

Rapid automatized naming (RAN) as a kindergarten predictor of future reading: A systematic review and meta-analysis

Journal:	<i>Reading Research Quarterly</i>
Manuscript ID	Draft
Manuscript Type:	Article
Keywords:	Theoretical perspectives, Policy, Fluency, Assessment, Struggling learners
Level of Learners:	1-Early childhood, 2-Childhood
Theoretical Perspectives:	
Areas of Expertise:	At-Risk Factors, Automaticity, Cognitive Processes, Dyslexia, Early Literacy
Method/Analysis:	Meta-Analysis
Abstract:	<p>Rapid automatized naming (RAN) tasks have been shown to be a strong correlate of reading abilities. RAN also predicts future reading across different ages, ability levels, and languages, and is often used in literacy screening. Thus, understanding the specific relations between early RAN and later reading difficulties is important. In this systematic review and meta-analysis (with N = 68 samples; k = 373 effect sizes; n = 10,513 participants), we test the extent to which measures of RAN assessed before grade school predict later reading performance in English-speaking children. We also test whether characteristics of the RAN tasks, reading measures, or sample demographics moderate this relationship. We found that kindergarten/preschool RAN is correlated with grade-school reading at $r = -.38$, similar in magnitude to previous meta-analyses that included various ages and languages. We found that alphanumeric RAN tasks are particularly strongly related to future reading, as compared with non-alphanumeric tasks ($p = .01$) but that other features of the RAN task, such as the number of items, do not alter its predictive significance. RAN predicts all types of reading measures, but more strongly predicts real word than nonword reading ($p < .001$). These results support a shared cognitive resource model in which the similarity between RAN and reading tasks accounts for their correlation. We provide practical guidelines based on these data for early screening for reading difficulties and dyslexia.</p>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Introduction

Reading is a complex process that requires the automatic integration of multiple cognitive and linguistic abilities. Reading-related skills such as rapid automatized naming (RAN), phonological awareness, and letter knowledge can all be measured at the pre-reading stage and predict later reading ability (Byrne et al., 1997; Pennington & Lefly, 2001; Scarborough, 1998; Schatschneider et al., 2004). However, it is currently a major challenge to accurately identify reading difficulties early in reading development, when intervention is likely more effective (Al Otaiba et al., 2014; Blachman et al., 2014; Cavanaugh et al., 2004; Lovett et al., 2017; Torgesen, 2004; Vellutino et al., 1998). Optimizing screening batteries that allow early identification of reading problems at the outset of schooling, and therefore earlier intervention, is critical to optimizing long-term outcomes for children with reading difficulties (Connor et al., 2014).

Numerous studies have examined pre-school and kindergarten-age predictors of later reading ability and how various factors can modify the relationship between predictors and reading outcomes (e.g., Hjetland et al., 2017). Across studies, the measures that are most commonly identified as strong predictors of later reading in English include phonological awareness, RAN, letter name and sound knowledge, and language ability (for reviews, see National Early Literacy Panel, 2008; Ozernov-Palchik & Gaab, 2016). Though RAN shares some processes with these other predictors, it has consistently been shown to uniquely relate to reading, beyond the contribution of phonological awareness (Manis et al., 2000; Wolf & Bowers, 1999), and beyond similar measures of general processing speed and single (discrete) item naming (Kirby et al., 2003; Logan et al., 2011). Further, whereas some measures such as letter knowledge are only predictive of reading for a short interval until they are mastered (Paris, 2005), RAN retains its concurrent and predictive relation with reading over time (Wagner et al.,

RAN AND READING: A LONGITUDINAL META-ANALYSIS

1997). Further, early RAN predicts reading over long time intervals, at least a decade into the future (Adlof et al., 2010; Mazzocco & Grimm, 2013). Importantly, the RAN-reading relationship persists across varying ages, reading abilities (Araújo et al., 2015; Araújo & Faísca, 2019), and languages with orthographies of varying depth (Caravolas et al., 2019; Furnes & Samuelsson, 2011).

Gaining a nuanced understanding of the relation between RAN and reading ability is important for two major reasons: informing educational/clinical practice and informing theory. In terms of informing practice, understanding the circumstances under which RAN best predicts later reading is crucial for screening and early identification of reading difficulties. For example, little is known about when the optimal time is to screen and whether the exact type of RAN test matters (in terms of number of items, type of items, use of raw or standardized score, and more). Identifying children with reading difficulties as early as possible, when intervention is more effective, would mitigate the compounding negative consequences that poor readers face under the predominant “wait to fail” model, such as reduced educational attainment, poorer socio-emotional well-being, and higher rates of entry into the juvenile justice system (Humphrey & Mullins, 2002; Richardson & Wydell, 2003; Svensson et al., 2001; Torgesen, 2004).

Understanding the nature of the RAN-reading relationship also informs understanding of the nature of reading ability and development and theory related to reading. Thirty years ago, Maryanne Wolf (1991) published a paper in *Reading Research Quarterly* summarizing the behavioral and cognitive neuroscience research on RAN as an example of how basic research might inform our understanding of why RAN and reading are related. Later, Wolf and Bowers (1999) presented the double deficit hypothesis of dyslexia, in which processes underlying RAN were a causal deficit in dyslexia, separate from but sometimes co-occurring with phonological

RAN AND READING: A LONGITUDINAL META-ANALYSIS

awareness deficits. Other multiple-deficit models also consider naming speed to play a causal role in reading (Menghini et al., 2010; Pennington, 2006; Pennington et al., 2012). However, in other prominent accounts such as the Simple View of Reading (Gough & Tunmer, 1986), the constructs of speed and automaticity as measured by RAN are considered to play a minor role at best (as part of the decoding component, Johnston & Kirby, 2006). Thus, the goal of this paper is not only to ask practical questions designed to inform how RAN can be used in practice, but to provide a clearer picture of the RAN-reading relationship over time, given powerful meta-analytic methods.

Defining RAN Tasks

RAN is measured by the time it takes a child to name an array of familiar items, such as objects, colors, numbers, or letters (Denckla & Rudel, 1976; Norton & Wolf, 2012), reflecting the automaticity of the multiple processes that are involved in this process (Wolf et al., 2000). There are several important parameters that define a true RAN task. First, the items to be named must be highly familiar or *automatized*. For example, when children are typically still learning their letters in kindergarten, the RAN letters task may not relate closely to reading because the naming is not automatized. However, once children have learned the names of letters and numbers with automaticity, these alphanumeric RAN tasks are completed faster than non-alphanumeric tasks (such as objects or colors), and are more strongly related to reading (Cardoso-Martins & Pennington, 2004; Schatschneider et al., 2004; Torgesen et al., 1997). Second, the items must be arranged in an array or grid and named in the left-to-right, row-by-row fashion that is analogous to reading in English. Naming items that are presented one at a time in a speeded manner (discrete naming) is not the same as the serial process of a true RAN task (Altani et al., 2020; de Jong, 2011; Logan et al., 2011; Protopapas et al., 2013), despite the fact

RAN AND READING: A LONGITUDINAL META-ANALYSIS

that some studies call this “discrete RAN.” Third, the RAN measure is usually based on time to complete the task. Some studies use the number of items/second or seconds/item (e.g., Schatschneider et al., 2004). Errors and self-corrections are not typically used in calculating a RAN score, but they may increase the time to name the array and thus be reflected in the naming time. Other factors can be calculated from a RAN task, such as pause time or change row-by-row, but these are less widely used in practice (Amtmann et al., 2007; Georgiou et al., 2006; Georgiou et al., 2008).

A longstanding question in the field is whether RAN should be subsumed under the larger construct of phonological awareness (PA) (Wagner et al., 1993; 1997). Many individual studies find that RAN is a unique predictor of reading, distinct from or beyond the contributions of PA and letter knowledge (Landerl et al., 2019; Norton & Wolf, 2012), and that they have distinct neural correlates (Norton et al., 2014, 2021). A previous meta-analysis from Swanson et al. (2003) provides a simple correlation matrix among these measures, from which partial correlations can be calculated. Using these values, we calculated that the unique contribution of RAN to reading above and beyond PA was $r = .34$, a moderate effect size.

Theories of Mechanisms Underlying the RAN-Reading Relationship

Many potential explanations for why RAN relates so strongly to reading have been posited, including their shared processes of global processing speed, orthographic access, and articulation (Wolf et al., 2000). Some explanations focus more on one underlying component than others; for example, Kail and Hall (1994) theorize that global processing speed is the causal link between age-related improvements in naming and reading. This view has been refuted by experiments showing that children with dyslexia had normal reaction times on a number of non-phonological or linguistic processing speed measures (Bonifacci & Snowling, 2008; Lervåg &

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Hulme, 2009). Others posit serial visual processing and orthography as the primary links between RAN and reading (Sunseth & Bowers, 2002). The RAN-reading relationship is driven by not only serial processing or left-to-right eye movements (Protopapas et al., 2013), but cascading processing (i.e., processing multiple items simultaneously in overlapping fashion), that is important for RAN and has been demonstrated with eye-tracking (Nayar et al., 2018).

More recent accounts have tried to reconcile these proposed causes. Some models suggest that RAN and reading are related because they share multiple underlying cognitive processes such as serial (as opposed to discrete) processing, object recognition, phonological retrieval, and articulation (Georgiou & Parrila, 2020; Norton & Wolf, 2012; Wolf et al., 2000). Thus, within an individual, a profile of strengths and weaknesses of underlying cognitive processes will similarly affect both RAN and reading, accounting for their correlation. Although the role of some specific processes such as articulation remain contentious (Georgiou & Parrila, 2020; Lervåg & Hulme, 2009), the view that multiple shared neural and cognitive processes underlie both RAN and reading (Cummine et al., 2015) is supported by a number of findings. For example, alphanumeric RAN is more correlated with reading than non-alphanumeric RAN, yet non-alphanumeric RAN still significantly predicts reading, indicating that the more overlapping processes that RAN and reading share, the more correlated they will be (Araújo et al., 2015; Lervåg & Hulme, 2009). Another example of this is that in skilled readers, RAN is more strongly correlated with reading fluency measures than other types of reading measures (Araújo et al., 2015), which occurs because of the shared timing demand.

Insights on how RAN Relates to Reading from Meta-Analyses

Previous meta-analyses have documented the significant correlation between RAN and reading across various reading constructs and languages. In the first published meta-analysis of

RAN AND READING: A LONGITUDINAL META-ANALYSIS

RAN and reading, Swanson et al. (2003) found a strong concurrent relationship between RAN and single word reading ($r = -.41$), when looking across a range of ages, reading abilities, and languages. (Note that here, we present all correlations as negative, despite factors like raw versus standard scores, indicating that faster RAN is associated with better reading, as this is always the observed direction of the relation.) Two subsequent meta-analyses have found a similar magnitude of relationship between RAN and reading, while providing new contributions. Araújo et al. (2015) found the overall concurrent RAN-reading relationship across languages to be $r = -.43$, with slightly higher correlations in opaque orthographies like English. Their analyses included substantially more studies, and thus provided greater statistical power than earlier work by Swanson and colleagues. In turn, Hjetland et al. (2017) found the longitudinal correlation from early RAN to later reading to range from $r = -.34$ to $-.37$, depending on the reading measures used. Thus, they demonstrated that longitudinal correlations with RAN have similar effect sizes to concurrent correlations.

Differences in RAN ability have also been identified in two meta-analyses of children with reading difficulties. In a meta-analysis of various cognitive and reading-related skills, Kudo et al. found that the effect size difference for RAN in children without vs. with reading difficulties was $d = 0.89$ (equivalent to $r = .41$), however only 10 samples were included in that analysis. In a much larger meta-analysis with 216 effect sizes analyzed, Araújo et al. (2019) documented an even larger RAN deficit in individuals with dyslexia ($d = 1.19$). These documented RAN deficits in children with reading difficulties/dyslexia support its use as an early screener.

In addition to demonstrating consistent correlations between RAN and reading, these meta-analyses also demonstrated that various factors (i.e. moderators), such as the type of stimuli

RAN AND READING: A LONGITUDINAL META-ANALYSIS

used, the orthographic depth of the language studied, and the type of reading measure, affect the strength of the RAN-reading correlation. Swanson et al. (2003) found that of 11 possible moderators, children's grade when RAN and reading was assessed was the only significant moderator, with older children showing a stronger relationship between RAN and reading. However, these analyses were likely underpowered due to the limited published literature available in 2003. As a result of methodological improvements and more available literature, Araújo et al. (2015) found another key moderator: opaque alphabetic orthographies have a stronger RAN-reading relationship than do transparent orthographies. They also found that the concurrent RAN-reading correlation was moderated by the type of RAN stimuli (such that alphanumeric stimuli had a stronger relationship with reading than non-alphanumeric), and by the type of reading measure (i.e., RAN had a stronger relationship with fluency-based measures than accuracy-based measures).

As noted above, only one meta-analysis has examined some aspects of the longitudinal RAN-reading relationship. The broader focus of the approach of Hjetland et al. (2017) was to assess a variety of longitudinal predictors of reading comprehension, such as vocabulary and grammar, as well as RAN, across various languages. As a result, they did not assess many key potential moderators of the RAN-reading relationship. They found mean effect sizes for RAN measures predicting later single word reading of $r = -.37$ and predicting reading comprehension of $r = -.34$. These correlations are slightly lower than those found by Araújo and colleagues (2015), perhaps due to Hjetland and colleagues' inclusion of only studies with reading comprehension measures and much smaller sample size overall, or the fact that this analysis included only longitudinal studies. Furthermore, in Hjetland et al.'s analyses, one study was an

RAN AND READING: A LONGITUDINAL META-ANALYSIS

extreme outlier and was included with a positive rather than negative correlation with RAN¹; thus, the effect sizes from this study may even be under-estimated.

Motivations and Goals for the Current Study

The purpose of this meta-analysis is to assess the longitudinal relationship from RAN measured in kindergarten or preschool to later reading abilities in English. We consider a variety of reading constructs, including measures of nonword decoding (i.e., reading nonsense words like “sorp”), sight word reading (i.e., reading single words that can be recognized without decoding), and reading fluency (i.e., reading sentences or paragraphs aloud as accurately and quickly as possible). This work thus extends a previous meta-analysis (Hjetland et al., 2017) to include articles that use all reading constructs rather than only reading comprehension as an outcome. We also perform extensive forward and backward snowball searching, as more papers were available to include beyond those identified in the Hjetland et al. (2017) dataset.

Practical Motivations

The key considerations for this design, including its focus on work in English-speaking children, early measures of RAN, and longitudinal relationships, are driven by a goal for this meta-analysis to inform specific policy recommendations for educators and administrators. It is clear that state- and local-level policymakers are looking for ways to best implement RAN in screening, as evidenced by the creation of measures such as the Arkansas Rapid Naming Screener and its use by other states (Arkansas Department of Education, 2017). As in previous meta-analyses examining the concurrent RAN-reading relationship, we also test several potential

¹ Bishop & League (2006) reported a positive correlation between RAN time and reading ability (it appears the authors used raw time measures of RAN). However, in an earlier report from the same sample, Bishop (2003) reported positive correlations using standard scores (in the expected direction of this relationship). The RAN-reading correlations from this paper should likely have been treated as negative in this case for Hjetland’s analyses, as all other measures in this and other meta-analyses were negative.

RAN AND READING: A LONGITUDINAL META-ANALYSIS

moderators, which address key practical questions. Practical questions, such as “how many items should a RAN task include?” and “at what age should I evaluate RAN?”, may help educators and clinicians choose effective screening measures.

Theoretical Motivations

Most meta-analyses of RAN focus on documenting the relationship between RAN and reading while generally not trying to explain why RAN and reading are related. Here, we will test several questions related to why RAN and reading are correlated. Theoretical questions, such as “do timed reading measures more strongly relate to RAN than untimed reading measures?” and “do nonword decoding tasks relate less strongly to RAN than sight word tasks?” may help researchers further converge on theory for why RAN and reading relate.

Summary

Because RAN ability develops considerably during the school-age years (Denckla & Rudel, 1974; Georgiou et al., 2006), its relationship to later reading ability may be different than the concurrent relations between RAN and reading at older ages. However, if early RAN reliably predicts later reading, it further increases the motivation to include RAN in kindergarten or preschool literacy screening. However, there is a lack of understanding of the theoretical and practical questions about how early RAN task performance relates to later reading abilities. As such, quantifying the average relationship between early RAN and later reading is the primary research question in this meta-analysis. Secondary questions are whether factors related to the RAN task, reading measure, or child participant sample, moderate the RAN-reading relationship. These specific questions and their rationale are explained in depth, and specific analyses proposed in the Method section.

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Method

This study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009). Data collection and extraction processes are described in text and in **Figure 1**. The PRISMA checklist is provided as Supplemental Material. Our data, protocols, processing and analyses scripts, and other related documents are available via Open Science Framework: [link is anonymized] https://osf.io/awpqq/?view_only=8e0022753d1a456e9127ce87b3241e30. This meta-analysis was considered exempt by the Institutional Review Board at [authors' institution].

Study Inclusion Criteria

For the present study, we focused on articles in which English was the primary language of the participants, as consistency of orthography can moderate the RAN-reading relationship (Araújo et al., 2015) and the largest number of published studies are in English. We acknowledge that English is not a representative orthography (Share, 2008), but that this analysis serves as a starting point and allows specific conclusions to be drawn in at least this one language. As we were interested in early predictors of reading ability, we only included articles in which the initial timepoint with RAN assessment was in (the US equivalent of) kindergarten or preschool (the earliest stage at which RAN can be measured reliably) and reading was subsequently measured at some point in Grades 1-5. Thus, we only included studies that spanned at least one school year. For studies that only reported the sample's age rather than grade, we included the study if the mean age was ≤ 78 months (age 6.5 years, or the middle of first grade in the US). Studies with children who spoke other languages were excluded; however, studies with bilingual children were included if a) the language of instruction was English and b) the children were described as fluent in English. All eligibility criteria can be found in **Table 1**.

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Data Collection

On September 26, 2019, we identified possible sources through full-text database searching using EBSCO (PsychINFO, PsycARTICLES, and ERIC) and PubMed. We used the search terms: (reading OR dyslexia) AND ("rapid naming" OR "naming speed" OR "rapid automat* naming" OR "RAN" OR "rapid serial naming") AND ("preschool*" OR "kindergart*" OR "pre-school*" OR "pre k*" OR "pre-k*" OR "prek*" OR "child*"), see **Table 1**. This search returned 4497 titles, 4088 of which were unique. **Figure 1** shows the number of articles at each stage.

Abstract and Title Screening

As a first step, one of two authors reviewed the title of each article from the database search; titles that were deemed to be clearly irrelevant were screened out. This title screening step resulted in 2014 potentially relevant articles with abstracts to be screened. These abstracts were then each reviewed by two different screeners. Three individuals contributed to abstract screening and consensus was reached in all cases of conflict. (Some articles were triple-screened during training, but all other articles were double screened.) Abstract screening for full-text inclusion agreement was 85% and all disagreements were resolved with consensus of three coders. 416 of these articles were deemed relevant and were then full-text screened. Seven trained coders screened full text for inclusion with 89% agreement and resolved all disagreements. From these, 92 articles met the eligibility criteria and 52 had relevant effect sizes (after contacting authors to obtain some that were not included in articles). These articles were then double coded for various measures of interest by five trained coders, exceeding 94% agreement across all variables. All disagreements were reviewed by the first and the second author and resolved through consensus.

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Snowball Search

After the database search and screening, a snowball search was conducted using references and citations of the 52 included studies with relevant effect sizes. For this snowball search, we used Microsoft Academic Graph (Wang et al., 2019), which is a database that tracks connections between published papers, such that every backward reference is also a forward citation, similar to Web of Science. All articles that were identified by the snowball search were title and abstract screened using the same processes as those described above. Snowball searching returned 43 articles that met the eligibility criteria. 15 of these studies had relevant effect sizes (after contacting authors) and added 10 unique samples. The search also returned 28 studies without relevant effect sizes, 14 of which were related to samples already contained in the corpus.

Contacting Authors

Authors from either the database search or snowball search whose paper had no relevant effect sizes (e.g., because of reporting regressions or grouped analyses rather than correlations) were contacted via email to request raw data or correlation matrices so that the information could be included in the current analysis. Of the papers where this was the case, 9 authors returned our request, providing data on 10 unique samples.

Data extraction

Data for this study were collected and managed using REDCap (Research Electronic Data Capture) tools hosted at [authors'] University (Harris et al., 2009, 2019). REDCap is a secure, web-based software platform designed to support data capture for research studies.

Data (including relevant information on the sample, tasks, and Pearson correlations) from each paper/sample were entered in REDCap by two independent coders, and consensus was

RAN AND READING: A LONGITUDINAL META-ANALYSIS

reached in case of any discrepancy. For longitudinal studies that measured RAN and/or reading at multiple timepoints, we extracted only one kindergarten/preschool time point and only one grade school timepoint. This design consideration intentionally minimizes variance, as our primary question is focused on the utility of RAN as an early screener. However, a side effect of this approach is that it limits the variability that can be explained by age of testing. The only cases in which we collected multiple grade school timepoints were where the measures used at each timepoint differed considerably and did not overlap. For example, Badian (1994) collected only decoding measures in fall of Grade 1, and collected comprehension measures in the spring. In this case, two timepoints were extracted, but only so that all possible types of reading measures were included. When multiple outcome points were available, we coded and analyzed the one closest to the end of grade 2, as this is the period in which children typically develop automaticity in word reading (Chall, 1983; Wolf et al., 2000) and when dyslexia is commonly diagnosed in the US.

We also prioritized extracting effect sizes from whole samples if they were available. This design consideration was also intended to minimize variability to focus on the utility of RAN as an early screener. However, there were many cases in which authors reported only data from subsamples. For example, Heath and Hogben (2004) report correlations separately for groups with Good PA and Poor PA.

Timing of initial and follow-up assessments were coded in terms of the sample's grade, as papers predominantly reported grade rather than age. Grade was further specified as fall (July-December) versus spring (January-June) semesters when available. Fall vs. spring specifications from Australia were flipped to match the US/UK/Canadian school year. UK and Australia samples were coded as kindergarten if they were called Year 1 and as Grade 1 if they were called

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Year 2. In the cases that age was reported without grade, we used the following guidelines to derive a code for grade: mean age <60 months was coded as preschool fall, and mean age between 60 and 66 months was coded as preschool spring. This 6-month progression was used all the way through grade school. For example, a sample with a mean age of 100 months at follow-up would be coded as Grade 2 spring if grade was not reported.

Effect Size Extraction

The scoring of the RAN task affected whether the Pearson correlation with reading would be positive or negative. If a raw score (i.e., time) or rate (time/item) was used, the correlation was entered as negative. If a standard score or rate (item/time) was used, this value was multiplied by -1. There were a few exceptions to this rule, in which a reading measure was either based on time or rate (e.g., Wolf et al., 1986) or expressed as a chronological age lag (Heath & Hogben, 2004). In addition, there were several ambiguous cases that were carefully considered, see details in Supplemental Materials.

Many studies assess RAN as part of a large battery of reading-related measures that potentially predict later reading. Due to the many constructs measured in these large and longitudinal studies, many researchers created latent RAN or reading measures through factor or principal components analysis (Dally, 2006; Macdonald et al., 2013). We decided to extract these correlations between one or two latent variables as they qualify as Pearson correlations, and later test whether including them would change our results.

RAN Measure Categories

The stimuli used in a RAN task are typically restricted to one of five types: colors, objects, letters, digits, or occasionally animals. Even more rarely, studies have used colored animals (e.g., Catts et al. 1999). The colored animals task (e.g., naming “blue cow,” “red dog,”

RAN AND READING: A LONGITUDINAL META-ANALYSIS

etc.) is included here as a RAN task, but not compared with other stimulus types in moderator analyses due to the very few studies that employed it. We also excluded tasks with multiple stimulus types in the array, such as letters and numbers, in order to focus on the classic RAN task. Previous meta-analyses have found that the relationship with reading is stronger between alphanumeric (i.e., letters or numbers) than non-alphanumeric stimuli (such as colors or objects; Araújo et al., 2015). However, this was assessed concurrently, whereas different results may be seen with early RAN predicting later reading. Further, many children do not know their letters accurately or automatically in kindergarten or preschool, making a RAN letters task inappropriate for these younger children. Thus, in the current study **we quantified each RAN task's relationship with later reading and whether alphanumeric RAN tasks are a stronger predictor of later reading than non-alphanumeric RAN.**

Reading Measure Categories

Reading can be assessed at different levels, from single word to connected text, and in terms of different constructs, including nonword decoding, accuracy, speed or efficiency, and comprehension. Previous meta-analysis of children of all ages indicates that RAN is associated with single word reading accuracy (i.e., word ID) at $r = -.41$ and reading comprehension at $r = -.45$ (Swanson et al., 2003). Hjetland and colleagues (2017) found mean effect sizes of $r = -.37$ for word reading and $r = -.34$ for reading comprehension with earlier RAN measures. However, the specific correlations between RAN and reading vary considerably between and within studies. For example, in one study (Cronin & Carver, 1998), kindergarten RAN scores related to Grade 1 Word ID scores at $r = -.37$ to $-.60$, depending on the RAN task, and to passage comprehension at $r = -.31$ to $-.57$. Thus, **we quantified RAN's relationship with 3 primary types of reading: fluency, comprehension, and single word reading.**

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Timed measures. Because RAN is a speeded task, it is typically more closely related to timed or speeded reading measures (Savage & Frederickson, 2005; Schatschneider et al., 2004). This is evident in studies of older students; for example, RAN speed in grade 3 significantly predicted performance on a timed single word reading task in grades 3, 4, and 5, but did not reliably predict untimed single word reading (Georgiou et al., 2009). Further, one theoretical account posits that processes underlying RAN constrain the development of reading fluency (Lervåg & Hulme, 2009). Thus, **we quantified RAN's relationship with timed and untimed reading measures.**

Nonword Reading. Nonword reading tasks have extra phonological demands that sight words do not. Previous meta-analyses (Araújo et al., 2015) found a weaker correlation between non-word reading and RAN than real word reading and RAN. This difference may exist because nonword reading is much less automatic than real word reading, even early in reading development. **Therefore, we quantified RAN's relationship with real-word reading and nonword reading, with the prediction that the relationship between RAN and nonword measures would be weaker than RAN and real word reading.**

Participant Characteristics

Reading ability. Among older students, there is mixed evidence regarding whether RAN is a stronger correlate or predictor of reading ability among children who are poor readers than typical or skilled readers. Some studies find a stronger concurrent RAN-reading relation in poor readers (Araújo et al., 2011; Bowers et al., 1988; Felton & Brown, 1990). One study found that RAN in 3rd grade significantly predicted later single word reading in 8th grade among poor readers, but that there was no such significant relation in good readers (Meyer et al., 1998). On the other hand, meta-analyses of concurrent RAN-reading relations in older children reveal that

RAN AND READING: A LONGITUDINAL META-ANALYSIS

the correlation between RAN and reading is similar in samples of typical readers and poor readers; Swanson et al. (2003) found correlations of $r = -.41$ for typical readers and $-.43$ for poor readers, and Araújo et al. (2015) found no significant differences in the magnitude of the concurrent relations between RAN and reading whether the sample of readers was poor/impaired ($r = -.49$), typical/average ($r = -.45$), or unselected ($r = -.43$). It is not known whether these differences across studies are due to a restricted range or “ceiling” effect in RAN among good readers with greater variability among poor readers (McBride-Chang & Manis, 1996) or whether differential relations truly exist in good versus poor readers.

Further, studies have examined children at risk for dyslexia (based on family history of dyslexia or behavioral characteristics) who are not yet poor readers. Children at family risk for dyslexia tend to have poorer RAN skills than their peers (Pennington & Lefly, 2001; van Bergen et al., 2012), yet not all children with risk or poor RAN scores go on to poor reading. Some studies find the RAN-reading relationship to be significantly weaker in those at risk for dyslexia; for example, Heath and Hogben (2004) found that children with poor PA skills had a pre-kindergarten RAN to Grade 2 Word ID correlation of $r = -.03$ compared to $-.38$ for children with good PA skills. Other studies have found highly similar effect sizes regardless of children’s abilities; for example, Hulme et al. (2015) found children at risk for dyslexia had a kindergarten RAN Objects to Grade 3 reading correlation of $r = -.22$ and children not at risk had a correlation of $r = -.21$. **Thus, we tested whether early RAN is a better predictor in samples of primarily typically developing children as opposed to samples with larger proportions of children identified as at-risk for reading difficulties.**

Practical Considerations

RAN AND READING: A LONGITUDINAL META-ANALYSIS

RAN Task Publication, Standardization and Test Length. There are a number of published, standardized and normed RAN measures that are used widely, including the Comprehensive Test of Phonological Processing (CTOPP/CTOPP-2; Wagner et al., 2013) and the RAN/RAS Tests (Wolf & Denckla, 2005), among others. However, many studies use researcher-created RAN tasks that have not necessarily been standardized or normed. Among these tests, the format of the RAN task, including how many different unique items (types) and total number of items included (tokens), also varies. A previous meta-analysis found no moderating effect for the total number of items in a RAN task on concurrent relations with reading (Araújo et al., 2015). Thus, **we tested whether using a published, standardized measure influenced the RAN-reading relationship, as well as whether RAN measures with different numbers of items per set or total items, were more strongly related to reading.**

Timing of initial RAN assessment and later reading assessment. Dyslexia is typically not diagnosed before the end of grade 2 because the heterogeneity of reading development profiles makes it difficult to reliably identify children who will have ongoing reading difficulty. Thus, it would be helpful to know when RAN assessment is effective for predicting later reading. In the US, kindergarten is a typical time for early reading screening; thus, many studies that investigate the longitudinal relations with RAN measure it at the start of kindergarten. However, some studies have assessed RAN in children as young as age 3.5 (McBride-Chang & Kail, 2002; Su et al., 2017), and widely-used normed measures of RAN are available for children ages 4 and up (e.g., CTOPP-2, Wagner et al., 2013). Thus, **we tested how the timing of RAN assessment (i.e., preschool vs. kindergarten) differentially impacts the RAN-reading correlation.**

Another important consideration is the timing of the later or “outcome” reading measure, as the nature of the relations between RAN and reading may change over the course of reading

RAN AND READING: A LONGITUDINAL META-ANALYSIS

development. For example, early in reading development, children are developing accuracy in reading, and over time, they become accurate and build automaticity; thus, RAN may relate to fluency-based reading more strongly when reading is more automatized. In a practical sense, for early identification of reading problems, it may be important to know at what outcome point this relation is stable. Wolf and colleagues (2000) suggested that RAN may play an attenuated role in predicting reading for typical readers after grade 2, because so many children achieve automaticity in naming. Thus, we **tested the extent to which the timing of reading assessment moderated the RAN-reading relationship.**

Outlier Handling

Due to the nature of nested effect sizes, we examined outliers at the study level. We did this by taking the mean of all numeric variables at the study level and then testing whether any observations fell above the 97.5%ile or below 2.5%ile. If a study fell outside of these values, it was further investigated and considered for inclusion on a case-by-case basis; importantly, this was done before analysis so as not to bias results. All studies/samples were retained for intercept-only models. For moderator analyses, several studies were excluded as they were outliers for the variable of interest. These cases are described in Supplemental Materials.

Study Quality and Risk of Bias

Study quality measures can be helpful in identifying whether certain design considerations, such as double-blind randomized control trials, yield less biased estimates of effect sizes. These features that reflect study quality are less clear for correlational, longitudinal research designs. Here, we use three measures of study quality and risk of bias: the use of a published, standardized RAN test, the use of latent variables, and the study's sample size. These were all separately analyzed as moderators of the RAN-reading relationship; as there is no gold-

RAN AND READING: A LONGITUDINAL META-ANALYSIS

standard or guidance for doing so, we did not feel it was appropriate to create a composite study quality and risk of bias measure.

Statistical Power

Power was calculated for each moderator analysis and is reported alongside each moderator analysis. As in Araújo et al. (2015), we used the value of 0.1 difference between Fisher's z values as the smallest difference that would be meaningful. For the sample risk proportion analysis (e.g., low, medium, and high risk proportion), we used .1 Fisher's z difference on either side of $z = .4$, as this is a typical RAN-reading correlation reported in other meta-analyses. As there is no widely accepted methodology for calculating moderator analyses' power in RVE models, we used the degrees of freedom from each moderator analysis (rounded to the nearest integer), which is an effective sample size. Using the mean sample size of $n = 176$ and an I^2 value of 75%, we used the *metapower* package (Griffin, 2020, 2021) to calculate power for each moderator tested. Because we are using an *a priori* effect size estimate, this is not a *post hoc* power calculation. Power values for each analysis are presented alongside each model in

Table 4.**Analysis Process and Plan**

Reported effects were transformed from Pearson correlations to Fisher's z -scores, which normalizes their distribution for analysis. They were then transformed back to Pearson correlations in reported results for ease of interpretation and comparison with other meta-analyses. 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; Hedges et al., 2010; R Core Team, 2013). These models allow for correlated effects within a study, maximizing data retention. Furthermore, these models allow the grouping of multiple

RAN AND READING: A LONGITUDINAL META-ANALYSIS

studies that share a sample (e.g., the International Longitudinal Twin Study; Furnes & Samuelsson, 2009, 2011). Intercept-only and moderator analyses were performed using the *robu* function. Moderators were tested in separate meta-regression models (e.g., one model testing alphanumeric stimuli as a moderator and a separate model testing dyslexia risk as a moderator), except in the case of time of assessment, where the initial and outcome timepoints were considered together.

To test for funnel plot asymmetry, which is indicative of publication or reporting bias, we used a technique that allows for multiple effect sizes per study. Traditional methods for examining funnel plot asymmetry, such as Egger's Regression or trim-and-fill analyses, only accommodate one effect size per study. Recently, these traditional methods have been expanded to correlated effects models with "sandwich" estimators (Rodgers & Pustejovsky, 2020). We therefore used an "Egger's Sandwich Regression" to test for funnel plot asymmetry. As our data came from a variety of sources, we also ran a moderator analysis to test whether published effect sizes were larger than unpublished effect sizes (e.g., an unpublished dissertation, data emailed from authors).

Results

Sample Description

The final analytic sample ($n = 10,513$) was drawn from 60 independent samples across 67 papers. The largest sample size in the Hjetland et al. (2017) longitudinal RAN analyses was 3746; the current study sample is thus nearly three times greater, even though we restricted the language of the participants to English and the initial timepoint to before first grade. For studies that reported age of participants at the initial timepoint, the mean age was 67.51 months (SD of 4.02) and a range of mean ages from 54-75 months across studies. The mean interval between

RAN AND READING: A LONGITUDINAL META-ANALYSIS

initial and final timepoints was 27.41 months, which is consistent with our prioritization of the Grade 2 timepoint. Other descriptive statistics for the samples included are presented in **Table 2**.

Intercept-only Models

We calculated an intercept-only model to assess our main research question, the overall correlation between preschool/kindergarten RAN and later reading. The intercept-only model yielded an average effect size of $z = -.40$ (95% CI: $-.37$ to $-.44$, $p < .001$), equivalent to a Pearson correlation of $r = -.38$. This indicates that on average, children with faster RAN time before grade school have stronger grade school reading performance. The forest plot for the overall intercept-only model is presented in Supplemental Material. Excluding studies that reported latent variables for RAN or reading resulted in nearly identical model results ($r = .38$). There was considerable variability in the effect sizes ($I^2 = 74.09$; $\tau^2 = .018$), indicating that analysis of moderators may further clarify this relationship. We also tested intercept-only models including only a subset of studies based on what types of RAN tasks and reading measures the study used. These results are presented in **Table 3**. All models were significant at $p < .001$, indicating that the relationship between various RAN and reading measures is quite robust.

Moderators

Practical and theoretically motivated moderators were analyzed and are presented in **Table 4**.

Practical Moderators

Unique RAN Items and Total RAN Items. We tested whether specific features of the RAN task administered in each study, such as the number of total items or the number of unique items, were differentially predictive of later reading. We found that neither the number of total items, nor the number of unique items moderated the RAN-reading relationship (all $ps > .26$).

RAN AND READING: A LONGITUDINAL META-ANALYSIS

This indicates that RAN test length and item composition does not meaningfully modify the RAN-reading relationship.

Standardized RAN Test. Next, we tested whether using published assessments that are standardized and normed, such as the RAN/RAS Tests or the RAN subtests from the CTOPP, affected the RAN-reading relationship. We found that using a published assessment had no effect ($\Delta r = .06$; $p = .18$) on the strength of the RAN-reading relationship. This also was an indicator of risk of study bias, indicating that study quality may be less likely to bias these results.

Age at Assessments. We tested whether the timing of the RAN or reading assessments (e.g., earlier or later than initial assessment at early kindergarten for RAN assessment or than Grade 2 for reading assessment) moderated the RAN-reading relationship. We found that age at reading assessment had no moderating effect ($\Delta r = 0.00$; $p = .97$), but that age at RAN assessment did have a marginally significant effect ($\Delta r = -.01$; $p = .07$), in the direction of later assessment having a stronger RAN-reading relationship. We considered that this result may be conflated with the RAN alphanumeric analysis, as younger children are less likely to be able to complete alphanumeric RAN. After controlling for whether the RAN task was alphanumeric or not, there was no effect of age at initial assessment ($\Delta r = 0.00$; $p = .15$). This result indicates that the exact timing of early RAN measurement does not differentially affect the RAN-reading relationship.

Theoretical Moderators

Alphanumeric vs. non-alphanumeric RAN. We directly tested whether alphanumeric RAN was a better predictor of reading than non-alphanumeric RAN. We found that alphanumeric RAN is a significantly stronger predictor of reading ($\Delta r = .13$; $p = .01$), meaning that RAN tasks with letters or numbers had a stronger correlation with reading than did tasks

RAN AND READING: A LONGITUDINAL META-ANALYSIS

with colors or objects. To consider the possibility that this relationship was conflated with initial age (because younger children may be less likely to have completed an alphanumeric task successfully), we ran the same analysis controlling for initial age, and the effect was unchanged ($\Delta r = .13$; $p = .01$).

Real vs. non-word reading. Next, we directly tested whether measures of nonword reading had a weaker relationship with RAN than measures of single, real word reading. We found a significant effect ($\Delta r = -.09$; $p < .001$), with measures of nonword reading having a weaker relationship with RAN than measures of single, real word reading.

Timed vs. Untimed reading. We then tested whether timed reading measures were more related to RAN than untimed measures. We found no difference ($\Delta r = .00$; $p = .90$) between timed and untimed reading measures as they relate to RAN.

Efficiency vs. Accuracy reading. As there were no differences in timed vs. untimed reading measures, we also tested whether measures of reading efficiency were more related to RAN than measures of reading accuracy only. We found no difference ($\Delta r = -.01$; $p = .73$) between how measures of reading efficiency and reading accuracy relate to RAN.

Dyslexia risk proportion in the sample. Using the three-level classification of dyslexia risk of the sample (low, medium, or high proportion of children at risk) in a single model, we tested whether the RAN-reading relationship was affected by dyslexia risk. There was no significant moderating effect of level of dyslexia risk (all $ps > .46$). This result indicates that the RAN is a similar predictor of reading across samples of children that vary in risk for dyslexia.

Risk of Bias Analysis

To assess risk of bias, we ran an Egger's Sandwich Regression, in which the standard deviation estimates from each study were used as the moderator. We found no risk of bias in our

RAN AND READING: A LONGITUDINAL META-ANALYSIS

effect size estimates ($p = .32$). Sample size is often used as a study quality measure as well; this result indicates that sample size has no significant effect on effect size estimates. However, because our data were composed of peer-reviewed studies, unpublished theses, and emailed data from published studies, we also ran moderator analyses with whether data were from a published paper or not (i.e., an unpublished dissertation or emailed data). These analyses revealed strong evidence of reporting bias, with published effect sizes being stronger than unpublished effect sizes ($\Delta r = .09$; $p = .02$). This effect was not driven by dissertations ($\Delta r = .003$; $p = .97$), but rather by other types of unpublished data (e.g., emailed data). Due to the highly nested nature of these data, a funnel plot visualization is not provided, given that plotting up to 27 effect sizes with the same standard error would result in essentially a horizontal line on the funnel plot and be difficult to interpret.

Discussion

This meta-analysis expands on previous findings by documenting the longitudinal relationship between early RAN and various measures of later reading abilities in English-speaking children. Consistent with previous research and meta-analyses, RAN tasks were found to be a strong predictor of all types of reading. The mean effect size found here for RAN predicting reading overall ($r = -.38$) is similar to meta-analyses of concurrent RAN-reading correlations, with r ranging from $-.38$ to $-.45$ depending on reading measure in Swanson et al. (2003), $r = -.43$ Araújo et al. (2015), $r = -.34$ for reading comprehension, and $r = -.37$ for Word ID in Hjetland et al. (2017).

Our meta-analysis adds uniquely to the literature assessing the links between RAN and reading by highlighting the relevance of assessing RAN in kindergarten or preschool, and the robustness of this relationship across various RAN and reading measures. The only existing

RAN AND READING: A LONGITUDINAL META-ANALYSIS

longitudinal meta-analysis between RAN and reading was limited in its coverage of the literature and theoretical scope, with no moderators assessed (Hjetland et al., 2017). Our database searching, in conjunction with a snowball search strategy, yielded many more included articles, resulting in a sample size nearly three times larger. This much larger sample was ascertained despite restricting our age range to kindergarten and preschool and restricting our language to English.

A major contribution of the present study is the analysis of a variety of potential practical and theoretical moderators of the relationship between early RAN and later reading. For practical moderators, our study was the first, to our knowledge, to examine RAN tasks that were published and standardized vs. researcher-created; these variations also did not significantly alter the predictive relation of RAN with reading. Our analyses that revealed that number of total items, and how many unique items were included in each set did not moderate the RAN-reading relationship align with and extend previous concurrent findings from Araújo and colleagues (2015).

For theoretical moderators, we found that RAN has a significantly stronger relation with reading when alphanumeric stimuli are used, replicating and extending a previous concurrent meta-analysis (Araújo et al., 2015); this pattern was also observed in our data, despite the young age of the RAN assessment here. In considering different reading measures as outcomes, we found only a significant difference for RAN better predicting real word than nonword reading. In contrast, Araújo and colleagues found differences between timed and untimed measures across orthographies, whereas we did not. We discuss the implications of these findings for theory and for practice, below.

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Insights to the Nature of the RAN-Reading Relationship

Our moderator analyses showed that alphanumeric RAN has a significantly stronger relationship with later reading than does non-alphanumeric RAN, as well as that nonword reading is significantly less related to RAN than real word reading. These two results, taken together, support a shared cognitive resource model, in that the more similar the processes that RAN and a given reading task tap, the more that they will correlate (Georgiou & Parrila, 2020). In the case of nonword reading, there is a heavy phonological decoding (letter-to-sound correspondence) component that RAN does not share. In the case of alphanumeric RAN, symbolic representation is required for both alphanumeric RAN and reading. At first glance, our results showing that timed measures did not correlate more strongly with RAN than untimed measures may seem contrary to this model; however, in the early years of reading development, accuracy-based and time-based measures correlate highly. A difference emerges in intermediate and advanced readers once children build automaticity, but it is not present in beginning readers in either our sample or in the beginning and pre-readers included in the meta-analysis from Araújo et al. (2015). These findings are consistent with the idea that reading is not yet automatic in early grades (Chall, 1983; Samuels & Flor, 1997), and as a result, various reading measures may be more highly correlated early in schooling (i.e., less differentiable) than they are at later stages when most children have developed automaticity. More highly correlated reading measures in our earlier outcome timepoint (centered around 2nd grade) would likely result in weaker moderating effects when comparing different types of reading measures.

Consistent with other meta-analyses' findings of no differences in relations with RAN between good vs. poor readers, we found no difference between samples with a large proportion of children at-risk for dyslexia and those with very few at risk. This may indicate that children at-

RAN AND READING: A LONGITUDINAL META-ANALYSIS

risk and children not at risk are using similar cognitive processes, even if these processes are impaired in children at risk. This further supports the idea of RAN being a continuous ability and having dimensional prediction of reading, rather than a dichotomous “present or absent” skill.

Practical Insights for Using RAN as a Screener

These results provide practical insights into using RAN (as part of a larger battery) for effective screening for later reading difficulties. Overall, our results indicate that the relation between early RAN and later reading is remarkably consistent. The particular characteristics of the RAN measure, such as number of items and whether the task was from a published test, did not significantly alter the strength of the RAN-reading relationship. These facets of RAN as a predictor had not been assessed in previous meta-analyses, yet they provide concrete guidance for researchers and educators in planning RAN measures for screening. There was not a significant difference between RAN measures conducted in preschool vs. kindergarten in terms of their relationship with later reading; there was a trend toward stronger predictive power, but the trend was reduced when controlling for alphanumeric RAN, which is often administered at later years. The advantage of earlier identification of potential reading difficulties, so that earlier intervention can be provided, suggest that it would be optimal to employ RAN tasks in screening in pre-school or pre-kindergarten, as soon as RAN can be assessed validly.

The stimulus type used in early RAN assessment is a relevant consideration, as alphanumeric RAN measures were more strongly related to later reading than were non-alphanumeric measures. An important caveat is that RAN tasks, by definition, depend on the child being able to name items with automaticity, and many articles noted that many children could not perform a RAN Letters task in kindergarten, as their letter name knowledge was not yet accurate and automatic (e.g., Catts et al., 1999). Thus, for children in kindergarten or

RAN AND READING: A LONGITUDINAL META-ANALYSIS

preschool who do not yet know the names of letters or digits automatically, a RAN task using colors or objects would be a better choice.

The question that frequently follows after RAN screening is “what RAN time or score is worrisome?” Unfortunately, research has not yet determined a single cutoff score for “dyslexia risk” or what is “good” versus “poor” RAN; in fact, this may not be possible given that RAN is both a continuous measure and one aspect of the constellation of reading-related abilities. At this point, using a published, standardized RAN measure that provides standard scores or percentiles provides the advantage that it may help educators and clinicians understand where a child’s RAN ability falls relative to their peers as an indicator of risk for dyslexia, even though our data showed that researcher-created measures equally predicted later RAN. It is important to note that administering a RAN task according to any standardized instructions and minimizing distractions so as to obtain the child’s best performance is crucial to obtaining a valid score.

Educators and clinicians should also recognize that an effective screening battery for dyslexia and reading difficulties must include RAN alongside other indicators such as phonological awareness (see Petscher et al., 2019, for recommendations). Even using the most evidence-based screening tools in combination with assessment of the child’s family or neuroimaging measures, there is still uncertainty about which children will develop reading difficulty (Norton et al., 2019; Zuk et al., 2020). As the field moves forward in understanding early indicators of reading difficulties, RAN will undoubtedly play a role, given its universal and robust relation with reading.

Limitations

There are several limitations of this study to consider. One potential limitation of our study was our analysis of the quality of the literature analyzed. We chose not to pursue

RAN AND READING: A LONGITUDINAL META-ANALYSIS

composite measures of study quality and instead chose to analyze study quality at the moderator level, based on the concern over validity of using simple sums to describe study quality.

Similarly, Hjetland et al. (2017) found no effect of study quality in an overlapping sample of papers, which aligns with our results that sample size, latent variables, and use of published/standardized tests do not predict variation in effect sizes.

Another limitation is the limited statistical power for moderator analyses. Although we found no differences for unique RAN items or total RAN items, we had limited power to detect possible effects for a multitude of reasons. Araújo et al. (2015) noted similar difficulties, even with a larger corpus of sources and subjects. We offer the same caution in interpreting our moderator analysis results with low power.

Other limitations relate to the RAN tasks themselves. One limitation was the fact that there were incomplete descriptions of the measures in many studies, which was particularly common for researcher-created RAN tasks. Despite our effort to carefully review all available information in the published papers (and in many cases, request additional details from authors via email), many papers had incomplete descriptions of their RAN tasks, particularly relating to how many unique items and how many total items the task had. Furthermore, there was not much variability in the number of unique items, as many articles used Denckla and Rudel's (1976) version or the updated RAN-RAS tests (Wolf & Denckla, 2005) each with 5 unique items per task, or the CTOPP that has 6 unique items. Despite the incomplete information from a number of studies, we believe we had sufficient power to detect these effects if they truly existed, as 288 (of 373) effect sizes were analyzed for the model that tested unique and total items as moderators.

RAN AND READING: A LONGITUDINAL META-ANALYSIS

The definition of at-risk in samples also varied greatly across studies and could limit interpretation of our results. For example, Cardoso-Martins and Pennington (2004) recruited a high-risk group from the children whose one of the parents has reading problems and a low-risk group from the children with no family history of reading problems. Hulme et al. (2015) also divided groups based on family history; however, they included another criterion of whether children have language impairment or not. In contrast, Heath and Hogben (2004) divided groups only based on poor and good phonological awareness abilities. Felton (1992) used teacher ratings of children's expected reading ability. On the other hand, this heterogeneity reflects the real-world heterogeneity of risk definitions, and we designed our categories to reflect that.

Future Directions

We chose to focus on only traditional RAN tasks at certain timepoints in the English language in order to maximize practical and policy impact. As a result, there are several clear directions for future research to expand upon our study by broadening the scope. Future studies may consider different designs, such as meta-analytic path modeling of the relationships among reading, RAN, PA and other predictors of reading. Though the majority of studies and all published tests focus on RAN total time, aspects of RAN such as analyses of inter-item pause times as a predictor would be promising to investigate, as pause times have been shown to relate highly with reading fluency (Lervåg & Hulme, 2009).

Given that we focused on a single outcome timepoint in each study that was close to the end of Grade 2, another potential future direction would be to test how longitudinal RAN-reading relationships change within studies and more broadly over time. As we prioritized collecting only one time point per study, we were not able to analyze whether correlations from early RAN to later reading changed over time within a study, as is suggested by a number of

RAN AND READING: A LONGITUDINAL META-ANALYSIS

authors (de Jong & van der Leij, 2002; Wagner et al., 1997). To our knowledge, correlated effects RVE models have not been used to analyze longitudinal, within-study data. Many of the papers collected for the present analysis would be ideal to use in testing whether RVE is suitable for longitudinally dependent effect sizes and provide further insight into how RAN relates to reading over time.

Another clear direction for future research is to include multiple languages to assess similarities and differences of RAN as a predictor reading ability. In the past, other authors had suggested that RAN is a better predictor in more transparent languages (see Georgiou et al., 2008). In their meta-analysis, Araújo et al. (2015) reported that orthographically opaque orthographies such as English have a stronger concurrent correlation between RAN and reading than do transparent orthographies, but we do not know how this pattern would look longitudinally.

Finally, given that a major focus was the utility of using RAN as a screener, future research should endeavor to provide concrete recommendations of what RAN performance indicates meaningful risk for reading difficulties and dyslexia. Few studies have provided clear formulas or cutoffs about which children are at greatest risk (Catts et al., 2001 is a notable exception). Even fewer studies have examined how best to provide intervention specific to children who have RAN difficulties that impact their reading, as it seems that training RAN itself is not possible (de Jong & Vrielink, 2004; Kirby et al., 2010). Indeed, early measures of RAN may be a robust early indicator of reading problems, akin to a “check engine light” that signals the need for further assessment and monitoring (Norton, 2020).

RAN AND READING: A LONGITUDINAL META-ANALYSIS

References

* indicates paper included in meta-analysis

- *Adlof, S. M., Catts, H. W., & Lee, J. (2010). Kindergarten predictors of second versus eighth grade reading comprehension impairments. *Journal of Learning Disabilities, 43*(4), 332–345. <https://doi.org/10.1177/0022219410369067>
- Al Otaiba, S., Connor, C. M., Folsom, J. S., Wanzek, J., Greulich, L., Schatschneider, C., & Wagner, R. K. (2014). To wait in tier 1 or intervene immediately: A randomized experiment examining first-grade response to intervention in reading. *Exceptional Children, 81*(1), 11–27. <https://doi.org/10.1177/0014402914532234>
- Altani, A., Protopapas, A., Katopodi, K., & Georgiou, G. K. (2020). Tracking the serial advantage in the naming rate of multiple over isolated stimulus displays. *Reading and Writing, 33*(2), 349–375. <https://doi.org/10.1007/s11145-019-09962-7>
- Amtmann, D., Abbott, R. D., & Berninger, V. (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., & Faisca, L. (2019). A meta-analytic review of naming-speed deficits in developmental dyslexia. *Scientific Studies of Reading, 23*(5), 349–368. <https://doi.org/10.1080/10888438.2019.1572758>
- Araújo, S., Inacio, F., Francisco, A., Faisca, L., Petersson, K. M., & Reis, A. (2011). Component processes subserving rapid automatized naming in dyslexic and non-dyslexic readers. *Dyslexia, 17*(3), 242–255. <https://doi.org/https://doi.org/10.1002/dys.433>
- Araújo, S., Reis, A., Petersson, K. M., & Faisca, 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>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Arkansas Department of Education. (2017). *Dyslexia Resource Guide*. Arkansas Department of Education.

http://dese.ade.arkansas.gov/public/userfiles/Learning_Services/Dyslexia/DRG-Final-12-13-17-JS1.pdf

*Badian, N. A. (1994). Preschool prediction: Orthographic and phonological skills, and reading. *Annals of Dyslexia*, 44(1), 1–25. <https://doi.org/10.1007/BF02648153>

*Badian, N. A. (1998). A validation of the role of preschool phonological and orthographic skills in the prediction of reading. *Journal of Learning Disabilities*, 31(5), 472–481. <https://doi.org/10.1177/002221949803100505>

*Badian, N. A., Duffy, F. H., Als, H., & McAnulty, G. B. (1991). Linguistic profiles of dyslexic and good readers. *Annals of Dyslexia*, 41, 221–245. <https://doi.org/10.1007/BF02648088>

*Badian, N. A., McAnulty, G. B., Duffy, F. H., & Als, H. (1990). Prediction of dyslexia in kindergarten boys. *Annals of Dyslexia*, 40, 152–169. <https://doi.org/10.1007/BF02648146>

*Biddle, K. R. (1996). *Timing deficits in impaired readers: An investigation of visual naming speed and verbal fluency*. Thesis.

Blachman, B. A., Schatschneider, C., Fletcher, J. M., Murray, M. S., Munger, K. A., & Vaughn, M. G. (2014). Intensive reading remediation in grade 2 or 3: Are there effects a decade later? *Journal of Educational Psychology*, 106(1), 46–57. <https://doi.org/10.1037/a0033663>

Bonifacci, P., & Snowling, M. J. (2008). Speed of processing and reading disability: A cross-linguistic investigation of dyslexia and borderline intellectual functioning. *Cognition*, 107(3), 999–1017. <https://doi.org/10.1016/j.cognition.2007.12.006>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- Bowers, P. G., Steffy, R., & Tate, E. (1988). Comparison of the effects of IQ control methods on memory and naming speed predictors of reading disability. *Reading Research Quarterly*, 23(3), 304–319. <https://doi.org/10.2307/748044>
- *Burgoyne, K., Malone, S., Lervag, A., & Hulme, C. (2019). Pattern understanding is a predictor of early reading and arithmetic skills. *Early Childhood Research Quarterly*, 49, 69–80. <https://doi.org/10.1016/j.ecresq.2019.06.006>
- Byrne, B., Fielding-Barnsley, R., Ashley, L., & Larsen, K. (1997). Assessing the child's and the environment's contribution to reading acquisition: What we know and what we don't know. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 265–285). Lawrence Erlbaum.
- *Caravolas, M., Lervag, A., Defior, S., Seidlova Malkova, 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>
- *Cardoso-Martins, C., & Pennington, B. F. (2004). The relationship between phoneme awareness and rapid serial naming skills and literacy acquisition: The role of developmental period and reading ability. *Scientific Studies of Reading*, 8(1), 27–52. https://doi.org/10.1207/s1532799xssr0801_3

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- *Catts, H. W. (1991). Early identification of dyslexia: Evidence from a follow-up study of speech-language impaired children. *Annals of Dyslexia*, *41*(1), 163–177.
<https://doi.org/10.1007/BF02648084>
- *Catts, H. W. (1993). The relationship between speech-language impairments and reading disabilities. *Journal of Speech and Hearing Research*, *36*(5), 948–958.
<https://doi.org/10.1044/jshr.3605.948>
- *Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading*, *3*(4), 331–361. https://doi.org/10.1207/s1532799xssr0304_2
- *Catts, H. W., Herrera, S., Nielsen, D. C., & Bridges, M. S. (2015). Early prediction of reading comprehension within the simple view framework. *Reading and Writing: An Interdisciplinary Journal*, *28*(9), 1407–1425. psych. <https://doi.org/10.1007/s11145-015-9576-x>
- *Catts, H. W., McIlraith, A., Bridges, M. S., & Nielsen, D. C. (2017). Viewing a phonological deficit within a multifactorial model of dyslexia. *Reading and Writing: An Interdisciplinary Journal*, *30*(3), 613–629. psych. <https://doi.org/10.1007/s11145-016-9692-2>
- *Catts, H. W., Nielsen, D. C., Bridges, M. S., Liu, Y. S., & Bontempo, D. E. (2015). Early identification of reading disabilities within an RTI framework. *Journal of Learning Disabilities*, *48*(3), 281–297. <https://doi.org/10.1177/0022219413498115>
- Cavanaugh, C. L., Kim, A.-H., Wanzek, J., & Vaughn, S. (2004). Kindergarten reading interventions for at-risk students: Twenty years of research. *Learning Disabilities: A Contemporary Journal*, *2*(1), 13.

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Chall, J. (1983). *Stages of Reading Development*. McGraw-Hill.

*Chiu, Y. D. (2018). The Simple View of Reading Across Development: Prediction of Grade 3 Reading Comprehension From Prekindergarten Skills. *Remedial and Special Education*, 39(5), 289–303. <https://doi.org/10.1177/0741932518762055>

*Christopher, M. E., Hulslander, J., Byrne, B., Samuelsson, S., Keenan, J. M., Pennington, B. F., DeFries, J. C., Wadsworth, S. J., Willcutt, E., & Olson, R. K. (2015). Genetic and environmental etiologies of the longitudinal relations between prereading skills and reading. *Child Development*, 86(2), 342–361. <https://doi.org/10.1111/cdev.12295>

*Cirino, P. T., Child, A. E., & Macdonald, K. (2018). Longitudinal Predictors of the Overlap between Reading and Math Skills. *Contemporary Educational Psychology*, 54, 99–111. <https://doi.org/10.1016/j.cedpsych.2018.06.002>

Connor, C. M., Alberto, P. A., Compton, D. L., & O'Connor, R. E. (2014). *Improving Reading Outcomes for Students with or at Risk for Reading Disabilities: A Synthesis of the Contributions from the Institute of Education Sciences Research Centers*. NCSER 2014-3000. National Center for Special Education Research. <https://eric.ed.gov/?id=ED544759>

*Cronin, V., & Carver, P. (1998). Phonological sensitivity, rapid naming, and beginning reading. *Applied Psycholinguistics*, 19(3), 447–461. <https://doi.org/10.1017/S0142716400010262>

*Cronin, V. S. (2013). RAN and double-deficit theory. *Journal of Learning Disabilities*, 46(2), 182–190. <https://doi.org/10.1177/0022219411413544>

Cummine, J., Chouinard, B., Szepesvari, E., & Georgiou, G. K. (2015). An examination of the rapid automatized naming–reading relationship using functional magnetic resonance imaging. *Neuroscience*, 305, 49–66. <https://doi.org/10.1016/j.neuroscience.2015.07.071>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- *Dally, K. (2006). The influence of phonological processing and inattentive behavior on reading acquisition. *Journal of Educational Psychology, 98*(2), 420–437. pdh.
<https://doi.org/10.1037/0022-0663.98.2.420>
- de Jong, P. F. (2011). What discrete and serial rapid automatized naming can reveal about reading. *Scientific Studies of Reading, 15*(4), 314–337.
<https://doi.org/10.1080/10888438.2010.485624>
- de Jong, P. F., & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific Studies of Reading, 6*(1), 51–77.
https://doi.org/10.1207/S1532799XSSR0601_03
- de Jong, P. F., & Vrieling, L. O. (2004). Rapid automatic naming: Easy to measure, hard to improve (quickly). *Annals of Dyslexia, 54*(1), 65–88. <https://doi.org/10.1007/s11881-004-0004-1>
- Denckla, M. B., & Rudel, R. G. (1976). Rapid automatized naming (R.A.N): Dyslexia differentiated from other learning disabilities. *Neuropsychologia, 14*(4), 471–479.
- *Dittman, C. K. (2016). The Impact of Early Classroom Inattention on Phonological Processing and Word-Reading Development. *Journal of Attention Disorders, 20*(8), 653–664.
<https://doi.org/10.1177/1087054713478979>
- *Ellis, N., & Large, B. (1988). The early stages of reading: A longitudinal study. *Applied Cognitive Psychology, 2*(1), 47–76. <https://doi.org/10.1002/acp.2350020106>
- *Evans, M. A., Bell, M., Shaw, D., Moretti, S., & Page, J. (2006). Letter Names, Letter Sounds and Phonological Awareness: An Examination of Kindergarten Children across Letters and of Letters across Children. *Reading and Writing: An Interdisciplinary Journal, 19*(9), 959–989. eric.

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- *Evans, M. A., Shaw, D., & Bell, M. (2000). Home literacy activities and their influence on early literacy skills. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Experimentale*, 54(2), 65–75.
- *Felton, R. H., & Brown, I. S. (1990). Phonological processes as predictors of specific reading skills in children at risk for reading failure. *Reading and Writing*, 2(1), 39–59.
<https://doi.org/10.1007/BF00383373>
- *Felton, R. H. (1992). Early identification of children at risk for reading disabilities. *Topics in Early Childhood Special Education*, 12(2), 212–229.
doi.org/10.1177/027112149201200206
- Fisher, Z., Tipton, E., & Zhipeng, H. (2017). *robumeta: An R-package for robust variance estimation in meta-analysis*.
- *Furnes, B., & Samuelsson, S. (2009). Preschool cognitive and language skills predicting kindergarten and grade 1 reading and spelling: A cross-linguistic comparison. *Journal of Research in Reading*, 32(3), 275–292. <https://doi.org/10.1111/j.1467-9817.2009.01393.x>
- *Furnes, B., & Samuelsson, S. (2011). Phonological awareness and rapid automatized naming predicting early development in reading and spelling: Results from a cross-linguistic longitudinal study. *Learning and Individual Differences*, 21(1), 85–95.
<https://doi.org/10.1016/j.lindif.2010.10.005>
- Georgiou, G. K., & Parrila, R. (2020). What mechanism underlies the rapid automatized naming–reading relation? *Journal of Experimental Child Psychology*, 194, 104840.
<https://doi.org/10.1016/j.jecp.2020.104840>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

*Georgiou, G. K., Parrila, R., & Kirby, J. (2006). Rapid naming speed components and early reading acquisition. *Scientific Studies of Reading, 10*(2), 199–220.

https://doi.org/10.1207/s1532799xssr1002_4

Georgiou, G. K., Parrila, R., Kirby, J. R., & Stephenson, K. (2008). Rapid naming components and their relationship with phonological awareness, orthographic knowledge, speed of processing, and different reading outcomes. *Scientific Studies of Reading, 12*(4), 325–

350. <https://doi.org/10.1080/10888430802378518>

Georgiou, G. K., Parrila, R., & Liao, C. H. (2008). Rapid naming speed and reading across languages that vary in orthographic consistency. *Reading and Writing, 21*(9), 885–903.

<https://doi.org/10.1007/s11145-007-9096-4>

*Georgiou, G. K., Torppa, M., Manolitsis, G., Lyytinen, H., & Parrila, R. (2012). Longitudinal predictors of reading and spelling across languages varying in orthographic consistency.

Reading and Writing: An Interdisciplinary Journal, 25(2), 321–346. psych.

<https://doi.org/10.1007/s11145-010-9271-x>

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6–10. <https://doi.org/10.1177/074193258600700104>

Griffin, J. W. (2020). *metapower: Power Analysis for Meta-Analysis* (0.2.1) [Computer software]. <https://CRAN.R-project.org/package=metapower>

Griffin, J. W. (2021). Calculating statistical power for meta-analysis using metapower. *The Quantitative Methods for Psychology, 17*(1), 24–39.

<https://doi.org/10.20982/tqmp.17.1.p024>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- *Gunn, B., Smolkowski, K., & Vadasy, P. (2010). Evaluating the Effectiveness of Read Well Kindergarten. *Journal of Research on Educational Effectiveness*, 4(1), 53–86.
<https://doi.org/10.1080/19345747.2010.488716>
- Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O’Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., & Duda, S. N. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, 95, 103208. <https://doi.org/10.1016/j.jbi.2019.103208>
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, 42(2), 377–381. <https://doi.org/10.1016/j.jbi.2008.08.010>
- *Hart, S. A., Petrill, S. A., Thompson, L. A., & Plomin, R. (2009). The ABCs of math: A genetic analysis of mathematics and its links with reading ability and general cognitive ability. *Journal of Educational Psychology*, 101(2), 388–402. pdh.
<https://doi.org/10.1037/a0015115>
- *Heath, S. M., & Hogben, J. H. (2004). Cost-effective prediction of reading difficulties. *Journal of Speech, Language, and Hearing Research*, 47(4), 751–765.
[https://doi.org/10.1044/1092-4388\(2004/057\)](https://doi.org/10.1044/1092-4388(2004/057))
- *Hecht, S. A., Burgess, S. R., Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2000). Explaining social class differences in growth of reading skills from beginning kindergarten through fourth-grade: The role of phonological awareness, rate of access, and print knowledge. *Reading and Writing*, 12(1), 99–128.
<https://doi.org/10.1023/A:1008033824385>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*, 39–65. <https://doi.org/10.1002/jrsm.5>
- Hjetland, H. N., Brinchmann, E. I., Scherer, R., & Melby-Lervåg, M. (2017). Preschool predictors of later reading comprehension ability: A systematic review. *Campbell Systematic Reviews, 13*(1), 1–155. <https://doi.org/10.4073/csr.2017.14>
- Hood, M. H. (2005). *The Role of Temporal and Phonological Processing In Early Reading Development: A Longitudinal Study*. Thesis. <https://doi.org/10.25904/1912/2673>
- *Hulme, C., Nash, H. M., Gooch, D., Lervåg, A., & Snowling, M. J. (2015). The foundations of literacy development in children at familial risk of dyslexia. *Psychological Science, 26*(12), 1877–1886. <https://doi.org/10.1177/0956797615603702>
- Humphrey, N., & Mullins, P. M. (2002). Self-concept and self-esteem in developmental dyslexia. *Journal of Research in Special Educational Needs, 2*(2), 1–13. <https://doi.org/10.1111/j.1471-3802.2002.00163.x>
- *Inoue, T., Georgiou, G. K., Parrila, R., & Kirby, J. R. (2018). Examining an extended home literacy model: The mediating roles of emergent literacy skills and reading fluency. *Scientific Studies of Reading, 22*(4), 273–288. doi.org/10.1080/10888438.2018.1435663
- Johnston, T. C., & Kirby, J. R. (2006). The contribution of naming speed to the simple view of reading. *Reading and Writing, 19*(4), 339–361. <https://doi.org/10.1007/s11145-005-4644-2>
- 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>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- *Kirby, J. R., Georgiou, G. K., Martinussen, R., Parrila, R., Bowers, P. G., & Landerl, K. (2010). Naming speed and reading: From prediction to instruction. *Reading Research Quarterly*, 45(3), 341–362. <https://doi.org/10.1598/RRQ.45.3.4>
- *Kirby, J. R., Parrila, R. K., & Pfeiffer, S. L. (2003). Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology*, 95(3), 453–464. <https://doi.org/10.1037/0022-0663.95.3.453>
- *Lachance, J. A., & Mazzocco, M. M. M. (2006). A longitudinal analysis of sex differences in math and spatial skills in primary school age children. *Learning and Individual Differences*, 16(3), 195–216. <https://doi.org/10.1016/j.lindif.2005.12.001>
- 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>
- *Ledesma, H. M. L. (2002). *Language factors influencing early reading development in bilingual (Filipino-English) boys*. Thesis.
- Lervåg, A., & Hulme, C. (2009). Rapid Automatized Naming (RAN) Taps a Mechanism That Places Constraints on the Development of Early Reading Fluency. *Psychological Science*, 20(8), 1040–1048. <https://doi.org/10.1111/j.1467-9280.2009.02405.x>
- *Lewis, B. A., Avrich, A. A., Freebairn, L. A., Hansen, A. J., Sucheston, L. E., Kuo, I., Taylor, H. G., Iyengar, S. K., & Steina, C. M. (2011). Literacy outcomes of children with early childhood speech sound disorders: Impact of endophenotypes. *Journal of Speech*,

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- Language, and Hearing Research*, 54(6), 1628–1643. psych. [https://doi.org/10.1044/1092-4388\(2011/10-0124\)](https://doi.org/10.1044/1092-4388(2011/10-0124))
- Logan, J. A. R., Schatschneider, C., & Wagner, R. K. (2011). Rapid serial naming and reading ability: The role of lexical access. *Reading and Writing*, 24(1), 1–25.
<https://doi.org/10.1007/s11145-009-9199-1>
- Lovett, M. W., Frijters, J. C., Wolf, M., Steinbach, K. A., Sevcik, R. A., & Morris, R. D. (2017). Early intervention for children at risk for reading disabilities: The impact of grade at intervention and individual differences on intervention outcomes. *Journal of Educational Psychology*, 109(7), 889–914. <https://doi.org/10.1037/edu0000181>
- *Macdonald, H. H., Sullivan, A. L., & Watkins, M. W. (2013). Multivariate screening model for later word reading achievement: Predictive utility of prereading skills and cognitive ability. *Journal of Applied School Psychology*, 29(1), 52–71.
<https://doi.org/10.1080/15377903.2013.751476>
- Manis, F. R., Doi, L. M., & Bhadha, B. (2000). Naming speed, phonological awareness, and orthographic knowledge in second graders. *Journal of Learning Disabilities*, 33(4), 325–333. <https://doi.org/10.1177/002221940003300405>
- *Mazzocco, M. M. M., & Grimm, K. J. (2013). Growth in rapid automatized naming from grades K to 8 in children with math or reading disabilities. *Journal of Learning Disabilities*, 46(6), 517–533. <https://doi.org/10.1177/0022219413477475>
- McBride–Chang, C., & Kail, R. V. (2002). Cross–Cultural Similarities in the Predictors of Reading Acquisition. *Child Development*, 73(5), 1392–1407.
<https://doi.org/10.1111/1467-8624.00479>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- McBride-Chang, C., & Manis, F. R. (1996). Structural invariance in the associations of naming speed, phonological awareness, and verbal reasoning in good and poor readers: A test of the double deficit hypothesis. *Reading and Writing: An Interdisciplinary Journal*, 8(4), 323–339. <https://doi.org/10.1007/BF00395112>
- *McIlraith, A. L. (2018). Predicting word reading ability: A quantile regression study. *Journal of Research in Reading*, 41(1), 79–96. <https://doi.org/10.1111/1467-9817.12089>
- *McMillen, S. M. (2018). Pescado or fish? Rapid automatic naming performance for young Spanish-speaking English language learners. Thesis.
- Menghini, D., Finzi, A., Benassi, M., Bolzani, R., Facoetti, A., Giovagnoli, S., Ruffino, M., & Vicari, S. (2010). Different underlying neurocognitive deficits in developmental dyslexia: A comparative study. *Neuropsychologia*, 48(4), 863–872. <https://doi.org/10.1016/j.neuropsychologia.2009.11.003>
- Meyer, M. S., Wood, F. B., Hart, L. A., & Felton, R. H. (1998). Selective predictive value of rapid automatized naming in poor readers. *Journal of Learning Disabilities*, 31, 106–117. <https://doi.org/10.1177/002221949803100201>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). Preferred reporting Items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- National Early Literacy Panel. (2008). *Developing Early Literacy: Report of the National Early Literacy Panel: A Scientific Synthesis of Early Literacy Development and Implications of Intervention*. National Institute for Literacy.
- Nayar, K., Gordon, P. C., Martin, G. E., Hogan, A. L., La Valle, C., McKinney, W., Lee, M., Norton, E. S., & Losh, M. (2018). Links between looking and speaking in autism and

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- first-degree relatives: Insights into the expression of genetic liability to autism. *Molecular Autism*, 9(1), 51. <https://doi.org/10.1186/s13229-018-0233-5>
- Norton, E. S. (2020). What educators need to know about rapid automatized naming (RAN). *LDA Bulletin*, 52(1), 25–28.
- Norton, E. S., Beach, S. D., Eddy, M. D., McWeeny, S., Ozernov-Palchik, O., Gaab, N., & Gabrieli, J. D. E. (2021). ERP Mismatch Negativity Amplitude and Asymmetry Reflect Phonological and Rapid Automatized Naming Skills in English-Speaking Kindergartners. *Frontiers in Human Neuroscience*, 15, 624617. <https://doi.org/10.3389/fnhum.2021.624617>
- Norton, E. S., Black, J. M., Stanley, L. M., Tanaka, H., Gabrieli, J. D. E., Sawyer, C., & Hoefft, F. (2014). Functional neuroanatomical evidence for the double-deficit hypothesis of developmental dyslexia. *Neuropsychologia*, 61, 235–246. <https://doi.org/10.1016/j.neuropsychologia.2014.06.015>
- Norton, E. S., Gabrieli, J. D. E., & Gaab, N. (2019). Neural predictors of dyslexia. In L. Verhoeven, C. Perfetti, & K. Pugh (Eds.), *Developmental Dyslexia across Languages and Writing Systems: A Handbook* (pp. 253–276). Cambridge University Press.
- Norton, E. S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology*, 63(1), 427–452. <https://doi.org/10.1146/annurev-psych-120710-100431>
- Ozernov-Palchik, O., & Gaab, N. (2016). Tackling the ‘dyslexia paradox’: Reading brain and behavior for early markers of developmental dyslexia. *WIREs Cognitive Science*, 7(2), 156–176. <https://doi.org/10.1002/wcs.1383>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- *Ozernov-Palchik, O., Norton, E. S., Sideridis, G., Beach, S. D., Wolf, M., Gabrieli, J. D. E., & Gaab, N. (2017). Longitudinal stability of pre-reading skill profiles of kindergarten children: Implications for early screening and theories of reading. *Developmental Science*, *20*(5). <https://doi.org/10.1111/desc.12471>
- Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, *40*(2), 184–202. <https://doi.org/10.1598/RRQ.40.2.3>
- *Parrila, R., Kirby, J. R., & McQuarrie, L. (2004). Articulation rate, naming speed, verbal short-term memory, and phonological awareness: Longitudinal predictors of early reading development? *Scientific Studies of Reading*, *8*(1), 3–26. https://doi.org/10.1207/s1532799xssr0801_2
- Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, *101*(2), 385–413. <https://doi.org/10.1016/j.cognition.2006.04.008>
- Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development*, *72*(3), 816–833. <https://doi.org/10.1111/1467-8624.00317>
- Pennington, B. F., Santerre-Lemmon, L., Rosenberg, J., MacDonald, B., Boada, R., Friend, A., Leopold, D., Samuelsson, S., Byrne, B., Willcutt, E. G., & Olson, R. K. (2012). Individual prediction of dyslexia by single vs. Multiple deficit models. *Journal of Abnormal Psychology*, *121*(1), 212–224. <https://doi.org/10.1037/a0025823>
- *Peterson, R. L., Arnett, A. B., Pennington, B. F., Byrne, B., Samuelsson, S., & Olson, R. K. (2018). Literacy acquisition influences children’s rapid automatized naming. *Developmental Science*, *21*(3), 1–9. [psych. https://doi.org/10.1111/desc.12589](https://doi.org/10.1111/desc.12589)

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- *Petersen, D. B., & Gillam, R. B. (2013). Accurately predicting future reading difficulty for bilingual Latino children at risk for language impairment. *Learning Disabilities Research & Practice, 28*(3), 113–128. <https://doi.org/10.1111/ldrp.12014>
- Petscher, Y., Fien, H., Stanley, C., Gearin, B., Gaab, N., Fletcher, J. M., & Johnson, E. (2019). *Screening for Dyslexia*. Improvingliteracy.Org.
- Protopapas, A., Angeli, A., & Georgiou, G. (2013). RAN backward: A test of the visual scanning hypothesis. *Scientific Studies of Reading, 17*(6), 453–461. <https://doi.org/10.1080/10888438.2013.769556>
- R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <http://www.R-project.org/>
- Richardson, J. T. E., & Wydell, T. N. (2003). The representation and attainment of students with dyslexia in UK higher education. *Reading and Writing, 16*(5), 475–503. <https://doi.org/10.1023/A:1024261927214>
- Rodgers, M., & Pustejovsky, J. (2020). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *MetaArXiv*, 1–25. <https://doi.org/10.31222/osf.io/vqp8u>
- Samuels, S. J., & Flor, R. F. (1997). The importance of automaticity for developing expertise in reading. *Reading & Writing Quarterly, 13*(2), 107–121. <https://doi.org/10.1080/1057356970130202>
- Savage, R., & Frederickson, N. (2005). Evidence of a highly specific relationship between rapid automatic naming of digits and text-reading speed. *Brain and Language, 93*(2), 152–159. <https://doi.org/10.1016/j.bandl.2004.09.005>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- *Scanlon, D. M., & Vellutino, F. R. (1996). Prerequisite skills, early instruction, and success in first-grade reading: Selected results from a longitudinal study. *Mental Retardation and Developmental Disabilities Research Reviews*, 2(1), 54–63.
[https://doi.org/10.1002/\(SICI\)1098-2779\(1996\)2:1<54::AID-MRDD9>3.0.CO;2-X](https://doi.org/10.1002/(SICI)1098-2779(1996)2:1<54::AID-MRDD9>3.0.CO;2-X)
- Scarborough, H. (1998). Predicting the future achievement of second graders with reading disabilities: Contributions of phonemic awareness, verbal memory, rapid naming, and IQ. *Annals of Dyslexia*, 48(1), 115–136. <https://doi.org/10.1007/s11881-998-0006-5>
- *Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology*, 96(2), 265–282. <https://doi.org/10.1037/0022-0663.96.2.265>
- Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an “outlier” orthography. *Psychological Bulletin*, 134(4), 584–615. [pdh. https://doi.org/10.1037/0033-2909.134.4.584](https://doi.org/10.1037/0033-2909.134.4.584)
- *Smith, S. L., Scott, K. A., Roberts, J., & Locke, J. L. (2008). Disabled Readers’ Performance on Tasks of Phonological Processing, Rapid Naming, and Letter Knowledge Before and After Kindergarten. *Learning Disabilities Research & Practice*, 23(3), 113–124.
<https://doi.org/10.1111/j.1540-5826.2008.00269.x>
- *Snowling, M. J., Lervåg, A., Nash, H. M., & Hulme, C. (2019). Longitudinal Relationships between Speech Perception, Phonological Skills and Reading in Children at High-Risk of Dyslexia. *Developmental Science*, 22(1). eric.
<http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1199876&site=ehost-live>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- Su, M., Peyre, H., Song, S., McBride, C., Tardif, T., Li, H., Zhang, Y., Liang, W., Zhang, Z., Ramus, F., & Shu, H. (2017). The influence of early linguistic skills and family factors on literacy acquisition in Chinese children: Follow-up from age 3 to age 11. *Learning and Instruction, 49*, 54–63. <https://doi.org/10.1016/j.learninstruc.2016.12.003>
- Sunseth, K., & Bowers, P. G. (2002). Rapid naming and phonemic awareness: Contributions to reading, spelling, and orthographic knowledge. *Scientific Studies of Reading, 6*(4), 401–429.
- Svensson, I., Lundberg, I., & Jacobson, C. (2001). The prevalence of reading and spelling difficulties among inmates of institutions for compulsory care of juvenile delinquents. *Dyslexia, 7*(2), 62–76. <https://doi.org/10.1002/dys.178>
- Swanson, H. L., Trainin, G., Necochea, 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. (2004). Preventing early reading failure. *American Educator, 28*(3), 6–9.
- *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>
- Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Burgess, S., & Hecht, S. (1997). Contributions of phonological awareness and rapid automatic naming ability to the growth of word-reading skills in second- to fifth-grade children. *Scientific Studies of Reading, 1*(2), 161–185. https://doi.org/10.1207/s1532799xssr0102_4

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- *Uhry, J. K. (2002). Kindergarten phonological awareness and rapid serial naming as predictors of Grade 2 reading and spelling. In E. Witruk, A. D. Friederici, & T. Lachmann (Eds.), *Basic functions of language, reading and reading disability*. 299–313. Kluwer Academic Publishers. https://doi.org/10.1007/978-1-4615-1011-6_18
- *Vadasy, P. F., & Sanders, E. A. (2008). Code-oriented instruction for kindergarten students at risk for reading difficulties: A replication and comparison of instructional groupings. *Reading and Writing, 21*(9), 929–963. <https://doi.org/10.1007/s11145-008-9119-9>
- van Bergen, E., de Jong, P. F., Plakas, A., Maassen, B., & van der Leij, A. (2012). Child and parental literacy levels within families with a history of dyslexia. *Journal of Child Psychology and Psychiatry, 53*(1), 28–36. <https://doi.org/10.1111/j.1469-7610.2011.02418.x>
- Vellutino, F. R., Scanlon, D. M., & Tanzman, M. S. (1998). The case for early intervention in diagnosing specific reading disability. *Journal of School Psychology, 36*(4), 367–397.
- *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>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive test of phonological processing, 2nd edition (CTOPP-2)*. Pro-Ed.
- Wang, K., Shen, Z., Huang, C., Wu, C.-H., Eide, D., Dong, Y., Qian, J., Kanakia, A., Chen, A., & Rogahn, R. (2019). A Review of Microsoft Academic Services for Science of Science Studies. *Frontiers in Big Data, 2*. <https://doi.org/10.3389/fdata.2019.00045>
- *Wolf, M. (1984). Naming, reading, and the dyslexias: A longitudinal overview. *Annals of Dyslexia, 34*(1), 87–115. <https://doi.org/10.1007/BF02663615>
- Wolf, M. (1991). Naming speed and reading: The contribution of the cognitive neurosciences. *Reading Research Quarterly, 123*-141. <https://doi.org/10.2307/747978>
- *Wolf, M., Bally, H., & Morris, R. (1986). Automaticity, retrieval processes, and reading: A longitudinal study in average and impaired readers. *Child Development, 57*(4), 988–1000. <https://doi.org/10.1111/j.1467-8624.1986.tb00260.x>
- Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91*(3), 415. <https://doi.org/10.1037/0022-0663.91.3.415>
- Wolf, M., Bowers, P. G., & 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., & Denckla, M. B. (2005). *RAN/RAS: Rapid Automatized Naming and Rapid Alternating Stimulus Tests*. Pro-Ed.
- *Wolter, J. A., Self, T., & Apel, K. (2011). Initial mental graphemic representation acquisition and later literacy achievement in children with language impairment: A longitudinal

RAN AND READING: A LONGITUDINAL META-ANALYSIS

study. *Journal of Learning Disabilities*, 44(6), 543–555.

<https://doi.org/10.1177/0022219410392042>

Zuk, J., Dunstan, J., Norton, E. S., Yu, X., Ozernov-Palchik, O., Wang, Y., Hogan, T. P., Gabrieli, J. D. E., & Gaab, N. (2020). Multifactorial pathways facilitate resilience among kindergarteners at risk for dyslexia: A longitudinal behavioral and neuroimaging study. *Developmental Science*, e12983. <https://doi.org/10.1111/desc.12983>

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Table 1*Search Terms and Study Eligibility Criteria*

Variable	Search Terms	Inclusion Criteria	Exclusion Criteria
Initial Sample Point	"preschool*" OR "kindergart*" OR "pre-school*" OR "pre k*" OR "pre-k*" OR "prek*" OR "child*"	If no grade listed, ≤ 78 months mean age	If US/CAN sample: Called Grade 1 (or later) If UK/AUS: Called Year 2 (or later)
RAN Measure	"rapid naming" OR "naming speed" OR "rapid automat* naming" OR "RAN" OR "rapid serial naming"		If measure was labeled RAN or rapid naming, but tested naming of all 26 letters
Reading Outcome	"reading" OR "dyslexia"		If reading measure was assessed before Grade 1
Language	<i>Search criteria were not restricted by language</i>	If sample was English L1 or early/simultaneous bilinguals	If sample was drawn from L2 English immersion school
Study Design	<i>Search criteria were not restricted by study design</i>	Longitudinal study, minimum 3 months	If the study was not longitudinal OR conducted for less than 3 months OR If the study was described as a case study

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Table 2*Descriptive Statistics for Samples Included in the Meta-Analysis*

Model	N	k	n	Mean	SD	Range
Initial Timepoint						
K	47	295	8552			
Pre-K	13	51	2508			
Mixed	1	27	139			
Final Timepoint						
Grade 1	27	134	5972			
Grade 2	28	164	3902			
Grade 3	11	51	1621			
Grade 4	8	24	2050			
Time between measures	60	373	10513	27.38	11.16	12-57
RAN Task						
<i>Publication</i>						
Published/Standardized	16	86	4526			
Not	46	287	6305			
Published/Standardized						
<i>Stimuli</i>						
Alphanumeric	22	109	4425			
Non-Alphanumeric	50	255	9068			
RAN Colors	22	69	4044			
RAN Objects	29	118	5689			
RAN Letters	16	63	3196			
RAN Numbers	12	35	3232			
<i>Composition</i>						
RAN Total Items	48	297	8457	72.28	43.77	24-216
RAN Unique Items	46	288	7136	5.84	2.50	4-20
Sample Risk Proportion						
Low Risk	42	238	8528			
Medium Risk	7	72	1579			
High Risk	12	63	487			
Latent Variable(s) Used						
Yes	5	12	1809			
No	58	361	9879			

Note. N = number of samples/studies; k = number of effect sizes; n = number of participants
 The N for some sections may not sum to 10513 as a result of these factors not being mutually exclusive within a study.

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Table 3. *Intercept-only Models*

Model	N	k	I²	τ^2	r	t	df	95% CI
All Studies/Samples	60	373	74.09	.018	-.38	-22.35	50.21	[-.44 -.37]
RAN Type								
<i>Colors</i>	22	69	66.39	.012	-.32	-11.60	19.69	[-.40 -.27]
<i>Objects</i>	29	118	74.21	.012	-.34	-15.67	25.75	[-.41 -.31]
<i>Letters</i>	16	63	68.11	.017	-.46	-15.01	10.81	[-.57 -.42]
<i>Digits</i>	12	35	76.94	.015	-.45	-11.60	10.42	[-.58 -.39]
Reading Measure Types								
<i>Reading Comprehension</i>	39	87	74.43	.021	-.38	-15.91	31.05	[-.46 -.35]
<i>Reading Fluency</i>	23	54	77.84	.036	-.35	-7.95	17.60	[-.47 -.28]
<i>Single Word Reading</i>	50	193	69.30	.015	-.38	-22.28	40.59	[-.44 -.36]
Reading Measure Splits								
Single Word Reading								
<i>Real Word Reading</i>	45	109	70.24	.015	-.41	-24.43	38.85	[-.46 -.39]
<i>Nonword Reading</i>	38	84	66.59	.013	-.33	-16.05	28.48	[-.38 -.29]
Timing								
<i>Timed Reading</i>	33	137	81.27	.032	-.37	-11.70	26.86	[-.46 -.32]
<i>Untimed Reading</i>	57	223	70.25	.014	-.37	-21.78	48.48	[-.43 -.36]
Efficiency and Accuracy								
<i>Efficiency</i>	22	57	52.52	.009	-.40	-19.15	14.12	[-.47 -.38]
<i>Accuracy</i>	48	155	70.07	.015	-.37	-20.48	41.42	[-.43 -.35]

Note. N = number of samples/studies; k = number of effect sizes

All models were significant at $p < .001$.

RAN AND READING: A LONGITUDINAL META-ANALYSIS

Table 4. Moderator Effects for Practical and Theoretical Considerations

Model	N	k	I ²	τ^2	r	t	df	p	95% CI	Power
Practical Considerations										
Unique RAN Tokens	45	285	68.49	.016						
Intercept					-.34	-3.84	5.94			
Unique Tokens					-.01	-.81	4.52	.46	[-.04 .03]	.11
Total RAN Items	46	290	63.80	.012						
Intercept					-.46	-13.5	25.6			
Total Items					.00	1.19	10.7	.26	[-.00 .00]	.19
Standardized RAN Test	60	373	73.94	.019						
Intercept					-.40	-21.07	37.81			
Published/Std Test					.06	1.37	23.52	.18	[-.03 .15]	.36
Age at Assessments	60	373	73.97	.019						
Intercept					-.31	-3.28	21.25			
Initial (RAN) Age (mos.)					-.01	-1.95	19.52	.07	[-.01 .00]	.31
Final (Reading) Age (mos.)					.00	0.04	19.51	.97	[-.00 .00]	
Theoretical Considerations										
Alphanumeric vs. Non-Alphanumeric	58	364	69.45	.015						
Intercept					-.46	-11.05	14.22			
Non-Alphanumeric					.13	2.78	21.83	.01	[.03 .23]	.33
Non-Word vs. Real Word Reading	50	193	66.55	.013						
Intercept					-.33	-15.62	28.50			
Real Word Measure					-.09	-3.73	37.09	<.001	[-.14 -.04]	.51
Timed vs. Untimed Reading	58	360	72.56	.017						
Intercept					-.37	-23.08	40.97			
Timed Reading					.00	.13	31.74	.90	[-.06 .07]	.45
Efficiency vs. Accuracy	56	212	69.45	.015						
Intercept					-.38	-20.23	39.61			
Efficiency					-.01	-.035	18.83	.73	[-.09 .06]	.30
Sample Risk Proportion	60	373	74.50	.019						
Intercept					-.35	-5.62	7.18			
Low Risk					-.05	-0.78	9.19	.46	[-.20 .10]	.36
Medium Risk					-.01	-0.17	11.78	.87	[-.19 .16]	

Note. N = number of studies; k = number of effect sizes

All intercepts were significant at $p < .01$. Moderator effects indicated in **bold** are $p < .05$.

