Individual differences in reading: Separable effects of reading experience and processing skill

Peter C. Gordon[1] · Mariah Moore[2] · Wonil Choi[3] · Renske S. Hoedemaker[1] · Matthew W. Lowder[4]

Abstract

A large-scale eye-tracking study examined individual variability in measures of word recognition during reading among 546 college students, focusing on two established individual-differences measures: the Author Recognition Test (ART) and Rapid Automatized Naming (RAN). ART and RAN were only slightly correlated, suggesting that the two tasks reflect independent cognitive abilities in this large sample of participants. Further, individual variability in ART and RAN scores were related to distinct facets of word-recognition processes. Higher ART scores were associated with increased skipping rates, shorter gaze duration, and reduced effects of word frequency on gaze duration, suggesting that this measure reflects efficiency of basic processes of word recognition during reading. In contrast, faster times on RAN were associated with enhanced foveal-on-parafoveal effects, fewer first-pass regressions, and shorter second-pass reading times, suggesting that this measure reflects efficient coordination of perceptual-motor and attentional processing during reading. These results demonstrate that ART and RAN tasks make independent contributions to predicting variability in word-recognition processes during reading.

Keywords Reading · Individual differences · Eye movements · ART · RAN

Reading requires many different types of processing. At the perceptual-motor level, recognition of written text depends critically on movement of the eyes and allocation of attention in the visual field. At the language level, readers must recognize words, determine contextually appropriate meanings, and integrate those meanings according to complex syntactic, semantic, and pragmatic rules. For skilled readers, rapid performance and effective coordination of these processes allows reading to occur at the rate of four to five words a second. Less skilled readers read more slowly, and across individuals faster reading is associated with better comprehension (Ashby, Rayner, & Clifton, 2005; Perfetti, 1985, 2007). The reasons for this variation in reading skill are of both theoretical and practical interest.

At the level of the eyes, reading consists of fixations, where the eyes are relatively still for intervals ranging from about 200 ms to a good deal longer, interspersed with saccades, where the eyes move very rapidly to another position. The measurement of eye movements has yielded a great deal of fine-grained evidence about the nature of reading processes on a moment-to-moment basis (Rayner, 1998). That evidence has provided the basis for sophisticated theories about the ways in which perception, attention, motor control, and language processing are coordinated (Engbert, Nuthmann, Richter, & Kliegl, 2005; Pollatsek, Reichle, & Rayner, 2006; Reilly & Radach, 2006). While there are important differences between these theories, they agree on many conclusions. Readers move their eyes in order to bring regions of text into the fovea, the small region of the visual field with the greatest acuity. While saccadic eye movements themselves are very rapid, the eyes must fixate for a period of time while the next saccade is being programmed, a constraint that limits the overall rate of eye movements. For skilled readers, the visual word-processing system overcomes these oculomotor constraints on reading rate through attentional mechanisms that allow processing of words that are near the one that is being fixated (Rayner, 2009). Depending on the specific model of eye movement during reading, attention is thought to be distributed across

✉ Peter C. Gordon
   pcg@unc.edu

[1] Department of Psychology, CB#3270, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3270, USA

[2] Department of Psychology, University of Minnesota, Minneapolis, MN, USA

[3] Division of Liberal Arts and Sciences, GIST College, Gwangju, South Korea

[4] Department of Psychology, University of Richmond, Richmond, VA, USA

multiple adjacent words (Engbert et al., 2005) or allocated serially, preceding each saccade to a given target location (Pollatsek et al., 2006). Regardless of the specific mechanism, skilled readers are able to acquire information from upcoming words ahead of fixation, facilitating recognition of the upcoming word once it is fixated. In some cases, short or highly predictable words can be fully identified before they are fixated, in which case a planned saccade may be cancelled, causing the eyes to skip over the initial target and land on the following word (Rayner, 1998, 2009).

Close coordination between the oculomotor programming, allocation of attention and lexical processing allows the eyes to move through text quickly and efficiently. However, in most readers this rapid forward progress is interrupted at least occasionally by saccades toward earlier words in the text, with regressive eye movements making up about 10%–15% of all saccades in skilled readers (Rayner, 2009). Most regressions are simple returns to the immediately preceding word, which likely are necessitated by oculomotor targeting errors (Rayner, 1998; Vitu, McConkie, & Zola, 1998). However, in cases where reading comprehension fails to keep up with rate of forward saccades, regressions to earlier portions of the sentences or discourse may be necessary (Frazier & Rayner, 1982; Rayner, Chace, Slattery, & Ashby, 2006). Skilled readers must be able to closely coordinate reading comprehension and reading rate in order to sustain consistent forward progress through the text.

Although much has been learned about eye-movement control during reading on a global level, there exists substantial variability in the speed and efficiency with which individual readers can coordinate their eye movements to process text. Over the past decade, researchers have begun to explore the nature of this variability, typically taking the approach of correlating reading-time measures with scores on a selection of cognitive tests. Although this approach has had some success, a major challenge in conducting experiments like this is that some of the most common measures of individual differences (e.g., working-memory capacity, vocabulary, reading comprehension) tend to be highly correlated with one another (see, e.g., Kuperman & Van Dyke, 2011). This high degree of intercorrelation makes it difficult to isolate the cognitive construct of interest when attempting to explain individual variability in eye-movement control. Accordingly, the current study adopts a different approach. We collected eye-tracking data from 546 participants, along with scores on two individual-differences tasks: Rapid Automatized Naming (RAN) tasks and an Author Recognition Test (ART). These measures were chosen because they have proven quite useful in the study of reading in children and young adults, yet they appear to depend on different cognitive abilities. Whereas RAN tasks assess speeded performance in a multiprocess vocal-control and ocular-control task, the ART is a nonspeeded assessment of knowledge about literature.

In RAN tasks (Denckla & Rudel, 1974), participants are presented with a grid of familiar stimuli (drawn from small sets of letters, numbers, colors, or objects) and must name them aloud in order as quickly and as accurately as possible. RAN performance was developed to predict children's future literacy gains (Lonigan, Schatschneider, & Westberg, 2008) and to diagnose reading disorders (Denckla & Rudel, 1974; Felton, Naylor, & Wood, 1990; Wolf, 1991). In addition, performance on this task is strongly related to reading ability in typically developing children, adolescents, and adults (Arnell, Joanisse, Klein, Busseri, & Rannock, 2009; Gordon & Hoedemaker, 2016; Kuperman & Van Dyke, 2011; Powell, Stainthorp, Stuart, Garwood, & Quinlan, 2007; Swanson, Trainin, Necoechea, & Hammill, 2003). RAN performance depends heavily on sustained attention as shown by its relation to performance on other attention-demanding tasks (Arnell et al., 2009). The predictive power of RAN tasks likely results from its similarity to reading in terms of the demands associated with sequential processing of simultaneously presented stimuli, as there is little to no relationship between single item naming and reading skill (Georgiou, Parrila, Cui, & Papadopoulos, 2013; Perfetti, Finger, & Hogaboam, 1978; Protopapas, Altani, & Georgiou, 2013; Stanovich, 1981).

Fine-grained analyses of the temporal relationship between eye movements and vocalization during RAN tasks have shown that variation in performance is heavily dependent on the coordination of eye movements, perceptual encoding, and working memory. Pan, Yan, Laubrock, Shu, and Kliegl (2013) found that the eye-voice span (EVS)—the number of items that the eyes are ahead of the voice (e.g., Buswell, 1921; Levin & Addis, 1979)—was related to children's performance on the RAN digit task. Gordon and Hoedemaker (2016) further detailed the relationship of eye-voice coordination to variation in RAN performance. They found that RAN completion times were consistently predicted by a model that included both EVS and the likelihood of the eyes regressing to an earlier item. This pattern suggests that fast RAN performance is facilitated by consistently having the eyes one or more items ahead of the voice, provided that the eyes do not advance so far as to result in confusion and a need to regress. Furthermore, the combination of EVS and regressions was shown to be a better predictor of individual differences in eye-movement measures of reading than the more traditional measure of time to complete a RAN trial. Across individuals, the combination of high EVS and few regressions during RAN tasks was related to sentence reading in the form of more first-pass skipping of words, shorter gaze durations, reduced effects of word frequency, and fewer first-pass regressions (Gordon & Hoedemaker, 2016).

In the ART (Stanovich & West, 1989), participants are presented with a list of names and asked to indicate which ones they recognize as authors (half the names are authors and half are foils). While overtly a test of a very narrow kind

of knowledge, the purpose of the ART is to assess print expo-sure, the amount of *practice* that an individual has in reading (or being read to in the case of young children). The ART and other literature recognition tests such as the Title Recognition Test, assess print exposure in a way that avoids the self-presentation biases that accompany self-report measures (Mol & Bus, 2011; Stanovich & West, 1989). The premise is that knowledge of literature comes primarily from reading and therefore that knowledge of authors' names is a good indication of the amount of reading that a person does. This premise is supported by the finding that there is a strong in-verse relationship between psychometrically measured diffi-culty of author names on the ART and the frequency with which those names appear in print (Moore & Gordon, 2015), a pattern that is consistent with the idea that amount of print exposure required in order to become familiar with an author's name decreases as a function of the frequency of that name in print. With this experience, readers gain orthographic and lexical knowledge that supports fast and efficient word recognition that is based on high-quality lexical representa-tions that play an important role in skilled reading (Perfetti, 1985, 2007).

Scores on the ART are correlated with orthographic pro-cessing (Stanovich & West, 1989), vocabulary (Beech, 2002; Lewellen, Goldinger, Pisoni, & Greene, 1993; Stanovich, West, & Harrison, 1995), speed of word recognition (Chateau & Jared, 2000; Sears, Campbell, & Lupker, 2006), reading comprehension (Martin-Chang & Gould, 2008; Stanovich & Cunningham, 1992, 1993), and standardized tests of verbal ability (Acheson, Wells, & MacDonald, 2008; Hall, Chiarello, & Edmonson, 1996; Lewellen et al., 1993; Stanovich et al., 1995). The amount of variance in individual reading ability that ART accounts for within an age group increases with level of schooling from elementary school through middle school, high school, and even to students at selective colleges. In this way, print exposure is part of a *virtuous spiral* where practice at reading leads to greater read-ing skill, which in turn leads to more time spent reading (Mol & Bus, 2011).

Research has shown a number of associations between ART scores and eye-movement behavior during reading. For example, readers with higher versus lower ART scores tend to show shorter first-pass reading times (Choi, Lowder, Ferreira, & Henderson, 2015; Lowder & Gordon, 2017; Moore & Gordon, 2015; Sears et al., 2006), reduced word-frequency effects (Moore & Gordon, 2015; Sears et al., 2006), reduced lexical-repetition effects (Lowder & Gordon, 2017), and larg-er perceptual spans (Choi et al., 2015).

As tests, RAN tasks and the ART are very different, with good performance on RAN depending on sustained attention in a rapid sequential task, whereas good performance on the ART requires knowledge of literature. The present study is a large-scale effort to assess how the combination of these very different abilities contributes to variation in word-recognition during reading among college-aged adults. Specifically, eye-movement patterns during sentence reading are assessed in relation to both ART and RAN, so that individual variation in each type of skill may be linked to distinct components of lexical processing. The focus on lexical processes is not meant to deny that higher-level factors, such as sentence complexity, also influence reading comprehension and that a full account of individual differences in reading would have to take those factors into account as well. Rather, it is a choice that reflects the greater understanding of eye movements in relation to words than in relation to factors such as sentence complexity, and the heterogeneity of the language materials, aggregated across multiple individual experiments, used in the study.

Method

Participants

Five hundred forty-six students at the University of North Carolina at Chapel Hill participated in one of 16 experiments in exchange for course credit, with 32 to 52 participants per experiment. They were all native English speakers and had normal or corrected-to-normal vision. Data were collected during the regular academic year beginning in spring 2011 and ending in spring 2014. The typical interquartile ranges for SAT Critical Reading and ACT English scores for students entering the school during those years were 590–700 and 26–33, respectively ("Common Data Set 2010–11," 2011).

Tests of individual differences

Each participant completed both an ART and RAN tasks. The ART (Moore & Gordon, 2015) listed the names of 65 authors along with 65 additional names that did not refer to known authors (foils) in alphabetical order. Participants were asked to circle those names that they recognized as referring to authors, with their score calculated as the number of authors correctly selected minus the number of foils incorrectly selected. RAN tasks were adapted from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999). The letter, digit, and color RANs matched the sequence of items from of CTOPP, but were generated on a word processor so there were slight differences in spacing, fonts, and colors used. The object RANs were scanned from CTOPP and then printed out. There were two trials (Forms A and B) each for letters, digits, colors and objects. Participants were asked to read aloud the names of the 36 items on each trial as quickly and accurately as possible. The experimenter used a stopwatch to measure the completion time for each trial and noted the number of errors on each trial. Error rates on RAN tests were low ($M = 0.22$ per trial). There was one trial

which would be considered spoiled by the CTOPPs standard (greater than four errors), but this trial, which had five errors, was included in the analyses that will be reported.

Materials in sentence-reading experiments

Each experiment included 30–60 experimental sentences that were unique to that study. Eleven of the 16 experiments included a common set of 30 filler sentences, which 391 participants read. Six of the experiments included an additional 55 common filler sentences.

Procedure

Participants' eye movements were recorded with an EyeLink 1000 system (SR Research) at a sampling rate of 1000 Hz. Stimuli were presented on a 20-inch ViewSonic G225f Monitor at a distance of 61 cm with a display resolution of 1,024 × 768. The tracker was calibrated at the start of each session and recalibrated as necessary throughout the session. A chin rest was used to minimize head movements. At the start of each trial, a fixation point was presented near the left edge of the monitor, marking the location where the first word of the sentence would appear. Once the participant's gaze on this point was steady, the experimenter presented the sentence. After reading the sentence, the participant pressed a key, which made the sentence disappear and a true–false comprehension question appear. Participants pressed one key to answer "true" and another key to answer "false." Within each experiment, the sentences were presented in a different random order for each participant.

Results

Analysis

Data analysis focused on several standard eye-movement measures (see Clifton, Staub, & Rayner, 2007; Rayner, 1998). These measures are associated with different levels of language processing, with some reflecting relatively early stages of processing (word recognition and lexical access) and others relatively late stages of processing (involving difficulty integrating a word with the unfolding meaning of the sentence or corrective processing). Measures of early processing included first-pass skipping rate, gaze duration, foveal-on-parafoveal effects, and parafoveal-on-foveal effects. *First-pass skipping rate* is the proportion of trials in which a given word is not fixated at all or is only fixated after a subsequent word has been fixated. *Gaze duration* is the sum of all first-pass fixations on a word; it begins when the word is first fixated and ends when gaze is directed away from the region, either to the left or right. We also tested for potential *foveal-on-*

parafoveal and *parafoveal-on-foveal* effects by examining *single-fixation duration* on a word (i.e., the average of the durations of the initial, first-pass fixation on a word, provided that the word received only one first-pass fixation) as a function of the frequency of the preceding word (i.e., *foveal-on-parafoveal*) and the frequency of the following word (i.e., *parafoveal-on-foveal*).

Measures of later processing included first-pass regression rate and second-pass duration. *First-pass regression rate* is the proportion of trials in which a reader's first pass on a word ends with a regressive saccade to an earlier portion of the sentence instead of a progressive saccade. *Second-pass reading* refers to reading a word after the eyes have exited the right boundary of the word. There is not consensus on the best way to calculate this measure. Calculating it as second-pass duration means including many zeroes (i.e., trials where the reader did not reread the word). Alternatively, it can be treated as a binary measure depending on whether the word is reread or not. We present the results of both analyses.

The first word of each line was excluded from all analyses as this word replaced the fixation cross when the sentence was displayed; this means that there was no entry saccade to the word. The last word of each line was also excluded because it is often not fixated. In addition, for sentences that spanned two lines, only the first line was analyzed because readers tend to undershoot when executing a return saccade to the beginning of the next line, making it difficult to determine what is a first-pass fixations near the beginning of a second line; given that most of the sentences did not have many words on the second line we decided that it was best to exclude the second line completely. All function words (i.e., determiners, prepositions, conjunctions, auxiliary verbs, and pronouns) and words with four or fewer characters were excluded because they are very frequently skipped. Long words (i.e., words with more than 10 letters) were recoded as having a length of 10 letters. Word frequency was estimated using the log frequency values from the SUBTLEXus database (Brysbaert & New, 2009); proper names were excluded because estimates of their frequency are not accurate (Lowder, Choi, & Gordon, 2013). Finally, values for gaze duration that were less than 81 ms were eliminated, and values for gaze duration and second-pass duration that were greater than 2.5 standard deviations from the participant's mean were eliminated.

The effects of lexical characteristics were assessed through multilevel models (MLMs), using SAS® software (SAS Institute Inc., 2016), with random slopes by participants for word frequency and length as one level and random slopes by item for ART and RAN as a second level. The procedure PROC HPMIXED was used for continuous dependent variables (gaze duration, single-fixation duration, and second-pass duration) and PROC GLIMMIX for logistic regression on binary variables (skipping and regressions). The models included

word frequency, word length in number of letters, word position in the sentence (as a proportion of number of words in the sentence), and the nominal factor of experiment (i.e., which of the 16 experiments the measures came from).[1] They were estimated with the maximal random effects structure allowable.[2] All predictors were grand-mean centered for these and subsequent models.

Individual differences in ART and RAN

The mean score on ART (number of authors correctly identified minus number of false alarms) was 14.48 (*SD* = 6.89) with reliability of .87 by Cronbach's alpha. Table 1 shows the means, standard deviations, and percentile cut points for each of the RAN subtests for the summed times on both forms for the present data set and for the same measures for the noncollege affiliated sample studied by Kuperman and Van Dyke (2011). Comparison of these measures indicates that RAN performance was very similar for the college students in the current study and the noncollege affiliated sample in Kuperman and Van Dyke.

Table 2 displays correlations between ART scores, mean RAN completion times, and each of the four RAN subtests. As indicated in the table, consistency across RAN types was highest for Letter and Digit RANs (*r* = .84) and Object and Color RANs (*r* = .64), with lower correlations for the other pairwise comparisons.

Figure 1 shows the relationship between ART and average completion time across RAN tasks. The two measures show a very small, though statistically significant, negative correlation, $r(546) = -.086$, $p = .044$; 95% CI [−.002, −.17], indicating that there was only a very slight association between better performance on the ART and the RAN. Removal of outliers from the mean ART and RAN score did not meaningfully alter this result, and correction for attenuation in the observed correlation due to measurement error had only a small effect given the high reliability of the two tests, $r(546) = -.097$; 95% CI [−.002, −.19].

The finding that ART and RAN were unrelated in the current study is consistent with other data from our lab, but conflicts to some extent with findings from other

labs. [3] Kukona et al. (2016) tested 70 community members (mean number of years of education = 11.23) and found significant associations between a different version of the ART and some subtests of the RAN (Digit: $r = -0.43$, Letter: $r = -0.27$, and Color: $r = 0.03$). This breakdown does not correspond to the present study where a significant, though very small, correlation between ART and RAN was found only for the Letter RAN. One possible explanation for this disparity is the substantial difference in ART performance (mean of 2.07 in Kukona et al., 2016, vs. 14.48 in the present study). Moore and Gordon (2015) performed an item response theory (IRT) analysis on the Acheson et al. (2008) ART and found that the test was skewed toward distinguishing print exposure at the higher end of reading ability and was less sensitive at the lower end of performance. While Kukona et al. (2016) used a different version of the ART, the low score that they found for the test provides cause for concern when interpreting its correlation with some RAN scores. Finally, Kukona et al. (2016) found that a composite measure of print exposure (ART plus the Magazine Recognition Test) was not significantly related to a composite measure of the RAN ($r = 0.11$, *ns*).

Matsuki, Kuperman, and Van Dyke (2016) studied 51 college students and found ART scores that were similar to the ones reported here ($M = 10.47$), though their RAN times were longer than ours and those of Kuperman and Van Dyke (2011). They reported a marginally significant association of Letter RAN with ART ($r = .26$, $p = .065$) but not of Digit RAN with ART ($r = .13$, p = .363). As noted above, we also found a statistically significant, though smaller, effect for the relationship between Letter RAN and ART, $r(542) = .10$, $p = .05$. The reasons for this smaller effect in our study are not clear, but

---

[1] Individual experiments had small to moderate effects on dependent variables. Word position had large to moderate effects on most variables: gaze duration and proportion of regressions increased, whereas skipping, single fixation duration (slightly), and rereading decreased as word position increased.

[2] If not otherwise stated, the maximal random effects structure for each model included a random intercept for participants with random slopes for word frequency and length, and a random intercept for items with random slopes for ART and RAN.

[3] In total, we have collected data on RAN and this version of the ART from more than 1,000 participants (including the participants who participated in the reading experiments used in the current study). The results for this large sample are given in the Table 3. The correlation between overall RAN score and ART is now significant as are the correlations between three of the individual RAN scores and ART. The greatest magnitude of correlation is still for the Letter RAN, but this correlation accounts only for 0.86% of the variance in the relationship. We believe that this is consistent with our contention that the association between these tests is negligible in the sampled population.

Evidence that the association between RAN and ART is very weak in a different population comes from a study of cognitive aging (Gordon, Islam, & Wright, 2019), which used the same RAN task used in this study (presented on a computer screen rather than on paper), and more importantly, a different ART. The items tested in the new ART were selected with two goals: (1) to include authors who were prominent in different time periods so that some authors would likely be known to older adults (OAs) and others to younger adults (YAs), and (2) to make the test easier because an IRT analysis of the Acheson et al. (2008) authors used here showed that the test was less effective at discriminating differences at the low end of ability than at the high end (Moore & Gordon, 2015). Table 4 shows the correlations between the "new" ART and the different measures of RAN separately for YAs and for OAs. The only significant correlation was between ART score and letter RAN for the OAs (−.255).

Table 1 Performance on RAN subtests for the current (college) sample and for the noncollege sample studied by Kuperman and Van Dyke's (2011) Table 1

| Time (secs) to complete both forms	College	M	SD	25%	50%	75%	Reliability
Digit (sec)	Yes	23.3	4.4	20.2	23.0	26.1	.89
	No	23.7	4.7	23.3	26.2	30.1	
Letter (sec)	Yes	25.6	4.5	22.4	25.2	28.5	.85
	No	25.2	4.9	21.8	24.3	29.2	
Object (sec)	Yes	42.7	6.4	38.6	41.8	45.7	.72
	No	43.0	7.6	37.6	41.6	48.6	
Color (sec)	Yes	37.1	5.9	33.0	36.2	40.1	.78
	No	38.7	8.3	33.5	37.0	42.1	

Notes. The values are given as the sum of Forms A and B (the two trials) for each type of RAN, so as to be consistent with Table 1 of Kuperman and Van Dyke (2011), which tested 81 participants. Reliability was assessed using test–retest reliability across Forms A and B of each RAN subtest. Test–retest reliability on the means of the four subtests was .91

Table 2 Correlations between ART score, mean RAN completion time, and completion time for the four RAN subtests

ART score	ART score	Mean RAN time	Digit	Letter	Object	Color
	—					
Mean RAN time	−.09[*]	—				
Digit	−.08	.81[**]	—			
Letter	−.10[*]	.82[**]	.84[**]	—		
Object	−.08	.78[**]	.39[**]	.40[**]	—	
Color	−.02	.86[**]	.54[**]	.57[**]	.64[**]	—

Note. [*] $p < .05$, [**] $p < .001$, $N = 542$ for those including Color subtest (due to exclusion of color-blind participants) and 546 for all other intercorrelations

even on the basis of the Matsuki et al. (2016) results, the correlation of ART with Letter RAN accounts for less than 7% of the variance shared between the two measures; it is
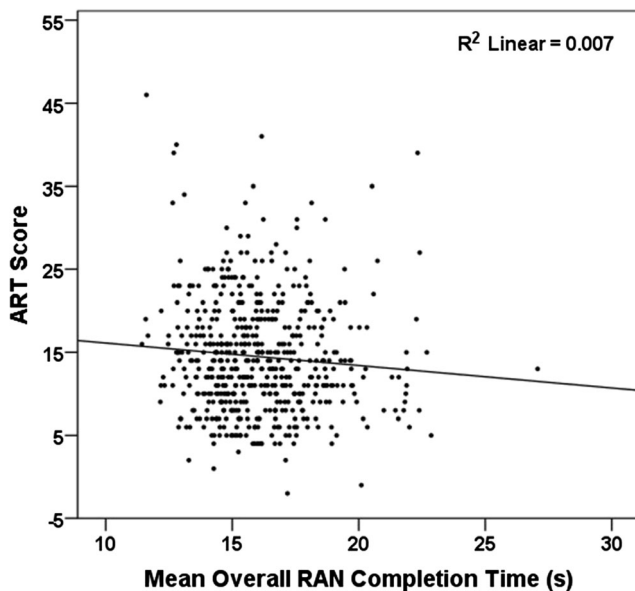
Fig. 1 Scatterplot showing relationship between ART and RAN performance per form averaged across types of RAN

unlikely that the relationship of reading measures with this shared variance would make much of a difference for the relationship between the reading measures that the ART and RAN taken separately. For those reasons we believe that it is useful to characterize the relationship between the ART and RAN as negligible for purposes of predicting individual differences in eye movements during reading, at least in the sampled population of a selective, public university.

Individual differences and eye movements during reading

Analysis of individual differences in eye movements during reading included the word-level factors of frequency and length and the subject-level factors of ART and RAN.[4] These four factors were included as fixed effects in all models regardless of whether

___
[4] Analyses performed with individual RAN tasks showed a tendency for somewhat stronger relations to eye movements for the symbolic RAN tasks (Letter and Digit) as compared with the nonsymbolic RAN tasks (Object and Color), but the pattern of differences was not consistent across eye-movement measures, and it is likely that more data would be required from each participant on eye-movements during reading and on each RAN task in order to determine whether different types of RAN tasks are associated with different patterns of eye movements during reading.

Table 3 Correlations between ART score and RAN score for a larger set of participants drawn from the same population as those tested in the present study

	Mn RAN	Letter RAN	Digit RAN	Color RAN	Object RAN
ART Score	−.075*	−.093**	−.065*	−.030	−.062*

Note. For Letter, Digit, and Object RANs, correlations were calculated over 1,009 participants including the 546 participants whose data were analyzed in the current study. For Color RAN, data were calculated for 1,002 participants after excluding those who reported being color-blind. $^*p \leq .01$, $^{**}p \leq .005$

Table 4 Correlations between RAN score and score on a new ART for a group of 79 (one color-blind) OAs (M_{age} 73.7 years, SD = 5.6) recruited from the communities surrounding Chapel Hill, NC, and Greenville, NC, and for a group of 100 YAs sampled from the same population as those tested in the present study

		M RAN	Letter RAN	Digit RAN	Color RAN	Object RAN
Score on "new" ART	OAs	−.105	−.255*	−.188	−.162	−.190
	YAs	−.095	−.043	−.081	−.112	−.085

Note. $^*p \leq .01$

their effects were significant. Interactions involving these factors were only included in the final analysis if their effects were significant ($p < .01$). For purposes of clarity, estimates of the contributions of the control variables (experiment and position of the word in a sentence) are omitted from the tables presenting the model fits. Models presented below were run with all 546 participants. Additional models run on the common filler sentences with 391 participants can be found in the Appendix. Overall, the pattern of results is similar, but many effects were less substantial most likely due to reduced ranges of lexical characteristics. Separate models are run on different eye-movement measures, making it necessary to implement precautions against the inflation of Type I error rate (von der Malsburg & Angele, 2017). We take a conservative approach and report the Bonferroni-adjusted significance level for all 12 models, including the six in the Appendix.

As shown in Table 5, skipping rate increased significantly as a function of word frequency, an effect that has been attributed to eye-movement control processes that cause a word to be skipped if it is recognized through parafoveal processing while the preceding word is being fixated (Inhoff & Rayner, 1986; White, 2008). In addition, skipping rate decreased significantly as a function of word length, an effect that has been attributed to the

limits of visual acuity in the parafovea and to systematic undershoot for long saccades (Brysbaert, Drieghe, & Vitu, 2005; Brysbaert & Vitu, 1998). Critically, there was a highly significant increase in skipping rate in association with higher ART scores, suggesting that individuals with high print exposure are better at recognizing words in the parafovea. The association between skipping rate and RAN was not significant.

As shown in Table 6, gaze duration decreased with increasing word frequency and increased with increasing word length; these consistently reported effects can be attributed to the dependence of the initiation of saccades on ease of word recognition. Gaze duration decreased as ART increased. The association with ART was moderated by highly significant interactions with the lexical factors of frequency and length; the direction of these interactions was such that an increase in ART scores was associated with a reduction in the processing difficulty caused by low-frequency and long words. Gaze duration decreased with improved (shorter) RAN completion time, but there were no significant interactions of RAN with the two lexical factors (frequency and length) for either measure.

Skipping rate and gaze duration have been consistently interpreted as reflecting the efficiency of word recognition

Table 5 Rate of word skipping

Effect	Estimate	Standard Error	t-value
Intercept	−0.23	0.16	−1.40
Log frequency	0.14	0.0165	8.32***
Length	−0.35	0.0088	−40.19***
ART	0.029	0.0042	6.75***
RAN	−0.022	0.013	−1.62

Note. $^{***}p \leq .001$, corrected for multiple comparisons. Random effects include random intercepts for participants and items

Table 6 Gaze duration

Effect	Estimate	Standard error	t value
Intercept	229.41	5.17	44.38***
Log frequency	−13.83	0.76	−18.19***
Length	5.98	0.41	14.51***
ART	−1.34	0.19	−6.96***
RAN	2.45	0.55	4.44***
Log Freq × ART	0.31	0.060	5.03***
Length × ART	−0.13	0.038	−3.33*

Note. $^*p \leq .05$, $^{***}p \leq .001$, corrected for multiple comparisons

(Clifton et al., 2007; Rayner, 1998). Therefore, this pattern suggests that RAN performance may be related to the efficiency of word recognition during reading, but that this relationship is not as strong as that of ART.

Table 7 shows the effects on single fixation duration of a word's frequency and length, as well as the frequency and length of the immediately adjacent words. Effects of the frequency of adjacent words were estimated through an analysis that additionally included both the frequency and length of the preceding and subsequent words in cases where a word and its adjacent neighbors each successively received only a single first-pass fixation. These analyses showed a significant foveal-on-parafoveal effect in that single-fixation duration on a word decreased with increasing frequency of the preceding word and further showed a significant parafoveal-on-foveal effect in that single-fixation duration decreased with increasing frequency of the subsequent word. Foveal-on-parafoveal effects have been consistently observed using both experimental manipulations of word frequency and regression methods (Henderson & Ferreira, 1990; White, Rayner, & Liversedge, 2005) and are thought to reflect the opportunity to begin recognition of a parafoveal word when recognition of the fixated word is easy, as when it is high frequency. In contrast, parafoveal-on-foveal effects have primarily been observed using regression methods (Kennedy & Pynte, 2005), as in the present study, but not experimental methods (Henderson & Ferreira, 1993), and there is disagreement over whether the effect reflects processing of the parafoveal word at the same time that the fixated word is being processed. It should be noted that for the present study the foveal-on-parafoveal effect is almost four times as large as the parafoveal-on-foveal effect.

Of primary concern are interactions between the individual differences measures and the characteristics of the preceding and following words, as such interactions would indicate individual differences in foveal-on-parafoveal processing and parafoveal-on-foveal processing. The only significant interaction of this sort was for RAN performance with the frequency

Table 7 Effects of adjacent words on single fixation duration

Effect	Estimate	Standard error	t value
Intercept	224.32	5.60	40.08***
Log frequency	−12.66	0.94	−13.33***
Length	0.65	0.45	1.44
ART	−0.70	0.20	−3.59**
RAN	2.86	0.58	4.92***
Last word freq	−10.11	0.39	−25.81***
Next word freq	−2.59	0.38	−6.61***
Length × ART	−0.14	0.43	−3.21*
Last Freq × RAN	0.43	0.13	3.23*

Note. $^{*}p \leq .05$, $^{**}p \leq .01$, $^{***}p \leq .001$, corrected for multiple comparisons

Table 8 First-pass regressions to earlier word

Effect	Estimate	Standard error	t value
Intercept	−2.29	0.15	−15.37***
Log frequency	−0.01	0.02	−0.60
Length	0.01	.01	1.05
ART	.006	.004	1.48
RAN	.066	.012	5.63***

Note. $^{***}p \leq .001$, corrected for multiple comparisons. Random effects include random intercepts for participants and items

of the preceding (last) word, with the direction of the interaction being such that better RAN performance (shorter completion time) was associated with an increase in the extent to which the time spent fixating the current word was reduced by increases in the frequency of the preceding word. In other words, better RAN performance was associated with greater beneficial foveal-on-parafoveal spillover.

Word frequency and length also had significant effects on eye-movement measures that are associated with later levels of processing. First-pass regression rate decreased with increasing word frequency ($B = -.055$, $SE = .0070$, $t = -7.81$), a decrease that can plausibly be attributed to a process wherein readers regress to earlier portions of a sentence in search of contextual or interpretive support when they encounter a difficult word. Second-pass reading time decreased with increases in word frequency ($B = -5.40$, $SE = 0.77$, $t = -7.02$) and increased with increases in word length ($B = 8.82$, $SE = 0.52$, $t = 16.84$), effects that are consistent with greater reprocessing of difficult words.

The results of the MLM analyses for the eye-movement measures of later processing are shown in Tables 8 through 10. Both the rate of first-pass regressions and second-pass reading (measured either as time spent looking at the word or the probability that the word was fixated) increased with slower RAN completion times. For second-pass reading time, the association with RAN was moderated by a significant interaction with frequency such that the extent to which second-pass time was longer for low-frequency as compared to high-frequency words was greater for those who had longer RAN completion times (see Table 9), but there was no

Table 9 Second-pass reading time

Effect	Estimate	Standard error	t value
Intercept	280.24	13.94	20.11***
Log frequency	−6.80	1.69	−4.03***
Length	11.66	0.95	12.34***
ART	0.58	0.46	1.27
RAN	5.78	1.63	3.53**
Freq × RAN	−1.00	0.32	−3.08*

Note. $^{*}p \leq .05$, $^{**}p \leq .01$, $^{***}p \leq .001$, corrected for multiple comparisons

Table 10 Second-pass reading analyzed as a binary variable

Effect	Estimate	Standard error	t value
Intercept	−0.387	0.183	−2.12
Log frequency	0.035	0.016	2.20
Length	−0.092	0.009	−10.47[***]
ART	−0.007	0.005	−1.50
RAN	−0.050	0.015	−3.39[**]

Note. [**] $p \leq .01$, [***] $p \leq .001$, corrected for multiple comparisons

significant interaction between word frequency and RAN completion times when second-pass reading was treated as a binary variable (see Table 10). Neither the rate of first-pass regressions nor second-pass reading was significantly related to ART. Together these results show that the RAN task taps variation in abilities that relate to later levels of lexical processing in the sentence-reading task which ART does not.

Discussion

The research reported here used data from a large sample of college students ($n = 546$) to examine the relationship between two established correlates of reading ability—the ART (Stanovich & West, 1989) and RAN (Denckla & Rudel, 1974)—and to determine how they are related to eye-movement measures of word recognition during reading. The results show that the correlation between scores on the ART and RAN is negligible, suggesting that the cognitive abilities assessed by the two tasks are virtually independent in the sampled population. For the ART, higher scores were associated with higher mean skipping rates, shorter mean gaze duration, and reduced effects of word frequency on gaze duration. ART scores were not significantly related to the effects of adjacent words or to the rate of first-pass regressions or the duration of second-pass reading. For RAN, better scores were associated with shorter mean gaze durations, greater influence of the frequency of the previously fixated word, fewer first-pass regressions, shorter mean second-pass reading times, and a reduced effect of word frequency on reading time. Thus, while ART and RAN scores are effectively unrelated in the current study, they both contribute successfully to the prediction of eye movements during reading. We interpret this pattern as showing the contribution to skilled reading of two types of cognitive processes, one related to efficiency in recognizing patterns of language and a second related to the coordination of pattern recognition with perception, attention, and motor control.

While we argue that there is a negligible relation between ART and RAN for our data, we do not wish to claim that this is always true, particularly when considered across development. Nonsymbolic RANs (color and object) were created so that versions of the test could be used with preliterate children (Denckla

& Rudel, 1974), and on average children ages 5 and 6 are faster on nonsymbolic than on symbolic RANs, with that relationship reversing subsequently (see Norton & Wolf, 2012, for a review). A nonnegligible relation between symbolic RAN and an appropriate measure of print exposure would be expected in a selection of children that included some who knew letters and numbers well and others who did not. At the other end of development, less data are available, but Gordon, Islam, and Wright (2019) found only a modest correlation between Letter RAN and ART in a sample of older adults (see footnote 3).

The ART is generally interpreted as measuring print exposure (amount of reading) and accordingly to reflect the knowledge of language that an individual has acquired by reading (e.g., Mol & Bus, 2011; Moore & Gordon, 2015; Stanovich & West, 1989). As discussed above, ART scores are correlated with vocabulary and with efficiency in recognizing orthographic and lexical patterns. The current study further showed that ART scores are strongly related to measures (skipping rate, mean gaze duration, and the effect of word frequency on gaze duration) that are commonly interpreted as reflecting the ease or efficiency of basic word recognition during reading (Rayner, 1998) after controlling for performance on the RAN task. The efficiency of word recognition that is found in readers with high ART scores is consistent with the idea that experience allows readers to develop high-quality lexical representations that effectively integrate orthographic knowledge with higher levels of language comprehension (Perfetti, 1985, 2007).

RAN tasks are surprisingly complex, but their value as a measure of cognitive ability depends on their sequential nature, as discrete naming tasks do not consistently predict reading performance (Georgiou et al., 2013; Perfetti et al., 1978; Stanovich, 1981). Of course, reading also requires processing a sequence of items, and some of the current results point to sources of the RAN-reading relationship. Regressive saccades during reading reflect a breakdown in the forward sequence of lexical processing, and second-pass reading time can only occur if there are regressive saccades. The current study (see also Kuperman & Van Dyke, 2011) showed that readers with poor RAN scores showed higher rates of first-pass regressions and second-pass reading durations than did readers with good RAN scores. In their analyses of eye-voice relations during RAN performance, Gordon and Hoedemaker (2016) found that rate of regressive saccades during RAN performance made significant contributions to eye-voice models developed to account for RAN completion time. As such, at least part of the success of RAN completion time in predicting the rate of regressive saccades may be due to a task-independent propensity to make regressive saccades. However, Gordon and Hoedemaker further found that an individual's rate of regressive saccades during RAN performance contributed significantly to the success of eye-voice models even when the predicted outcome was the speed of RAN performance over intervals in which no regressive saccades occurred. As such, the

rate of regressive saccades during RAN performance task can be taken as indication of an individual's susceptibility to disruptions in smooth processing even in cases where the disruption does not result in an observable regressive saccade.

The current study also found that better RAN scores were associated with shorter mean gaze durations, which is consistent with the findings of Kuperman and Van Dyke (2011), but it did not find, as they did, that RAN scores interacted with both word frequency and word length for measures of early word-recognition processes (e.g., gaze duration) and for measures of later word processing (e.g., second-pass reading). It is likely that these discrepancies are due to differences in statistical approaches taken in the two studies. As discussed in the introduction, the intercorrelations among the individual difference variables in Kuperman and Van Dyke (2011) were very high; further, Kuperman and Van Dyke assessed the relationship of individual predictors to reading measures one at a time rather than entering them all into a single model, as was done with the RAN and ART in this study. In addition, it is possible that the differences in the populations being studied contributed to the discrepancy in findings. The noncollege sample studied by Kuperman and Van Dyke (2011) may have exhibited a larger range of performance for online measures of reading than did the college sample studied here. However, performance on RAN itself was remarkably similar for the two groups in terms of measures of central tendency and distribution (see Table 1).

In addition, RAN scores showed a significant interaction on single-fixation durations with the frequency of the word on the previous fixation (see Table 7), a finding similar to the one reported by Choi et al. (2015). The interaction was such that better RAN performance was associated with greater use of the opportunity to process the word in the parafovea when the fixated word was easily recognized (i.e., a greater level of beneficial foveal-on-parafoveal spillover). This effect has been interpreted as reflecting the ability to allocate attention to the processing of the word in the parafovea, either through a serial attention-switching mechanism (Pollatsek et al., 2006; Reichle et al., 2003) or the parallel spread of attention (Engbert et al., 2005). Our detailed analysis of eye-voice relations during performance of RAN tasks has shown that the ability to keep the eyes ahead of the voice contributes to better RAN performance (Gordon & Hoedemaker, 2016), a pattern that suggests that RAN in part measures the ability to allocate perceptual and motor processing to different information as a way of increasing processing efficiency. Parafoveal preview processing requires allocating attention away from the center of fixation. The relationship between RAN performance and the extent of beneficial foveal-on-parafoveal spillover suggests that effective allocation of attention independently of the direct perceptual and motor components of eye-movement control during reading may draw on the same ability that coordinates perceptual and motor processing in RAN tasks.

The study reported here examines the correlation between ART scores and RAN performance, and the correlation of those measures with various measures of eye-movement control during reading. The finding of a negligible correlation between RAN and ART in this sample, while both had high reliability, suggests that the two measures are a good choice for explaining individual differences in the process of reading, at least in college students. Despite this, use of just these two measures is limited by the fact that performance on each is a mixture of assessment of underlying ability with task-specific behavior. For example, the ART is a test that is bound to a particular culture at a particular point in time, meaning that the version used in a study must be targeted toward the specific group being studied. As the test becomes less suitable for the individuals in a study, the portion of task specific behavior likely increases, thereby lessening the validity of the task. Even if the task is perfectly targeted for the population being studied, problems remain in interpreting correlations between ART and reading. We have for good reason (Mol & Bus, 2011; Moore & Gordon, 2015; Stanovich & West, 1989) treated the ART as a measure of print exposure, and we believe that it measures this. But ART is also correlated with various types of knowledge, such as vocabulary size (Stanovich & Cunningham, 1992), making unclear which is the effective variable during reading (or whether print exposure and vocabulary size can even be differentiated). Considerations such as these lead to the use of latent-variable approaches such as confirmatory factor analysis and structural equation modeling where patterns of the convergence and divergence of correlational patterns are discovered among multiple tests. The correlation between these latent variables to reading measures offers evidence about underlying abilities that is less contaminated by idiosyncrasies of any single individual-differences task. However, the tradeoff is clear. Current SEM models require extensive testing on a large battery of tasks (e.g., an hour and half in McVay & Kane, 2012, to 6 hours in Kane et al., 2016) as opposed to the 10 minutes required to administer the ART and RAN.

Conclusion

The results of this large-scale study are consistent with the idea that there is a great deal of variation in reading ability even among college students enrolled at a selective university, and that this variation can be fruitfully examined in relation to variation in measures of more basic cognitive processes. Two established measures of cognitive abilities, the ART and RAN, were found to be virtually unrelated in this sample, but each contributed independently to predicting characteristics of eye movements during reading. Practice at reading, as measured by ART scores, affects the speed of recognizing words and perhaps other patterns in language, whereas the processing skill measured by RAN reflects the ability to coordinate the recognition of patterns of language with the processes of perception, attention, and motor control that are essential to skilled reading.

Appendix

Table 11 Rate of word skipping with common fillers

Effect	Estimate	Standard error	*t* value
Intercept	−0.013	0.206	0.06
Log frequency	0.13	0.059	2.16
Length	−0.282	0.032	−8.71***
ART	0.025	0.0058	4.22***
RAN	−0.014	0.006	−0.81

Note. *** $p \leq .001$, corrected for multiple comparisons. Random effects include random intercepts for participants and items

Table 12 Gaze duration with common fillers

Effect	Estimate	Standard error	*t* value
Intercept	250.30	8.70	28.79***
Log frequency	−8.64	2.60	−3.32***
Length	3.33	1.37	2.43
ART	−1.50	0.25	−5.93***
RAN	3.70	0.70	5.26***
Log Freq × ART	0.36	0.13	2.69

Note. *** $p \leq .001$, corrected for multiple comparisons

Table 13 Effects of adjacent words on single fixation duration with common fillers

Effect	Estimate	Standard error	*t* value
Intercept	225.71	6.86	32.89***
Log frequency	−7.62	2.74	−2.78*
Length	−0.50	1.39	−0.36
ART	−1.07	0.27	−4.00***
RAN	3.18	0.79	4.01***
Last word freq	−8.47	1.37	−6.81***
Next word freq	−3.61	1.40	−2.58

Note. * $p \leq .05$, *** $p \leq .001$, corrected for multiple comparisons

Table 14 First-pass regressions to earlier word with common fillers

Effect	Estimate	Standard error	*t* value
Intercept	−2.22	0.23	−9.59***
Log frequency	−0.17	0.98	−1.73
Length	0.13	0.05	−2.50
ART	0.01	0.005	1.85
RAN	0.07	−0.015	4.62***

Note. *** $p \leq .001$, corrected for multiple comparisons. Random effects include random intercepts for participants and items

Table 15 Second-pass reading time with common fillers

Effect	Estimate	Standard error	*t* value
Intercept	179.48	18.89	9.50***
Log frequency	−4.05	4.49	−0.90
Length	3.89	2.43	1.60
ART	0.50	0.54	0.92
RAN	7.06	1.62	4.35***

Note. *** $p \leq .001$, corrected for multiple comparisons

Table 16 Second-pass reading analyzed as a binary variable with common fillers

Effect	Estimate	Standard error	*t* value
Intercept	−0.09242	0.222	−0.42
Log frequency	0.010	0.064	0.87
Length	−0.029	0.035	0.41
ART	−0.011	0.006	0.06
RAN	−0.068	0.018	−3.81**

Note. ** $p \leq .01$, corrected for multiple comparisons

References

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods, 40*, 278–289.

Arnell, K. M., Joanisse, M. F., Klein, R. M., Busseri, M. A., & Tannock, R. (2009). Decomposing the relation between rapid automatized naming (RAN) and reading ability. *Canadian Journal of Experimental Psychology, 63*, 173–184.

Ashby, J., Rayner, K., & Clifton, C., Jr. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology, 58A*, 1065–1086.

Beech, J. R. (2002). Individual differences in mature readers in reading, spelling, and grapheme-phoneme conversion. *Current Psychology, 21*, 121–132.

Brysbaert, M., Drieghe, D., & Vitu, F. (2005). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), Cognitive processes in eye guidance (pp. 53–77). Oxford, England: Oxford University Press.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American-English. *Behavior Research Methods, 41*, 977–990.

Brysbaert, M., & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), Eye guidance in reading and scene perception (pp. 125–148). Oxford, England: Elsevier.

Buswell, G. T. (1921). The relationship between eye-perception and voice-response in reading. *Journal of Educational Psychology, 12*, 217–227.

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory & Cognition, 28*, 143–153.

Choi, W., Lowder, M. W., Ferreira, F., & Henderson, J. M. (2015). Individual differences in the perceptual span during reading: Evidence from the moving window technique. *Attention, Perception, & Psychophysics, 77*, 2463–2475.

Clifton, C., Jr., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), Eye movements: A window on mind and brain (pp. 341–371). Amsterdam, Netherlands: Elsevier.

Common Data Set 2010–11. (2011). Retrieved from http://oira.unc.edu/files/2012/03/cds_2010_2011.pdf

Denckla, M. B., & Rudel, R. (1974). Rapid "automatized" naming of pictured objects, colors, letters and numbers by normal children. *Cortex, 10*, 186–202.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*, 777–813.

Felton, R. H., Naylor, C. E., & Wood, F. B. (1990). Neuropsychological profile of adult dyslexics. *Brain and Language, 39*, 485–497.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178–210.

Georgiou, G. K., Parrila, R., Cui, Y., & Papadopoulos, T. C. (2013). Why is rapid automatized naming related to reading? *Journal of Experimental Child Psychology, 115*, 218–225.

Gordon, P. C., & Hoedemaker, R. S. (2016). Effective scheduling of looking and talking during rapid automatized naming. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 742–760.

Gordon, P. C., Islam, A. T., & Wright, H. H. (2019). *Rapid automatized naming (RAN): Effects of aging on a predictor of reading skill.* Manuscript under review.

Hall, V. C., Chiarello, K. S., & Edmonson, B. (1996). Deciding where knowledge comes from depends on where you look. *Journal of Educational Psychology, 88*, 305–313.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 417–429.

Henderson, J. M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47*, 201–221.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40*, 431–439.

Kane, M. J., Meier, M. E., Smeekens, B. A., Gross, G. M., Chun, C. A., Silvia, P. J., & Kwapil, T. R. (2016). Individual differences in the executive control of attention, memory, and thought, and their associations with schizotypy. *Journal of Experimental Psychology: General, 145*(8), 1017–1048.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45*, 153–168.

Kukona, A., Braze, D., Johns, C. L., Mencl, W. E., Van Dyke, J. A., Magnuson, J. S., . . . Tabor, W. (2016). The real-time prediction and inhibition of linguistic outcomes: Effects of language and literacy skill. *Acta Psychologica, 171*, 72–84.

Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language, 65*, 42–73.

Levin, H., & Addis, A. B. (1979). The eye-voice span. Cambridge, MA: MIT Press.

Lewellen, M. J., Goldinger, S. D., Pisoni, D. B., & Greene, B. G. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization. *Journal of Experimental Psychology: General, 122*, 316–330.

Lonigan, C., Schatschneider, C., & Westberg, L. (2008). Identification of children's skills and abilities linked to later outcomes in reading, writing, and spelling. In *Developing early literacy: Report of the National Early Literacy Panel* (pp. 55–106) Washington, DC: National Institute for Literacy.

Lowder, M. W., Choi, W., & Gordon, P. C. (2013). Word recognition during reading: The interaction between lexical repetition and frequency. *Memory & Cognition, 41*, 738–751.

Lowder, M. W., & Gordon, P. C. (2017). Print exposure modulates the effects of repetition priming during sentence reading. *Psychonomic Bulletin & Review, 24*, 1935–1942.

Martin-Chang, S. L., & Gould, O. N. (2008). Revisiting print exposure: Exploring differential links to vocabulary, comprehension and reading rate. *Journal of Research in Reading, 31*, 273–284.

Matsuki, K., Kuperman V., & Van Dyke, J. A. (2016). The Random Forests statistical technique: An examination of its value for the study of reading. *Scientific Studies of Reading, 20*(1), 20–33.

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General, 141*(2), 302–320.

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin, 137*, 267–296.

Moore, M., & Gordon, P. C. (2015). Reading ability and print exposure: Item response theory analysis of the Author Recognition Test. *Behavior Research Methods, 47*, 1095–1109.

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*, 427–452.

Pan, J., Yan, M., Laubrock, J., Shu, H., & Kliegl, R. (2013). Eye-voice span during rapid automatized naming of digits and dice in Chinese normal and dyslexic children. *Developmental Science, 16*, 967–979.

Perfetti, C. A. (1985). Reading ability. New York, NY: Oxford University Press.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*, 357–383.

Perfetti, C. A., Finger, E., & Hogaboam, T. (1978). Sources of vocalization latency differences between skilled and less skilled young readers. *Journal of Educational Psychology, 70*, 730–739.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the EZ reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*, 1–56.

Powell, D., Stainthorp, R., Stuart, M., Garwood, H., & Quinlan, P. (2007). An experimental comparison between rival theories of rapid automatized naming performance and its relationship to reading. *Journal of Experimental Child Psychology, 98*, 46–68.

Protopapas, A., Altani, A., & Georgiou, G. K. (2013). Development of serial processing in reading and rapid naming. *Journal of Experimental Child Psychology, 116,* 914–929.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457–1506.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading, 10*, 241–255.

Reilly, R. G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research, 7*, 34–55.

SAS Institute Inc. (2016). Base SAS® 9.4 procedures guide. Cary, NC: SAS Institute Inc.

Sears, C. R., Campbell, C. R., & Lupker, S. J. (2006). Is there a neighborhood frequency effect in English? Evidence from reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1040–1062.

Stanovich, K. E. (1981). Relationships between word decoding speed, general name-retrieval ability, and reading progress in first-grade children. *Journal of Educational Psychology, 73*, 809–815.

Stanovich, K. E., & Cunningham, A.E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition, 20*, 51–68.

Stanovich, K. E., & West R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*, 402–433.

Stanovich, K. E., West, R. F., & Harrison, M. R. (1995). Knowledge growth and maintenance across the life span: The role of print exposure. *Developmental Psychology, 31*, 811–826.

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

Vitu, F., McConkie, G. W., & Zola, D. (1998). About regressive saccades in reading and their relation to word identification. In G. Underwood (Ed.), Eye guidance in reading and scene perception (pp. 101–124). Oxford, England: Elsevier Science.

von der Malsburg, T., & Angele, B. (2017). False positives and other statistical errors in standard analyses of eye movements in reading. *Journal of Memory and Language*, *94*, 119–133.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). Comprehensive test of phonological processing: CTOPP. Austin, TX: PRO-ED.

White, S. J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 205–223.

White, S. J., Rayner, K., & Liversedge, S. P. (2005). The influence of parafoveal word length and contextual constraint on fixation durations and word skipping in reading. *Psychonomic Bulletin & Review, 12*, 466–471.

Wolf, M. (1991). Naming speed and reading: The contribution of the cognitive neurosciences. *Reading Research Quarterly, 26*, 123–141.