2022

RAPID AUTOMATIZED NAMING (RAN) AND REAL WORD READING FLUENCY: AN EYE-TRACKING STUDY

Alexia Martins
University of Rhode Island, alexia_martins@uri.edu

RAPID AUTOMATIZED NAMING (RAN) AND REAL WORD READING FLUENCY: AN EYE-TRACKING STUDY

BY

ALEXIA MARTINS

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

SPEECH-LANGUAGE PATHOLOGY

UNIVERSITY OF RHODE ISLAND

2022

MASTER OF SCIENCE THESIS

OF

ALEXIA MARTINS

APPROVED:

Thesis Committee:

Major Professor Alisa Baron

Vanessa Harwood

LeAnne Spino-Seijas

Rabia Hos

Brenton DeBoef
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2022

ABSTRACT

Although literacy skills, as typically measured by reading accuracy and fluency, are known to greatly influence a child's later academic and social success, national literacy rates continue to decrease across grade levels. Research has shown that serial rapid automatized naming (RAN) is one of the best predictors of reading fluency; however, there is significant variability regarding the specific relationship between RAN and reading as noted between meta-analyses, age groups, RAN stimuli, and measures of socioeconomic status (SES). This study aims to investigate the relationship between serial RAN and 1) real word reading fluency and 2) socioeconomic status (SES) across thirty-three first and second grade participants. Specifically, RAN was measured by rapid letter naming (RLN) and rapid digit naming (RDN) raw scores, and real word reading fluency was determined utilizing the eye-tracking measure of gaze duration. To prevent bias in determining SES based on one measure alone, SES was calculated in three ways: mother's education, free and/or reduced lunch services, and a composite measure of parental education and occupation. The results indicated that RAN and gaze duration are not correlated when the first and second grade participants are combined in analyses. However, when analyzed separately, gaze duration and RAN were correlated for the second grade only. It is possible that correlations were not significant in first grade as there was a smaller sample size and their reading skills were weaker than second grade students. It is possible that RAN may be a more useful measure of reading fluency only when students become more skilled readers. In research, it is common that first and second grades are grouped together and labeled as "beginning readers." To gain a better understanding of RAN's correlation with reading

fluency across grade levels it may be best to separate first and second grades in future studies as there is significant variability in reading skills between these two grades. As both RLN and RDN were significantly correlated with our measure of reading fluency (gaze duration) for second grade participants, RLN and RDN may be useful diagnostic tools within a comprehensive evaluation of reading fluency for students who are more skilled readers. Lastly, the results indicate that RAN and SES were not correlated across all three indicators of SES. It is possible that the small and relatively homogenous sample size may negatively impact the correlation strength. The results of the insignificant RAN-SES correlations across all three measures suggest that RAN may be unrelated to SES. Therefore, it is possible that RAN is a more innate ability rather than a skill that can be influenced by external factors like SES. However, future research with larger and more heterogeneous sample sizes is necessary.

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to my major professor, Dr. Alisa Baron for imparting her knowledge and expertise throughout the entirety of this project. Her reassuring presence, gentle guidance, and continuous dedication have been invaluable to my professional development as a clinician-researcher. I would also like to acknowledge my thesis committee members, Dr. Vanessa Harwood, Dr. LeAnne Spino-Seijas, and Dr. Rabia Hos for their dedication and collaboration throughout this project. I would also like to thank my parents, avós, and fiancé for never letting me doubt myself through their continuous love and encouragement. Lastly, I want to thank the participants and families without whom this project would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE PAGE

CHAPTER 1

INTRODUCTION

Childhood language and literacy skills are crucial in predicting later educational success and social outcomes (Sheridan et al., 2011). In Rhode Island, 48% of third graders did not meet the English Language Arts expectations on the Rhode Island Comprehensive Assessment System in 2019 (Greenwood, 2020). The reading scaled scores on the National Assessment of Educational Progress decreased from 2017 to 2019 in Rhode Island along with sixteen additional states for fourth graders and thirty additional states for eighth graders (National Center for Education Statistics, 2020). Since national literacy rates continue to decrease across grade levels, implementing reliable and successful literacy diagnostic and intervention strategies is critical for educational and social success.

Literacy assessment and intervention tend to focus on reading accuracy and reading fluency. Haager et al. (2014) define reading accuracy as a student's ability to recognize words automatically or with little effort in decoding. In contrast, reading fluency is the student's ability to decode with "speed and quality, involving minimal conscious effort from cognitive and linguistic processes" (Norton & Wolf, 2012, p. 429). There are multiple predictors of reading accuracy and fluency including phonological awareness, phonological working memory, and rapid automatized naming (RAN) (Kirby et al., 2003).

Phonological processing, which includes phonological awareness, phonological working memory, and RAN, involves one's ability to analyze, recall, and manipulate phonological units (Rezaei & Mousanezhad Jeddi, 2019). More specifically, phonological awareness consists of detecting, discriminating, and manipulating sounds of words through tasks including but not limited to blending sounds, rhyming, and segmenting words into onsets and rimes (Oakhill & Kyle, 2000). Phonological working memory is one's ability to hold phonological information in short-term memory, commonly assessed through nonword repetition tasks (White, 2021). Phonological awareness and working memory have been linked to reading development and achievement (e.g., Ehri et al., 2001); however, researchers have argued that "…phonological awareness and naming speed have different predictive roles at different points in reading development" (Kirby et al., 2003, pp. 1). Therefore, the focus of this study is on the relationship between naming speed and reading fluency to better understand RAN's relationship in reading development in the first and second grades.

RAN

RAN is the ability to accurately name familiar items as quickly as possible (Denckla & Rudel, 1976). RAN is also considered to be under the umbrella of phonological processing but independent of phonological awareness (Swanson et al., 2003). Items within a RAN task can be presented in an array (i.e., serial RAN) or one at a time (i.e., discrete-naming). Although naming speed has been assessed in both formats, discrete-naming tasks are seen as inconsistent predictors of reading skills (Georgiou et al., 2013; Logan et al., 2011). Furthermore, serial RAN tasks mimic the left-to-right serial reading process and are strongly related to reading (Araújo & Faísca, 2019). Therefore, serial RAN will be the focus of this study to determine the nature of its relationship with reading fluency.

Norton (2012) explains that a serial RAN task must include: 1) an array of systemically repeated items rather than single-item naming, 2) familiar items to the reader, such as letters, digits, colors, or objects, and 3) accuracy measured through naming time. RAN has been established as a strong predictor of reading across symbolic and non-symbolic systems (e.g., Cohen-Mimran et al., 2021; Kirby et al., 2003; Pan et al., 2011; Vassen et al., 2010; Verhagen et al., 2010); however, the specifics of how RAN best predicts reading are still unclear. It is relevant to note that improving RAN should not be a treatment method. In a longitudinal study, Lervag and Hulme (2009) concluded that there is no evidence suggesting that reading fluency improves from direct growth of RAN, reinforcing that RAN should not be implemented in intervention. A meta-analysis by Araújo et al. (2019) suggests that difficulties in naming speed are permanent in those with reading disorders and challenging to improve. Although RAN might not be useful as a reading intervention strategy, current research specific to RAN's relationship with reading fluency skills can serve as a diagnostic tool.

Eye-tracking

Eye-tracking allows researchers to analyze forward eye movements (saccades), backward eye movements (regressions), and stops (fixations) when reading (see Rayner, 1978; 1998). Conklin and Pellicer-Sánchez (2016) explain how eye-tracking technology is versatile and reliably investigates visual word processing beyond the level of the word. Both eye-tracking reading experiments and RAN tasks provide naturalistic examples of reading that require similar cognitive and linguistic processing, including visual object recognition and speech production processes (Norton, 2012). More specifically, Gordon & Hoedemaker (2016) explain that RAN is a measure of lexical retrieval because successful performance depends on the speed and accuracy of "visual recognition of individual stimuli, access to phonological codes, and rapid articulation" (p. 742). It is suggestive that serial RAN and oral reading tasks share fundamental neurolinguistic, visual to verbal processing (Gordon & Hoedemaker, 2016). Therefore, this study appropriately analyzes the unique relationship between serial RAN and eye-tracking measures of oral reading.

CHAPTER 2

REVIEW OF LITERATURE

RAN is one of the best predictors of reading fluency and an early identifier of reading difficulties (Norton & Wolf, 2012); however, there is significant variability regarding the specific relationship between RAN and reading as noted between meta-analyses and across readers' ages, measures of socioeconomic status (SES), and measures of RAN stimuli (either letters, digits, colors, objects, or a combination of more than one; e.g., Araújo et al., 2015; Kirby et al., 2003; Ozernov-Palchik et al., 2017).

Meta-analyses

A meta-analysis by Araújo et al. (2015) states that a moderate-to-strong relationship exists between RAN and real word reading fluency. On the contrary, the Swanson et al. (2003) meta-analysis concludes that RAN best predicts pseudoword reading and spelling rather than real word reading and reading comprehension. The discrepancies between these two meta-analyses may result from differences in focus. Swanson et al. (2003) included 35 studies that targeted both phonological awareness and RAN, whereas Araújo et al. (2015) included 137 studies and focused strictly on RAN. More recently, another meta-analysis by Araújo and colleagues (2019), states that there is no significant difference between reading measures (word reading fluency, pseudoword reading fluency, and reading accuracy) and RAN abilities in predicting reading disorders (i.e., dyslexia vs. typical reading). The inconsistencies across conclusions from meta-analyses presents the need to further study the correlation between RAN and real word reading fluency to determine their true relationship.

Age

The reader's age is one variable to consider when assessing the relationship between RAN and reading fluency. Araújo et al. (2015) state that age did not significantly impact the relationship between RAN and real word reading fluency; however, Swanson et al. (2003) state that there is a positive correlation between grade level and RAN/real-word reading suggesting that the RAN to real word reading fluency relationship is stronger with increased reading education. More specifically, Wolff (2014) reports that explicit phonics training (i.e., training in the alphabetic principle, or the relationship between letters and spoken sounds) plays a role in increasing RAN's reliability in predicting reading fluency in later reading years. Similarly, Kirby et al. (2003) suggest that naming speed best predicts reading skills in later years possibly due to the change in reading from phonetic to orthographic processes. On the contrary, researchers conclude that 60 to 75% of children with reading and/or learning difficulties demonstrate RAN task deficits (Katzir et al., 2008; Waber et al., 2004; Wolf et al., 2002) with the strongest correlation between kindergarten and second grade. Since the first and second grades are often grouped together in the current RAN literature (e.g., Araújo et al., 2015; Katzir et al., 2008; Waber et al., 2004), additional research focusing on the first and second grades individually is necessary to determine any potential differences in the RAN to real-word reading fluency relationship across critical reading fluency grade levels.

Socioeconomic Status (SES)

Another variable to consider is SES. SES can be defined as “an individual’s or household’s relative position in a social hierarchy” (Hoff & Ribot, 2015, p. 324) with a focus on access to wealth, power, or prestige (Mueller & Parcel, 1981). SES is often considered a measure of language input quality (De Cat, 2021). SES can be reported through various measures including parental education, parental occupation, free and/or reduced lunch services, family income, or a combination of more than one. Researchers often use mother’s education as the sole measure of SES and have argued that maternal education is the most stable and reliable parental measure in predicting child outcomes (e.g., Bornstein et al., 2003; Hoff et al., 2012). However, Vanderauwera et al. (2019) more recently suggest that including both maternal and paternal highest levels of education is the best predictor of reading success. Lastly, financial capital, commonly measured by parental occupation, is a reliable measure of SES by providing an understanding of a family’s social network and available resources (Bradley & Corwyn, 2002). Although maternal education, paternal education, and parental occupation are individually associated with a child’s developmental success, Hoff and Ribot (2015) suggest that multiple indicators of SES lead to greater predictive ability. Similarly, De Cat (2021) concludes that a composite measure of SES including both parental education and occupation is the most effective in predicting language proficiency in bilingual school-aged children in comparison to individual measures of parental education and occupation. Therefore, similar to the procedure completed by Gathercole et al. (2016), a composite score of education level and occupation for both parents will be calculated to measure one measure of SES in this study.

Although SES factors directly correlate with a child's linguistic performance (Ginsborg, 2006), there are inconsistencies across RAN literature regarding which SES measures have been used. Liu and Georgiou (2017) used a Home Literacy Environment questionnaire, mother's education, and mother's occupation as separate measures to conclude that informal home literacy measures do not show a strong correlation with RAN abilities in their 141 Chinese kindergarteners. In a New England study, Ozernov-Palchik et al. (2017) grouped participants based on free and/or reduced lunch provided at school. Researchers concluded that the majority of students who show deficits in RAN were only in the low SES group in comparison to those with RAN and phonological awareness deficits. Due to the inconsistencies across individual and composite measures, both individual measures of SES (mother's education and free and/or reduced lunch services) and a composite score of parental education and occupation will be targeted in this study.

RAN Stimuli

Lastly, differences in RAN stimuli may contribute to the inconsistencies seen in the RAN to real word reading fluency relationship. Wood et al. (2017) used letters, digits, colors, and objects stimuli to determine that RAN abilities correlated with language and literacy skills. However, Araújo et al. (2015) conclude that alphanumeric/symbolic stimuli (letters and digits) correlate best with reading fluency rather than non-alphanumeric/non-symbolic stimuli (colors and objects). Furthermore, Gordon and Hoedemaker (2016) conclude through an eye-tracking experiment that serial alphanumeric stimuli are more strongly associated with automaticity and the encoding to articulation reading process, possibly due to the smaller and more closed sets from which letters and digits come from. Specific to letters and digits, Vaessen and Blomert (2010) suggest that rapid digit naming (RDN) best predicts reading fluency between first and sixth grade rather than rapid letter naming (RLN). On the contrary, Araújo et al. (2015) conclude that the correlation between RLN and reading performance was slightly, yet not significantly higher than RDN. Therefore, letters and digits will be the stimuli used in this study to determine how they are individually associated with reading fluency within our first and second grade population.

The commonly used standardized measures to test RAN abilities are *Rapid Automatized Naming and Rapid Alternating Stimulus Tests* (RAN-RAS; Wolf & Denckla, 2005), as well as subtests within the following tests: the *Comprehensive Test of Phonological Processing-2nd Edition* (CTOPP-2; Wagner et al., 2013), the *Clinical Evaluation of Language Fundamentals-5th Edition* (CELF-5; Wiig et al., 2013), the *Kaufman Test of Educational Achievement-Third Edition (*KTEA-3*;* Kaufman A. & Kaufman N., 2014), the *Woodcock-Johnson IV Tests of Achievement* (Schrank, 2014), and the *Woodcock Reading Mastery Tests* (WRMT-3; Woodcock, 2011). Araújo et al. (2015) state that the format of the RAN task, including the total number of RAN items or the number of different tokens presented, did not have a significant impact on the correlation between RAN and reading fluency. Therefore, the CTOPP-2, which contains both RLN and RDN subtests, was the measure chosen for this study.

The inconsistencies discussed across meta-analyses as well as the variability in reader's age, SES, and RAN stimuli might mediate the relationship between RAN and real word reading fluency. For example, the differences between Liu and Georgiou (2017) and Ozernov-Palchik et al. (2017) regarding the RAN/SES relationship may stem from the inconsistencies in how SES had been measured (mother's education and occupation vs. free and/or reduced lunch services) and not their intended outcomes; therefore, suggesting that additional research is necessary to determine RAN's true association with SES and real word reading fluency skills. Therefore, this current study will uniquely incorporate rapid letter and digit naming raw scores, multiple measures of SES, and an eye-tracking experiment.

Real word reading fluency will be assessed through an eye-tracking experiment that gathers a more fine-grained, natural, and quantifiable record of word reading behavior in comparison to traditional measures (Conklin & Pellicer-Sánchez, 2016). Eye-tracking technology has been widely used to examine both child and adult reading skills (e.g., Araújo et al., 2020; Kim et al., 2020; Tiffin-Richards & Schroeder, 2015); however, this study is unique by individually comparing an eye-tracking measure of real word reading fluency to rapid letter and digit naming across the first and second grade population.

In a recent eye-tracking study examining RAN abilities in adult readers with dyslexia, researchers used *gaze duration*, or the total duration of fixations in a target region before progressing or regressing to another target region as their eye-tracking measure of fluency due to its association with recognition and activation of phonological codes (Araújo et al., 2020). In another recent study analyzing RAN abilities in child readers, researchers also used *gaze duration* as a measure of reading fluency (Kim et al., 2020). Kim et al. (2020) explain that *gaze duration* "is thought to reflect (earlier) processes including orthographic processing up to lexical access…" (p. 4). Measurements of early processing are essential to gain a better understanding of literacy development in beginning readers (Kim et al., 2020). Therefore, *gaze duration* will be the eye-tracking measure used to analyze real word reading fluency in this study.

Research Questions:

1. Regarding *real word reading fluency* measured within an eye-tracking task in first and second graders:
    a. What is the relationship with *rapid letter naming*?
    b. What is the relationship with *rapid digit naming*?
2. Does the relationship between *real word reading fluency* and:
    a. *rapid letter naming* differ from first to second grade?
    b. *rapid digit naming* differ from first to second grade?
3. What is the relationship between *socioeconomic status* (SES) and:
    a. *rapid letter naming*?
    b. *rapid digit naming*?

CHAPTER 3

METHODOLOGY

Participants

With the support of the Rhode Island IDeA Network of Biomedical Research Excellence (RI-INBRE), on-site research was conducted in Newport, Rhode Island in the Fall of 2019. The study included 44 monolingual ($n=35$) and bilingual ($n=9$) students in first and second grade. All students participated in behavioral language and literacy assessments and an eye-tracking reading experiment. The bilingual participants were not included in the analyses. The 35 monolingual ($n=27$) and functionally monolingual ($n=8$) students are the focus of this study with an average age of 6.74 years ($SD = 0.65$). A functionally monolingual student is operationalized as one who has less than 20% exposure and use of another language (Spanish in this case) and therefore was unable to complete behavioral language testing in Spanish (Bedore et al., 2006). One participant was excluded due to participation in special education services and an additional participant was excluded due to the inability to complete the RAN task per standardized protocol. Therefore, a total of 33 students in the first ($n=14$) and second ($n=19$) grades were included in the study.

Behavioral Procedure

Students participated in behavioral assessments to measure nonverbal IQ, language, and literacy skills. The tests administered were the *Clinical Evaluation of Language Fundamentals-5th Edition* (CELF-5; Wiig et al., 2013), the *Comprehensive*

Test of Phonological Processing-2nd Edition (CTOPP-2; Wagner et al., 2013), *the Dynamic Indicators of Basic Early Literacy Skills* (DIBELS; University of Oregon, 2018), the *Receptive One-Word Picture Vocabulary Test* (ROWPVT-4; Martin & Brownell, 2011), and the *Wechsler Abbreviated Scale Intelligence-2nd Edition* (WASI-2; Wechsler, 2011).

Comprehensive Test of Phonological Processing-2nd Edition (CTOPP-2)

The CTOPP-2 is a standardized assessment that targets phonological awareness, phonological memory, and phonological access/naming for reading development between the ages of five and twenty-four. The subtests administered were: Elision, Blending Words, Sound Matching, Phoneme Isolation, Nonword Repetition, Memory for Digits, Rapid Digit Naming, and Rapid Letter Naming. There are two different forms, one for participants between the ages of five and six (Form A) and one for ages seven through twenty-four (Form B). Thirteen participants fell between the ages of five and six and used Form A. Twenty participants fell between the ages of seven through twenty-four and used Form B. The focus of this study is phonological access through the Rapid Letter Naming (RLN) and Rapid Digit Naming (RDN) subtests that are identical in Forms A and B.

RLN and RDN subtests include four rows of nine letters/digits each, for a total of thirty-six items. A raw score is determined by the number of seconds needed to name all thirty-six items. Raw scores for both subtests are then converted into scaled scores. RLN and RDN scaled scores are then added to create a composite score of phonological access/naming. The task is no longer automatic nor accurate with four or more errors (Norton, 2019), so trials with four or more errors in accuracy of letter/digit naming were not included in the analysis, resulting in the exclusion of 1 participant in Grade 2. Although RLN and RDN can be analyzed with composite scores (e.g., Wood et al., 2017), raw scores are more often used (e.g., Araújo et al., 2015; Kirby et al., 2003; Lervåg & Hulme, 2009; Liu & Georgiou, 2017; Ozernov-Palchik et al., 2017) because they allow for more fine-grained analysis and do not group naming times. Therefore, RLN and RDN raw scores will be used in this study.

Socioeconomic Status (SES) Procedure

One parent per child (*n = 30*) completed a language questionnaire (Peña et al., 2018) to gather information regarding demographics, including the education and occupation statuses of both parents. The Hollingshead Four-Factor Index of Socioeconomic Status (Hollingshead, 1975) was used to code parental education on a seven point scale (7 = graduate/professional training, 6 = standard college or university graduation, 5 = partial college, at least one year of specialized training, 4 = high school graduate, 3 = partial high school, 10th or 11th grade, 2 = junior high school, including 9th grade, 1 = less than 7th grade, 0 = not applicable or unknown) and parental occupation on a nine point scale (9 = higher executive, proprietor of large businesses, major professional, 8 = administrators, lesser professionals, proprietor of medium-sized business, 7 = smaller business owners, farm owners, managers, minor professionals, 6 = technicians, semi-professionals, small business owners (business valued at $50,000-70,000), 5 = clerical and sales workers, small farm and business owners (business valued at $25,000-50,000), 4 = smaller business owners (< $25,000), skilled manual laborers, craftsmen, tenant farmers, 3 = machine operators and semi-skilled workers, 2 = unskilled workers, 1 = farm laborers, menial service workers, students, housewives, (dependent on welfare, no regular occupation), 0 = not applicable or unknown). Due to the inconsistencies in measuring SES across the current research, this study will use three separate measures of SES: mother's education, free and/or reduced lunch services, and a composite measure of SES. These three measures will be individually correlated with rapid letter and digit naming to determine any potential differences in the relationships between the SES measures and RAN.

Mother's Education Procedure

The first individual measure used to calculate SES is mother's education. The Hollingshead Four-Factor Index of Socioeconomic Status (Hollingshead, 1975) education levels for birth mothers ($n = 28$) will be individually correlated with rapid letter and digit naming. Conventionally, the education levels of birth mothers are exclusively calculated as mother's education. Therefore, the two nontraditional families within this study were excluded only from this analysis.

Free and/or Reduced Lunch Services Procedure

The second individual measure used to calculate SES is free and/or reduced lunch services provided at school. Parents were asked to report the lunch service provided at school, if applicable within the previously mentioned language questionnaire (Peña et al., 2018). Students that received free lunch ($n = 7$) were scored as a (1). Students that received reduced lunch ($n = 1$) were scored as a (2). Students that received neither free nor reduced lunch ($n = 22$) were scored as a (3). These scores were then individually correlated with RLN and RDN.

Composite Measure Procedure

Similar to the procedure completed by Gathercole et al. (2016), the composite measure of SES was classified by the addition of the Hollingshead Four-Factor Index of Socioeconomic Status (Hollingshead, 1975) education and occupation levels. Specifically, education and occupation levels were added to create a composite score and averaged per family. Each parent's education level and occupation scores were added together to create a composite score with a maximum of sixteen per parent. For the 7 participants with single parents, SES was measured by the one parent's composite score of education and occupation. For the remaining families with two parents, SES was measured by averaging the composite scores of both parents.

Apparatus

Eye movements were recorded using an EyeLink Portable Duo with a sampling rate of 500 Hz. Stimuli sentences were presented on a 17.3 inch PC computer. Stimuli sentences were presented in Times New Roman in black, size 20, on a white background. Participants sat at a viewing distance between 500 and 700 mm.

Eye-Tracking Task

A sentence-level reading task was researcher created. Twenty sentences with embedded target words that were familiar to early school-age children were created. The target words did vary in word frequency, but none were so infrequent that children were unfamiliar with them (*M* = 6374436.75 words, *SD* = 4206701.88). One practice trial, with the target word "book" was followed by twenty sentences with embedded target words. Sentences were either five or ten words in length. Two versions were created. Target words in list 1 included: king, cow, rain, sea, window, space, tree, umbrella, farm, cotton, dog, horse, cloud, lake, corn, fish, ear, cat, plane, and nose. Target words in list 2 included: bed, bell, ice, paint, sun, oil, mouth, apple, soil, wave, mountain, bear, farmer, factory, clock, bird, yard, hat, milk, and grass. The average word frequencies across versions did not differ. Each target was embedded within a five- or ten-word sentence but only appeared once in each version of the experiment. One version was administered to each participant and versions were counterbalanced across participants. Real word reading fluency was measured by the gaze duration of the target words listed above.

Eye-Tracking Procedure

The experiment was conducted in a windowless classroom. Participants were tested in one session after school hours following the completion of language and literacy testing to ensure a quiet environment. The experiment was created using Experiment Builder software (SR Research Experiment Builder 1.10.165). A nine-point calibration and validation of the eye tracker were conducted for each participant prior to testing. Recalibration and validation were completed if there was an x- or y-axis drift detected. Participants were instructed to read the sentences aloud and were audio recorded. Each trial was preceded by a drift check and fixation point which triggered the presentation of the stimulus sentence and was followed by a yes-no comprehension question presented auditorily through headphones. Participants responded to the question by pressing the right or left trigger button on a gamepad.

Data Cleaning Procedure

Frequently, data cleaning is not discussed within the methods section of peer-reviewed articles. Conklin and Pellicer-Sánchez (2016) explain the importance of including data cleaning procedure information for transparency and replicability. The data cleaning was conducted in Data Viewer (EyeLink Data Viewer 3.2.1) and was cleaned in several stages. Trials were removed if the participant did not correctly read the target word aloud, resulting in the exclusion of 15.15% of trials. The data set was then reparsed with the "PSYCHOPHYSICAL" parameter settings. Event reparsing modifies the settings with more "conservative thresholds" for reading (EyeLink® data viewer user's manual, section 5.1). The data was then cleaned using a 3-stage cleaning. Godfroid (2020) explains "…that short fixations [50 to 100 ms] do not reflect cognitive processing … [and] it is better to merge or remove short fixations in applied research" (pp. 261). Therefore, fixations of less than 80 ms and within 0.5° were merged with neighboring fixations, fixations of less than 40 ms and within 1.25° were merged with neighboring fixations, and any remaining fixations less than 80 ms were excluded from the analysis. The data was then drift-corrected to their average Y position to account for poor calibrations, participant movements, or drifts in gaze positions (Godfroid, 2020). Drift corrections were first completed automatically and then manually, if necessary. Manual drift corrections were required if one or more fixation positions exceeded the batch drift correction threshold of 30. Any trials with less than three fixations were removed from the analysis, which resulted in the exclusion of 7.29% of trials. Lastly, any trials with data loss were excluded, resulting in the exclusion of 2.43% of trials. Hessels and Hooge (2019) refer to “data loss” as track loss due to participants’ head turns, blinks, and/or the eye tracker’s “technical difficulties” (p. 2).

For this current study, we focused on one eye-tracking measure of gaze duration as the metric for real word reading fluency of the target words. Gaze duration was compared individually with RLN and RDN raw scores.

CHAPTER 4

FINDINGS

Prior to conducting any statistical analyses regarding the research questions, preliminary analyses were completed. First, a histogram of the gaze duration raw data revealed that the data did not meet the normal distribution criteria needed to conduct traditional analyses due to a positive skew (see Figure 1). As Godfroid (2020) explains, eye-tracking reading data tends to be skewed and researchers must "...address skew in their data in order to satisfy the normality assumption of parametric tests and safeguard statistical power" (p. 263). Therefore, a logarithmic transformation, common in eye-tracking research (e.g., Eder et al., 2020; Parshina et al., 2021) was utilized to meet the normality assumption (see Figure 2) and prepared the data for correlation analyses. All proceeding correlation analyses were completed using the log transformed data. Means and standard deviations of all behavioral and eye-tracking measures for each grade are presented in Table 1.

Figure 1. Histogram of raw gaze duration data

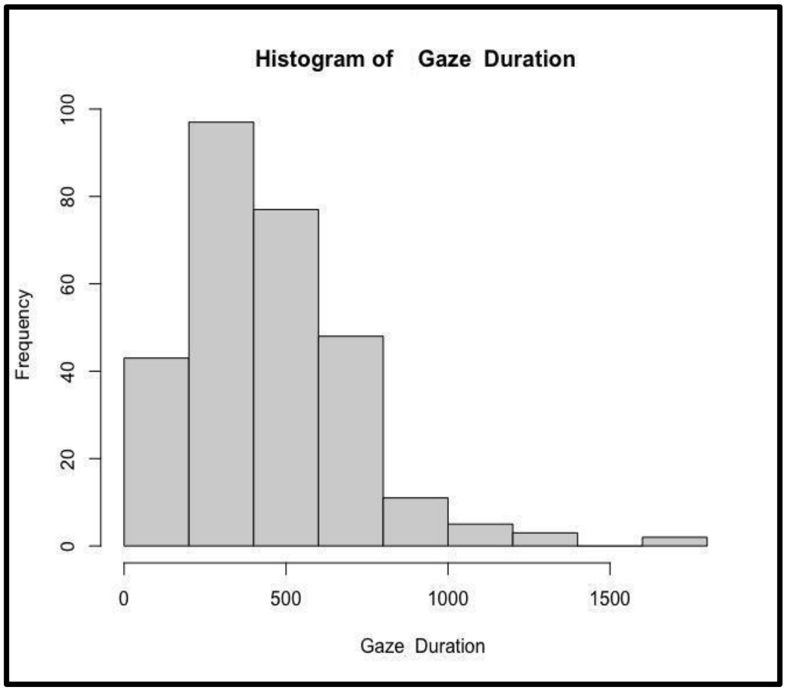

Figure 2. Histogram of log transformed gaze duration data

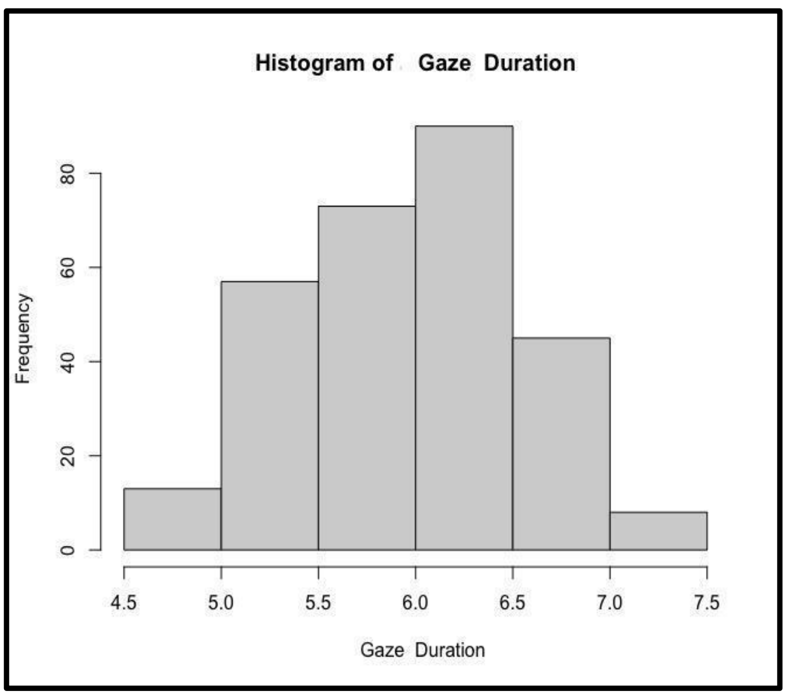

Table 1. Means (M) and standard deviations (SD) for all Rapid Automatized Naming and eye-tracking measures by grade

	RLN		*RDN*		*Gaze Duration Raw Data*		*Log transformed Gaze Duration*	
Grade	M	SD	M	SD	M	SD	M	SD
1st	28.93	7.84	25.79	5.49	409.79	98.35	5.84	.27
2nd	24.05	5.93	21.00	4.58	468.34	87.86	5.96	.17

Note. RLN = Rapid Letter Naming; RDN = Rapid Digit Naming; RLN and RDN measured in seconds; Gaze Duration measured in milliseconds

Secondly, the correlation between RLN and RDN was calculated. The results of the Pearson correlation indicate that RLN and RDN are highly correlated ($r(31) = .79$, $p < .001$). In addition, a t-test was conducted to determine any potential differences between gaze duration and RAN across target words embedded within five- and ten-word sentences. The results indicate that there is no statistical difference in target word gaze duration between five- and ten-word sentences ($t(61.94) = -.56$, $p = .57$). Therefore, five- and ten-word sentences were combined in subsequent gaze duration correlation analyses. Lastly, a series of t-tests were completed to better understand the relationship between the first and second grades across the measured variables. A RDN t-test reveals that there is a significant difference in naming speed between the first and second grades ($t(25.00) = 2.65$, $p = .01$). However, a t-test reveals that the difference in RLN between the first and second grades is only approaching significance ($t(23.28) = 1.95$, $p = .06$). Lastly, a t-test reveals that the difference in gaze duration between the first and second grades is also approaching significance ($t(26.23) = -1.77$, $p = .09$). Results of all correlation analyses for the following research questions are presented and summarized in Table 2.

Research Question 1

1a - What is the relationship between rapid letter naming and real word reading fluency in first and second graders?

To determine the relationship between RLN and real word reading fluency across both first and second grades, a Pearson correlation was conducted. The results indicate that the correlation between gaze duration and RLN is insignificant (r(31) = -.04, *p* = .81).

1b - What is the relationship between rapid digit naming and real word reading fluency in first and second graders?

Similar to RLN, the relationship between RDN and real word reading fluency was calculated through a Pearson correlation. The results indicate that the correlation between RDN and gaze duration is insignificant (r(31) = -.08, *p* = .65).

Research Question 2

2a - Does the relationship between real word reading fluency and rapid letter naming differ from first to second grade?

The relationship between real word reading fluency and RLN across grade levels was calculated through Pearson correlations. The results indicate that the relationship between RLN and gaze duration differs from first to second grade. Specifically, the correlation between RLN and gaze duration for first grade participants is insignificant (r(12) = -.33, *p* = .25). However, the correlation between RLN and gaze duration for second grade participants is strong (r(17) = .64, *p* = .003).

2b - Does the relationship between real word reading fluency and rapid digit naming differ from first to second grade?

Similar to RLN, the correlations between gaze duration and RDN differ from first to second grade. Specifically, the correlation between gaze duration and RDN for first grade participants is insignificant (r(12) = -.34, p = .23). However, the correlation between gaze duration and RDN for second grade participants is moderate-strong (r(17) = .58, p = .009).

Research Question 3

The correlation between mother's education and the composite measure of SES is strong (r(26) = .73, $p < .001$). Similarly, the composite measure of SES and lunch services provided at school are moderately correlated (r(28) = .48, p = .007). In addition, mother's education and lunch services provided at school are moderately correlated (r(26) = .58, p = .001). To determine if there is a difference in SES between the first and second grades, a t-test was conducted. The results indicate that there is no significant difference in SES, as measured by the composite measure between the first and second grades (t(24.97) = 1.18, p = .25).

3a - What is the relationship between socioeconomic status (SES) and rapid letter naming?

To determine the relationship between RLN and mother's education, a Pearson correlation was calculated. The results indicate that the correlation between RLN and mother's education is insignificant (r(26) = -.18, p = .35). Similarly, a correlation was calculated to determine the relationship between RLN and lunch services provided at school. The results indicate that the correlation between RLN and lunch services is insignificant (r(28) = -.10, p = .59). Lastly, a correlation was calculated to determine the relationship between RLN and the composite measure of SES. The results indicate that the correlation between RLN and SES is also insignificant (r(28) = -.07, p = .70).

3b - What is the relationship between socioeconomic status (SES) and rapid digit naming?

Similar to RLN, a correlation was calculated to determine the relationship between RDN and mother's education. The results indicate that the correlation between RDN and mother's education is insignificant (r(26) = .02, p = .92). Similarly, a correlation was calculated to determine the relationship between RDN and lunch services at school. The results indicate that the correlation is also insignificant (r(28) = -.02, p = .68). Lastly, a correlation was calculated to determine the relationship between RDN and the composite measure of SES. The results indicate that the correlation is also insignificant (r(28) = .05, p = .77).

Table 2. Correlations between RLN/RDN and measures of real word reading fluency and SES

	RLN		RDN	
Grade	1st	2nd	1st	2nd
Gaze Duration	r(12) = -.33	r(17) = .64*	r(12) = -.34	r(17) = .58*
Mother's Education	r(26) = -.18		r(26) = .02	

Lunch Services	r(28) = -.10	r(28) = -.02
SES Composite	r(28) = -.07	r(28) = .05

Note. RLN = Rapid Letter Naming; RDN = Rapid Digit Naming; *$p < .01$.

CHAPTER 5

CONCLUSION

Discussion

This study aims to investigate the relationship between serial rapid automatized naming (RAN) and 1) real word reading fluency and 2) measures of socioeconomic status (SES) across thirty-three first and second grade participants. RAN is measured by rapid letter naming (RLN) and rapid digit naming (RDN), and real word reading fluency is determined utilizing an eye-tracking measure of gaze duration. In an attempt to prevent bias in determining SES based on one measure alone, SES is calculated in three ways: mother's education, free and/or reduced lunch services, and a composite measure of parental education and occupation.

Research Question 1

The results indicate that RAN, represented by RLN and RDN, did not correlate with how long it took first and second grade participants (*combined*) to read a target word, as measured by gaze duration. As discussed in Research Question 2, there were significant differences in correlations when the first and second grades analyzed individually. However, when combined, these results do not support the conclusions from the Araújo et al. (2015) meta-analysis that showed a moderate-to-strong correlation between RAN and real word reading fluency. Araújo et al. (2015) argues "…there are large variations between studies in the magnitude of the reported RAN–reading correlations. The extent to which differences in measures and samples across studies may act as moderators of this association is as yet poorly understood…[and] the variability in the RAN–reading association observed in the literature is likely related to certain specifics of the studies" (p. 878). Araújo et al. (2015) included 137 studies in their meta-analysis, allowing for large variability, possibly explaining why our results do not align with their conclusions. In addition, the discrepancies between our results and Araújo et al. (2015) may be due to the specific differences between studies, including but not limited to the types of readers included (e.g., pre-readers, intermediate readers, and/or advanced readers), methodology in measuring real word reading fluency (e.g., behavioral measures or eye tracking technology), and/or the low number of preserved total trials. It is suggested that future studies replicate previous methodologies with larger sample sizes to determine reliability of results.

Along with gaze duration correlation analyses, a qualitative count of the most frequent target words reread was determined. The nine target words that were reread in 25% or more of trials are grass, hat, soil, paint, ice, yard, factory, bed, and clock (see Table 3). Although these are high-frequency words, they proved to be more difficult for participants to read in the first attempt and required rereading to gain sentence meaning. It is suggested that future RAN-reading fluency studies utilize an eye-tracking measure of *rereading duration* in analyses. Kim et al. (2020) explain that *gaze duration* "is thought to reflect (earlier) processes including orthographic processing up to lexical access, whereas *rereading time* … is an indicator for (later) processes related to higher-level processing like syntactic integration" (p. 4). Measurements of both early and later processing are essential to include in order to gain a better understanding of literacy development in beginning readers (Kim et al., 2020). The complexity in calculating rereading duration (due to the 0 inflated nature of the raw data) goes beyond the scope of this thesis; however, should be included in future studies.

Table 3. Count of most common target words re-read

Target Word	grass	hat	soil	paint	ice	yard	factory	bed	clock
Times reread	6/10	6/13	4/11	9/27	4/12	5/16	4/13	4/14	3/12

Note. Times re-read = Number of times word was re-read/total presentations of the word (e.g., "grass; 6/10" = grass was re-read six times out of the ten eye-tracking trials preserved utilizing the target word)

Research Question 2

The results indicate that RLN and RDN were not correlated with how long it took *first* grade students to read a target word, as measured by gaze duration; however, RLN and RDN were correlated with how long it took *second* grade students to read a target word. There are a few potential reasons why differences across first and second grades occur. The first grade sample size (*n=14*) was smaller than the second grade sample size (*n=19*). Not only was the first grade group smaller, but it was also a weaker, less skilled reading group. Specifically, the first grade students had an average of 5.50 trials per participant included in gaze duration analyses. However, the second grade students had an average of 10.47 trials per participant. Therefore, raw eye-tracking data for first grade participants was not retained as well as for second grade participants. These intergroup differences may result in differences in correlation strength and significance.

Along with the correlational analyses, t-tests suggest differences across first and second grades, specifically with RDN. One of the t-test conducted comparing RDN in first and second graders suggests that second grade students can significantly name digits faster than first grade students. On the other hand, a t-test conducted comparing RLN between grades is only approaching significance. A t-test conducted comparing gaze duration between grades is also approaching significance; however, the direction contradicts the t-tests for RLN and RDN. Therefore, second grade students are behaviorally faster but slower when utilizing eye-tracking technology. It is possible that the previously discussed differences in reading skills across the first and second grades impacts the directionality of the gaze duration difference between grades. More specifically, most second graders were included in eye-tracking reading trials; however, only the stronger first grade trials were retained. Therefore, it is possible that the negative direction seen in the t-test comparing gaze duration across grades is due to an imbalance in the number of retained trials containing a range of reading skills across the grade levels. A larger sample size is necessary in future studies to better understand the RAN-gaze duration relationship across the first and second grades. In our sample, there is not a significant difference in RLN and gaze duration between the first and second grades; however, with a larger sample size, based on previous literature, it's predicted that the analyses would have been significant.

The correlation data suggests that it is possible that RAN may be a useful measure of reading fluency only when participants become more automatic readers, as seen in Swanson et al. (2003). The Swanson et al.'s (2003) meta-analysis states that there is a positive correlation between grade level and RAN/real-word reading, suggesting that the RAN to real word reading fluency relationship is stronger with increased reading education. In research, it is common that first and second grades are grouped together (e.g., Araújo et al., 2015; Katzir et al., 2008; Waber et al., 2004). More specifically, Araújo et al. (2015) combined first and second grades as "beginning readers" (pp. 872) as part of their methodology and concluded that grade did not significantly impact the relationship between reading and RAN. Therefore, it is possible that their conclusion was due to the combination of first and second grades in their analyses. To gain a better understanding of RAN's correlation with reading fluency across grade levels, the first and second grades should be separated in future studies due to the potential for reading skills to significantly vary between those grades. For example, first grade students who are still learning to decode may be considered "poor readers", or "early readers" while second grade students who can automatically decode, may be considered more "skilled readers." Therefore, first and second grades containing both poor and skilled readers should not be combined in analyses given that grade level may significantly impact reading skills (Swanson et al., 2003).

Similar to findings from Araújo et al. (2015), the relationship between gaze duration and RLN ($r = .64$) was slightly stronger than the relationship between gaze duration and RDN ($r = .58$) across second grade readers. Since the differences in correlation strength between RLN and RDN are minor, it is suggestive that both are appropriate measures to utilize when correlating reading fluency for more automatic readers. Therefore, RLN and RDN may be useful diagnostic tools within a comprehensive evaluation of reading fluency for students who are more automatic readers although future research is necessary to determine this more conclusively.

Research Question 3

The results indicate that the three measures of socioeconomic status (SES): mother's education, free and/or reduced lunch services provided at school, and a composite measure of parental education and occupation are highly correlated with each other. The results indicate that none of the measures of SES are correlated with RLN and RDN. It is possible that the small sample size (*n=30)* may negatively impact the correlation strength. In addition, there is a lack of heterogeneity within the sample. More specifically, there were twenty-two participants that received regular lunch services at school, one participant that received reduced lunch, and only seven that received free lunch. Therefore, if looking at lunch services only, roughly 73% of the sample was within the highest SES group, around 3% was in the middle, and only 23% of the sample was within the lowest SES group. These significant differences may negatively impact the correlation strength. It is recommended that future studies include larger, more heterogeneous sample sizes. It is important to note that there is no significant difference in SES between the first and second grades, unlike RAN and eye-tracking analyses. The results of the insignificant RAN-SES correlations across all three measures suggest that RAN was truly unrelated to SES in our study and conclusions are not mediated by the methodology in determining SES. Therefore, it is possible that RAN is a more innate ability rather than a skill that can be influenced by external factors (e.g., SES). However, future research with larger sample sizes is necessary.

Limitations

Due to COVID-19, the sample size for this study was limited to thirty-five monolingual and functionally monolingual participants. Researchers had intended to conduct this study as a longitudinal study with four data points across two school years, with a larger cohort. However, given the necessary modifications due to the rise of the COVID-19 pandemic, and the in-person nature of the eye-tracking experiment, the study ended after the first data point.

An additional limitation is the low number of eye-tracking trials retained in gaze duration analyses, especially for first grade participants. It is recommended that future studies that include struggling readers utilize more high-frequency target words and sentences to obtain a higher success rate using eye-tracking technology.

Future Directions

In this study, researchers chose to utilize standardized behavioral measures to calculate RAN and eye-tracking technology to analyze the fine-grained reading movements of our participants, as seen in previous studies (Gordon & Hoedemaker, 2016). However, to the best of our knowledge, eye-tracking technology has not been utilized to measure RAN while reading skills are measured by standardized behavioral tests like the *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS). The DIBLES is a clinically used standardized assessment that determines a student's reading status (University of Oregon). Some of the DIBLES subtests include oral reading fluency (ORF), word reading fluency (WRF), nonsense word fluency (NWF), and letter naming fluency (LNF). Future studies may want to validate commonly used eye-tracking measures (e.g., gaze duration) with standardized behavioral assessments, like the DIBLES to determine if reading fluency conclusions align across both. Each measure has its strengths and weaknesses (e.g., retention of eye-tracking trials, scoring on behavioral measures, etc.), and it may benefit researchers to consider how to uniquely utilize behavioral and eye-tracking measures in future studies.

Although SES was not directly correlated with RAN in this study, the true relationship between SES and RAN is still unclear. A more consistent procedure in how to best measure SES within language and literacy research is critical for reliability and validity of results. There is significant variability in how SES has been operationalized across RAN studies (e.g., Liu and Georgiou, 2017; Ozernov-Palchik et al., 2017) and reading literature in general. The differences in *how* SES had been operationalized may mediate the true conclusions made. As seen in Table 4, there are individual differences in how a participant in this study would be categorized (e.g., high SES vs. low SES) across the three indicators of SES. For example, participant 9 has a lunch services score of 3 (out of 3) which falls in the highest SES group but a composite score of 4.25 (out of 7.50) which falls in the lowest SES group within our participant sample. In this study, researchers chose to utilize all three indictors of SES in the correlation analyses to prevent bias; however, many studies only utilize one indicator of SES (e.g., Bornstein et al., 2003; Liu and Georgiou, 2017; Ozernov-Palchik et al., 2017). Therefore, if each study operationalizes SES in a unique way, how can we truly conclude results are not influenced by *how* SES is operationalized? In this study, we chose to remove the two nontraditional families when conducting mother's education analyses due to uncertainty in how to include those participants. However, American household composition continues to change, including an increasing number of nontraditional families (VanOrman & Jacobsen, 2020). Therefore, how can we include each participant in future SES research with respect to the diversity in family structure? A comprehensive analysis of SES measures in relation to language and literacy proficiency with special consideration to cultural and linguistic diversities would help to answer these questions and create a more consistent procedure for measuring SES. De Cat (2021) argues that future research is critically necessary to "unveil the actual SES-related dimensions that affect children's language development" (p. 321). Hoff et al. (2012) also discuss the negative consequences of variability in SES methodology and that "...large-scale longitudinal databases, combined with sophisticated modeling techniques are beginning to untangle these intertwined and bidirectional influences [of SES]" (p. 603). However, to conduct reliable and valid SES research moving forward, more research attempts to determine the most appropriate way in determining SES methodology is essential.

Table 4. SES scores for all participants across mother's education, lunch services, and a composite measure of parental education and occupation

Participant	Mother's Education (out of 7)	Lunch Services (out of 3)	Composite Measure of Parental Education and Occupation (out of 7.50)
1	5	1	4.25
2	6	3	5.25
3	--	--	--
4	--	--	--
5	7	3	7.25

6	6	3	5.50
7	6	1	4.00
8	6	3	6.75
9	5	3	4.25
10	5	3	5.25
11	7	3	7.00
12	6	3	4.75
13	5	2	3.25
14	5	3	6.00
15	7	3	7.50
16	7	3	5.75
17	7	3	7.50
18	5	3	3.50
19	7	3	7.50
20	6	3	4.75
21	6	3	4.75
22	6	3	5.25
23	N/A	3	5.75
24	6	3	5.25
25	6	3	6.50
26	5	3	5.50
27	5	1	6.50
28	3	1	3.00
29	4	1	4.50
30	6	1	4.50
31	4	1	3.50
32	--	--	--
33	N/A	3	4.25

Note. Mother's education (3 = partial high school/10th or 11th grade, 4 = high school graduate, 5 = partial college or specialized training, 6 = standard college graduate, 7 = graduate professional training). Lunch services (1 = free lunch, 2 = reduced lunch, 3 = full priced lunch). -- = SES data was not collected. N/A = nontraditional families.

Possible Clinical Implications

Previous studies have established RAN as a strong predictor of reading (e.g., Cohen-Mimran et al., 2021; Kirby et al., 2003; Pan et al., 2011; Vassen et al., 2010; Verhagen et al., 2010) and early identifier of reading difficulties (Norton & Wolf, 2012). However, there are significant inconsistencies across the literature regarding the strength of the RAN-to-reading relationship across grade levels (e.g., Araújo et al., 2015; Swanson et al., 2003; Katzir et al., 2008; Waber et al., 2004; Wolf et al., 2002). The results of this study revealed that RAN is significantly correlated with reading fluency only for second grade students. Therefore, it is possible that RAN should only be utilized as a reading fluency diagnostic tool for older, more skilled readers. Given that phonological awareness and phonological working memory have been established as a prelinguistic skill linked to reading development and achievement for younger students (e.g., Ehri et al., 2001), it is possible that phonological awareness should be utilized as reading diagnostic tool for younger, less skilled readers (e.g., kindergarten and first grade) while RAN should be utilized for older, more skilled readers (e.g., second grade). This study's results are correlational and not causal; therefore, more advanced statistical analyses are needed to determine RAN's *predicative* ability and further strengthen the clinical implications discussed here. However, this study's correlation analyses provide some practical implications that align with previous research. Results from Kirby et al. (2003) indicate "…that phonological awareness and naming speed, measured in kindergarten, make independent contributions to the prediction of reading. Phonological awareness is the more powerful predictor in kindergarten and grade 1, whereas naming speed is more powerful in the later grades" (pp. 4). Given that reading skills are continuing to decrease across grade levels (National Center for Education Statistics, 2020), improving reading diagnostics to identify struggling readers as early as possible is critical for our students' academic and social successes. RAN may play a unique and significant role in reading diagnostics and future research that improves RAN's clinical implications is essential at this time.

BIBLIOGRAPHY

Araújo, S., & Faísca, L. (2019). A meta-analytic review of naming-speed deficits in developmental dyslexia. *Scientific Studies of Reading, 23*(5), 349-368. doi:10.1080/10888438.2019.1572758

Araújo, S., Huettig, F., & Meyer, A. (2020). What underlies the deficit in rapid automatized naming (RAN) in adults with dyslexia? Evidence from eye movements. *Scientific Studies of Reading*, 1-17. doi:10.1080/10888438.2020.1867863

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

Bedore, L., Fiestas, C., Peña, E., & Nagy, V. (2006). Cross-language comparisons of maze use in Spanish and English in functionally monolingual and bilingual children. *Bilingualism: Language and Cognition, 9*(3), 233-247. doi:10.1017/S1366728906002604

Bornstein, M. H., Hahn, C. S., Suwalsky, J. T. D., Haynes, O. M. (2003). Socioeconomic status, parenting, and child development: The Hollingshead four-factor index of social status and the socioeconomic index of occupations. In: Bornstein, M.H., Bradley, R.H. (Eds.), *Socioeconomic Status, Parenting, and Child Development.* Erlbaum, Mahwah, NJ, pp. 29–82.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*(1), 371-399.

doi:10.1146/annurev.psych.53.100901.135233

Cohen-Mimran, Yifat, R., & Banai, K. (2021). Size matters? Rapid automatized naming of shape sizes, reading accuracy and reading speed. *Journal of Research in Reading*, *44*(4), 882–896. doi:10.1111/1467-9817.12376

Conklin, K., & Pellicer-Sánchez, A. (2016). Using eye-tracking in applied linguistics and second language research. *Second Language Research, 32*(3)*,* 453-467. doi:10.1177/0267658316637401

De Cat, Cécile. (2021). Socioeconomic status as a proxy for input quality in bilingual children? *Applied Psycholinguistics, 42(*2), 301–324. doi:10.1017/S014271642000079X

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

Document Version, 3. 2. 1. EyeLink® data viewer user's manual.

Eder, Richter, J., Scheiter, K., Keutel, C., Castner, N., Kasneci, E., & Huettig, F. (2020). How to support dental students in reading radiographs: Effects of a gaze-based compare-and-contrast intervention. *Advances in Health Sciences Education: Theory and Practice*, *26*(1), 159–181. doi: 10.1007/s10459-020-09975-w

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36*, 250–287. doi:10.1598/RRQ.36.3.2

EyeLink Data Viewer 3.2.1 [Computer software]. (2018). Mississauga, Ontario, Canada: SR Research Ltd.

Gathercole, V. C. M., Kennedy, I., & Thomas, E. M. (2016). Socioeconomic level and bilinguals' performance on language and cognitive measures. *Bilingualism: Language and Cognition, 19*(5), 1057-1078. doi:10.1017/S1366728915000504

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*(1), 218–225. doi:10.1016/j.jecp.2012.10.015

Godfroid, A. (2020). Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide (1st ed.). Routledge.

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*(5), 742–760. doi:10.1037/xhp0000171

Greenwood, G. (2020). *2020 Rhode Island Kids Count Factbook.* Greenwood Associates.

Haager, D., Dimino J., Windmueller M. (2014). *Interventions for Reading Success: Vol. Second edition.* Brookes Publishing.

Hessels, R. S., & Hooge, I. T. C. (2019). Eye tracking in developmental cognitive neuroscience –the good, the bad and the ugly. *Developmental Cognitive Neuroscience, 40,* 1-11. doi:10.1016/j.dcn.2019.100710

Hoff, E., Laursen, B., Bridges, K. (2012). Measurement and model building in studying the influence of socioeconomic status on child development. In:

Lewis, M., Mayes, L. (Eds.), *A Developmental Environmental Measurement Handbook*. Cambridge University Press, Cambridge, England.

Hoff, E., & Ribot, K. M. (2015). Language development: Influence of socio-economic status. *International Encyclopedia of the Social & Behavioral Sciences*, *2*(13), 324-328. doi:10.1016/B978-0-08-097086-8.23132-2

Hollingshead, A. A. (1975). Four-factor index of social status. Unpublished manuscript, Yale University, New Haven, CT.

Katzir, T., Kim, Y., Wolf, M., Morris, R., & Lovett, M. W. (2008). Comparing subtypes of children with dyslexia letter, word, and connected text levels of reading. *Journal of Learning Disabilities*, *41*, 47–66. doi:10.1177/0022219407311325

Kaufman, A. S., & Kaufman, N. L. (2014). *Kaufman Test of Educational Achievement–Third Edition (KTEA-3).* Bloomington, MN: NCS Pearson.

Kim, Y., Petscher, Y., & Vorstius, C. (2020). The relations of online reading processes (eye movements) with working memory, emergent literacy skills, and reading proficiency. *Scientific Studies of Reading,* 1-20. doi:10.1080/10888438.2020.1791129

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. doi:10.1037/0022-0663.95.3.453

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. doi:10.1111/j.1467-9280.2009.02405.x

Liu, C., & Georgiou, G. K. (2017). Cognitive and environmental correlates of rapid automatized naming in Chinese kindergarten children. *Journal of Educational Psychology, 109*(4), 465–476. doi:10.1037/edu0000151

Logan, J. A. R., Schatschneider, C., & Wagner, R. K. (2011). Rapid serial naming and reading ability: The role of lexical access. *Reading and Writing: An Interdisciplinary Journal*, *24*(1), 1–25. doi:10.1007/s11145-009-9199-1

Martin, N., & Brownell, R. (2011). *Receptive One-Word Picture Vocabulary Test–Fourth Edition* (ROWPVT-4). Novato, CA: Academic Therapy Publications. Mueller, C.W., & Parcel, T.L. (1981). Measures of socioeconomic status: Alternatives and recommendations. *Child Development, 52*, 13–30. doi:10.1111/j.1467-8624.1981.tb03013.x

Norton, E. (2019). *Linking Rapid Automatized Naming (RAN), Reading, and the Brain*. [Webinar]. SPELL-Links.

Norton, E. (2020). What educators need to know about rapid automatized naming (RAN). *LDA Bulletin, 52*(1), 25-28.

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. doi:10.1146/annurev-psych-120710-100431

Oakhill, J., & Kyle, F. (2000). The relation between phonological awareness and working memory. *Journal of Experimental Child Psychology, 75*(2), 152–164. doi:10.1006/jecp.1999.2529

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), 1-18. doi:10.1111/desc.12471

Pan, J., McBride-Chang, C., Shu, H., Liu, H., Zhang, Y., & Li, H. (2011). What is in the naming? A 5-year longitudinal study of early rapid naming and phonological sensitivity in relation to subsequent reading skills in both native Chinese and English as a second language. *Journal of Educational Psychology, 103*, 897–908. doi:10.1037/ a0024344

Parshina, Laurinavichyute, A. K., & Sekerina, I. A. (2021). Eye-movement benchmarks in Heritage Language reading. *Bilingualism*, *24*(1), 69–82. doi: 10.1017/S136672892000019X

Peña, E. D., Gutierrez-Clellen, V. F., Iglesias, A., Goldstein, B. A., & Bedore, L. M. (2018). Bilingual Input/Output Survey (BIOS). Baltimore, MD: Brookes Publishing.

RStudio Team (2020). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85,* 618–60. doi:10.1037/0033-2909.85.3.618

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–442. doi:10.1037/0033-2909.124.3.372

Reading Performance. (2020). National Center for Education Statistics. Retrieved February 2021, from https://nces.ed.gov/programs/coe/indicator_cnb.asp

Rezaei, A., & Mousanezhad Jeddi, E. (2019). The contributions of attentional control components, phonological awareness, and working memory to reading ability. *Springer Science and Business Media LLC, 49*, 31-40. doi:10.1007/s10936-019-09669-4

Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV Tests of Achievement*. Rolling Meadows, IL: Riverside.

Sheridan, S. M., Knoche, L. L., Kupzyk, K. A., Edwards, C. P., & Marvin, C. A. (2011). A randomized trial examining the effects of parent engagement on early language and literacy: The getting ready intervention. *Journal of School Psychology, 49*(3), 361-383. doi:10.1016/j.jsp.2011.03.001

SR Research Experiment Builder 1.10.165 [Computer software]. (2011). Mississauga, Ontario, Canada: SR Research Ltd.

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

Tiffin-Richards, S. P., & Schroeder, S. (2015). Word length and frequency effects on children's eye movements during silent reading. *Vision Research (Oxford), 113*, 33-43. doi:10.1016/j.visres.2015.05.008

University of Oregon. (2018). *8th Edition of Dynamic Indicators of Basic Early Literacy Skills* (DIBELS®). Eugene, OR: University of Oregon.

Vaessen, A., Bertrand, D., Tóth, D., Csépe, V., Faísca, L., Reis, A., et al. (2010). Cognitive development of fluent word reading does not qualitatively differ between transparent and opaque orthographies. *Journal of Educational Psychology, 102*, 827–842. doi:10.1037/a0019465

Vaessen, A., & Blomert, L. (2010). Long-term cognitive dynamics of fluent reading development. *Journal of Experimental Child Psychology*, *105*, 213–231. doi:10.1016/j.jecp.2009.11.005

Vanderauwera, J., van Setten, E. R. H., Maurits, N. M., & Maassen, B. A. M. (2019). The interplay of socio-economic status represented by paternal educational level, white matter structure and reading. *PLoS ONE*, *14*(5), 1–18. doi:10.1371/journal.pone.0215560

VanOrman, A. & Jacobsen, L. A. (2020). *U.S. household composition shifts as the population grows older; More young adults live with parents.* Population Reference Bureau.

Verhagen, W. G. M., Aarnouste, C. A. J., & van Leeuwe, J. F. (2010). Spelling and word recognition in grades 1 and 2: Relations to phonological awareness and naming speed in Dutch children. *Applied Psycholinguistics, 31*(1), 59–80. doi:10.1017/S0142716409990166

Waber, D.P., Forbes, P. W., Wolff, P.H., Weiler, M.D. (2004). Neurodevelopmental characteristics of children with learning impairments classified according to the double-deficit hypothesis. *Journal of Learning Disabilities, 37*, 451–61. doi:10.1177/00222194040370050701

Wagner, R., Torgesen, J., Rashotte, C., & Pearson, N. (2013). CTOPP: *Comprehensive Test of Phonological Processing–Second Edition*. Austin, TX: PRO-ED.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence-Second Edition (WASI-II)*. NCS Pearson.

White, M. J. (2021). Phonological working memory and non-verbal complex working memory as predictors of future English outcomes in young ELLs. *The International Journal of Bilingualism, 25*(1), 318–337. doi: 10.1177/1367006920948136

Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals–Fifth Edition* (CELF-5). Bloomington, MN: NCS Pearson.

Wolf, M., & Denckla, M. B. (2005). *RAN/RAS: Rapid automatized naming and rapid alternating stimulus tests.* Austin, TX: Pro-ed.

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

Wolff, U. (2014). RAN as a predictor of reading skills, and vice versa: Results from a randomised reading intervention. *Annals of Dyslexia, 64*, 151–165. doi:10.1007/s11881-014-0091-6

Wood, C., Bustamante, K., Fitton, L., Brown, D., & Petscher, Y. (2017). Rapid automatic naming performance of young Spanish–English speaking children. *Languages (Basel), 2*(3), 13. doi:10.3390/languages2030013

Woodcock R. (2011). *Woodcock Reading Mastery Tests.* Third Edition Manual ed. Bloomington, MN: Pearson.