN-reading relationship (Semrud-Clikeman et al., 2000; Shao et al., 2013) emphasises response competition inherent in rapid naming, and so the role of effective response inhibition processes in fast production of correct responses.

The cascaded processing hypothesis introduced by Protopapas et al. (2013) postulates that the mechanisms of the RAN-reading relationship change over the course of development, but in advanced readers RAN is related uniquely to *reading fluency* (of connected text or of lists of words). This is because RAN and fluent reading both rely on cascaded processing, namely one where "Multiple items are processed simultaneously, so that one item may be processed phonologically while the preceding one is articulated and the next one is recognized visually" (Protopapas et al., 2013, p. 924).

The theoretical accounts listed here need not be mutually exclusive, as RAN tasks, by their very nature, must involve orchestration of several cognitive subprocesses: attentional (sustained attention, response inhibition), perceptual (feature analysis, pattern recognition), lexical (phonological, semantic) and motor (articulation), and the failure of any of these processes may compromise both fluent naming and fluent reading (McWeeny et al., 2022; Wolf et al., 2000). Indeed, the relative importance of each of these processes for naming as well as reading may vary as a function of reading proficiency (e.g., Georgiou et al., 2014).

RAN-reading relationship: review of evidence

Over the last two decades, several meta-analyses have explored RAN-reading relationships across different orthographies. Monolingual RAN studies identify it as a unique and unidirectional predictor of reading across languages with minor differences between orthographies (for a review, see Landerl et al., 2022). Four out of five meta-analyses included cross-sectional data (Araújo et al., 2015; Chen et al., 2021; Song et al., 2016; Swanson et al., 2003), whereas the one by McWeeny et al. (2022) focused on the longitudinal data only. Table 1 compares the scope and summarises the main findings of those studies.

Previous meta-analyses (Table 1) have identified a number of potential moderators of the RAN-reading relationship. RAN tasks (e.g., alphanumeric vs.

non-alphanumeric), reading domain (e.g., word reading vs. pseudoword decoding vs. reading comprehension), reading outcome measure (accuracy vs. fluency), writing system (alphabetic vs. non-alphabetic), orthographic consistency of the alphabetic writing (e.g., opaque vs. transparent), grade level, and reading level (e.g., average/skilled vs. at risk vs. dyslexic readers) have been shown as a source of variation across studies in monolingual children.

Given the above-mentioned moderators have been well-explored in previous meta-analyses, in the current study we have decided to check whether the patterns observed in research with monolingual children hold for bilingual children as well.

RAN and bilingualism

Research on L2 reading suggests considerable overlap between reading processes in L1 and L2 (e.g., August & Shanahan, 2006; Verhoeven et al., 2019). Theoretical accounts of this cross-linguistic relationship have traditionally relied on the notion of transfer (see Chung et al., 2019 for a review). Historically, transfer was a term in the contrastive analysis hypothesis as formulated by Lado (1957) and referred specifically to the transfer of linguistic forms from L1 to L2. An entirely different understanding of transfer was provided by Cummins (1979) in his common underlying proficiency (CUP) framework and, specifically, the interdependence hypothesis, which assumes the transfer of skills and conceptual knowledge between languages that takes place in the common underlying proficiency. Further research and theorising focused on establishing what is transferred in the common underlying proficiency, drawing on both above-mentioned approaches to varying degrees. For example, Koda (2008), in the transfer facilitation model, states that L1 metalinguistic awareness is transferred in learning to read in an L2 and this transfer is modulated by the typological distance between languages. Cummins' model was developed as the common underlying cognitive processes model by Geva and Ryan (1993) who, in addition to L2 proficiency, identify general cognitive processes, such as working memory and executive function, as significant predictors of L2 reading that do not merely transfer between languages, but underlie both L1 and L2 reading. The interactive transfer framework (Chung et al., 2018, 2019) acknowledges that both language-specific (e.g., orthographic processing, L1–L2 distance, L2 complexity) and domain-general factors interact, i.e., jointly affect bilingual reading outcomes. Since RAN has components that can be considered both domain-general, like general processing speed, cross-modal temporal integration and the capacity for cascaded processing, and language-specific, like retrieval of language-specific word forms that depends on the degree of automaticity that is, in turn, related to proficiency in that language, we consider the interactive transfer model to be the most appropriate theoretical framework to analyse the RAN-reading relationship in bilingual children.

In agreement with the predictions of the interactive transfer model, there is evidence that, due to its domain-general component, RAN measured in L1 might correlate with literacy skills in L1 and a typologically distant L2, regardless of the differences in orthography and writing systems (Shum et al., 2016). However, there is also evidence that bilingualism per se might affect basic processing mechanisms

involved in RAN tasks. Especially relevant for serial naming, as in RAN tasks, is the finding that bilinguals take longer to name pictures than monolinguals and this difference does not disappear with repetition (Ivanova & Costa, 2008). Bilingual lexical processing can be further affected by differences in executive function (Hartanto et al., 2019) and general processing speed (Hilchey & Klein, 2011), at least in some bilingual contexts, which might allow the bilinguals to compensate for slower lexical access. All this means that bilingual lexical processing is essentially different from monolingual lexical processing and the extent to which these differences affect the RAN-reading relationship is unknown, which means that the RAN-reading relationship in bilinguals merits systematic investigation.

Moreover, bilingualism is not a categorical variable (Luk & Bialystok, 2013), which means that bilingual outcomes are sensitive to a plethora of contextual factors which, unfortunately, are often underreported in empirical studies. Both RAN and reading outcomes might be modified by factors related to bilingualism such as language proficiency- and exposure-related factors (e.g., the amount and quality of input in either language, daily rates of L1/L2 use, length of L2 exposure). In the present meta-analysis, we define bilingualism broadly as an ability to use two languages, irrespectively of the level of proficiency in either language, which is incomparable across studies. We also decided to include the type of bilingualism as a moderator in our analysis in order to examine the role of the most basic exposure-related factors such as the order of acquisition (simultaneous vs. sequential) and country of residence (L1- or L2-speaking). *Simultaneous* bilinguals acquire both languages from birth. *Sequential* bilingualism refers to cases when the second language is acquired after the first language has been established. While sequential and simultaneous bilinguals acquire the L2 in naturalistic settings, L2 learners are explicitly taught the language in the classroom, while residing in the L1-speaking country.

Overall, in addition to RAN stimuli type (alphanumeric vs non-alphanumeric), reading outcomes (word reading, pseudoword decoding, text reading, and reading comprehension), writing systems (alphabetic vs other), the type of bilingualism is a potential moderator in the RAN-reading relationships in bilingual children we aim to explore.

The current study

The purpose of this study is to assess the strength of concurrent and longitudinal RAN-reading relationship in bilingual learners, both within one language (L1 RAN–L1 reading, and L2 RAN–L2 reading) and between languages (L1 RAN–L2 reading, and L2 RAN–L1 reading). Moreover, the study aims to explore the most likely moderators of that relationship. Our research questions are:

(1) What is the strength of the relationship between RAN and reading in bilingual children concurrently and longitudinally?
(2) Does that strength differ when measured within language (L1–L1 or L2–L2) as opposed to between languages (L1–L2 or L2–L1)?

(3) Does that strength vary as a function of (a) RAN task (alphanumeric vs non-alphanumeric); (b) reading domain (word reading; pseudoword decoding; reading comprehension; text reading); (c) Method of reading assessment (fluency vs accuracy); (d) reading group (typical, at risk, dyslexic); (e) reading level; (f) type of bilingualism; (g) writing system (alphabetic vs other); and (h) orthographic consistency of the alphabetic writing?

(4) Is the strength of the RAN-reading relationship similar or different for monolingual and bilingual children?

(5) Are the moderators of RAN-reading relationships any different in monolinguals than in bilinguals?

Method, search strategies and inclusion criteria

To identify the studies for inclusion in this meta-analysis, PsycINFO, Scopus, Web of Science databases were searched for publications from January 1970 up till September 2021. The procedure, inclusion and exclusion criteria are presented in Fig. 1. The following search terms were used: (RAN OR "rapid naming" OR "rapid automatized naming" OR "rapid serial naming" OR "naming speed") AND ("second language" OR "bilingual'') AND (reading). This search yielded a number of 334 articles, book chapters and dissertations. We also checked the reference lists of previous meta-analyses of bilingual research (Gunnerud et al., 2020; Melby-Lervåg & Lervåg, 2014) to identify studies to be included in the meta-analysis. This allowed us to add 5 additional references. We limited our search to published studies only. The first author assessed each title from the search list, obviously irrelevant articles were excluded. After removing duplicates, a list of 205 articles was compiled for the next stage of abstracts evaluation. Two screeners reviewed abstracts independently, 86 studies were agreed to be included for the full-article revision stage (see Fig. 1). At this level the inter-rater agreement was 92%, the consensus was reached by discussions. Finally, after the full articles screening procedure only 38 met the inclusion criteria and were examined in this meta-analysis (see Online Resource 1).

The following inclusion criteria were used:

(1) Study reported original empirical data;
(2) Study reported zero-order correlations between RAN and reading outcomes;
(3) Sample size was reported;
(4) Mean sample age below the age of 18;
(5) Study reported data on bilingual groups only or compared bilingual and monolingual samples;
(6) Standard RAN tasks were administered;
(7) Study was published in English.

We operationalized some broad constructs as follows. To be considered a *bilingual study*, it had to include one of the following groups: (1) equal or simultaneous bilinguals—children acquiring two languages simultaneously; (2) sequential
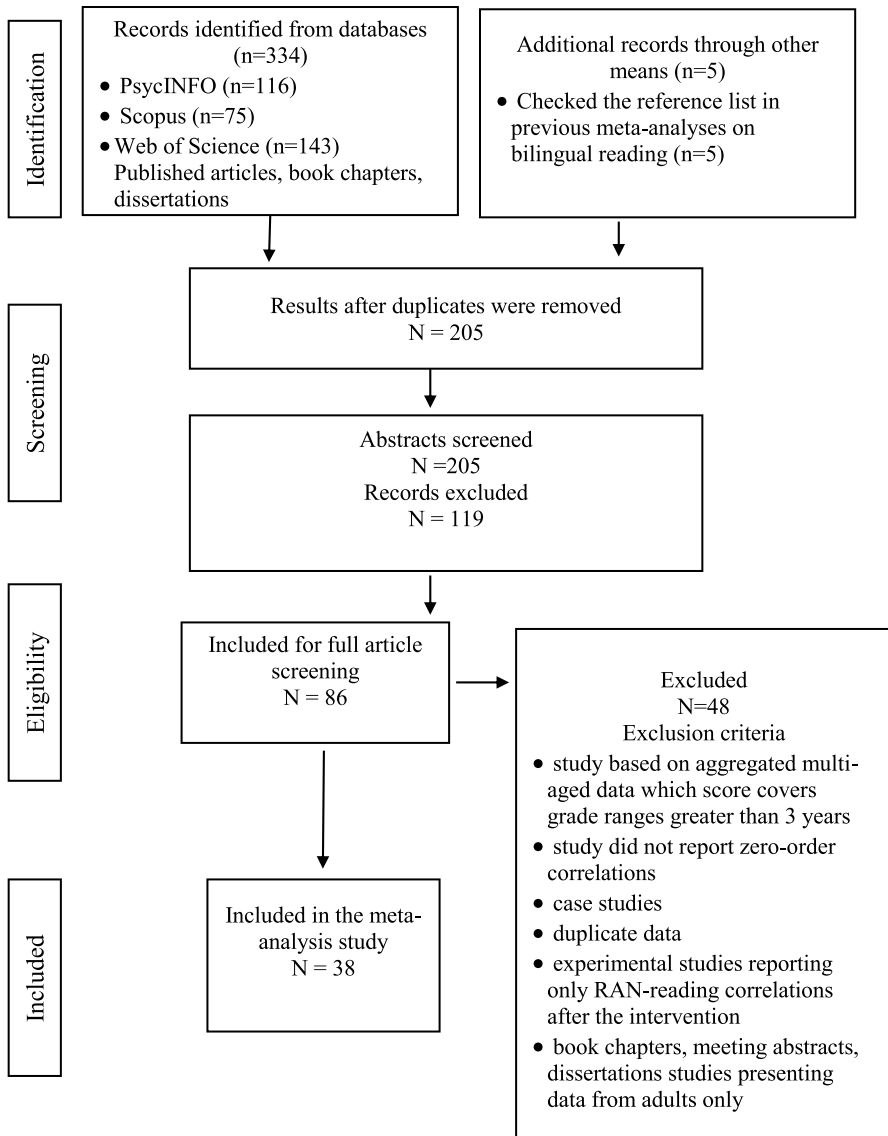
Fig. 1 Flow diagram for searching the studies in the meta-analysis (adapted from https://prisma-statement.org)

bilinguals, emergent bilinguals—children using L1 (one language) at home and L2 (another language) at school; (3) ESL learners, or second language learners—children learning a second language at school while resident in a country where their native language is spoken. The following tasks were considered *a measure of reading*: speed or accuracy of reading lists of words, pseudowords or connected text, or responding to comprehension checking questions (multiple choice, true–false, etc.).

To be considered *a RAN measure,* tasks had to include rapid naming of a set of colours, digits, letters, or objects. Discrete rapid naming tasks were not regarded as a measure of RAN in this study. After applying these criteria, we ended up with 38 studies, which include 313 effect sizes and 5312 participants from 47 independent samples.

Coding procedure and variables

In our meta-analysis, Pearson's correlation coefficient (r) was used as the effect size unit. Only zero-order correlations between RAN and reading outcomes were extracted and coded for the analyses. If subtest scores of RAN tasks were reported along with a composite score of RAN tasks, only individual scores were extracted and coded.

Depending on the scoring type of RAN tasks, the Pearson correlation coefficients reported in studies are positive or negative. When a study measured time of RAN tasks, the correlations were negative (e.g., less time spent on RAN tasks, higher reading outcome scores). Whereas, if a rate (a number of correctly named items per fixed time) was a RAN task measure, reported correlations were positive (e.g., more items are named on RAN tasks, better reading outcomes). At the first stage of coding, correlations were extracted as negative or positive as they were initially reported in the studies. Since most studies reported correlations between RAN time and reading scores, correlations concerning rate were multiplied by -1 to make the direction of all correlations consistent.

All selected studies were coded according to the following categories: (a) language pair, (b) RAN task, (c) reading domain, (d) method of reading assessment (accuracy, fluency), (e) sample features (reading level, reading group), (f) type of bilingualism, (g) writing system, and (h) orthographic consistency of the alphabetic writing. Studies that provided longitudinal data were also coded in accordance with the time interval between RAN and reading outcomes measures.

Language pair

All studies were assigned to four categories according to the language, in which RAN tasks and reading outcomes were measured. When both RAN and reading outcomes were measured in L1, we assigned studies to a within language (L1) category. Respectively, when both RAN and reading outcomes were measured in L2, studies were assigned to a group of within language (L2). If RAN was measured in L1 and reading outcomes in L2, studies were labelled as a cross language (L1–L2) group. Whereas, when RAN was measured in L2 and reading outcomes in L1, we assigned studies to a cross language (L2–L1) category.

RAN task

RAN tasks were coded into four groups: RAN letters, RAN numbers, RAN objects, and RAN colors. Some studies also reported composite scores of RAN (e.g., objects

and colors). These data were excluded from the moderator analyses but included in the overall mean effect size analysis.

Reading domain

In our meta-analysis, reading measures were divided into four domains: *word reading, pseudoword decoding, text reading, and reading comprehension.* Word reading category included measures of single real word reading tasks. Pseudoword decoding group comprises measures of pseudoword reading. Text reading category covers all tasks that measure sentence or passage reading accuracy (e.g., number of correctly read words in the sentence or text). Reading comprehension includes all the measures of passage, sentence or text reading tasks where the outcome were responses to comprehension checking questions (multiple-choice, close-ended or open-ended). When a study reported correlations for a reading task that did not fall into any of those categories, they were excluded from the moderator analysis (e.g., chain word or loanword reading tasks).

Method of reading assessment

We assigned each reading outcome into one of two categories: accuracy (e.g., number of correctly read words/pseudowords) or fluency (e.g., correctly read words/ pseudowords per unit of time).

Reading level

This variable was coded in four groups: pre-readers (kindergarten), novice readers (1–2 grade), intermediate readers (3–4 grade), proficient readers (5 grade and above). When a study provided participants' mean age, we assigned it to a respective group based on the provided information.

Reading group

Three categories were created to code studies according to the reading status of the sample: (a) typical readers—sample was described as not having reading difficulties, (b) at risk group—participants were from families with dyslexia history, (c) dyslexic—samples were identified as dyslexic/poor readers.

Type of bilingualism

Different terms are often used in bilingualism research to refer to different types of bilingualism. To make comparison across studies possible, we created three coding categories: *simultaneous bilinguals, sequential bilinguals, and L2 learners.* The category of simultaneous bilinguals includes participants who acquired two languages since birth. Sequential bilinguals' category includes participants who acquired L1 at home while their L2 was the language of instruction at school or it was the dominant language of their community. The category of L2 learners includes those children,

for whom L2 was only taught as a part of the school curriculum, while L1 was the language of instruction (as well as the home language and the dominant language of the community).

Writing systems

Two categories were used to code studies regarding their writing systems. Studies were categorised as "alphabetic" (e.g., English, French) or "other" (e.g., Chinese (logographic), Japanese (syllabic), depending on the language writing system, in which reading outcomes were measured.

Orthographic consistency

Three categories were used to code studies according to their writing systems: opaque (e.g., English), intermediate (e.g., Dutch), and shallow (e.g., Greek) (Seymour et al., 2003). Each study was assigned to a specific category based on the orthography of reading measures.

Time interval

Sixteen out of 38 selected studies reported not only concurrent but also longitudinal data on RAN-reading relationships, the latter including 109 effect sizes. However, methodological differences between studies such as frequency of testing sessions, different time-intervals between testing sessions made the classification of longitudinal data quite perplexing. Therefore, we extracted only correlations between RAN measured in the kindergarten and at the beginning of the first grade (pre-reading stage) and reading domains measured at the end of the first and second grades. Within that range two categories were created: "kindergarten—grade 1" and "kindergarten—grade 2". The first group includes all correlations between RAN measured at the pre-reading stage (kindergarten, beginning of first grade) and reading domains measured at the end of the first year. The second group includes correlations between RAN also measured at kindergarten and the beginning of the first grade, and reading domains measured in the second grade.

Statistical methods

To address each research question, three separate datasets were created (1) bilingual concurrent dataset includes correlations of RAN and reading domains measured at the same time point (e.g., RAN and reading outcome measured in the first grade); (2) bilingual longitudinal dataset contains effect sizes of RAN and reading domains measured at different time points (e.g., RAN measured in kindergarten and reading outcomes in the first grade or second grade); (3) paired concurrent dataset includes data for monolinguals and bilinguals compared directly within a study.

Reported effects were transformed from Pearson correlations to Fisher's *Z*-scores to normalise their distribution for analyses. Later, they were transformed back to Pearson correlations to facilitate the interpretation of the results.

To accommodate multiple effect sizes per study, we used correlated effects models using robust variance estimation (RVE) with the R package *robumeta* (Fisher et al., 2017; R Core Team, 2022). Using the correlated effects method, we assumed a general between-outcomes correlation of $\rho = 0.80$. A random-effects model was used for the overall model (i.e., it included all effect sizes). To assess the presence of effect size heterogeneity we computed Cochran's Q and I^2_{RVE}. The estimate of τ^2_{RVE} describes the variance of the true effect sizes, i.e., it represents true heterogeneity of effect sizes between studies, and not spurious heterogeneity due to sampling error. Related to this, I^2_{RVE} represents the percentage of variation between studies that is due to true heterogeneity rather than the sampling error. Cochran's Q-statistic (which follows Chi-square distribution) is the difference between the observed effect sizes and the fixed-effect model estimate of the effect size.

Moderators were tested in separate meta-regression models. In addition to an overall model, we examined whether effect sizes differed based on combinations of RAN-reading domains, language pairs, RAN tasks, measures of reading assessment, writing systems, reading levels, reading groups, type of bilingualism, time intervals. For these categorical moderators, we employed ANOVA-analogue models. To compare effect sizes for different groups Wald test was applied, which uses the *clubSandwich* package in R (Pustejovsky, 2022).

We included *p*-values and 95% confidence intervals for each point estimate. To detect potential publication bias we produced funnel plots and tested their asymmetry using the Trim-and-Fill analysis and the Vevea and Hedges weight-function model (Vevea & Hedges, 1995), which uses the *weightr* package (Coburn & Vevea, 2017) in R.

Results

The final sample for concurrent RAN-reading analyses in bilingual children was drawn from 27 independent samples with 182 effect sizes. An intercept-only model was calculated to assess our first research question about an overall strength of the relationship between RAN and reading in bilingual children. This model yielded an average effect size of $Z = -0.42$ [95% CI $(-.47, -.65)$, $p < .001$]. This corresponds to $r = -.39$ [95% CI $(-.44, -.35)$, see Table 2]. This may be interpreted as a weak to moderate effect (Schober et al., 2018). In addition, the degree of the between-study variability was found to be quite high $\tau^2_{RVE} = 0.028$, $Q(181) = 851.2$, $p < .001$, $I^2_{RVE} = 85\%$.

The longitudinal dataset for bilingual children included 16 independent studies with 109 effect sizes. An overall magnitude of RAN-reading relationship across two-year period was not significantly different from the concurrent effect size and produced the result of $Z = -0.40$ [95% CI $(-.46, -.34)$, $p < .001$], which corresponds to $r = -.38$ [95% CI $(-.42, -.32)$, see Table 3]. The degree of the between-study

Table 2 Concurrent moderator analyses for RAN-reading relationship in bilingual children

	n	K	Z	p	r	95% CI of Z	95% CI of r	I^2_{RVE}	τ^2_{RVE}	Q
Null model										
Intercept	27	182	−0.42	<.001	−.39	[−.47, −.36]	[−.44, −.35]	84.75	0.0284	851.2
Moderators										
Language pair	27	182						82.96	0.05	1061.95
Between language (L1–L2)		45	−0.36	<.001	−.34	[−.50, −.23]	[−.46, −.23]			
Between language (L2–L1)		30	−0.38	<.001	−.36	[−.47, −.28]	[−.44, −.28]			
Within language L1		44	−0.43	<.001	−.40	[−.50, −.37]	[−.46, −.35]			
Within language L2		63	−0.48	<.001	−.44	[−.58, −.41]	[−.52, −.38]			
RAN task	27	182						84.9	0.06	1198.77
Colours		6	−0.27	<.001	−.25	[−.33, −.21]	[−.32, −.20]			
Digits		102	−0.41	<.001	−.39	[−.49, −.33]	[−.45, −.32]			
Letters		52	−0.45	<.001	−.42	[−.56, −.34]	[−.51, −.33]			
Objects		22	−0.40	<.001	−.38	[−.43, −.37]	[−.41, −.35]			
Reading domain	26	178						82.27	0.05	1129
Word reading		109	−0.42	<.001	−.39	[−.47, −.36]	[−.44, −.34]			
Pseudoword decoding		16	−0.55	<.001	−.50	[−.70, −.40]	[−.61, −.38]			
Text reading		19	−0.62	<.001	−.54	[−.68, −.56]	[−.59, −.50]			
Reading comprehension		34	−0.25	<.001	−.25	[−.34, −.16]	[−.33, −.16]			
Measure of reading assessment	25	148						57.79	0.02	348.26
Accuracy		107	−0.40	<.001	−.38	[−.46, −.35]	[−.43, −.33]			
Fluency		41	−0.63	<.001	−.56	[−.72, −.54]	[−.61, −.50]			
Reading group	27	182						84.73	0.05	1185.43
Typical		156	−0.40	<.001	−.37	[−.46, −.34]	[−.43, −.33]			
At risk		12	−0.28	<.001	−.27	[−.31, −.26]	[−.30, −.26]			
Dyslexia		14	−0.59	<.001	−.53	[−.75, −.44]	[−.63, −.41]			
Reading level	27	182						85.45	0.06	1244.04

Table 2 (continued)

	n	K	Z	p	r	95% CI of Z	95% CI of r	I^2_{RVE}	τ^2_{RVE}	Q
Pre-readers		28	−0.35	<.001	−.34	[−.39, −.31]	[−.37, −.30]			
Novice readers		55	−0.42	<.001	−.39	[−.52, −.32]	[−.48, −.31]			
Intermediate readers		81	−0.50	<.001	−.47	[−.63, −.38]	[−.56, −.36]			
Proficient readers		18	−0.37	<.001	−.36	[−.50, −.24]	[−.46, −.24]			
Type of bilingualism	27	182						84.40	0.06	1160.07
L2 learners		86	−0.43	<.001	−.41	[−.51, −.36]	[−.47, −.34]			
Sequential bilinguals		82	−0.36	<.001	−.34	[−.44, −.27]	[−.41, −.26]			
Simultaneous bilinguals		14	−0.57	<.001	−.51	[−.65, −.49]	[−.57, −.45]			
Writing system	27	182						84.90	0.06	1198.77
Alphabetic		144	−0.41	<.001	−.39	[−.49, −.34]	[−.45, −.33]			
Other		38	−0.42	<.001	−.40	[−.49, −.36]	[−.45, −.35]			

n Number of samples; K Number of correlations; Z Fisher's Z; p p-values; r Bivariate correlation; CI Confidence interval; I^2_{RVE} Percentage of variation which is beyond the sampling error; τ^2_{RVE} Between-studies variance; Q Cochran's Q

Table 3 Longitudinal moderator analyses for RAN-reading relationship in bilingual children

	n	K	Z	p	r	95% CI of Z	95% CI of r	I^2_{RVE}	τ^2_{RVE}	Q
Null model	16	109								
Intercept			−0.40	<.001	−.38	[−.46, −.34]	[−.42, −.32]	60.42	0.01	272.86
Moderators										
Language pair	16	109						60.72	0.01	274.91
Between language (L1–L2)		31	−0.40	<.001	−.38	[−.48, −.32]	[−.44, −.31]			
Between language (L2–L1)		19	−0.34	<.001	−.33	[−.54, −.14]	[−.49, −.14]			
Within language (L1)		22	−0.42	<.001	−.40	[−.56, −.29]	[−.51, −.28]			
Within language (L2)		37	−0.40	<.001	−.38	[−.63, −.16]	[−.56, −.16]			
RAN task	16	109						56.46	0.01	248.03
Colors		4	−0.31	<.001	−.30	[−.31, −.31]	[−.30, −.30]			
Digits		27	−0.43	<.001	−.41	[−.51, −.36]	[−.47, −.34]			
Letters		17	−0.43	<.001	−.41	[−.67, −.20]	[−.58, −.19]			
Objects		61	−0.33	<.001	−.32	[−.42, −.25]	[−.39, −.24]			
Reading domain	15	107						57.94	0.01	252.02
Word reading		74	−0.38	<.001	−.36	[−.43, −.32]	[−.40, −.31]			
Pseudoword decoding		15	−0.32	.048	−.31	[−.63, −.01]	[−.56, −.01]			
Text reading		3	−0.51	<.001	−.47	[−.51, −.51]	[−.47, −.47]			
Reading comprehension		15	−0.43	.001	−.40	[−.57, −.29]	[−.52, −.28]			
Measure of reading assessment	14	94						54.55	0.01	204.61
Accuracy		54	−0.37	<.001	−.36	[−.45, −.30]	[−.42, −.29]			
Fluency		40	−0.42	<.001	−.39	[−.47, −.36]	[−.44, −.34]			
Reading group	16	109						62.24	0.01	286.02
Typical		101	−0.39	<.001	−.37	[−.46, −.33]	[−.43, −.32]			
At risk		8	−0.47	<.001	−.44	[−.47, −.47]	[−.44, −.44]			
Reading level	16	109						62.33	0.01	286.7
Pre-readers		63	−0.39	<.001	−.37	[−.46, −.32]	[−.43, −.31]			

Table 3 (continued)

	n	K	Z	p	r	95% CI of Z	95% CI of r	I^2_{RVE}	τ^2_{RVE}	Q
Novice readers		46	− 0.41	<.001	−.39	[−.55, −.27]	[−.50, −.26]			
Time interval	16	109						59.29	0.01	260.06
Kindergarten—grade 1		55	− 0.43	<.001	−.40	[−.49, −.36]	[−.46, −.35]			
Kindergarten—grade 2		54	− 0.36	<.001	−.35	[−.47, −.26]	[−.43, −.26]			
Type of bilingualism	16	109						58.47	0.01	261
L2 learners		21	− 0.42	<.001	−.40	[−.53, −.32]	[−.48, −.31]			
Sequential bilinguals		86	− 0.37	<.001	−.35	[−.46, −.28]	[−.43, −.27]			
Simultaneous bilinguals		2	− 0.57	<.001	−.52	[−.57, −.57]	[−.52, −.52]			
Writing systems	16	109						61.79	0.01	282.65
Alphabetic		88	− 0.40	<.001	−.38	[−.47, −.33]	[−.43, −.32]			
Other		21	− 0.39	<.001	−.38	[−.59, −.20]	[−.53, −.20]			

n Number of samples; K Number of correlations; Z Fisher's Z; p p-values; r Bivariate correlation; CI Confidence interval; I^2_{RVE} Percentage of variation which is beyond the sampling error; τ^2_{RVE} Between-studies variance; Q Cochran's Q

variability was lower than in the concurrent analysis but still relatively high $\tau^2_{RVE}=0.01$, $Q(108)=272.86$, $p<.001$, $I^2_{RVE}=60\%$.

To answer the fourth and fifth research questions concerning the magnitude of RAN-reading relationship for monolinguals and bilinguals, a dataset from 3 studies with 50 concurrent effect sizes for bilinguals and 22 effect sizes for monolinguals was analysed, which directly compared monolingual and bilingual groups. An average effect size for bilinguals was found to be $Z=-0.49$, $p=.067$ [95% CI $(-1.07, 0.09)$], equivalent to $r=-.45$ [95% CI $(-.79, .09)$] and $Z=-0.53$, $p<.05$ [95% CI $(-.80, -.25)$] with $r=-.48$ [95% CI $(-.66, -.25)$] for the monolingual group. There was a considerable between-study variability for both groups: $\tau^2_{RVE}=0.025$, $Q(49)=161.9$, $p<.001$, $I^2_{RVE}=69\%$ for bilinguals and $\tau^2_{RVE}=0.018$, $Q(21)=34.2$, $p<.001$, $I^2_{RVE}=54\%$ for monolinguals, which together with a small number of available effect sizes severely limited the opportunities for meaningful analysis. Therefore, we relegated the analyses of these results to Online Resource 2.

Moderator analyses

Considering the presence of such variability, moderator analyses for concurrent (see Table 2) and longitudinal (see Table 3) datasets in bilinguals were conducted to identify the sources of that variability. Wald-tests within moderators' results for the bilingual populations are presented in Table 4.

Language pair

In the concurrent analyses RAN correlations within languages were somewhat stronger than between languages, though the difference was not statistically significant. Longitudinal analyses showed quite similar (and so also not statistically different) correlations within languages and between languages.

RAN task

Our meta-analysis broadly replicated previous findings indicating that alphanumeric RAN (letters and digits) tends to predict reading better than non-alphanumeric RAN (colors and objects). This was the case for both concurrent and longitudinal analyses. However, most differences between alphanumeric and non-alphanumeric conditions failed to reach statistical significance. Only differences between RAN digits and RAN colours $F(1, 6.2) = 15.4$, $p = .007$ and RAN digits and RAN objects $F(1, 11) = 4.84$, $p = .05$ reached statistical significance, and for longitudinal data only.

Measure of reading assessment

The analyses of concurrent data revealed that RAN correlated stronger with *fluency* than with *accuracy* $F(1, 9.9) = 21.5$, $p = .001$. In the longitudinal dataset the correlation was also stronger for fluency than accuracy, yet the difference was small and not statistically significant.

Table 4 Wald-test within moderators' analysis

Moderators	Comparison	F Concurrent	df	p	F Longitudinal	df	p
Language pair	Between L2–L1 versus between L1 and L2	0.037	(1, 17)	0.849	0.765	(1, 4.7)	0.425
	Within L1 versus between L1 and L2	1.05	(1, 18)	0.319	0.219	(1, 6.2)	0.656
	Within L2 versus between L1 and L2	3.08	(1, 19)	0.095	0.003	(1, 7.6)	0.955
	Within L1 versus between L2 and L1	1.58	(1, 17)	0.226	1.24	(1, 5.6)	0.311
	Within L2 versus between L2 and L1	3.39	(1, 15)	0.087	0.346	(1, 4.7)	0.583
	Within L2 versus within L1	1.41	(1, 19)	0.249	0.094	(1, 6.5)	0.768
RAN task	Digits versus colors	5.79	(1, 11)	0.238	15.4	(1, 6.2)	0.007
	Letters versus colors	8.04	(1, 12)	0.188	2.91	(1, 2.7)	0.197
	Objects versus colors	8.93	(1, 13)	0.152	0.302	(1, 5.6)	0.604
	Letters versus digits	0.46	(1, 14)	0.505	<0.001	(1, 5.2)	0.998
	Objects versus digits	0.059	(1, 5.5)	0.817	4.84	(1, 11)	0.05
	Objects versus letters	0.951	(1, 7.2)	0.361	2.46	(1, 5.3)	0.175
Reading domain	RC versus PSW	12.7	(1, 7.5)	0.010	1.3	(1, 3.8)	0.259
	TR versus PSW	0.845	(1, 3.4)	0.418	4.94	(1, 2.4)	0.122
	WR versus PSW	3.41	(1, 5.3)	0.121	0.528	(1, 3.2)	0.416
	TR versus RC	43.5	(1, 2.7)	0.010	2.56	(1, 3.7)	0.209
	WR versus RC	9.3	(1, 5.9)	0.023	1.11	(1, 7.7)	0.299
	WR versus TR	28.4	(1, 2.0)	0.032	34.8	(1, 10)	<.001
Measure of reading assessment	Fluency versus accuracy	21.5	(1, 9.9)	0.001	1.07	(1, 5)	0.348
Writing system	Other versus alphabetic	0.067	(1, 17)	0.798	<0.001	(1, 4.1)	0.980

Table 4 (continued)

Moderators	Comparison	F Concurrent	df Concurrent	p	F Longitudinal	df Longitudinal	p
Reading level	Novice versus pre-readers	1.57	(1, 9.5)	0.240	0.091	(1, 10)	0.769
	Intermediate versus pre-readers	5.48	(1, 11)	0.004	–	–	–
	Proficient versus pre-readers	0.086	(1, 7.4)	0.778	–	–	–
	Intermediate versus novice	1.26	(1, 12)	0.282	–	–	–
	Proficient versus novice	0.306	(1, 6.2)	0.600	–	–	–
	Proficient versus intermediate	2.21	(1, 7.6)	0.177	–	–	–
Type of bilingualism	Sequential versus L2 learners	2.02	(1, 12)	0.180	1.16	(1, 8.5)	0.311
	Simultaneous versus L2 learners	3.39	(1, 1.1)	0.295	17.2	(1, 3.8)	0.016
	Simultaneous versus sequential	7.09	(1, 1.3)	0.180	28.6	(1, 8.1)	0.001
Reading group	At risk versus typical	13.54	(1, 1.2)	0.141	7.49	(1, 13)	0.0178
	Dyslexia versus typical	5.22	(1, 4)	0.084	–	–	–
	Dyslexia versus at risk	14.1	(1, 2)	0.064	–	–	–
Time interval	Kindergarten—grade 1 versus Kindergarten—grade 2	–	–	–	1.63	(1, 13)	0.224

RC Reading comprehension, *PSW* Pseudoword decoding, *WR* Word reading, *TR* Text reading

Reading domain

In the concurrent analysis RAN demonstrated significantly stronger associations with *text reading tasks F* (1, 2.7) = 43.5, *p* = .010*, pseudoword decoding F* (1, 7.5) = 12.7, *p* = .010 and *word reading* tasks *F* (1, 5.9) = 9.3, *p* = .023, compared to *reading comprehension* tasks, which showed the weakest association. Moreover, the association was stronger for *text reading* than *word reading* tasks: *F* (1, 2.0) = 28.4, *p* = .032. In the longitudinal analysis, RAN showed the strongest correlation with *text reading tasks*, followed by *reading comprehension tasks*, then *word reading* and *pseudoword decoding* tasks. The only significant difference was, again, between text reading and word reading tasks: *F* (1, 10) = 34.8, *p* < .001

Type of bilingualism

Concurrent RAN-reading relationships appeared to be the strongest for simultaneous bilinguals, followed by L2 learners, and sequential bilinguals. No significant differences were found between the three groups. Longitudinal analysis showed the same rank order of correlation strength (highest for simultaneous bilinguals, weakest for sequential ones). The Wald test revealed significantly stronger RAN-reading correlations for *simultaneous* over *sequential* bilinguals *F* (1, 8.1) = 28.6, *p* = .001 and over *L2 learners F* (1, 3.8) = 17.2, *p* = .016.

Reading group

Concurrent analyses revealed the strongest RAN-reading relationship for the *dyslexic group*, followed by *typical*, and *at risk* groups. Two of these contrasts (dyslexic vs. at risk and dyslexic vs. typical) approached significance (*p* < .10). In the longitudinal analysis, RAN correlated stronger with reading in an at *risk* group than in a *typical* group. Longitudinal data for *a dyslexic* group was not available in the dataset. The Wald test showed a significant difference between *at risk* and *typical* groups *F* (1, 13) = 7.49, *p* = .018.

Reading level

Concurrent RAN-reading correlations for intermediate and novice readers were stronger than for pre-readers and proficient readers. The difference was significant between intermediate and pre-readers group only *F* (1, 11) = 5.48, *p* = .004. In the longitudinal analysis, only two groups (pre-readers and novice readers) were included, showing similar strength of the RAN-reading relationship.

Writing system

RAN-reading correlations in *alphabetic* and *other* languages groups were similar, both concurrently and longitudinally.

Time interval

RAN-reading correlations were somewhat stronger for the shorter time interval (*kindergarten-grade 1*) than the longer one (*kindergarten-grade 2*). The difference was not significant, however.

In relation to *orthographic depth*, our bilingual dataset contained studies with data predominantly in opaque languages (e.g. English) with a few studies reporting data in intermediate orthographies (e.g., Dutch, Norwegian). Considering the dominance of opaque orthography (English) in our database, we were not able to include orthographic depth in the moderator analyses.

Publication bias

To assess whether our results were driven by publication bias, we ran the Trim-and-Fill analysis and Vevea and Hedges' likelihood ratio test. The results revealed that our meta-analyses for RAN-reading relationships were at minimal risk of publication bias. The funnel plot for the bilingual concurrent dataset (see Fig. 2) showed that no studies were missing to reach the funnel plot symmetry. The Vevea and Hedges' likelihood ratio test results $chi2(6) = 6.78$, $p = .342$ were not statistically
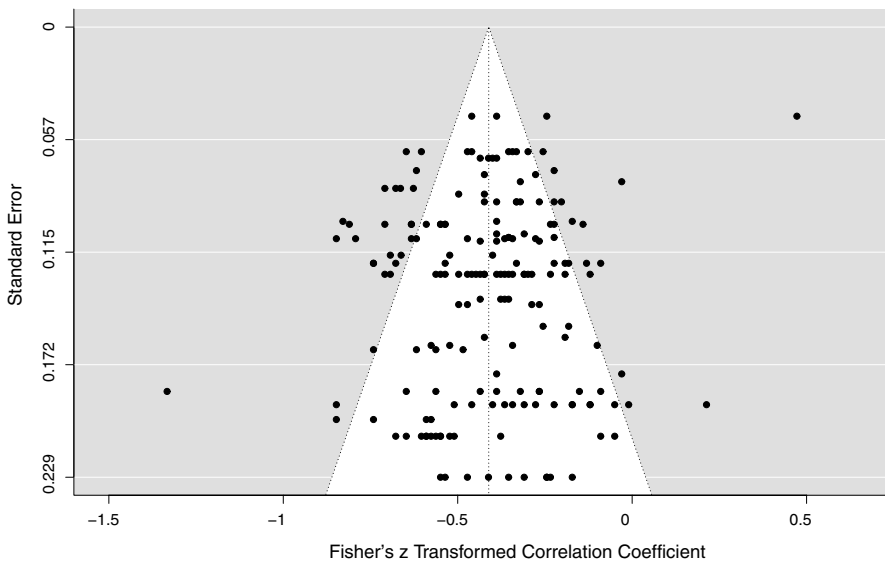
Fig. 2 Funnel plot for RAN-reading concurrent relationship. Black circles indicate observed effect sizes
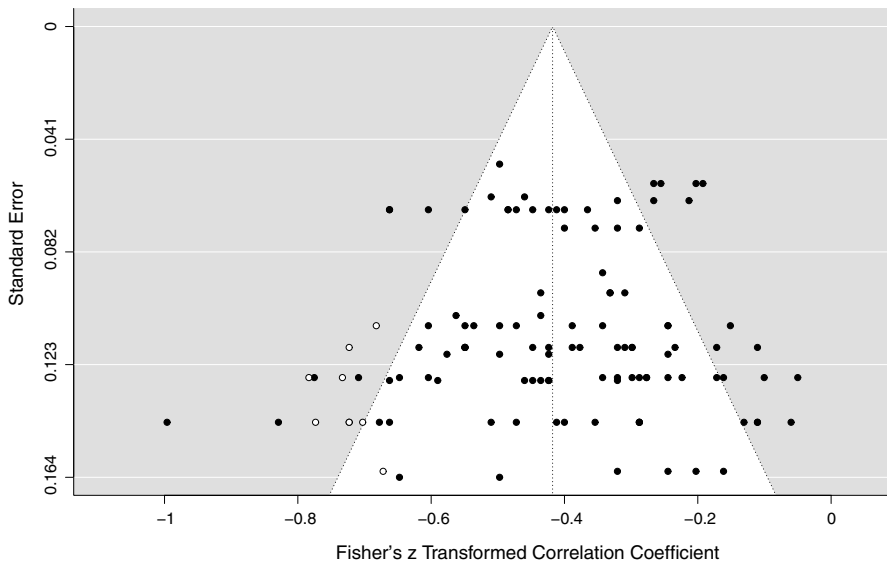
Fig. 3 Funnel plot for RAN-reading longitudinal relationship. Black circles indicate observed effect sizes, white circles indicate imputed effect sizes

significant, indicating low risk of publication bias. The funnel plot for the longitudinal data (see Fig. 3) showed that only 9 imputations were needed to reach the funnel plot symmetry. Having added them, the overall effect size stayed effectively the same [changed from $Z=-0.3968$ ($r=-.38$) to $Z=-0.4184$ ($r=-.40$)]. The Vevea and Hedges' likelihood ratio test results $chi2(5)=2.13$, $p=.831$ were not statistically significant either, indicating low risk of publication bias.

Discussion

The present study is the first meta-analysis of the RAN-reading relationship in bilingual children, thus filling in the important gap in previous research. It adds to previous research by reporting both concurrent and longitudinal RAN-reading relationships.

Our *first research question* was to determine the strength of concurrent and longitudinal RAN-reading relationship in bilingual children. Our findings revealed the average correlation between RAN and reading to be $r=-.39$, which is comparable to those reported in previous meta-analyses in monolingual children ($r=.43$: Araújo et al., 2015; $r=.44$: Chen et al., 2021; see Table 1). The magnitude of longitudinal RAN-reading association was $r=-.38$, that is in line with previous longitudinal meta-analyses results in English-speaking monolinguals ($r=-.38$: McWeeny et al., 2022) in the English language. Our findings indicate that RAN tasks can predict reading ability in bilingual learners just as effectively as in monolinguals.

Our *second research question* was to compare the strength of RAN-reading relationships measured within language and between languages. To our knowledge, this study is the first systematic review to explore this. We found that the RAN-reading correlations were comparable whether measured within or between languages. The same was the case whether those relationships were analysed concurrently and longitudinally. Thus, RAN tasks have similar potential to predict individual differences in reading not only within language but also across languages. Within the interactive transfer model (Chung et al., 2019), we can explain the RAN-reading relationship across languages, irrespective of their typological difference, by interaction of both domain-general and language-specific components of RAN.

Considering between-study variability, our *third research question* aimed to analyse moderators to gain some insights into RAN and reading relationships in bilingual children. Our analysis broadly replicated previous meta-analyses (Araújo et al., 2015; Chen et al., 2021) insofar as *alphanumeric* RAN tasks showed stronger correlations with reading than *non-alphanumeric* RAN tasks. However, in our analysis most of those differences were fairly small and fell short of statistical significance. In addition, RAN correlated significantly more strongly with *fluency* than with *accuracy* measures, though in the concurrent analysis only. This finding is in line with previous reports (Araújo et al., 2015; Song et al., 2016—though note McWeeney et al. (2022), where no significant difference was detected) and support the claim that the RAN-reading correlation can be partly explained by the shared processing speed requirement (Araújo et al., 2015). We also examined whether *reading domain* (the type of reading tasks) had a moderator effect on the RAN-reading relationship. Here, the findings are harder to interpret as they were not fully consistent across concurrent and longitudinal analyses. In the concurrent analysis, text reading emerged as the strongest correlate of RAN, while reading comprehension clearly the weakest. This pattern broadly replicated Araújo et al. (2015) analyses (where the differences were smaller though). In the longitudinal analysis it was, once more, the *text reading* that emerged as the strongest correlate, while pseudoword reading was the weakest. The same moderator analysis allowed us to compare *word* and *pseudoword reading*—a contrast analysed, with mixed results, in several previous meta-analyses (Araújo et. al., 2015; McWeeney et al., 2022: significant difference in favour of real words; Chen et al, 2021: no difference whatsoever). Our results showed no clear outcome: in the concurrent analysis, correlations were stronger for pseudoword decoding than real word reading, while reverse was the case in the longitudinal analysis, with neither difference being statistically significant. Thus, the moderating role of the reading domain still requires further clarification.

A major contribution of the current study is the analysis of *the type of bilingualism* as a moderator of the RAN-reading relationship, as no previous meta-analytic review has explored the role of bilingualism-related factors. Our analyses produced some (albeit rather weak) evidence that the type of bilingualism may be a moderator of the RAN-reading relationship. While RAN and reading correlated significantly in all three bilingual groups, the correlations were the strongest for the simultaneous bilinguals, and weakest for the sequential bilinguals, with L2 learners falling in between. This pattern was observed in concurrent and longitudinal analyses alike. However, none of those differences were significant in the concurrent analyses.

Although the longitudinal analysis revealed a significant difference between the simultaneous group and two other groups (L2 learners and sequential), we have to treat this result with caution as only two effect sizes from one study were available for the simultaneous group.

We also found (both concurrently and longitudinally) the mean effect size for the dyslexic group to be larger relative to the typical readers and at risk groups. Yet, the differences between groups were not significant concurrently. Only longitudinal comparison between at risk and typical groups reached significance ($p = .018$). However, an unproportionally smaller number of effect sizes for at risk and dyslexic groups drawn from a few studies compared to typical readers does not allow us to make strong conclusions as to the moderator effect of bilinguals' reading proficiency.

Exploration of the *reading level* moderator revealed no significant differences between groups, except for intermediate and pre-readers groups. Also, we did not find any difference between the parameters of the writing system, and time interval (for longitudinal data).

Our *fourth and fifth research questions* were to compare the strength of RAN-reading correlations for monolingual and bilingual participants directly, i.e., holding the methods of assessment constant. Unfortunately, a small number of such direct comparison studies did not allow moderator analyses. Though we computed an average effect size for the monolingual group ($r = -.48$, $p = .014$, $n = 22$) and for the bilingual group ($r = -.45$, $p = .067$, $n = 50$) (the latter effect failing to reach significance), we could not answer our research questions. One recommendation for future research, which follows from our analysis, is simply that more studies comparing mono- and bilingual participants are needed.

By and large, not only does RAN predict reading equally well both within and across languages in bilingual children, but also the effects of moderators studied in previous meta-analyses in monolingual children have been found to be similar in bilingual children. Thus, our findings regarding the moderators' effects are broadly consistent with the findings from previous meta-analyses (e.g., Araújo et al, 2015). These findings provide support for the validity of RAN measures in the assessment of reading difficulties in the bilingual context.

Limitations

Some limitations of this study should be noted. First, we limited our meta-analysis only to published studies, which could potentially reflect the publication bias. Yet, statistical methods aimed at identifying one revealed that our study was at a low level of the publication bias risk. Second, the description of bilingualism (e.g., languages spoken at home, age of second language acquisition, etc.) varied considerably across studies. The information on language status and proficiency of the participants was often incomplete and presented inconsistently, making classification (into bilingualism types) difficult. While language status could be deduced from the study description, insufficient information about the participants' language proficiency did not allow any meaningful comparisons across studies. Moreover, the majority

of the studies in our dataset dealt with either sequential bilinguals or L2 learners, with simultaneous bilinguals being grossly underrepresented. A low number of the reported effect sizes might skew the interpretation of the results, thus pointing to a need for further research into the effect of the different bilingual contexts on the RAN-reading relationship. A small number of samples and reported effect sizes also influenced the analyses of the categories of reading level, writing system, language pairs, and time interval.

Another issue of this meta-analysis is methodological differences between studies. Some studies focusing specifically on RAN reported a larger number of effect sizes with detailed description of the methodology, RAN tasks and reading outcomes, whereas others reported RAN effect sizes only as a part of an assessment battery providing more general information on how RAN was measured. Therefore, while some of our results can be considered preliminary, they clearly identify areas where there is a need for further research.

Conclusion

This first meta-analysis of the RAN-reading relationship in bilingual children builds on the body of research that already exists in monolingual children. Our findings demonstrate that, in bilingual children, significant RAN-reading relationships (of weak to moderate strength) exist, both concurrently and longitudinally. The strength of those relationships is comparable to those observed in monolingual participants. Not only was the strength of the RAN-reading relationship robust within one language, but it also proved to be a strong predictor between languages, demonstrating a cross-language effect. We suggest this can be accounted for if we assume that RAN performance relies on some universal, domain-general processes, which also underlie reading. Our analysis also suggests that the type of bilingualism may be a significant moderator of the RAN-reading relationship (at least longitudinally), though here the relevant data are insufficient to reach a firm conclusion. We also identified other moderators the role of which remains unclear, suggesting areas for further research.

Declarations

Conflict of interest The authors have no conflict of interest to declare that are relevant to the content of this article.

References

Amtmann, D., Abbott, R. D., & Berninger, V. W. (2007). Mixture growth models of RAN and RAS row by row: Insight into the reading system at work over time. *Reading and Writing, 20*(8), 785–813. https://doi.org/10.1007/s11145-006-9041-y

Araújo, S., Reis, A., Petersson, K. M., & Faísca, L. (2015). Rapid automatized naming and reading performance: A meta-analysis. *Journal of Educational Psychology, 107*(3), 868–883. https://doi.org/10.1037/edu0000006

August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth.* Lawrence Erlbaum.

Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science, 24*(8), 1398–1407. https://doi.org/10.1177/0956797612473122

Caravolas, M., Lervåg, A., Mikulajová, M., Defior, S., Seidlová-Málková, G., & Hulme, C. (2019). A cross-linguistic, longitudinal study of the foundations of decoding and reading comprehension ability. *Scientific Studies of Reading, 23*(5), 386–402. https://doi.org/10.1080/10888438.2019.1580284

Chen, Y.-J.I., Thompson, C. G., Xu, Z., Irey, R. C., & Georgiou, G. K. (2021). Rapid automatized naming and spelling performance in alphabetic languages: A meta-analysis. *Reading and Writing, 34*(10), 2559–2580. https://doi.org/10.1007/s11145-021-10160-7

Chung, S. C., Chen, X., & Deacon, S. H. (2018). The relationship between orthographic processing and spelling in Grade 1 French immersion children. *Journal of Research in Reading, 41*(2), 290–311. https://doi.org/10.1111/1467-9817.12104

Chung, S. C., Chen, X., & Geva, E. (2019). Deconstructing and reconstructing cross-language transfer in bilingual reading development: An interactive framework. *Journal of Neurolinguistics, 50*, 149–161. https://doi.org/10.1016/j.jneuroling.2018.01.003

Chung, S. C., Geva, E., Chen, X., & Deacon, S. H. (2021). Do we "laugh" or "la8gh"? Early print knowledge and its relation to learning to read in English and French. *Scientific Studies of Reading, 25*(6), 519–533. https://doi.org/10.1080/10888438.2020.1863970

Coburn, K. M., & Vevea, J. L. (2017). *weightr: Estimating weight-function models for publication bias. R package version 1.1.2.* https://CRAN.Rproject.org/package=weightr.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*, 222–251. https://doi.org/10.3102/00346543049002222

Denckla, M. B., & Rudel, R. (1974). Rapid "automatized" naming of pictured objects, colors, letters and numbers by normal children. *Cortex, 10*(2), 186–202. https://doi.org/10.1016/S0010-9452(74)80009-2

Denckla, M. B., & Rudel, R. G. (1976). Rapid 'automatized' naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia, 14*(4), 471–479. https://doi.org/10.1016/0028-3932(76)90075-0

Fernandes, T., Araújo, S., Sucena, A., Reis, A., & Castro, S. L. (2017). The 1-min screening test for reading problems in college students: Psychometric properties of the 1-min TIL. *Dyslexia, 23*(1), 66–87. https://doi.org/10.1002/dys.1548

Fisher, Z., Tipton, E., & Zhipeng, H. (2017). *Robumeta: Robust variance meta-regression. R package version 2.0.* Retrieved from https://cran.rproject.org/web/packages/robumeta/index.html.

Fleury, F. O., & de Avila, C. R. B. (2015). Rapid naming, phonological memory and reading fluency in Brazilian bilingual students. *CoDAS, 27*(1), 65–72. https://doi.org/10.1590/2317-1782/20152014091

Georgiou, G. K., Papadopoulos, T. C., & Kaizer, E. L. (2014). Different RAN components relate to reading at different points in time. *Reading and Writing, 27*(8), 1379–1394. https://doi.org/10.1007/s11145-014-9496-1

Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *Journal of Educational Psychology, 100*(3), 566–580. https://doi.org/10.1037/0022-0663.100.3.566

Geva, E., & Ryan, E. B. (1993). Linguistic and cognitive correlates of academic skills in first and second languages. *Language Learning, 43*(1), 5–42. https://doi.org/10.1111/j.1467-1770.1993.tb00171.x

Gunnerud, H. L., ten Braak, D., Reikerås, E. K. L., Donolato, E., & Melby-Lervåg, M. (2020). Is bilingualism related to a cognitive advantage in children? A systematic review and meta-analysis. *Psychological Bulletin, 146*(12), 1059–1083. https://doi.org/10.1037/bul0000301

Hartanto, A., Toh, W. X., & Yang, H. (2019). Bilingualism narrows socioeconomic disparities in executive functions and self-regulatory behaviors during early childhood: Evidence from the early childhood longitudinal study. *Child Development, 90*(4), 1215–1235. https://doi.org/10.1111/cdev.13032

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review, 18*, 625–658. https://doi.org/10.3758/s13423-011-0116-7

Hornung, C., Martin, R., & Fayol, M. (2017). General and specific contributions of RAN to reading and arithmetic fluency in first graders: A longitudinal latent variable approach. *Frontiers in Psychology, 8*, 1746. https://doi.org/10.3389/fpsyg.2017.01746

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

Kail, R., & Hall, L. K. (1994). Processing speed, naming speed, and reading. *Developmental Psychology, 30*(6), 949–954. https://doi.org/10.1037/0012-1649.30.6.949

Kail, R., & Hall, L. K. (1999). Sources of developmental change in children's word-problem performance. *Journal of Educational Psychology, 91*(4), 660–668. https://doi.org/10.1037/0022-0663.91.4.660

Kirby, J. R., Desrochers, A., Roth, L., & Lai, S. S. V. (2008). Longitudinal predictors of word reading development. *Canadian Psychology/psychologie Canadienne, 49*(2), 103–110. https://doi.org/10.1037/0708-5591.49.2.103

Koda, K. (2008). Impacts of prior literacy experience on second-language learning to read. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first- and second-language literacy development* (pp. 68–96). Routledge.

Lado, R. (1957). *Linguistics across cultures*. University of Michigan Press.

Landerl, K., Castles, A., & Parrila, R. (2022). Cognitive precursors of reading: A cross-linguistic perspective. *Scientific Studies of Reading, 26*(2), 111–124. https://doi.org/10.1080/10888438.2021.1983820

Landerl, K., Freudenthaler, H. H., Heene, M., De Jong, P. F., Desrochers, A., Manolitsis, G., Parrila, R., & Georgiou, G. K. (2019). Phonological awareness and rapid automatized naming as longitudinal predictors of reading in five alphabetic orthographies with varying degrees of consistency. *Scientific Studies of Reading, 23*(3), 220–234. https://doi.org/10.1080/10888438.2018.1510936

Li, M., Kirby, J., & Georgiou, G. K. (2011). Rapid naming speed components and reading comprehension in bilingual children: RAN components and reading across languages. *Journal of Research in Reading, 34*(1), 6–22. https://doi.org/10.1111/j.1467-9817.2010.01476.x

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

McBride-Chang, C., & Suk-Han Ho, C. (2005). Predictors of beginning reading in Chinese and English: A 2-year longitudinal study of Chinese kindergartners. *Scientific Studies of Reading, 9*(2), 117–144. https://doi.org/10.1207/s1532799xssr0902_2

McWeeny, S., Choi, S., Choe, J., LaTourrette, A., Roberts, M. Y., & Norton, E. S. (2022). Rapid automatized naming (RAN) as a kindergarten predictor of future reading in English: A systematic review and meta-analysis. *Reading Research Quarterly, 57*(4), 1187–1211. https://doi.org/10.1002/rrq.467

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first- and second-language learners. *Psychological Bulletin, 140*(2), 409–433. https://doi.org/10.1037/a0033890

Protopapas, A., Altani, A., & Georgiou, G. K. (2013). Development of serial processing in reading and rapid naming. *Journal of Experimental Child Psychology, 116*, 914–929. https://doi.org/10.1016/j.jecp.2013.08.004

Pustejovsky, J. (2022). *ClubSandwich: cluster-robust (sandwich) variance estimators with small-sample corrections. R package version 0.5.5.* https://CRAN.R-project.org/package=clubSandwich.

R Core Team. (2022). *R: A language and environment for statistical computing (version 4.1.3). R Foundation for Statistical Computing*. Available from https://www.R-project.org.

Ryan, M., Jacobson, L. A., Hague, C., Bellows, A., Denckla, M. B., & Mahone, E. M. (2017). Rapid automatized naming (RAN) in children with ADHD: An ex-Gaussian analysis. *Child Neuropsychology, 23*(5), 571–587. https://doi.org/10.1080/09297049.2016.1172560

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia, 126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Semrud-Clikeman, M., Guy, K., Griffin, J. D., & Hynd, G. W. (2000). Rapid naming deficits in children and adolescents with reading disabilities and attention deficit hyperactivity disorder. *Brain and Language, 74*(1), 70–83. https://doi.org/10.1006/brln.2000.2337

Seymour, P. H. K., Aro, M., Erskine, J. M., Collaboration with COST Action A8 network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*(2), 143–174. https://doi.org/10.1348/000712603321661859

Shao, Z., Meyer, A. S., & Roelofs, A. (2013). Selective and nonselective inhibition of competitors in picture naming. *Memory & Cognition, 41*(8), 1200–1211. https://doi.org/10.3758/s13421-013-0332-7

Shum, K. K., Ho, C. S., Siegel, L. S., Au, T., & K. (2016). First-language longitudinal predictors of second-language literacy in young L2 learners. *Reading Research Quarterly, 51*(3), 323–344. https://doi.org/10.1002/rrq.139

Song, S., Georgiou, G. K., Su, M., & Hua, S. (2016). How well do phonological awareness and rapid automatized naming correlate with Chinese reading accuracy and fluency? A meta-analysis. *Scientific Studies of Reading, 20*(2), 99–123. https://doi.org/10.1080/10888438.2015.1088543

Swanson, H. L., Trainin, G., Necoechea, D. M., & Hammill, D. D. (2003). Rapid naming, phonological awareness, and reading: A meta-analysis of the correlation evidence. *Review of Educational Research, 73*(4), 407–440. https://doi.org/10.3102/00346543073004407

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities, 27*(5), 276–286. https://doi.org/10.1177/002221949402700503

Vataja, P., Lerkkanen, M.-K., Aro, M., Westerholm, J., Risberg, A.-K., & Salmi, P. (2021). The predictors of literacy skills among monolingual and bilingual Finnish-Swedish children during first grade. *Scandinavian Journal of Educational Research*. https://doi.org/10.1080/00313831.2021.1942191

Verhoeven, L., Perfetti, C., & Pugh, K. (2019). Cross-linguistic perspectives on second language reading. *Journal of Neurolinguistics, 50*, 1–6. https://doi.org/10.1016/j.jneuroling.2019.02.001

Vevea, J. L., & Hedges, L. V. (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika, 60*(3), 419–435. https://doi.org/10.1007/BF02294384

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology, 30*(1), 73–87. https://doi.org/10.1037/0012-1649.30.1.73

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., Donahue, J., & Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology, 33*(3), 468–479. https://doi.org/10.1037/0012-1649.33.3.468

Wolf, M., Bowers, P., & Biddle, K. (2000). Naming-speed processes, timing, and reading: A conceptual review. *Journal of Learning Disabilities, 33*(4), 387–407. https://doi.org/10.1177/002221940003300409

Wolf, M., O'Rourke, A. G., Gidney, C., Lovett, M., Cirino, P., & Morris, R. (2002). The second deficit: An investigation of the independence of phonological and naming-speed deficits in developmental dyslexia. *Reading and Writing, 15*(1/2), 43–72. https://doi.org/10.1023/A:1013816320290

Authors and Affiliations

Victoria Kishchak[1] · **Anna Ewert[1]** · **Paulina Halczak[1]** · **Paweł Kleka[2]** · **Marcin Szczerbiński[3]**

Anna Ewert
aewert@amu.edu.pl

Paulina Halczak
paudzi1@amu.edu.pl

Paweł Kleka
pawel.kleka@amu.edu.pl

Marcin Szczerbiński
m.szczerbinski@ucc.ie

[1]	Faculty of English, Adam Mickiewicz University, Poznań, Poland

[2]	Faculty of Psychology and Cognitive Studies, Adam Mickiewicz University, Poznań, Poland

[3]	School of Applied Psychology, University College Cork, Cork, Ireland