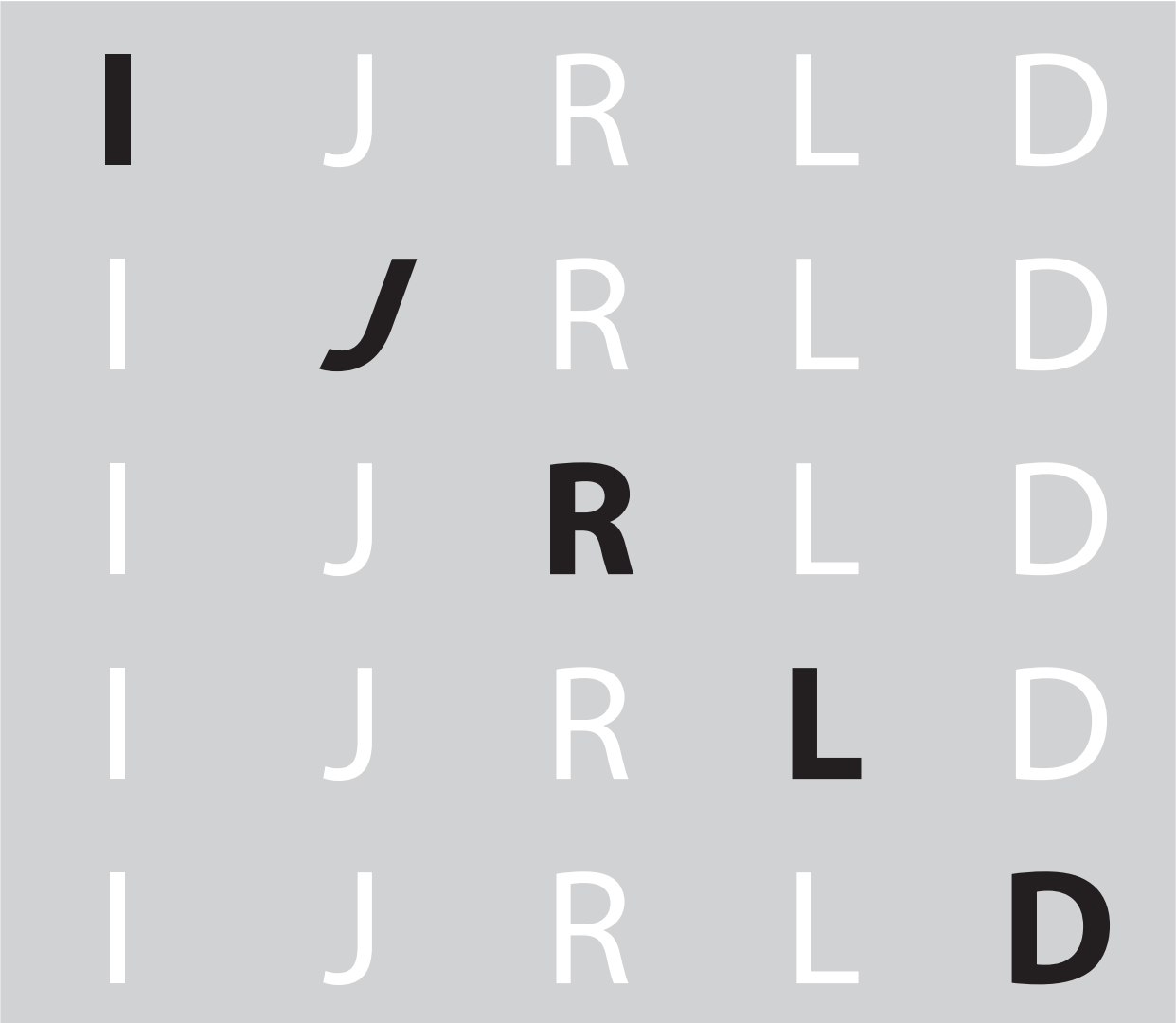
International Journal for Research in Learning Disabilities

volume 5, issue 2, 2022

formerly *Thalamus*



Journal of the
International Academy for Research in Learning Disabilities

Editor: David Scanlon, Boston College, United States

International Academy for Research in Learning Disabilities

International **J**ournal for **R**esearch in **L**earning **D**isabilities
formerly *Thalamus*
journal of the *International Academy for Research in Learning Disabilities*

Editor's Note: Due to an editing error, the citation appearing on page 3 of IJRLD, Volume 5, Issue 1, should correctly read:

McBride, C., & Cheah, Z.R.E. (2021). The "Write Stuff": What Do We Know About Developmental Dysgraphia? *International Journal for Research in Learning Disabilities, 5*(1), 3-12. https://doi.org/10.28987/ijrld.5.1.3

IJRLD
International Journal for Research in Learning Disabilities

Volume 5, No. 2, 2022

Table of Contents

Guest Reviewers

Annemie Desoete, *Universiteit Gent*
Michael Dunn, *Washington State University Vancouver*
Anna Maria Hintz, *Carl von Ossietzky Universität Oldenburg*
Li-Yu Hung, *National Taiwan Normal University*
Michael Paal, *Carl von Ossietzky Universität Oldenburg*
Susana Padeliadu, *Aristotle University of Thessaloniki*
Karen Waldie, *University of Auckland*

From the Editor

Do you remember what you were doing in 2012 even though that already seems like "the old days"? One thing you were likely doing was reading our "new" journal for the first time. With this issue, the IJRLD marks its 10th anniversary. Leading up to 2012, the Executive Board and the Publications Committee committed to revitalizing *Thalamus*, the original incarnation of this journal created by Academy founders William Cruickshank and Jacob Valk. We agreed that our mission for the journal would be to share the best work of Academy members with each other and with our global colleagues. Further, we sought to advance knowledge about learning disabilities. We have succeeded on both accounts, responding to the mission of informing and leading – novel research and scholarly commentary have appeared in these pages and have been shared with our global audience.

Journal publications play an important role in communicating scholarship. In ways different from conference presentations, social media exchanges, civic and civil service, and the contributions of public intellectuals, journal publications require detailed explications of science that are rigorously peer-reviewed before ever being published. Journals are repositories that are relied upon by scholars and others to advance knowledge – and practice – such as what we provide here about learning disabilities. Journal publications do not change the world in a day, but they do make vital contributions to its future.

Much has transpired in our world in the past 10 years, including within our field of learning disabilities. Thanks to the contributions of many Academy members and friends, we have built up the IJRLD to fulfill its mission in its first 10 years and, as such, have laid a strong a foundation for going forward.

With this issue, I am completing my time as editor. Serving as editor has been a remarkable experience. I have been introduced to scholars and scholarship from around the world. I have been challenged to contemplate what is science and what is knowledge, and I have had a front row seat to the emerging future of our field. I have also been rewarded with the collaboration of many colleagues in the Academy. Several deserve special mention, but I dare not try to name them as there have been so many and I would not know where to stop. So, I offer a thank-you to all who have partnered with me in any way to produce and disseminate our journal.

Our mission to inform and lead will never be completed. Welcome to our new editors, Drs. Deborah Reed and Matthias Grünke. They will shape the journal into its next decade. We are fortunate that these two Academy fellows and leaders in our field have volunteered to provide their services. I look forward to what's ahead for our Academy, our field, and our world, and to the IJRLD's continuing role.

In Fellowship,
David Scanlon

Identifying Students at Risk for Reading Difficulties in an Indian Context

Sunaina Shenoy,[1] Richard Wagner,[2] Deeksha Suresh,[3] and Nisha Menzies Rao[4]
[1]University of New Mexico
[2]Florida State University
[3]University of Edinburgh, UK
[4]St. Joseph's College, Bangalore, India

Abstract

The purpose of this study was to triangulate four data sources – Dynamic Indicators of Basic Early Literacy Skills (DIBELSNext), Easy Curriculum-Based Measures (easyCBM), Test of Silent Reading Efficiency and Comprehension (TOSREC), and Classroom-Based Written Assessment – to identify students at risk for reading difficulties. Participants were 1,025 students in Grades 1, 3, and 5 attending low-cost, middle-cost, and high-cost private schools in an urban city center in Southern India. In general, we found that a larger percentage of students were identified as being at risk when we used the reading measures alone, but the numbers reduced considerably when we triangulated those scores with the written assessment. For students in Grade 1, we found that a median of 5% of students were identified as at risk across all school sites using the four data sources; for students in Grades 3 and 5, this percentage increased to 20%. Implications for reading assessment and intervention in the Indian context are discussed.

Keywords: Triangulation of data, reading difficulties, curriculum-based measures, at risk, India

Linguistic Background and the Role of English

English and Hindi are the two official languages of India, but English is used as the link language for most people who speak different state languages (National Council of Educational Research and Training, 2011). It is the primary language of business, and is viewed as the language for economic and social mobility (Ramanathan & Bruning, 2003). It is estimated that 90 million children in India are being formally schooled in English (Kalia, 2007). However, English and the privileges associated with it remain inaccessible to those who are from a lower SES, with the middle class assuming a position of power through its access to English (Ramanathan & Atkinson, 1999).

In India, schools typically follow a three-language formula (Aggarwal, 1991) that is ratified by the National Curriculum Framework 2005 (Ramachandran et al., 2005). The first language is the language of instruction (one of the official languages, English or Hindi); the second is the official language; that is, not the language of instruction (English or Hindi), which is introduced by Grade 5; and the third language is the state language, which is introduced by Grade 7 (Ramachandran et al., 2005). Unfortunately, measuring students' proficiency in these languages is difficult, given the inconsistent time frames in which schools opt to introduce them. An additional complexity is that in most urban centers, a child's home language may differ from the national or state languages introduced in school.

A typical child in India is exposed to at least four languages from ages 0-13 years: a home language (L1); school language 1 (L2) which is the language of instruction (English, in our sample); school language 2 (L3) which is the national language Hindi; and school language 3 (L4) which is the state language (Kannada, in our sample). (Shenoy et al., 2020, p. 2) Eighty percent of Indian schools are government schools, but because of the poor quality of education offered there, 27% of Indian children are privately educated (Annual Status of Education Report, 2016). In urban centers, more

Shenoy, S., Wagner, R., Suresh, D., & Rao, N. M. (2022). Identifying Students at Risk for Reading Difficulties in an Indian Context. *International Journal for Research in Learning Disabilities, 5*(2), 4-17. https://doi.org/10.28987/ijrld.5.2.4

than 50% of children (27 million) attend private schools (Annual Status of Education Report [ASER], 2016). These schools follow a state, national, or international standardized curriculum, and the language of instruction is English (Kurrien, 2005). Government schools, on the other hand, follow a state-level curriculum, and the language of instruction is the state language.

There is a push towards English medium instruction in the private schools in order to promote social and economic mobility for students attending these schools. According to Kalyanpur (2020), this is creating a new group of marginalized students in India – a group that attends low-cost private schools and for whom English still seems inaccessible because of the poor quality of instruction in these schools.

In schools where English is the language of instruction, in a seven-hour school day, a student is exposed to six hours of instruction in English and one hour of instruction in Hindi and/or the state language (Shenoy et al., 2020). By middle or high school, a large number of students are proficient in all domains of speaking, listening, reading, and writing in English. These students consider the English language to be their dominant language since they have been exposed to it more than the other languages, and it is pervasive across academic content areas.

Focus on Written Assessments

Students in India are typically assessed only on their written content-area skills, based on the assumption that such assessments indirectly also measure students' reading abilities. The assessment method at most schools appears to be primarily concerned with testing written skills and does not account for other aspects such as oral language or reading. The assessments are structured as opportunities for students to reproduce content that has been extensively covered in class (Ramanathan, 2001), and questions are limited to those for which teachers have given students appropriate responses; as a result, the assessment pattern is restricted to testing students' short-term memory, and there is no learning or assessment of non-written English literacy skills, potentially leading to lower learning outcomes in primary education (Vyas, 2014).

The predominant format for testing is targeted at students producing short answers and essays, and it taps into their rote memorization skills alone (Ramanathan, 2008). Linguistic creativity, which is often influenced by the school and home environment, is restricted to the teacher's interpretation of the textbook, and students are not directly tested on other aspects of language, such as speaking, listening, and reading in the elementary grades. This may impede the development of overall English literacy and have significant ramifications for the general learning process (Kumar & Rani, 2016).

Reading Instruction

The predominant method used to teach reading in India is the alphabet-spelling method (Gupta, 2014). Students are taught letter names and how to spell out words and, therefore, bypass the sound structure of the language, acquiring new words by sight-word recognition instead. Students are expected to learn "common" words as a whole and to recognize new, unfamiliar words by rote memorization (Annamalai, 2004). In a similar way, students move from learning letter names and words to learning sentences by rote (Dixon et al., 2011). Thus, they are not taught how to blend or segment letter sounds into words and can only read words that are familiar to them, with limited comprehension.

Further, it is very common for teachers in Indian classrooms to teach reading by focusing on written products, such as copying from the board and choral recitation, rather than comprehension. One teacher in Gupta's (2014) study reported: "These children are not reading because they are not copying the letters. In class, teachers used terms that are central to initial reading – picture, word, letter, sound and spelling – interchangeably" (p. 3912).

Dixon et al. (2011) attempted to introduce phonics-based instruction in English-medium low-income private schools in Hyderabad, India. A control group received traditional English instruction involving rote learning and whole-word recognition and the experimental group received phonics-based instruction. Their findings showed a statistically significant difference between the experimental and control groups, with the experimental group performing better on measures of reading, spelling, and sounding out letters and words (Dixon et al., 2011).

Similar findings have been reported for students attending rural schools in India (Gupta, 2014). For example, Nishanimut et al. (2013) introduced a phonics approach in L2 English, where letter sounds were represented by the symbols used in the child's L1 (Kannada) and found that tapping into their L1 reading instruction helped students learn English better than phonics-based instruction programs in English alone.

Only a limited number of studies have been conducted in the area of reading instruction in India in the last few years, and Shenoy et al. (in review) have recently published a paper to address this gap. Using DIBELSNext (Good et al., 2011) to observe reading progress, we found that students who received both

one and two years of phonics instruction in preschool significantly outperformed those who did not receive any phonics instruction on the literacy skills assessed. Moreover, the incidence of students being at risk for reading difficulties reduced significantly with an increase in years of phonics instruction. Beyond this, as far as we know, no other reading instruction programs have been researched within the Indian context.

Context of the Present Study and Research Questions

We were interested in measuring L2 English reading skills because it was the language of instruction for our sample of students in Bangalore, and represented their access to literacy. But at the time of the study, we could not find any measures that were developed in the Indian context. We, therefore, decided to adapt and use curriculum-based measures developed in the United States (US), namely DIBELSNext (Good et al., 2011), easyCBM (Anderson et al., 2014), and TOSREC (Wagner et al., 2019), and established reliability and validity for these measures (Shenoy et al., 2020).

DIBELSNext (Good et al., 2011) is a widely used tool to measure reading and literacy skills in the US. Measuring students' reading skills is an important component that educators consider while making intervention decisions for their students. Researchers at the University of Oregon developed and revised the easyCBM measures (Anderson et al., 2014). The focus has been to facilitate "data-driven instructional decision making through enhanced reporting options" (Anderson et al., 2014, p. 4), in order to promote progress-monitoring and universal screening in schools (Deno, 2003; Keller-Margulis et al., 2008). These curriculum-based measures were developed in line with the reading areas that were deemed important by the National Reading Panel (National Reading Panel et al., 2000), including phonological awareness, phonics (alphabetic principle), fluency, and comprehension (Riedel & Samuels, 2007). Moreover, these reading skills represent the developmental continuum, and the changes in subtests across grade levels parallel student development (Adams, 1990; Chall, 1996; Ehri, 2005; Paris & Hamilton, 2009). In addition, the National Education Policy of India (NEP; Ministry of Human Resource Development, Government of India, 2020), the latest government mandate, states that there is an urgent national need for students to attain universal foundational literacy and foundational numeracy by Grade 3. Our rationale for using these measures was precisely because they represented a universal trajectory of reading acquisition in English and followed a developmental continuum. We were also interested in aligning our study with the NEP goal and providing schools in our sample a means of measuring foundational reading skills in addition to the measures of foundational writing skills they were using.

Our rationale for using an English measure and not a bilingual or multilingual assessment was as follows: (a) Even though our sample of students were bilingual, they were not biliterate; they were only literate in English and not in their native languages; and (b) though they came from different home language backgrounds, English was their link language in the classroom, and they used it to communicate with their teachers and peers.

Given the context and focus on written assessments and that reading is not explicitly taught, we also wanted to introduce reading assessments and progress-monitoring tools that could not only keep track of student progress but also help guide instruction for teachers. Because teachers in Bangalore, India, did not follow a phonics-based curriculum, we expected to see overall low scores on measures of decoding skills such as letter sounds, nonsense word fluency, and phoneme segmenting, but were still interested in learning how students performed on these critical reading subtests. Additionally, we wanted to explore how students would perform on fluency and comprehension measures and whether the reading instruction they were currently receiving – namely the alphabet-spelling method (Gupta, 2014) – would impact their scores. Our rationale for utilizing both the DIBELSNext and easyCBM was to be able to capture a wide variety of subtests measuring reading in elementary grades that followed a universal developmental continuum, as well as to observe their efficacy and reliability as assessments of L2 English reading development within the Indian context. For consistency, we maintained the content and administration procedures employed in the US. The comprehension passages were modified to reflect names that are common within the Indian context (e.g., *Abby* was replaced with *Asha*), and some words were changed to reflect common usage in the culture (e.g., *jump rope* was replaced with *skipping rope*), but the essence of the passages in terms of meaning and comprehension was not changed.

We were cognizant that our reading measures were measuring skills not taught in the Indian context, especially for students in Grade 1, so we collected data on an equivalent classroom-based written assessment. This served as our primary data source in order to triangulate our reading measures developed in the US with those developed by classroom teachers in India to be able to identify risk. The resulting research questions were the following:

1. Is there any variability in the number of students identified as at risk for reading difficulties across low-cost, middle-cost, and high-cost schools?
2. What is the difference in risk percentages when we

consider only the reading measures vs. the reading measures triangulated with the classroom-based written assessment?

Method

Participants

The sample consisted of 1,025 students from Grades 1, 3, and 5. Students came from different home language backgrounds and were enrolled in English-medium schools. They did not receive any additional bilingual support for the development of their home languages and were not expected to be biliterate in both languages. The demographic information of the students is presented in Table 1.

School Setting

The six participating school sites were located in an urban city center, Bangalore. Two schools were low-cost, two were middle-cost, and two were high-cost schools. Table 1 presents the school characteristics. For the purposes of this study, a low-cost school was defined as a private school in Bangalore, India, where the annual tuition cost per student was approximately Rupees 7,200 ($120); the middle-cost school was a private school where the annual tuition cost per student was approximately Rupees 40,000 ($667); and the high-cost school was a private school where the annual tuition cost per student was approximately Rupees 150,000 ($2,500). Moreover, low-income household was defined as families whose monthly income was between Rupees 0-20,000 ($0-275), middle-income household was defined as families whose monthly income was between Rupees 21,000-70,000 ($285-956) and high-income household was defined as families whose monthly income exceeded Rupees 71,000 (above $1,000).

The low-cost schools followed a state board curriculum that is prescribed by the state of Karnataka; the middle-cost schools followed a national board curriculum that is prescribed by the Central Board of Education in India; one high-cost school followed the national board curriculum, the other followed a Montessori curriculum. The national board curriculum is more rigorous, designed to prepare students to find national and international jobs. The state board curriculum is less rigorous, intended to prepare students to find jobs within the state of Karnataka only.

All schools in our sample introduced English in kindergarten and provided instruction in English in all content areas throughout the school day. Teachers' language proficiency varied considerably, with teachers from the state board schools being less fluent in English than teachers from the national board and Montessori schools. One low-cost school followed the alphabet-spelling method for 100% of their reading instruc-

Table 1
Demographic Data for the Students in the Sample

			Grade 1 (*N* = 346)		Grade 3 (*N* = 328)		Grade 5 (*N* = 329)	
Individual Characteristics			Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
	Gender	Male	171	49.42	179	54.57	189	57.45
		Female	175	50.58	149	45.43	140	42.55
	SES	Low-Income	46	13.29	40	12.20	45	13.68
		Middle-Income	175	50.58	220	67.07	210	63.83
		High-Income	125	36.13	68	20.73	74	22.49
School Characteristics	School Type	Low-Cost 1	37	10.69	36	10.98	37	11.25
		Low-Cost 2	9	2.60	4	1.22	8	2.43
		Middle-Cost 1	74	21.39	84	25.61	83	25. 23
		Middle-Cost 2	101	29.19	136	41.46	127	38.60
		High-Cost 1	107	30.92	51	15.55	71	21.58
		High-Cost 2	18	5.20	17	5.18	3	0.91
	Curriculum	State	46	13.29	40	12.20	45	13.68
		National	282	81.50	271	82.62	281	85.41
		Montessori	18	5.20	17	5.18	3	0.91

tion, one low-cost school, two middle-cost schools, and one high-cost school followed a combination of alphabet-spelling and phonics-based instruction; the Montessori school followed a 100% phonics-based approach.

Measures

The reading measures were administered across three phases during the 2017-18 academic year. The academic year in India is from June-March, so the reading measures were administered during three time periods,July-August, October-November, and January-February to correspond with benchmark assessments that are administered in fall, winter, and spring in the US. The total individual administration time was approximately 30 minutes per student.

DIBELSNext Subtests

All the DIBELSNext subtests were timed measures and were administered for 1 minute each. The following subtests were administered for students in Grade 1.

Letter Naming Fluency. This subtest measures the student's ability to name uppercase and lowercase letters; students earned 1 point for each letter that was correctly named.

Phoneme Segmenting Fluency. In this task, the student has to break up a word into its corresponding sound segments; for example, the word *cat* has three sound segments: /c /a/ /t/. The student earned 1 point for each correct sound produced.

Nonsense Word Fluency. This subtest consists of two parts: correct letter sounds (CLS) and whole words read (WWR). It measures students' knowledge of letter-sound correspondences and their ability to process CVC combinations that are non-words (e.g., /v/ /o/ /l/). Students earned 1 CLS for each correct letter sound read by itself or as part of a make-believe word. They also earned 1 WWR for each whole word read correctly without first being sounded out.

Oral Reading Fluency. The ORF subtest was used to measure the ability to accurately read an unknown passage; the student earned 1 point for each word that was read correctly.

Retell Fluency. After students completed the ORF subtest, they were asked to recall and retell the story that they just read as part of the ORF subtest. They earned 1 point for every word in their retell that was connected to the passage.

The following subtests were administered for students in Grades 3 and 5:

Oral Reading Fluency. The ORF subtest was used to measure the ability to accurately read an unknown passage; the student earned 1 point for each word read correctly. It was a timed test administered for 1 minute.

Retell Fluency. After students completed the ORF subtest, they were asked to recall and retell the story that they just read as part of the ORF subtest. They earned 1 point for every word in their retell that was connected to the passage. It was a timed test administered for 1 minute.

DAZE Comprehension. DIBELS Daze is a close comprehension measure that measures students' understanding of the meaning of a word within the context of a sentence. It was individually administered, and students were given 3 minutes to complete the test. According to the authors' directions, approximately every seventh word was replaced by a box containing the correct word and two distractor words. Students were asked to silently read a passage and circle their word choices. The scores represent the number of correct and incorrect words, and an adjusted score that compensates for guessing is calculated based on the number of correct and incorrect responses.

easyCBM Subtests

All easyCBM subtests were timed and were administered for 1 minute each. The following subtests were administered for students in Grade 1:

Letter Names. This subtest was equivalent to the Letter Naming Fluency subtest on the DIBELSNext.

Letter Sounds. Letters of the alphabet were presented in either upper- or lowercase format, and students were asked to produce the letter sounds. Students had to produce as many letter sounds as possible and earned 1 point per letter sound identified correctly.

Phoneme Segmenting. This subtest was equivalent to the Phoneme Segmentation Fluency subtest on the DIBELSNext.

Word Reading Fluency. A list of words was presented to students, who were asked to read as many words as possible; they earned 1 point for every word correctly read.

Passage Reading Fluency (PRF). The PRF subtest was equivalent to the ORF subtest on the DIBELSNext.

The following subtests were administered for students in both Grades 3 and 5 except for Word Reading Fluency, which was only administered to students in Grade 3.

Word Reading Fluency. A list of words was presented to students, who were asked to read as many words as possible; they earned 1 point for every word correctly read. It was a timed test administered for 1 minute.

Passage Reading Fluency (PRF). The PRF subtest was equivalent to the ORF subtest on the DIBELSNext. It was a timed test administered for 1 minute.

Multiple-Choice Reading Comprehension. Students were instructed to silently read a comprehension passage and answer 20 multiple-choice comprehension questions that followed. This subtest was group-admin-

istered by class sections in the schools, and typically took 30 minutes to complete. Scores were calculated as the number of correct responses out of the 20 questions.

We chose to use both the DIBELSNext and easyCBM measures because (a) we could record reliability and validity of measures within the Indian context; and (b) we wanted to assess which set of subtests was easier for teachers in our sample to access, administer, and score within this context.

Test of Silent Reading Efficiency and Comprehension (TOSREC)

We chose the TOSREC (Wagner et al., 2010) as an added progress-monitoring measure because the DIBELSNext and easyCBM had a reading fluency subtest but not a comprehension test for Grade 1. Students were expected to read various statements and conclude if they were true or false. For example, they read a statement such as "A lion can fly" and checked a box labeled "yes" or "no." The test was timed for 3 minutes, and raw scores were calculated by subtracting incorrect responses from correct ones. Some words that were written in American English were changed to Indian English so that students would comprehend them in this context (e.g., The word *cookies* was changed to *biscuits*). But otherwise, the meaning of the text was retained in all instances. We found that out of all the fluency and comprehension measures, the TOSREC represented a somewhat culture-free test, because of the generic statements, which were either true or false, rather than a passage or story that had many more cultural references. It seemed to be a preferred test for this context and captured comprehension at a sentence level.

Classroom-Based Written Assessment

In addition to the reading measures, we asked teachers to provide us with students' scores on classroom-based written assessments. These assessments, which are administered five times during the school year, play a critical role in the final decision regarding student promotion from one grade to the next. We only collected classroom-based assessment data from tests that were conducted within two weeks of when we collected data on the reading measures so as to capture students' progress in reading and writing skills simultaneously.

These written assessments were curriculum-based measures that reflected what was taught in class. Though students have to take tests in all content areas, we collected data on tests that measured their skills in English grammar (sentence structure, tenses, nouns, singular and plural), vocabulary (synonyms, antonyms, adjectives), comprehension (passage level: read a passage and answer short questions on it), and composition (short essay on a topic). It is important to note that all the content for the exams is covered in class and measures students' rote memorization skills rather than creativity or critical thinking skills. High-Cost School 2 was a Montessori school; they do not have exams, so we could not get these data from that school site.

Data Analysis

We followed the risk tables provided by the authors of the curriculum-based measures. For DIBELSNext, we calculated a composite score per grade level and then coded students' scores as "at or above benchmark," "below benchmark," and "well below benchmark." For easyCBM, we coded students' score per subtest as being "core instructional support," "strategic instructional support," and "intensive instructional support." We then averaged the scores across subtests and recorded student scores that fell below the recommended 20th percentile and coded those students as needing "intensive instructional support across all subtests." For TOSREC, we followed the ratings provided by the authors: "above average," "average," "below average," "poor," and "very poor." We then coded the scores that were below the 10th percentile; that is, the "poor" and "very poor" categories as students needing "intensive instructional support." Finally, for the classroom-based written assessment, we followed the cut-off percentages set by teachers at the school sites: above 60% or "above-average grades," between 41-59% or "passing grades," and below 40% or "failing grades." We then examined the classification of students across data sources and focused on the following ratings: "well below benchmark" on DIBELSNext, "intensive instructional support across all subtests" on easyCBM, "poor/very poor" on TOSREC, and "failing grades" on classroom-based written assessment. These students were classified as "at risk for reading difficulties."

Results

The study sought to ascertain the number of students at each of the school sites who could be identified as at risk for reading difficulties across all four data sources. The data are organized around grade level.

Grade 1

Table 2 and Figure 1 provide a representation of the percentage of students from Grade 1 who were identified as being at risk across all four data sources. A total of 368 students were evaluated across all school sites.

Students Identified as at Risk on the Three Reading Measures

If we were to consider only the reading measures, approximately 80% of students from low-cost schools (*n*

Table 2
Students Identified as at Risk for Reading Difficulties Across All Four Data Sources in Grade 1

N = 368	At risk on DIBELS	At risk on easyCBM	At risk on TOSREC	At risk on classroom-based assessment	At risk across all four data sources
Low-Cost School 1 (*n* = 37)	34 (92%)	35 (96%)	37 (100%)	04 (11%)	03 (8%)
Low-Cost School 2 (*n* = 09)	06 (67%)	06 (67%)	09 (100%)	0	0
Middle-Cost School 1 (*n* = 84)	26 (31%)	32 (38%)	47 (56%)	08 (10%)	06 (7%)
Middle-Cost School 2 (*n* = 105)	55 (52%)	63 (60%)	83 (79%)	08 (10%)	03 (3%)
High-Cost School 1 (*n* = 115)	55 (48%)	57 (50%)	77 (70%)	22 (19%)	18 (16%)
High-Cost School 2 (*n* = 18)	01 (5%)	01 (5%)	03 (13%)	N/A	0

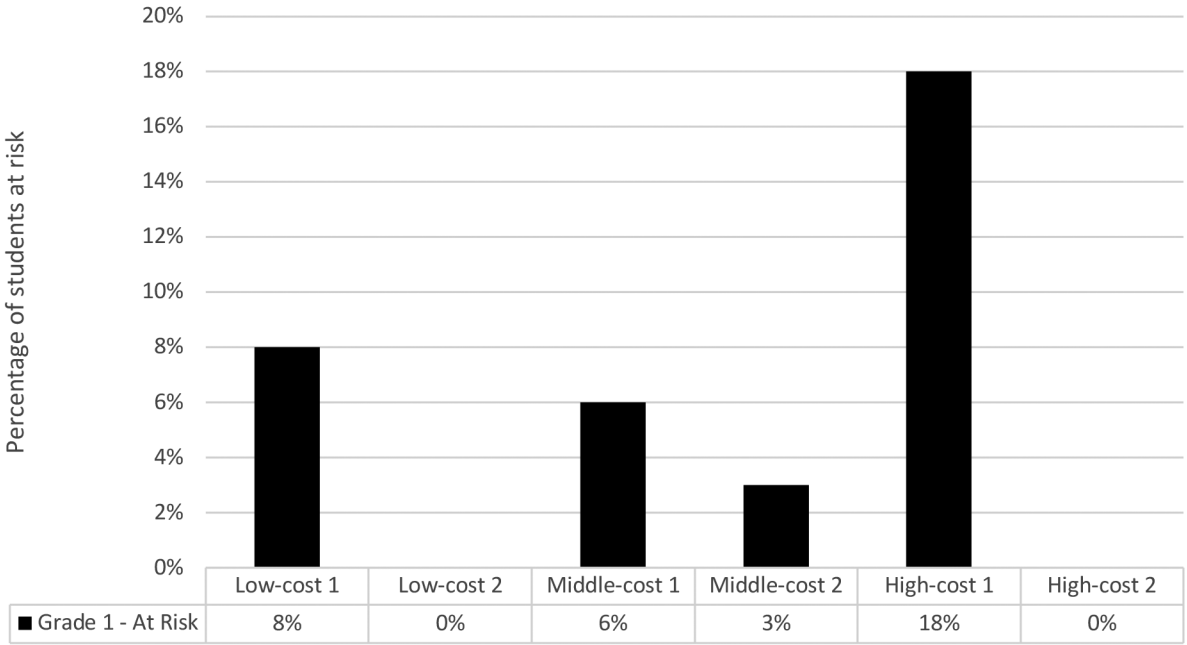


Figure 1
Students at Risk for Reading Difficulties in Grade 1

= 46) would be considered at risk on the DIBELSNext; 80% of students would be considered at risk on the easyCBM, and 100% would be considered at risk on the TOSREC. These numbers dropped for the middle-cost schools (*n* = 189) to approximately 43%, 49%, and 68% on the DIBELSNext, easyCBM, and TOSREC, respectively. We saw a similar trend for students in High-Cost School 1 (*n* = 115), where 48%, 50%, and 70% of students were identified as being at risk on the DIBELSNext, easyCBM, and TOSREC, respectively. In High-Cost School 2 (*n* = 18), however, 5% of students were identified as being at risk on DIBELS and easyCBM and 13% were at risk on TOSREC.

Students Identified as at Risk on the Classroom-Based Writing Measure

When we only considered classroom-based assessment, 11% of students at Low-Cost School 1 ($n = 37$) and no students at Low-Cost School 2 ($n = 9$) were considered at risk. Similarly, for the middle-cost schools, 10% of students at both school sites ($n = 84$; $n = 105$) were considered to be at risk. Finally, 19% of students at High-Cost School 1 ($n = 115$) were classified as being at risk when considering classroom assessment.

Thus, a significantly lower percentage of students were found to be at risk when considering performance on classroom assessment as opposed to performance on reading measures. A drop can be seen in the percentage of students at risk on reading measures, with numbers being higher in low-cost schools and lower in middle-/high-cost schools. The variation between DIBELS, easyCBM, and TOSREC is interesting and will be discussed further.

Students Identified as at Risk When Combining All Four Data Sources

Eight percent of students were at risk at Low-Cost School 1 ($n = 37$) and none at Low-Cost School 2 ($n = 9$) when all four data sources were considered. Similarly, for the middle-cost schools, 7% of students from Middle-Cost School 1 ($n = 84$) and 3% of students from Middle-Cost School 2 ($n = 105$) were found to be at risk. The high-cost schools differed on the number of students identified as at risk, with 16% of students from High-Cost School 1 ($n = 115$) being identified, and none from High-Cost School 2 ($n = 18$).

Grade 3

Table 3 and Figure 2 represent the percentage of students from Grade 3 who were identified as at risk across all four data sources. A total of 328 students were assessed across six school sites.

Students Identified as at Risk on the Three Reading Measures

Students' scores on the reading measures indicated that 100% of students at the low-cost schools ($n = 40$) were seen to be at risk on DIBELS and easyCBM, and 88% of students were observed to be at risk on TOSREC. We observed a difference in the middle-cost schools, with more students being reported as at risk on DIBELS and easyCBM than on the TOSREC. That is, 79% and 63% of students ($n = 220$) were considered at risk on DIBELS and easyCBM, respectively, and 32% of students on TOSREC. High-Cost School 1 ($n = 51$) presented similar trends with 45%, 35%, and 19% of students as at risk on DIBELS, easyCBM, and TOSREC, respectively. In High-Cost School 2 ($n = 17$), on the other hand, 11% of students were considered to be at risk across each of the measures. A drop can be seen in the percentage of students at risk on the reading measures, with numbers being higher in the low-cost schools and in lower in high-cost schools. The variation between DIBELS, easyCBM, and TOSREC is interesting, and will be discussed further.

Students Identified as at Risk on the Classroom-Based Writing Measure

A total of 50% of students from Low-Cost School 1 ($n = 36$) and 25% from Low-Cost School 2 ($n = 4$) were classified as being at risk when considering only classroom assessment. These numbers decreased in the middle-cost schools: 24% of students in Middle-Cost School 1 ($n = 84$) and 5% of students in Middle-Cost School 2 ($n = 136$) were seen to be at risk. These numbers decreased even further in High-Cost School 1, where 14% of students were identified as being at risk based on classroom assessment. Similar to Grade 1, a lower percentage of students would be considered at risk if we only considered classroom assessments.

Students Identified as at Risk When Combining All Four Data Sources

A total of 50% of students from Low-Cost School 1 ($n = 36$) and 25% from Low-Cost School 2 ($n = 4$) were identified as at risk across all four data sources. With respect to the middle-cost schools, 18% of students at School Site 1 ($n = 84$) and 4% at School Site 2 ($n = 136$) were identified to be at risk across all four data sources. Finally, 6% of students in High-Cost School 1 ($n = 51$) and 11% of students at High-Cost School 2 ($n = 17$) were identified to be at risk.

Grade 5

Table 4 and Figure 3 outline the number of students identified as at risk across all four data sources and all six school sites, for a total of 329 students.

Students Identified as at Risk on the Three Reading Measures

When we consider Grade 5 students' performance across reading measures alone, their performance on the DIBELSNext, easyCBM, and TOSREC does not vary greatly. For the low-cost schools ($n = 44$) scores indicate that 100% of students were considered at risk on easyCBM, 94% were at risk on the TOSREC, and 93% are seen to require intensive instructional support on DIBELSNext. For the middle-cost schools ($n = 210$), 60% of students were at risk on DIBELSNext, 69% on easyCBM, and 58% on TOSREC. For students in High-Cost School 1 ($n = 71$), we saw lower percentages, with 32% of students being identified as at risk when considering DIBELS, 34% on TOSREC, and 41% on easyCBM. In High-Cost School 2 ($n = 3$), no students were identified as being at risk on the reading measures.

Students Identified as at Risk on the Classroom-Based Writing Measure

In Low-Cost Schools 1 (n = 37) and 2 (n = 8), 62% of students were seen as at risk when considering only classroom-based assessment. The percentages were the same as students identified as at risk across all four data sources. For Middle-Cost School 1 (n = 83), 17% of students were at risk, and for Middle-Cost School 2 (n = 127) 12% of students were at risk on the classroom assessment. At High-Cost School 1, 28% of students were considered to be at risk.

Results indicate a difference in the number of students identified as at risk across all four data sources vs. when only the reading measures were considered. Mar-

Table 3
Students Identified as at Risk for Reading Difficulties Across All Four Data Sources in Grade 3

N = 328	At risk on DIBELS	At risk on easyCBM	At risk on TOSREC	At risk on classroom-based assessment	At risk across all four data sources
Low-Cost School 1 (*n* = 36)	36 (100%)	36 (100%)	36 (100%)	18 (50%)	18 (50%)
Low-Cost School 2 (*n* = 04)	04 (100%)	04 (100%)	03 (75%)	01 (25%)	01 (25%)
Middle-Cost School 1 (*n* = 84)	62 (74%)	53 (63%)	26 (31%)	20 (24%)	15 (18%)
Middle-Cost School 2 (*n* = 136)	114 (84%)	86 (63%)	44 (32%)	07 (5%)	6 (4%)
High-Cost School 1 (*n* = 51)	24 (45%)	18 (35%)	10 (19%)	07 (14%)	3 (6%)
High-Cost School 2 (*n* = 17)	2 (11%)	2 (11%)	2 (11%)	N/A	2 (11%)

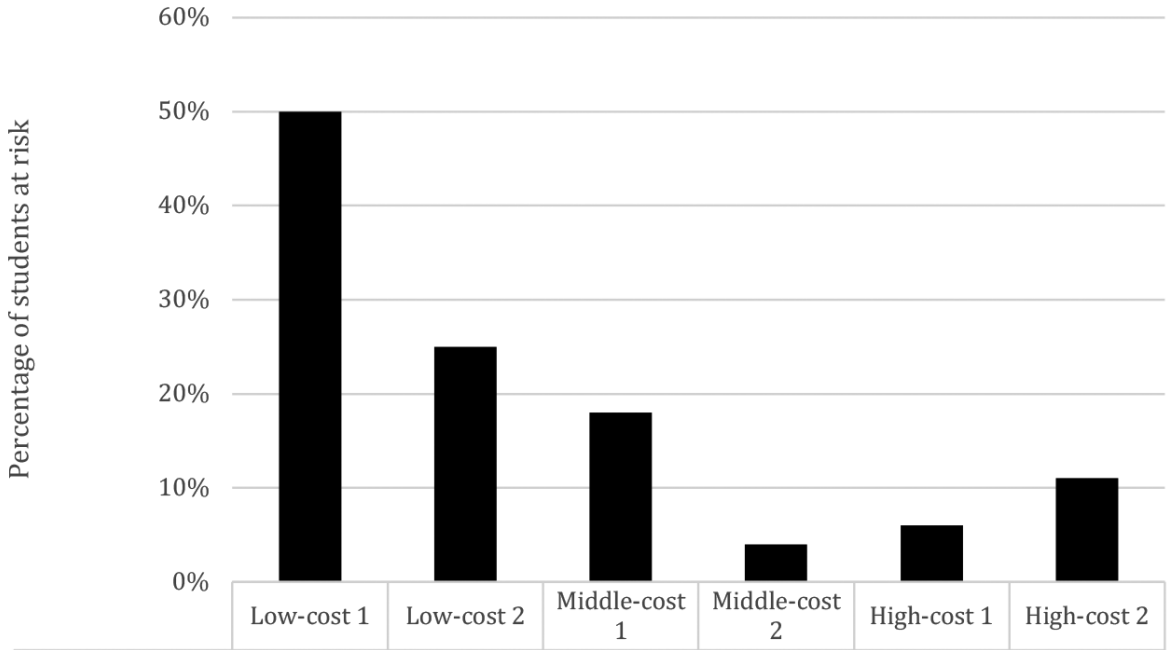


Figure 2
Students at Risk for Reading Difficulties in Grade 3

ginal variations can be noted between reading sources and its implications can be further elaborated. Additionally, the inclusion of written classroom-based assessment may influence these findings and will be further discussed.

Students Identified as at Risk When Combining All Four Data Sources

As evidenced by the analyses, 62% of students at Low-Cost School 1 (n = 37) and 25% of students at Low-Cost School 2 (n = 8) were identified as being at risk for reading difficulties across all four data sources. With respect to the middle-cost schools, 6% of students in School Site 1 (n = 83) and 9% of students in School Site 2 (n = 127) were identified as needing intensive support and were at risk across all four data sources. Additionally, 15% of students at High-Cost School 1 (n = 71) and none at High-Cost School 2 (n = 3) were identified as at risk across all four data sources.

Table 4
Students Identified as at Risk for Reading Difficulties Across All Four Data Sources in Grade 5

N = 329	At risk on DIBELS	At risk on easyCBM	At risk on TOSREC	At risk on class-room-based assessment	At risk across all four data sources
Low-Cost School 1 (*n* = 37)	36 (97%)	37 (100%)	37 (100%)	23 (62%)	23 (62%)
Low-Cost School 2 (*n* = 8)	07 (88%)	08 (100%)	07 (88%)	02 (25%)	02 (25%)
Middle-Cost School 1 (*n* = 83)	41 (49%)	48 (58%)	43 (52%)	14 (17%)	05 (6%)
Middle-Cost School 2 (*n* = 127)	91 (72%)	102 (80%)	81 (64%)	15 (12%)	11 (9%)
High-Cost School 1 (*n* = 71)	23 (32%)	29 (41%)	24 (34%)	21 (28%)	11 (15%)
High-Cost School 2 (*n* = 3)	0	0	0	N/A	0

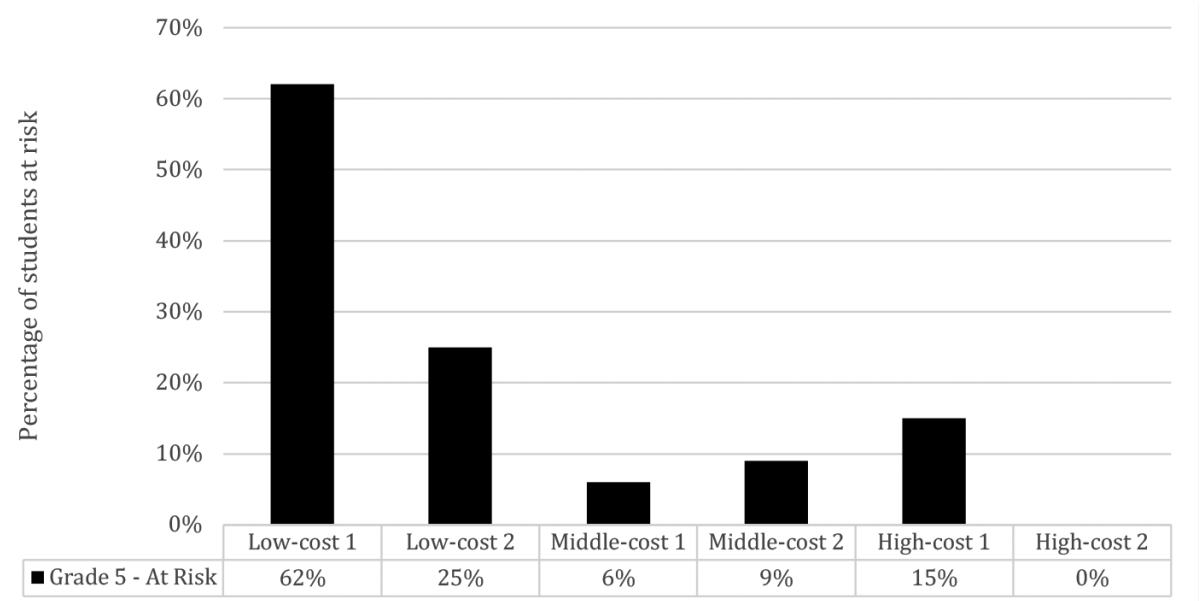


Figure 3
Students at Risk for Reading Difficulties in Grade 5

Discussion

The purpose of this study was to examine (a) the variability in students at risk for reading difficulties across low-cost, middle-cost, and high-cost private schools in India; and (b) the difference in risk percentages when we considered only the reading measures versus the reading measures triangulated with the classroom-based written assessment.

For students in Grade 1, if we examined the reading measures, the median percentages of students who were identified as at risk for reading difficulties varied from 94% in the low-cost school sites, to 54% in the middle-cost school sites, and 50% in High-Cost School 1 and 5% in High-Cost School 2. These median percentages reduced significantly when looking at both reading measures and classroom-based written assessments: 4% in the low-cost schools, 4% in the middle-cost schools, 16% in High-Cost School 1, and 0% in High-Cost School 2. The latter finding, which was established by examining both the reading measures and classroom-based written assessment, is closer to an expected prevalence rate in a given population of students compared to the former finding.

For students in Grade 3, if we examined reading measures, the median percentages of students who were identified as at risk for reading difficulties varied from 100% in the low-cost schools, to 63% in the middle-cost schools, to 35% in High-Cost School 1 and 11% in High-Cost School 2. These median percentages reduced to a great degree when we considered students on both reading measures and classroom-based written assessments: 37.5% in the low-cost schools, 11% in the middle-cost schools, 6% in High-Cost School 1, and 11% in High-Cost School 2. Interesting, the change in median percentages was more significant in Grade 1 than Grade 3, especially for students in low-cost and middle-cost schools. This may be attributed to students finding it more difficult to mask their inability to read by rote memorization of answers to questions on written exams.

For students in Grade 5, if we examined reading measures, the median percentages of students who were identified as at risk for reading difficulties varied from 98.5% in the low-cost schools, to 61% in the middle-cost schools, to 34% in High-Cost School 1 and 0% in High-Cost School 2. These percentages reduced to a great degree when we considered students on both reading measures and classroom-based written assessments: 74.5% in Low-Cost Schools, 7.5% in the Middle-Cost Schools, 15% in t High-Cost School 1, and 0% in High-Cost School 2. By triangulating the data sources, we still found that a quarter to half and more than half of the school population might be at risk for reading difficulties in Grades 3 and 5, respectively, in the low-cost schools. The percentages seem closer to an expected prevalence rate in a given population of students across middle-cost and high-cost schools.

Overall, our results are shocking, showing that a large percentage of students in Indian schools are not able to read at grade level. In schools where students predominantly followed the alphabet-spelling method, the percentage of students at risk in Grade 1 dropped from 87% in the low-cost schools to 54% in middle-/high-cost schools; the percentages of students at risk in Grade 3 dropped from 96% in low-cost schools to 50% in middle-/high-cost schools; and the percentage of students at risk in Grade 5 dropped from 96% in low-cost schools to 54% in middle-/high-cost schools. These numbers suggest that well over half the student population in the middle-/high-cost schools were not able to read at grade level. On the other hand, when these scores are compared to students following the Montessori curriculum in High-Cost School 2, we found that 8%, 11%, and 0% of students were at risk for reading disabilities in Grades 1, 3, and 5, respectively.

Another interesting trend in scores was that when we considered all four data sources, including the classroom-based written assessment, the overall percentage of students identified as being at risk increased from Grade 1 (~ 9%), to Grade 3 (~19%), to Grade 5 (~ 24%) across all schools in the sample. So not only were students not reading at grade level, their reading difficulties were also affecting their scores on written assessments, and more students were being identified as at risk in the later elementary years on classroom-based written assessments. That is, rote memorization of answers on written exams seems to be more difficult for students as they progress through the elementary grades.

Implications for Research

The identification of students at risk for reading difficulties in India is currently tied to underlying environmental factors such as a lack of progress-monitoring tools in reading and limited explicit and systematic reading instruction in schools, as well as limited access to literacy at home. Writing assessment is emphasized more than reading because writing is the medium of assessment for the Grades 10 and 12 national and state-level board exams that prepare students for college. The emphasis on writing is built into the curriculum right from kindergarten, and is entrenched in schooling practices in India as teachers and parents believe that it leads to better success on the school-leaving exams.

Reading, on the other hand, is not viewed as a required skill for future career prospects and is, therefore, not

emphasized within the school system and is assessed as a byproduct of writing. Thus, further research is warranted in areas of reading assessment. As mentioned, at the time when this study was conducted, no progress-monitoring tools were developed for the Indian context, so we opted to use tools that were normed in the US, but recently, Rao et al. (2021) have introduced the Dyslexia Assessment for the Languages of India (DALI), which includes both progress-monitoring tools and a dyslexia screener normed on the Indian population. Moreover, Misquitta et al. (2022) have developed a literacy-based application, FABLe, for the Indian context that complements the DIBELS but introduces reading passages from Indian textbooks. Given these new resources, more studies should incorporate these tools and focus on their efficacy in various settings in order to eventually give them traction needed to become universal screening tools across schools in India.

In addition to assessment, the curriculum does not focus on explicit reading instruction. Current research in the field of reading suggests that phonics-based instructional programs are the most effective at improving reading outcomes for both monolingual and bilingual students. A few recent studies have explored the efficacy of phonics-based instruction programs in the Indian context (Dixon et al., 2011; Patel et al., 2020; Shenoy et al., 2022), and more studies in this area will help establish better reading programs for students across India. Specifically, there is a need for introducing reading instruction as a systemic change to complement current writing instructional practices. Moreover, without reading assessment and intervention in place, we cannot distinguish students who might be at risk for dyslexia. This is especially complicated in a multilingual context with a push towards an English medium of instruction.

Finally, gaps in learning become more apparent as students move through the elementary grades. Whereas students in middle-/high-cost schools have access to English literacy in their homes, students in low-cost schools generally do not have access to English literacy outside of school. This creates a huge gap in learning for these students as they are not able to pick up reading at home in addition to not getting any reading instruction in school. Providing students with reading instruction is not only related to best practices in the field but is also a move towards equity in education – to improve outcomes for all students, irrespective of their socioeconomic status.

Implications for Practice

The practical implications of improving the reading scores of all students, including students at risk for reading difficulties, include moving away from the alphabet-spelling method that emphasizes rote memorization to phonics-based programs and improving family literacy for students from low-income backgrounds.

The teaching of English in India can be traced back to the British colonial rule, more specifically to a policy known as Macaulay's Minute on Education (Macaulay, 1835), which instigated a theme of rote memorization, an absence of inquiry and critical thinking, as well as a centrally imposed curriculum. The post-colonial and independence eras saw the evolution of the English language from being a mere colonial legacy to becoming a primary language of international commerce and communication, which can be attributed to the liberalization of the Indian economy and globalization (Meganathan, 2020). Unfortunately, the Indian education system in general is still geared towards rote memorization of all subject areas, including reading.

One of the key drivers of this problem is the foundational literacy curriculum in India. The complex and vast amount of content that students are expected to complete forces teachers to resort to superficial coverage of learning materials and rote memorization, instead of facilitating deeper thinking (Ministry of Human Resource Development, 2014, 2018). India has been too focused on advanced content rather than building foundational skills in a developmentally appropriate manner (Banerjee & Duflo, 2011; Glewwe et al., 2009; Pritchett & Beatty, 2012). The most empirically supported method for teaching foundational reading in English is systematic phonics (Ehri et al., 2001; Gersten & Baker, 2003; Johnston & Watson, 2005; National Reading Panel et al., 2000; Stuart, 1999, 2004; Torgesen, 2000). However, in India, the predominant method of teaching reading is the alphabet-spelling method (Gupta, 2014) in which students bypass letter-sound correspondences and are taught to read by rote memorization and sight-word recognition. Even the latest revision of The National Education Policy (Ministry of Human Resource Development, Government of India, 2020) does not mention integrating well-established, evidence-based, and developmentally appropriate, English reading instruction approaches like phonics that are highly effective in building foundational reading skills in the younger years.

This paper calls for a breakaway from these archaic policies and a change in the Indian education system to encompass more critical thinking skills and inquiry-based learning, especially in terms of reading instruction. There is an imminent need to teach the science of reading through decoding unfamiliar words rather than creating a large store of recognizing unfamiliar words by rote. This is especially critical for addressing the foundational literacy gap mentioned in the NEP (Ministry of Human Resource Development, Government of India, 2020).

The push towards English education in private schools in India has detrimental effects on students from

low-income backgrounds compared to middle- and high-income backgrounds, creating a new generation of marginalized youth (Kalyanpur, 2020). A protective factor in middle-/high-cost schools seems to be family literacy practices in English that motivate students to learn to read. Improving family literacy practices in either English or the home language in low-cost schools will also positively impact students' reading scores.

Limitations and Future Directions

First, this study reported on a sample of students from Bangalore, an urban city center in India. It would be helpful to conduct studies in other urban and rural areas in India as a comparison and be able to generalize our findings. Second, our sample of students in the Montessori school that received 100% phonics-based instruction is a very small comparison group. There is a need for intervention studies across various subpopulations to illustrate the efficacy of phonics-based instruction in the Indian context and improve the generalizability of our results. Third, the reading assessment tools that were used were normed on a population of students in the US, so there is a need for future studies to use tools that have recently been developed for the Indian context. Taken together, this will facilitate a stronger argument that a systemic change in reading assessment and instructional programs is needed, which, in turn, will improve reading outcomes for students.

References

Adams, M. J. (1990). *Beginning to read: Thinking and Learning about print.* MIT Press.

Aggarwal, S. (1991). *Three language formula: An educational problem*. Gyan Publishing House.

Anderson, D., Alonzo, J., Tindal, G., Farley, D., Irvin, P. S., Lai, C. F., ... & Wray, K. A. (2014). *Technical manual: easyCBM. Technical report #1408*. Behavioral Research and Teaching.

Annamalai, E. (2004). Medium of power: The question of English education in India. In J. W. Tollefson & A.B.M Tsui (Eds.), *Medium of instruction policies: Which agenda? Whose agenda* (pp. 177-193). Lawrence Erlbaum Associates.

ASER. (2016). *Annual status of education report* (2016). ASER Centre.

Banerjee, A. V., & Duflo, E. (2011). Why aren't children learning. *Development Outreach, 13*(1), 36-44. https://doi.org/10.1596/1020-797X_13_1_36

Chall, J. S. (1996). American reading achievement: Should we worry? *Research in the Teaching of English*, *30*(3), 303-310. Retrieved from https://www.jstor.org/stable/40171366

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, *37*(3), 184-192. https://doi.org/10.1177/00224669030370030801

Dixon, P., Schagen, I., & Seedhouse, P. (2011). The impact of an intervention on children's reading and spelling ability in low-income schools in India. *School Effectiveness and School Improvement*, *22*(4), 461-482. https://doi.org/10.1080/09243453.2011.625125

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, *9*(2), 167-188. https://doi.org/10.1207/s1532799xssr0902_4

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*, 393-447. doi:10.3102/00346543071003393.

Gersten, R., & Baker, S. (2003). English-language learners with learning disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 94-109). The Guilford Press.

Glewwe, P., Kremer, M., & Moulin, S. (2009). Many children left behind? Textbooks and test scores in Kenya. *American Economic Journal: Applied Economics*, *1*(1), 112-35. doi:10.1257/app.1.1.112

Good, R. H., Kaminski, R. A., & Cummings, K. (2011). *DIBELSnext*. Cambium Learning.

Gupta, R. (2014). Change in teaching practices: Case of phonics instruction in India. *Procedia-Social and Behavioral Sciences*, *116*, 3911-3915. https://doi.org/10.1016/j.sbspro.2014.01.865

Johnston, R., & Watson, J. (2005). *The effects of synthetic phonics teaching on reading and spelling attainment: A seven year longitudinal study*. Available at http://www.gov.scot/Publications/2005/02/20688/52449

Kalia, V. (2007). Assessing the role of book reading practices in Indian bilingual children's English language and literacy development. *Early Childhood Education Journal*, *35*(2), 149-153. https://doi.org/10.1007/s10643-007-0179-2

Kalyanpur, M. (2020). Disrupting the narrative of universality of inclusive education: The new marginalization of low-income, English language learners in India. *The Educational Forum*, *8*(4), 296-308. https://doi.org/10.1080/00131725.2020.1796071

Keller-Margulis, M. A., Shapiro, E. S., & Hintze, J. M. (2008). Long-term diagnostic accuracy of curriculum-based measures in reading and mathematics. *School Psychology Review*, *37*(3), 374-390. https://doi.org/10.1080/02796015.2008.12087884

Kumar, J., & Rani. (2016). *A study of language creativity among secondary school students in relation to their linguistic skills school environment and home environment*. Retrieved from http://shodhganga.inflibnet.ac.in:8080/jspui/handle/10603/112444

Kurrien, J. (2005). Notes for the meeting of the National Focus Group on Teaching of English, and note on introduction

of English at the primary stage. In *Minutes, National Focus Group-English*. National Council of Educational Research and Training.

Macaulay, T. B. (1835). *Minute on Education (1835) by Thomas Babington Macaulay*. Retrieved from http://www.columbia.edu/itc/mealac/pritchett/00generallinks/macaulay/txt_minute_education_1835.html

Meganathan, R. (2019). Research in English language education in India. *Indian Educational Review*, 57(2), 7-72.

Ministry of Human Resource Development, Department of School Education and Literacy. (2014). *All India survey on higher education: 2014*.

Ministry of Human Resource Development, Department of School Education and Literacy. (2018). *All India survey on higher education: 2018*.

Ministry of Human Resource Development, Government of India. (2020). *National education policy* (NEP). https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf

Misquitta, R., Shenoy, S., Ghosh, A., & Kotwal, N. (2022). Examining the technical adequacy and efficacy of FABLe to identify students at risk for reading difficulties in India. Asia P*acific Education Review.* https://doi.org/10.1007/s12564-022-09751-y

National Council of Educational Research and Training. (2014). *National curriculum framework 2005* (No. id: 1138).

National Reading Panel (US), National Institute of Child Health, Human Development (US), National Reading Excellence Initiative, National Institute for Literacy (US), United States. Public Health Service, & United States Department of Health. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.

Nishanimut, S. P., Johnston, R. S., Joshi, R. M., Thomas, P. J., & Padakannaya, P. (2013). Effect of synthetic phonics instruction on literacy skills in an ESL setting. *Learning and Individual Differences*, *27*, 47-53. https://doi.org/10.1016/j.lindif.2013.06.007

Paris, S., & Hamilton, E. E. (2009). The development of children's reading comprehension. In S. Israel & G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 32-53). Lawrence Erlbaum.

Patel, P., Torppa, M., Aro, M., Richardson, U., & Lyytinen, H. (2018). GraphoLearn India: The effectiveness of a computer-assisted reading intervention in supporting struggling readers of English. *Frontiers in Psychology*, *9*, 1045. https://doi.org/10.3389/fpsyg.2018.01045

Pritchett, L., & Beatty, A. (2012). *The negative consequences of overambitious curricula in developing countries.* Center for Global Development Working Paper No. 293, HKS Working Paper No. RWP12-035. Available at SSRN: https://ssrn.com/abstract=2102726 or http://dx.doi.org/10.2139/ssrn.2102726

Ramachandran, V., Pal, M., Jain, S., Shekar, S., & Sharma, J. (2005). *Teacher motivation in India*. *Discussion paper*. Azim Premji Foundation.

Ramanathan, H. (2001). *Assessment and testing in an English classroom in India*. Paper presented at the Midwestern Educational Research Association, Chicago, USA. Retrieved from https://files.eric.ed.gov/fulltext/ED459205.pdf

Ramanathan, H. (2008). Testing of English in India: A developing concept. *Language Testing*, *25*(1), 111-126. https://doi.org/10.1177/0265532207083747

Ramanathan, V., & Atkinson, D. (1999). Individualism, academic writing, and ESL writers. *Journal of Second Language Writing*, *8*(1), 45-75. https://doi.org/10.1016/S1060-3743(99)80112-X

Ramanathan, H., & Bruning, M. D. (2003). Reflection on teaching oral English skills in India: A research report. *Journal of the International Society for Teacher Education*, *7*(1), 48-55.

Rao, C., Sumathi, T. A., Midha, R., Oberoi, G., Kar, B., Khan, M., ... & Singh, N. C. (2021). Development and standardization of the DALI-DAB (dyslexia assessment for languages of India-dyslexia assessment battery). *Annals of Dyslexia*, 1-19. https://doi.org/10.1007/s11881-021-00227-z

Riedel, B. W., & Samuels, S. J. (2007). The relation between DIBELS, reading comprehension and vocabulary in urban first grade students. *Reading Research Quarterly, 42*, 546-567. Retrieved from http://www.jstor.org/stables/20068316

Shenoy, S., Wagner, R. K., & Rao, N. M. (2020). Factors that influence reading acquisition in L2 English for students in Bangalore, India. *Reading and Writing: An Interdisciplinary Journal, 33*(7), 1809-1838. https://doi.org/10.1007/s11145-020-10047-z

Shenoy, S., Iyer, A., & Zahedi, S. (2022). *Phonics-based instruction and improvement in foundational reading skills of kindergartners in the Indian schooling context* [Manuscript submitted for publication] University of New Mexico.

Stuart, M. (1999). Getting ready for reading: early phoneme awareness and phonics teaching improves reading and spelling in inner-city second language learners. *British Journal of Educational Psychology, 69*, 587-605. doi:10.1348/000709999157914

Stuart, M. (2004). Getting ready for reading: a follow-up study of inner city second language learners at the end of Key Stage 1. *British Journal of Educational Psychology, 74*, 15-36. doi:10.1348/000709904322848806

Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice, 15*(1), 55-64.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2010). *TOSREC: Test of silent reading efficiency and comprehension*. Pro-Ed.

Vyas, A. (2014). Low learning outcomes in primary schools in India. *The Indian Economist.* https://www.academia.edu/14568078/LOW_LEARNING_OUTCOMES_IN_PRIMARY_SCHOOLS_IN_INDIA

Error Patterns in the Knowledge-Based Inference-Making of Less Skilled Middle-Grade Readers: An Exploratory Study

Amy E. Barth,[1] Cathy Newman Thomas,[2] Heather Kincaid[1], Ethan Ankrum[1], Brandon Ruiz,[1] and Leslie Salazar[1]
[1]Buena Vista University
[2]Texas State University

Abstract

The primary aims of this mixed-method study were to (a) examine the effectiveness of a brief inference intervention, (b) compare the types of knowledge-based inferencing errors less skilled middle grade readers make, and (c) evaluate if self-reported cognitive load relates to inferencing. Participants ($N = 17$) were randomly assigned to a graphic organizer-inference intervention (GO-Inference) ($n = 9$) or business-as-usual (BAU) condition ($n = 8$), and differences between groups were explored for each study purpose. Quantitative and qualitative results suggest that while less skilled readers in the GO-Inference condition made modest progress in forming knowledge-based inferences, they continued to struggle to distinguish relevant from irrelevant information from text and/or retrieve the knowledge necessary to form inferences. Students in the BAU condition were more likely to make errors such as providing irrelevant information or failing to respond. Additionally, students in the GO-Inference condition reported lower cognitive load during inference-making tasks.

Keywords: Knowledge-based inference, reading comprehension, adolescent less skilled readers, cognitive load

More than 80% of eighth-grade students in the United States do not have the content-area knowledge and reading skills they need to enter high school and succeed (ACT, 2019). Further, randomized control trials examining the effectiveness of multi-component reading interventions designed to close this gap among middle-grade struggling readers report effects ranging from small to negligible on standardized measures of comprehension (Scammaca et al., 2016; Scammaca et al., 2015), with more significant effects requiring more than one year of intervention. These findings suggest that (a) current intervention approaches have not been intensive enough to budge the intractable deficits presented by middle-grade struggling readers (see Scammaca et al., 2016; Scammaca et al., 2015) or (b) the components do not adequately build the types of higher-order reading skills (i.e., inference-making, comprehension monitoring, and understanding of text structure) involved in comprehension itself. For this reason, systematic investigations examining the nature of higher order reading skills, such as inference-making, remain a priority in the field. Further, considering students' perceived cognitive load for these higher-order skills can provide important information about the potential effectiveness of interventions.

Knowledge-Based Inferencing

Among adolescent readers, inferencing is the strongest predictor of comprehension after controlling for other reading-related skills (Ahmed et al., 2016). One type of inferences, knowledge-based inferences, is particularly important for comprehension because it helps readers build a meaningful

Barth, A., Newman Thomas, C., Kincaid, H., Ankrum, E., Ruiz, B., & Salazar, L. (2022). Error Patterns in the Knowledge-Based Inference-Making of Less Skilled Middle-Grade Readers: An Exploratory Study. *International Journal for Research in Learning Disabilities, 5*(2), 18-35. https://doi.org/10.28987/ijrld.5.2.18

representation of the text (Kintsch, 1988; O'Brien et al., 2015). Knowledge-based inference-making involves integrating information across sentences and between information in the text with one's existing general knowledge of the topic (O'Brien et al., 2015).

Proficiency in knowledge-based inferencing differs significantly between skilled and less skilled adolescent readers (Barth et al., 2015). That is, less skilled adolescent readers have trouble forming knowledge-based inferences that support understanding of short sections of text (Barnes et al., 2015). Indeed, when required to construct inferences across larger text sections, their accuracy decreases to chance levels (Barth et al., 2015; Cain et al., 2004). Finally, less skilled adolescent readers are less likely, less accurate, and less efficient at forming inferences, particularly inferences that require integrating relevant general knowledge with important information in the text (Barnes et al., 2015; Barth et al., 2015).

Why Might Less Skilled Readers Fail to Form Knowledge-Based Inferences?

First, readers largely fail to form knowledge-based inferences because they lack relevant background knowledge, their knowledge is inaccurate, or their knowledge is incomplete (e.g., Ackerman et al., 1990; Barnes et al., 2015; Cain et al., 2001 ; Casteel, 1993). That is, a knowledge-based inference can only be made when the requisite knowledge to make that inference is available. Second, readers may fail to form knowledge-based inferences because they do not activate, access, or use relevant background knowledge that is stored and available in long-term memory (Barnes et al., 2015; Barnes et al., 1996; Cain et al., 2001). Third, readers may fail to retrieve relevant information from text (Barth et al., 2015) or integrate relevant knowledge with important information in the text (Barnes et al., 2015; Barth et al., 2015; Cain et al., 2001). Fourth, readers may not recognize that an inference is called for to fully understand the text (Whitehead, 1929). Finally, readers may fail to form an inference or may generate an incorrect inference because the cognitive load required to access, retrieve, and integrate relevant background knowledge from semantic memory with important information in the text exceeds their processing capabilities (de Jong, 2010).

Research Informing the Proposed Study

Little research has examined the common sources of inference failure among less skilled readers. To address this gap in the literature, Cain et al. (2001) explicitly trained students in background knowledge to criterion and then examined the effects of background knowledge on inference-making among skilled and less skilled comprehenders age 7 to 8 years old (N = 26). Findings revealed that the errors made by skilled and less skilled comprehenders occurred at different stages in the inference-making process. That is, less skilled comprehenders often failed to accurately recall information that had to be integrated to form the inference. Skilled comprehenders, on the other hand, often failed to accurately integrate the relevant textual premise and knowledge-based item. Finally, a small proportion of errors could be attributable to forming the wrong inference, suggesting that skilled and less skilled comprehenders did a poor job of selecting the relevant information from text and/or from their general knowledge.

Building on the Cain et al. (2001) study, Elbro and Buch-Iverson (2013) examined if use of a graphic organizer designed to help elucidate for readers the textual premise and knowledge-based information that must be integrated to generate an inference led to more accurate inferencing among typically developing sixth-grade students. Following two weeks of daily instruction, students in the graphic organizer condition made significant improvements in inference-making and reading comprehension.

Of importance to the current study is the use and effectiveness of the graphic organizer. The purpose of the graphic organizer was to help students understand that an inference was needed for comprehension and identify the knowledge needed to make the correct inference. The twofold purpose of the graphic organizer is important for several reasons. First, less skilled readers are not strategic in their reading (Pressley, 2002), and graphic organizers have been demonstrated to be an evidence-based practice for students with learning disabilities to improve comprehension, including identification of main idea and development of inferences (Dexter & Hughes, 2011). Second, graphic organizers can provide a scaffold that preserves working memory, thereby reducing cognitive load so that higher-level thinking and reasoning about text can be enabled (Dexter & Hughes, 2010). Third, even when less skilled readers have the knowledge needed to make inferences while reading, they do not generate them as readily as more skilled readers, and in such instances the graphic organizer serves as a prompt. For example, Cain and colleagues (2001) reported that in some cases, students directly answered a question about the knowledge needed to form the inference but did not go on to generate the inference using that knowledge when manded.

A more recent study by Daugaard et al. (2017) reported significant correlations between vocabu-

lary and both comprehension and inference-making among sixth-grade students (N = 53). This provides additional support for the notion that knowledge (i.e., word and world) is essential for inference-making among readers in the secondary grades (Ahmed et al., 2016; Cromley & Azevedo, 2007).

Finally, previous research has established the contributions of working memory to inference-making (e.g., Currie & Cain, 2015; Swanson et al., 2018). Inferencing is a complex process that draws on working memory because readers must search, find, retrieve, and then integrate relevant background knowledge from memory with new information extracted from text. Additionally, researchers and theorists are exploring the role of cognitive load and its relation to working memory and reading (de Jong, 2010; Paas & Merriënboer, 2020). Limited working memory suggests a higher cognitive load (Galy et al., 2012), whereas greater working memory capacity suggests a lower cognitive load. For less skilled older readers, considering their perceptions related to the cognitive load for tasks is important because it relates to their willingness to expend effort and engage in complex learning (Stevenson & Mussalow, 2018). Student perceptions of cognitive load can inform instructional design and screen to identify students who are struggling (Laurie-Rose et al., 2014).

The current study builds on this research base in several ways. First, the study used instructional materials frequently used by social studies classroom teachers to teach content but not regularly applied to teach inferencing. Texts were expository to build both content knowledge and word knowledge. This contrasts to prior work by Cain et al. (2001), who used an experimental procedure that taught students a knowledge base about a fictional world prior to measuring their inferencing skills about that planet. Next, like Elbro and Buch-Iverson (2013), a graphic organizer was incorporated into the knowledge-based inferencing instructional condition to help conditionalize the knowledge needed for inference-making among typically developing middle-grade students. Our study extends this work by focusing on less skilled middle-grade readers and identifying their common sources of inference failure.

Finally, we examined cognitive load across the intervention to determine if the intervention reduced readers' perceived burden of forming knowledge-based inferences. No previous inferencing study has addressed perceived cognitive load. Managing cognitive load is critical for freeing up cognitive space for learning, along with the social validity aspect of providing interventions that are perceived by the learners to be useful and acceptable (Wolf, 1978).

Study Purpose

The present study was designed to:

1. Determine the effectiveness of a brief inference intervention for rural, middle-grade struggling readers.
2. Identify whether the inference errors made by students who complete the inference intervention differ from those of students receiving business-as-usual (BAU) instruction.
3. Determine the extent to which the cognitive load of students who complete the inference intervention is lessened relative to students who complete BAU instruction.

To address the first purpose, intervention effectiveness was explored using quantitative methods, including descriptive statistics, to examine demographic characteristics of students, calculation of effect sizes to determine group differences on dependent variables, and use of Mann Whitney U, a non-parametric test to evaluate group differences in inferencing-error types. Given our small sample, case studies were also developed to demonstrate the relationships between standardized test scores, performance on inference-making, and cognitive load to illustrate variables that might be manipulated and guide future research.

To address the second study purpose, student responses were qualitatively coded for errors using a directed content analysis approach (Hsieh & Shannon, 2005) to look for common features based upon and extending the previous early/late staging of errors proposed by Cain et al. (2001). Common error types that emerged during analysis were named and defined to provide a way of interpreting students' skills and quality of inferencing. These were then ordered into a hierarchy of difficulty based on Cain et al.'s (2001) previous work.

Finally, to address the third study purpose, students were asked to complete an adapted form of the NASA Task Load Index (TLX; Hart, 2006), a measure of cognitive load, to learn about their perceptions of the ease or difficulties they experienced when trying to form inferences post intervention. The original quantitative index was developed by NASA to understand how to manage cognitive attention and maximize productivity. For our purposes, it was adapted to be used by children.

Method

Participants

Study participants were students from one middle school (N = 17) serving students in Grades 6-8, located in a rural, working-class community in the midwestern United States. The student population is <5% Asian, 10.8% African American, <5% Hispanic, 0% Indian, and 79.8% White. One hundred percent of students are economically disadvantaged, as measured by participation in the school's free and reduced-price lunch program (see Table 1).

Table 1
Demographics

Variables	BAU (n = 8)	GO-Inference (n = 9)	School District (n = 345)
Grade			
6	5	3	130
7	3	3	85
8	0	3	130
Gender			
Male	1	5	177
Female	7	4	168
Free/reduced lunch	8	9	345
Race			
White	8	8	244
African American	0	1	54
Hispanic	0	0	24
Other	0	0	23
Special education			
No	7	8	298
Yes	1	1	47

Note. BAU: Business-as-Usual condition. GO-Inference: Intervention condition.

Potential participants were identified through a two-step process. First, based on scores on the state-mandated reading comprehension competency test, the Missouri Assessment Program (MAP; Missouri Department of Elementary and Secondary Education, 2013), from the previous year; students who performed at or below Basic were identified as less skilled readers and recruited for participation. Second, recruited participants were required to return signed parental consent forms approved by the university's Institutional Review Board and their school district board. Students were excluded from participation if the school identified them as Limited English Proficient or as having a significant disability such as blindness, severe cognitive impairment, or severe behavioral/emotional disability.

All students who consented to participate in the study (N = 17) were randomly assigned to either the GO-Inference condition (n = 9) or the BAU condition (n = 8). The racial/ethnic composition of the final sample was 94% White and 6% African American. Sixty-five percent were female, 100% were eligible for free or reduced-price lunch, and no students received special education services. The final sample included a mix of less skilled readers from Grade 6 through Grade 8, with eight students in Grades 6, six in Grade 7, and three in Grade 8.

Instructional Characteristics and Training

Tutors

Instruction in both the GO-Inference and BAU conditions was provided by two special education doctoral research assistants. These tutors received six hours of training on key instructional elements and procedures, features of effective instruction and behavior management, and strategies for supporting student engagement in both instructional conditions. Additionally, they participated in weekly meetings to receive ongoing instructional support and daily check-ins to ensure high levels of fidelity of implementation.

Duration and Intensity

Students received 20 minutes of instruction 4 times per week for 8 sessions over a 2-week period, for a total of 160 minutes. Instruction was delivered in small, mixed-grade groups. Instructional time, study duration, and release from scaffolding mirror Elbro and Buch-Iverson's (2013) instructional delivery.

Graphic Organizer (GO-Inference) Condition

Tutors utilized a researcher-developed, scripted, explicit, instructional routine and graphic organizer (see Figure 1) to help students acquire knowledge and practice knowledge-based inferencing. Graphic organizers are an evidence-based, high-leverage practice for supporting the development of inference-making (Dexter & Hughes, 2013; Elbro & Buch-Iverson, 2013; McLesky et al., 2017). The intervention consisted of three components: (a) scaffolded instruction, (b) feedback, and (c) gradual release of support.

Component 1 – Scaffolded Instruction. Tutors scaffolded instruction to help students gain knowledge and independence in accurately forming

knowledge-based inferences through "think aloud-style" instruction (Kucan & Beck, 1997). The tutor thought out-loud about what they were (a) looking at in text, (b) thinking about relevant knowledge, and (c) doing to integrate information from the text with knowledge from memory. The tutor and students first read the text on ancient Egypt (see Figure 1). Then, the tutor modeled and guided students' formation of knowledge-based inferences using a graphic organizer. The graphic organizer helped to elucidate the information from the text and the general knowledge that should be recalled and then integrated to form knowledge-based inferences about ancient Egypt (see Figure 1).

Component 2 – Feedback. To increase accuracy when forming knowledge-based inferences, tutors gave students written feedback on their graphic organizers. Specifically, the purpose of the feedback was to improve students' accuracy in identifying accurate and relevant information from text or knowledge and then integrating these two pieces of information when forming knowledge-based inferences. For example, the tutor might provide feedback on the accuracy of the prior knowledge, but guide the student to re-examine the paragraph with relevant textual information as students often retrieved proximal information from text rather than the exact information called for. Alternatively, if a student provided no response to prior knowledge, the tutor might supply that knowledge. If the student provided no response to textual information, the tutor might provide cues for where/what to look for. Students used the written feedback to correct any part of their inference (i.e., identification of information from text or knowledge and the integration of text with knowledge).

Component 3 – Gradual Release of Support. The tutor gradually released the amount of instructional support. By the third intervention session, students independently read the passage and inference question, identified important information from text and knowledge required to form the inference, and generated their knowledge-based inference using the graphic organizer. In their small group, with the tutor, students discussed the text and knowledge components of their inference, shared their knowledge-based inference, and used peer feedback and/or written feedback from the tutor to revise any part of their inference.

Business-as-Usual (BAU) Condition

In this condition, tutors used a scripted instructional routine to help students identify the main ideas of informational texts on ancient Egypt. This instructional routine was widely used in the district's English language arts and social studies classrooms to support reading comprehension, was familiar to students, and, therefore, could be considered business-as-usual. In this condition, students were prompted to identify important details in the text and then synthesize them into a main idea statement. Research has supported the contribution of main ideas when forming inferences (Fritschmann et al., 2007), and main idea is a commonly implemented and effective intervention to improve comprehension and build general knowledge (Kim et al., 2012). Instruction consisted of two components: (a) fluent reading and understanding of text and (b) formation of main idea statements.

Component 1 – Fluent Reading and Understanding of Text. To support students' fluent reading and understanding of the text, tutors used an explicit instructional sequence (Archer & Hughes, 2011) of "I do," "we do," and "you do" for each instructional session. Tutors orally read the text, focusing on modeling appropriate fluency, prosody, and pronunciation of vocabulary. Next, they directed the students to whisper-read the passage aloud synchronously. Finally, students read the passage silently.

Component 2 – Main Idea of Text. Following the three readings of the text, the tutor guided students' identification of the main ideas of the text. Small-group discussion helped students to prioritize the most important ideas in the text.

Instructional Materials

For both the GO-Inference and BAU conditions, each instructional lesson consisted of the same 5-9 sentence informational paragraphs ranging in Lexile from 1,000L-1,220L, from the book *Egypt World* (Caldwell, 2013). Each passage included a researcher-developed knowledge-based inference question, which students in the GO-Inference instructional condition answered by completing a graphic organizer (see Figure 1). BAU students completed a worksheet identifying the main idea that would be required if asked to make an inference.

Fidelity of Implementation

Tutors audio-recorded all intervention sessions and completed an implementation fidelity checklist that was specific to their condition. Using the audio-recordings, fidelity checklists were also completed by the primary investigator for four of eight sessions (50%). Checklists documented completion of the intervention components in the GO-Inference and BAU conditions. Using percent agreement, implementation reliability was calculated as 97% for the

GO-Inference condition and 95% for the BAU condition, demonstrating that both interventions were implemented with high fidelity.

Procedures

Students completed a pretest battery designed to describe basic reading skills, verbal and content knowledge, word reading fluency, inference-making, nonverbal reasoning, and working memory. Pretest assessments were completed within a two-week window prior to the first day of intervention. Following the instructional period, students completed a posttest assessment battery designed to examine the types of inference errors middle-grade less skilled readers make. Reading comprehension, verbal knowledge, and content knowledge were also reassessed. Posttest assessments were completed in a two-week window following the last day of intervention. Three months later, a delayed posttest battery was administered to measure the retention of content knowledge and inference-making error types. The delayed posttest assessments were completed in a one-week window. Testing at all three timepoints occurred in the school library during a time identified by the school principal.

Following completion of an extensive training program on test administration and scoring, two graduate research assistants administered all assessments to individual students. After all participants had completed pretest or posttest, research assistants evaluated the fidelity of test administration using a two-step process. First, the research assistants double-checked their own item-level scoring and calculation of raw scores and standardized scores. Second, the research assistants verified each other's item-level scores, raw scores, and standardized scores.

Measures

Reading Comprehension. The Gates-MacGinitie Reading Tests-Fourth Edition (GMRT-4; MacGinitie, 2000) Reading Comprehension subtest is a timed (35-minute), group-administered assessment consisting of expository and narrative passages ranging in length from 3 to 15 sentences. Students read each passage silently and answer multiple-choice questions. Internal consistency reliability ranges from .91 to .93, and alternate form reliability is reported as .80 to .87 (MacGinitie, 2000). The Reading Comprehension subtest was administered at pretest and immediate posttest.

Word Reading Fluency. The Test of Word Reading Efficiency-2 (TOWRE-2; Torgesen et al., 2012) is an individually administered, standardized assessment of word reading fluency. For the Sight Word Efficiency subtest, the participant decodes a list of 104 real words as accurately and efficiently as possible within 45 seconds. For the Phonemic Decoding Efficiency subtest, the participant decodes a list of 63 nonwords as accurately and efficiently as possible within 45 seconds. Alternate form and test-retest reliability coefficients exceed .90 for students in the middle grades. The TOWRE-2 was administered at pretest.

Verbal Knowledge. The GMRT-4 Vocabulary Test (MacGinitie, 2000) is a standardized assessment of verbal knowledge. Students read and answer 45 multiple-choice items that assess word, world, and content knowledge. Internal consistency for students in Grades 6-8 is .83-.89. The GMRT Vocabulary Test was administered at pretest and immediate posttest.

Nonverbal Reasoning. The Kauffman Brief Intelligence Test-2 (K-BIT-2; Kaufman & Kaufman, 2004) Matrices Subtest is a standardized assessment designed to measure fluid thinking. It is individually administered and consists of 46 nonverbal items that involve visual stimuli, both meaningful (i.e., people and objects) and abstract (i.e., designs and symbols). All items are multiple-choice, requiring the participant to point to the correct response or to say its letter. Internal consistency coefficients (split-half) for the nonverbal scores for students in Grades 6 through 8 range from .86 to .91. The Matrices subtest was administered at pretest.

Inference-Making. The Test of Language Competence-2 (TLC-2; Bowers et al., 2009) Listening Comprehension-Inference-Making subtest requires that the student form two plausible inferences on the basis of two sentences that describe the beginning and end of a causal chain. The two inferences are selected from four statements. All 12 items and 2 practice items are read to the student. Internal consistency for the TLC-2 ranges from 0.59 to 0.70. The Listening Comprehension-Inference-Making subtest was administered at pretest.

Working Memory. The Woodcock-Johnson-III (Schrank et al., 2001) Numbers Reversed subtest and Memory for Words subtest are measures of short-term memory. For the Numbers Reversed subtest, students are asked to repeat a series of digits backwards. This subtest requires the ability to temporarily store and recode information presented orally by the examiner. For the Memory for Words subtest, students are asked to repeat a series of unrelated words. The subtest measures verbal memory span. Reliability for both subtests exceeds .90 for students in Grades 6-8. Both working-memory subtests were administered at pretest.

Cognitive Load. An adapted version of the NASA Task Load Index (TLX; Hart, 2006) was implemented to measure perceived cognitive load. The TLX is a self-re-

ported assessment of workload that allows workers to report the demands of tasks related to mental, physical, and temporal demand, performance, effort, and frustration (see https://humansystems.arc.nasa.gov/groups/tlx/). The instrument can be completed online and uses a sliding 7-point scale. For the purposes of this study, participants' ratings of their performance, effort, and frustration were of interest.

We adapted the scale to be completed by paper and pencil and to be age appropriate, creating a checklist for children to indicate the levels of demand they experienced each day of instruction. A 7-point fixed point scale was employed. For performance, students rated their daily work as *failure, not good, needs work, okay, pretty good, very good*, and *perfect*, with *failure* scored as 1 point, and *perfect* scored as 7. For effort and frustration, students rated the demands from *very low, pretty low, a little low, okay, a little high, pretty high,* and *very high*, with *very low* scored as 1 point, and *very high* scored as 7 points. Students were scheduled to complete this measure following daily instruction. Previous research (Laurie-Rose et al., 2014) demonstrated the validity of an adapted NASA TLX self-report of cognitive load with respect to academic tasks, even for very young children.

Content Knowledge. The Egyptian Content Knowledge Assessment (Barth & Elleman, 2017) is a 25-item custom assessment measuring basic knowledge of ancient Egypt. Of the 25 items, 11 are multiple-choice and 14 are constructed response; 6 items tap vocabulary, 6 items tap inferencing, and 13 items tap literal comprehension. Internal consistency, calculated using Cronbach's alpha, was .89 and .85 among students in Grade 5 and Grades 6-8, respectively. The Egyptian Content Knowledge Assessment was administered at pretest, immediate posttest, and delayed posttest.

Proximal Measure of Knowledge-Based Inferencing. Students read two texts (i.e., Nile and Building Pyramids) selected from the Qualitative Reading Inventory-5 (QRI-5; Leslie & Caldwell, 2010) and answered knowledge-based inference questions developed for each text. The Nile passage is 294 words in length, 850 Lexiles in difficulty, and includes 5 knowledge-based inference questions. The Building Pyramids passage is 304 words in length, 850 Lexiles in difficulty, and includes 4 knowledge-based inference questions.

For this task, student read the Nile passage and then orally answered the knowledge-based inference questions. After answering all questions, students received the Graphic Organizer-Inference and were prompted to complete it for each inference question. Next, students completed the Building Pyramids passage using the same format. Internal consistency for Nile and Building Pyramids Total score was .83 (Barth & Elleman, 2017). The Proximal Measure of Knowledge-Based Inferencing was administered at pretest, immediate posttest, and delayed posttest.

Curriculum-Based Measure of Knowledge-Based Inference-Making. The Curriculum-Based Measure (CBM) of knowledge-based inference-making consists of three passages drawn from the book *Egypt World* (Caldwell, 2013). Students read each passage aloud and answered a knowledge-based inference question using the Graphic Organizer-Inference. The graphic organizer permitted examination of the information for text, relevant knowledge, and integration of text with knowledge. The CBM Knowledge-Based Inference-Making measure was administered at immediate posttest and delayed posttest.

Analysis Plan

To address the first study purpose, descriptive statistics for the pretest assessments were first calculated. Second, Cohen's *d* and Hedges' *g* were calculated to quantify how much the treatment group differed from

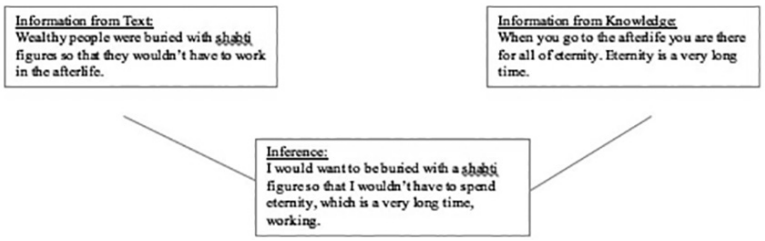


Figure 1

Sample Passage and Graphic Organizer for Knowledge-Based Inferencing

"Ancient Egyptians worried that they might be expected to work in the afterlife. For this reason, wealthy people were often buried with small shabti figures. It was thought that these model servants, inscribed with a spell that allowed them to spring to life when needed, would be able to perform the tasks instead" (Caldwell, 2013). Why would you want to be buried with a shabti figure?

the BAU group on the proximal measures of inference-making following the short intervention period. Both statistics are similar except when sample sizes are below 20 participants, then Hedges' *g* outperforms Cohen's *d* (Durlak, 2009; Hedges & Olkin, 1985).

To address the second study purpose regarding inferencing error types (see Table 2), the Mann Whitney *U*, a nonparametric test, evaluated differences between GO-Inference and BAU conditions on the QRI Knowledge-Based Inference Assessment and CBM Knowledge-Based Inference-Making. The Mann Whitney *U* is a robust test that is appropriate for analysis of ordinal data and accommodates the small sample sizes and non-normal distributions for some data in this study (Nachar, 2008).

A directed content analysis (Hsieh & Shannon, 2005) was conducted to explore the types of errors made by the middle-grade participants. Previous work by Cain et al. (2001) with younger children had staged inference-making errors into (a) early errors by less skilled comprehenders (e.g., failing to recall prior knowledge required) and (b) late errors by more skilled comprehenders (e.g., failure to accurately integrate textual information with prior knowledge). These stages were the starting point for coding in the current study, with the intent of exploring application to older students and also to further define error types. Each error was identified and grouped for similarities by the third author, including all individual student errors across both instructional conditions. Error types were given initial names/codes. The second author independently reviewed each error group for agreement about similarities and naming conventions. The second and third authors met to discuss any differences and complete data reduction and naming. This directed content analysis extended Cain et al.'s (2001) conceptualization of early and late errors to include five types of errors from text, five types of errors from knowledge, and nine types of errors that occurred during the integration of text and knowledge (see Figure 1).

Finally, to address the third study purpose related to the cognitive load associated with knowledge-based inference-making, we compared the means on the three dimensions of cognitive load – performance, effort, and frustration. Also, because no study to date has examined the relation between inferencing and cognitive load, data from individual case studies of student participants were examined to elucidate any trends that may inform future research and future intervention design.

Table 2
Knowledge-Based Inferencing Error Types

Category of Error	Specific Type of Error
Text	Inaccurate Information From Text
	Irrelevant Information From Text
	Text Information Replaced With Knowledge Information
	Omission of Information From Text
	Total Text Errors
Knowledge	Inaccurate Information From General Knowledge
	Irrelevant Information From General Knowledge
	General Knowledge Replaced With Text Information
	Omission of General Knowledge
	Total Knowledge Errors
Integration of Text and Knowledge	Inaccurate Text Information Used During Integration
	Irrelevant Text Information Used During Integration
	Omission of Text Information During Integration
	Inaccurate General Knowledge Used During Integration
	Irrelevant General Knowledge Used During Integration
	Omission of General Knowledge During Integration
	Total – Text Integration Errors
	Total – Knowledge Integration Errors
	Totals – Knowledge and Text Integration Errors

Results

Descriptive Statistics

Table 2 reports demographic information for the GO-Inference and BAU conditions and the participating school district.

Table 3 reports descriptive statistics on standardized measures of word reading fluency, vocabulary, reading comprehension, nonverbal intelligence, inference-making, and working memory for students randomized into GO-Inference and BAU conditions. As seen when examining pretest performance, no significant differences existed between the BAU and GO-Inference intervention groups at pretest. Participants performed approximately one standard deviation below

the mean on a standardized measure of word reading fluency and approximately two thirds of a standard deviation below the mean on a standardized measure of reading comprehension, nonverbal reasoning, and working memory. Students performed at the 50th percentile on the standardized measure of inference-making. Also, students performed close to the mean on a standardized measure of verbal knowledge but generally knew less than half of the Egyptian content knowledge before instruction began (see Table 3).

Table 3
Pretest Descriptive Statistics

Assessment	GO-Inference		BAU	
	Mean	*SD*	*Mean*	*SD*
TOWRE	84.78	12.03	80.25	4.68
GMRT Vocabulary	99.94	13.43	100.39	12.77
GMRT Comprehension	92.72	10.38	90.15	11.83
KBIT-2 Matrices	88.67	19.14	87.13	14.02
TLC-2 Listening Comp – Inference-Making	9.11	1.83	10.63	1.51
WJ Memory for Words_SS	89.0	11.19	88.88	22.56
WJ Memory for Words_W	484.11	12.30	480.50	32.85
WJ Numbers Reversed_SS	83.89	10.47	87.34	27.78
WJ Numbers Reversed_W	491.0	12.0	491.20	31.01

Note. *Group differences $p<.05$. SS = standardized score; W = w-score.

Study Purpose 1: To determine the effectiveness of a brief inference intervention for rural, middle-grade struggling readers.

Table 4 reports performance on the measures administered across the three testing time points (i.e., pretest, immediate posttest, and delayed posttest). Hedges' *g* suggests that at the immediate posttest, very small differences existed between the BAU and GO-Inference groups on the Egyptian Content Knowledge total score ($g = 0.04$). Looking closely at the literal items on the assessment, we can see that the BAU condition, which focused on generating main idea statements, outperformed the GO-Inference condition ($g = -0.39$). In contrast, on the inference items, the GO-Inference condition performed slightly higher than the BAU ($g = 0.14$). At the delayed posttest, the literal and inference scores favored the GO-Inference condition ($g = .18$ and $g = .14$, respectively).

On the Proximal Measure of Knowledge-Based Inference-Making, the GO-Inference condition outperformed the BAU condition on the total raw score ($g = .86$), text score ($g = .80$), knowledge score ($g = .39$), and integration score ($g = .62$). Students in both conditions performed near the floor on knowledge and integration at pretest. A similar pattern was found at the delayed posttest, with effect sizes ranging from $g = 0.09$ to 0.81.

The CBM Knowledge-Based Inference-Making, comprised of texts from the intervention, was administered at the posttest and delayed posttest. At posttest, students in the GO-Inference condition performed higher than those in the BAU condition on the total raw score ($g = 0.27$), text score ($g = 0.51$), and knowledge score ($g = 0.24$), but lower than BAU on the integration score ($g = -.60$). At the delayed posttest, the GO-Inference condition again outperformed the BAU on the raw score ($g = .55$), text ($g = .51$), and knowledge score ($g = 0.16$) but now also on the integration score ($g = 0.65$).

Study Purpose 2: To identify whether the inference errors made by students who complete the inference intervention differ from those of students receiving BAU instruction.

To examine whether the types of inference errors differed by condition, student responses on the Proximal Measure of Knowledge-Based Inference-Making and CBM Knowledge-Based Inferencing were combined and analyzed with respect to the 19 potential errors that could lead to an incorrect knowledge-based inference. Errors were classified as coming from text, knowledge, or an integration of the two (see Table 2). We determined if the text-based and knowledge-based information was inaccurate, irrelevant, omitted, or if students provided information from text when they should have provided general knowledge (and vice versa). Table 5 reports significant differences between GO-Inference and BAU conditions.

When the error scores from the two measures, Proximal Measure of Knowledge-Based Inference-Making and CBM Knowledge-Based Inferencing, were combined, four errors were found to be significantly different between the two groups: (a) Irrelevant Information From Knowledge, (b) Irrelevant Information From Text Used During Integration, (c) Omission of Information From Text Used During Integration, and (d) Irrelevant Information From Knowledge Used During Integration.

First, students in the GO-Inference condition tended to provide significantly more irrelevant information from knowledge when forming knowledge-based inferences than students in the BAU

Table 4
Summary of Descriptive Statistics for Pretest, Posttest, and Delayed Posttest

Measure	Condi-tion	Pretest		Posttest		Comparison		Delayed Posttest		Comparison	
		M	*SD*	*M*	*SD*	*d*	Hedges' *g*	*M*	*SD*	*d*	Hedges' *g*
Gates Vocab	GO	99.94	13.43	93.72	11.02						
	BAU	100.39	12.77	99.14	6.96						
						-0.617	-0.55				
Gates Comprehension	GO	92.72	10.38	92.5	10.07						
	BAU	90.15	11.83	92.01	12.88						
						0.045	0.041				
Content Knowledge RS	GO	10	2.35	11.22	3.15			10.89	3.18		
	BAU	11.13	3.98	11.63	2.83			9.25	4.65		
						-0.145	-0.13			0.444	0.396
Content Knowledge VS	GO	2.22	0.83	1.89	0.78			2.22	1.09		
	BAU	1.38	0.92	1.5	0.76			1.38	0.74		
						0.539	0.48			0.948	0.846
Content Knowledge LS	GO	5.33	3.25	6.11	2.56			5.67	2.55		
	BAU	7.13	2.33	7	1.6			5.13	3.23		
						-0.436	-0.389			0.199	0.178
Content Knowledge IS	GO	2.44	1.13	3.22	1.39			3	1		
	BAU	2.63	1.41	3	1.6			2.75	2.25		
						0.157	0.14			0.156	0.139
QRI Raw Score	GO	3.78	2.22	9.67	3.91			9.56	5.2		
	BAU	3	2.39	6.13	3.87			6.13	4.26		
						0.968	0.863			0.763	0.68
QRI Text Score	GO	3.11	1.7	5.67	2.06			5.22	3.5		
	BAU	2.16	5.67	3.63	2.77			2.39	3.07		
						0.898	0.801			0.911	0.812
QRI Knowledge Score	GO	3.11	0.71	2.22	1.48			2.33	2.13		
	BAU	0.88	1.46	1.63	1.41			2.13	2.36		
						0.434	0.387			0.095	0.085
QRI Integration Score	GO	0.33	0.71	1.78	1.4			2	0.5		
	BAU	0	0	0.85	1.46			1.73	0.76		
						0.693	0.618			0.453	0.404
CBM Raw Score	GO	n/a	n/a	2.44	1.74			2	1.5		
	BAU	n/a	n/a	2	1.31			1.25	1.04		
						0.301	0.269			0.612	0.545
CBM Text Score	GO	n/a	n/a	1.56	1.24			1.44	0.88		
	BAU	n/a	n/a	1	0.76			1	0.76		
						0.571	0.509			0.567	0.505
CBM Knowledge Score	GO	n/a	n/a	0.78	0.67			0.33	0.5		
	BAU	n/a	n/a	0.63	0.52			0.25	0.46		
						0.264	0.235			0.177	0.158
CBM Integration Score	GO	n/a	n/a	0.11	0.33			0.22	0.44		
	BAU	n/a	n/a	0.38	0.52			0	0		
						-0.67	-0.597			0.729	0.65

Note. Content Knowledge RS = Egyptian Content Knowledge Assessment Raw Score. Content Knowledge VS = Egyptian Content Knowledge Assessment Vocabulary Score. Content Knowledge LS = Egyptian Content Knowledge Assessment Literal Comprehension Score. Content Knowledge IS = Egyptian Content Knowledge Assessment Inferencing Score.

condition (U = 14, p = 0.018). Second, compared to the BAU condition, the GO-Inference condition integrated irrelevant information from text when forming knowledge-based inferences (U = 17, p = 0.037). Third, relative to the GO-Inference condition, the BAU condition failed to provide any information from text when forming knowledge-based inferences (U = 6.5, p = 0.001). Fourth, when asked to integrate information from knowledge, the students in the GO-Inference condition provided more irrelevant information than the students in the BAU condition (U = 13.5, p = 0.0135).

Table 5

Significant Differences in Errors Made While Forming Knowledge-Based Inferences by GO-Inference and BAU Condition

Error Type	MWU	*p*	MDN	Range	M.R. Graphic Organizer-Inference	M.R. BAU
Error 1	14	0.02	-1	-2 - 2	11.44	6.25
Error 2	17	0.04	0	-3 - 3	11.11	6.63
Error 3	6.5	0.001	2	-2 - 11	5.72	12.69
Error 4	13.5	0.01	0	-2 - 1	11.50	6.19

Note. $p < .05$ MWU = Mann Whitney *U*. MDN = median. M. R. = Mean Rank. Error 1: Irrelevant General Knowledge Used During Integration on Combined Measure (i.e., Proximal Measure of Knowledge-Based Inferencing Plus CBM Knowledge-Based Inference-Making). Error 2: Irrelevant Text Information Used During Integration on Combined Measure. Error 3: Omission of Text Information During Integration on Combined Measure. Error 4: Irrelevant General Knowledge Used During Integration on Combined Measure.

Study Purpose 3: To determine the extent to which the cognitive load of students who complete the inference intervention is lessened relative to students who complete BAU instruction.

To examine if the cognitive load associated with knowledge-based inference-making varied by condition, we examined the means of the self-reported measure of cognitive load. Descriptive results showed that students in the experimental GO-Inference condition were more confident in their inference-making skills, rating their performance an average of 6.67 on a 7-point scale, compared to students in the BAU condition, who rated their performance an average of 4.73. Students in the latter condition felt that making inferences was more effortful than their peers in the GO-Inference condition, ranking their effort at 3.28, between *a little low* and *okay*. Participants in the GO-Inference condition rated their effort between *pretty low* and *a little low* at 2.5. Students in the GO-Inference condition rated their frustration at 1.67, between *very low* and *pretty low*, while students in the BAU condition rated their frustration a bit higher at 3.32, between *a little low* and *okay*. (See Figures 2 through 4 for comparisons between BAU and GO-Inference groups on the cognitive load measures of performance, effort, and frustration, respectively. See Table 6 for mean scores by participant.)

Student Case Studies

Given the exploratory nature of this research, data from individual case studies of student participants may contribute to understanding of the patterns in the data. Specifically, we examined the reading, cognitive, and cognitive-load data for two students from each condition. In the Go-Inference condition, we

Table 6

Case Study Participant Scores by Condition for Cognitive Load, KBIT-2, TOWRE, and WJ-IV

Participant	Condition	Measure								
		NASA Task Load Index			KBIT	WJ		TOWRE	Gates-MacGinitie	
		Performance	Effort	Frustration	VK	MW	NR	Composite	Vocab	Comp
Participant 7	GO	7	2.71	1	88	82	84	80	87.39	87.4
Participant 15	GO	6.875	1.125	1	104	83	68	66	77.39	77.39
Participant 4	BAU	4.7	3.86	3.57	98	52	90	84	100.3875	87.4
Participant 14	BAU	5.14	5	5.14	97	117	118	84	115.3875	83.387

Note. *pretest scores. NASA Task Load Index of cognitive load provides raw-score estimates of Performance, Effort, and Frustration. KBIT VK: Kauffman Brief Intelligence-2 Verbal Knowledge standard score. WJ MW: Woodcock Johnson-III Memory for Words standard score. WJ NR: Woodcock Johnson-III Numbers Reversed standard score. TOWRE Composite is the Inference-Makings combined standard score for Sight Word Reading Efficiency and Phonemic Decoding Efficiency subtests. Gates MacGinitie Vocab represents the Vocabulary standard score. Gates-MacGinitie Comp represents the Comprehension standard score.

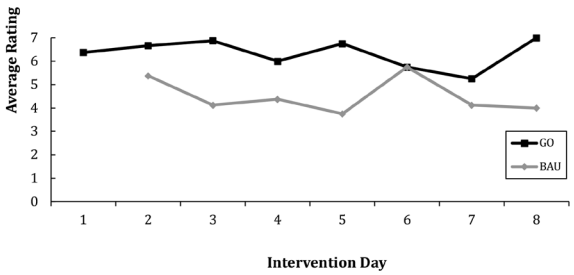


Figure 2
Changes in Average Performance Rating Through Intervention Days

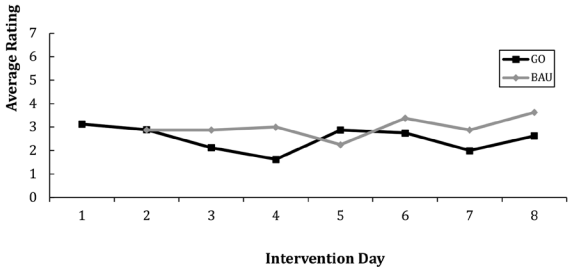


Figure 3
Changes in Average Effort Rating Through Intervention Days

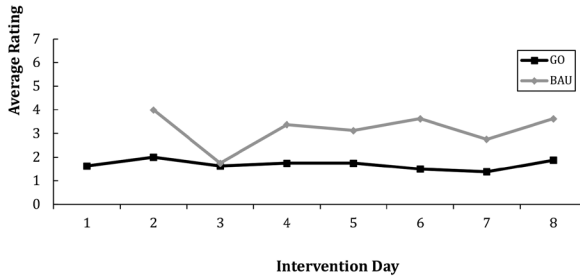


Figure 4
Changes in Average Frustration Rating Through Intervention Days

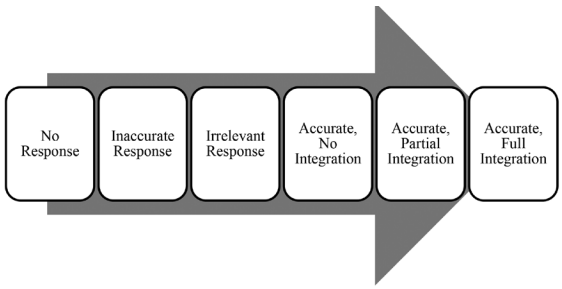


Figure 5
Hypothesized Stages of Errors Made When Forming Knowledge-Based Inferences

examined the performance of Participants 7 and 15. In the BAU condition, we examined the performance of Participants 4 and 14. Below we summarize important information for the selected participants.

Participant 7 (GO)

Participant 7 had a standard score of 88 on the KBIT-2 Verbal Knowledge, representing average potential and indicating average general knowledge. On tests of working memory, standard scores were 82 for Memory for Words and 84 for Numbers Reversed, falling just below average in each case, leading us to anticipate some weaknesses in working memory. The TOWRE standard score of 80 demonstrates sight word and non-sense word reading below average. The Gates-MacGinitie standard scores for vocabulary and comprehension were in the average range. On the cognitive load measure, this student reported confidence in their performance with a score of 7/7, pretty low effort with a score of 2.71/7, and little frustration with a score of 1/7. On Instructional Day 3, in response to a prompt about the Nile being the river of life, this student reported from text that "it had rich black mud," from knowledge that "fertilized land/soil helps to grow crops," making the inference that "the Nile flooded leaving the rich black soil."

Participant 15 (GO)

Participant 15 had a standard score of 104 on the KBIT-2 Verbal Knowledge, falling just above the mean. The participant's standard scores on the tests of working memory were 83 for Memory for Words, just slightly below average, and 68 for Numbers Reversed, indicating a more significant weakness for working memory. The below-average standard score of 66 on the TOWRE indicates that decoding and sight word reading skills are limited. Vocabulary and comprehension as measured by the Gates-MacGinitie were also low, each with scores of 77. In spite of average intelligence as measured by KBIT-2, limitations in working memory and general knowledge are expected to impact reading performance. On the self-report of cognitive load, this student rated their performance at 6.875 on a 7-point scale, almost *perfect*, and rated both effort (1.125/7) and frustration (1/7) as *very low*. On Instructional Day 5, prompted to make inferences

about why Thutmose took revenge on his mother, from the text, the student reported "so she would not be king," and from knowledge, "so her name would not be remembered in history," with the inference "so he could get power and rule Egypt."

Participant 4 (BAU)

For Participant 4, the KBIT-2 Verbal Knowledge score of 98 approached the mean. For working memory, there was a discrepancy between the subtests, with a very low standard score of 52 for Memory for Words and an average score of 90 for Numbers Reversed. Numbers Reversed requires the ability to hold and manipulate numbers in a sequence. Familiarity with numbers and order aids in this task. Memory for Words may compound demand because unrelated words must be manipulated and reported in a correct, new sequence. Both tasks rely on auditory attention and working memory. The weakness in Memory for Words might impact the ability to organize, store, and retrieve sight words and map fluently in decoding. The TOWRE-2 standard score of 84 is just below average for sight word and non-sense word reading. Based on Gates-MacGinitie standard scores, vocabulary skills were at the mean, and comprehension skills were average. On the self-report of cognitive load, this student reported their skills as falling between *okay* and *pretty good* (4.7/7), and reported effort (3.86/7) and frustration (3.57/7) between *a little low* and *okay*.

Participant 14 (BAU)

For Participant 14, the KBIT-2 Verbal Knowledge standard score of 97 approximates the mean. This student had the strongest scores across all participants for working memory, with scores just above average of 117 and 118, respectively, for Memory for Words and Numbers Reversed. Comparably, Gates-MacGinitie vocabulary was 115, at the top of the average range. However, comprehension was 83, falling just below average. TOWRE-2 standard scores estimating sight word reading and non-sense word reading just below average may be related to comprehension limitations. Further, this student reported confidence in their performance as *pretty good* (5/7) and rated their effort (5/7) and frustration (5.14/7) both as *a little high*.

Discussion

In this exploratory study, instruction was designed to promote acquisition of relevant knowledge and to provide opportunities to form knowledge-based inferences. Less skilled adolescent readers are less efficient at forming knowledge-based inferences relative to proficient adolescent readers (Barnes et al., 2015) even when they possess the knowledge needed to form the inference. Inference failure may be caused by several reasons, such as a failure to recall and/or integrate information from the text with relevant knowledge of the topic. In addition, readers may generate incorrect inferences rather than those intended or may fail to make an inference because they do not realize that one is necessary to maintain their understanding of text (Cain et al., 2001).

A failure to form knowledge-based inferences for any one of these reasons will result in a less detailed and integrated understanding of text and impair comprehension. As a result, it is important to understand how to improve the accuracy with which inferences are made, the types of errors students make following instruction, as well as students' perceptions of instruction and their performance. Information about perceived cognitive load can be informative for instructional design and delivery to reduce cognitive load. Further, managing cognitive load can increase student engagement and self-efficacy for learning, factors that are especially important for middle-grade struggling readers (Billingsley et al., 2018; Guthrie & Davis, 2013). Below, we summarize and discuss the results of this study as they relate to these points.

Key Findings

Effectiveness of a Brief Inference Intervention

Following eight intervention sessions (i.e., 160 minutes of instruction), students in the GO-Inference condition outperformed students in the BAU condition on multiple measures of knowledge-based inference-making. Very small but practically meaningfully differences were found on a proximal measure of Egyptian content knowledge and a standardized measure of reading comprehension, both favoring the GO-Inference condition. Results align with those of Elbro and Buch-Iverson (2013) showing large, positive effects of a graphic organizer-based inference intervention for typically developing middle-grade students following 240 minutes of instruction (i.e., eight, 30-minute sessions). It is important to note that the primary difference between Elbro and Buch-Iverson's and the present research is the magnitude of the effect sizes.

Our results suggest that a graphic organizer approach that proceduralizes the inferencing process holds potential for improving inference-making among less skilled middle-grade readers. An important next step is to determine the duration and intensity necessary to both automatize the inferencing process and ensure that less skilled readers form as many inferences while reading as more skilled readers do.

Inferencing Errors

Results of this study also suggest that after providing less skilled adolescent readers with the opportunity to learn a knowledge base on ancient Egypt and practice in the formation of knowledge-based inferences, they had trouble discriminating between relevant vs. irrelevant but related background knowledge needed to form knowledge-based inferences. That is, when failing to form knowledge-based inferences, students in the GO-Inference condition correctly identified related information from text but incorrectly identified the knowledge that should be integrated to form the inference. This is encouraging in that they recognized that an inference was called for, and their knowledge was conditionalized and activated, although they were not able to accurately select the exact information from memory that was needed. This is consonant with the general lack of specificity less skilled readers evidence in word knowledge and finding main ideas (Hogan et al., 2011; Scammaca et al., 2007). In comparison, students in the BAU condition often offered no response, either indicating that they had not learned the knowledge or it was not activated when called for.

Interesting, despite daily exposure, neither group grew significantly in their content knowledge of ancient Egypt, making the attempts by the GO-Inference group to provide the requisite knowledge to form inferences even more noteworthy. Evaluation of the knowledge provided by students in the GO group revealed that the knowledge was not completely wrong. It consisted of information that was topically related but irrelevant for an accurate inference to be made.

At the point of integration, students in the Go-Inference condition failed to form correct knowledge-based inferences for two reasons. First, if the knowledge they identified was irrelevant, they consistently integrated this irrelevant knowledge into their inference. Second, in some cases, students identified the correct textual premise only to integrate irrelevant information from text with general knowledge. Collectively, these results suggest that the difficulties less skilled adolescent readers face when forming knowledge-based inferences occur at the point of knowledge identification and continue throughout the integration process (Barnes et al., 2015; Cain et al., 2001). Further, they suggest that learning about ancient Egypt and the inferencing process was occurring across the intervention and that additional modeling of the inferencing process and opportunity to acquire the requisite knowledge and use it to form inferences has potential for improving inferencing and comprehension outcomes.

Results of this study also indicate that relative to the GO-Inference condition, students in the BAU condition failed to integrate relevant information from text needed to form knowledge-based inferences. That is, although students correctly identified important information from text, they omitted this information upon integration with general knowledge, often offering no response at all. These findings suggest that main idea approaches designed to teach readers how to identify important information in text, but not to understand when and why to use that knowledge should be useful. Thus, main idea instruction in this study did not generalize to inference-making among less skilled adolescent readers, who needed to both identify and integrate relevant textual premises with relevant knowledge when forming knowledge-based inferences. This is supported by recent research showing that main idea approaches do not generalize to improved inferencing among less skilled adolescent readers (Barth et al., 2016).

Based on the types of errors less skilled adolescent readers made when forming knowledge-based inferences, our qualitative data suggest that students appear to progress through stages as they learn to accurately form knowledge-based inferences (see Figure 5). This notion aligns with previous work by Cain et al. (2001) indicating that less skilled readers make inferencing errors at an earlier stage in the inferencing process than more skilled readers.

Our data suggest that regardless of condition, both groups of students progressed through the same stages. However, students in the GO-Inference condition transitioned across a greater number of stages than students in the BAU condition. This suggests that explicit instruction that specifically teaches the process of inferencing and provides targeted corrective feedback helps students to operationalize the process of inferencing. The proposed stages are based on the performance of a small sample of less skilled adolescent readers, and future quantitative and qualitative research is required to validate these stages with a larger, more diverse sample. Our findings do, however, provide a starting point for a finer-grained analysis of errors that can inform instruction and guide intervention.

Our work suggests that in Stage 1, students fail to provide information from text or general knowledge. Stage 2 students provide information from text or general knowledge, but one or both pieces are inaccurate. Stage 3 students provide information from text and knowledge, but one or both pieces are irrelevant. Stage 4 students accurately identify the relevant information from text and general knowledge but fail to accurately integrate both pieces of information. Stage 5 students accurately identify but only partially integrate information from text and general knowledge. Finally, Stage 6

students accurately identify and fully integrate information from text and general knowledge.

Cognitive Load

Additionally, our findings suggest that future inference-making research should consider the perceived cognitive load of this complex task for students and its relationship to learning outcomes. In this study, students in the experimental (GO-Inference) condition rated their performance higher and their effort and frustration as less than their peers in the BAU condition, suggesting that the explicit instruction in inference-making and the scaffold of the graphic organizer helped to reduce the demands on working memory and to manage extrinsic cognitive load.

Managing cognitive load is intended to optimize long-term memory (Chandler & Sweller, 1991). For inference-making, which relies on the efficient search and retrieval of specific information from long-term memory (O'Brien et al., 1998), considering cognitive load in instructional design may support improved learning. With this small sample, examining relationships among cognitive skills such as working memory and cognitive load was not possible, but should be studied in the future, along with additional studies to validate adapted measures of cognitive load for children (Laurie-Rose et al., 2014).

Practical Implications

This experimental trial, using randomization to condition, was designed to improve our understanding of the types of errors less skilled adolescent readers make when forming knowledge-based inferences. The practical implication of these findings suggests first that less skilled adolescent readers experience difficulty identifying relevant information from text or general knowledge needed to form knowledge-based inferences. Consequently, instructional approaches designed to improve inferencing need to explicitly teach students how to both identify and integrate relevant information. Graphic organizers demonstrate some evidence for capacity to scaffold this knowledge activation and integration.

Second, the results suggest that main idea approaches that help readers to identify important information in text, but do not explicitly show students how to integrate important information in text when forming knowledge-based inferences, may not be sufficient for supporting the formation of knowledge-based inferences among less skilled adolescent readers.

Third, as students are learning how to form knowledge-based inferences, they may transition in the types of errors that they make. Consequently, classroom teachers should examine the process of inferencing (i.e., types of errors) as much as the product of inferencing (i.e., was the inference correct or incorrect?). An important next step is to determine how much instruction and practice is required to move students through the error process more rapidly as well as the type of feedback that facilitates this transition. This step is clinically important because it will provide insight about the duration and intensity of instruction needed to help adolescent less skilled readers form knowledge-based inferences during reading as well as the type of feedback that teachers might use to move students across stages.

Fourth, for older struggling readers, who have often experienced substantial and persistent academic failures (Stevenson & Mussalow, 2018), and who have cognitive profiles that predict poor reading performance *and* relate to behaviors such as disengagement, avoidance, and other challenging behaviors, considering their perceptions of instructional conditions can (a) guide instructional design (Billingsley et al., 2018; Guthrie & Davis, 2013), (b) be useful in screening for early intervention (Laurie-Rose et al., 2014), and (c) serve as a potential antecedent preventative for behavior (Stevenson & Mussalow, 2018).

Study Limitations and Future Research

This study provides preliminary information about the nature of inferencing errors among adolescent less skilled readers. The exploratory nature of the study revealed a couple of important limitations to consider for future research. First, although two measures of knowledge-based inferencing were administered, future research should further develop and refine measures of knowledge-based inference-making. For example, readers make a variety of knowledge-based inferences while reading. Inferences may be generated to create causal, spatial, or temporal relations or be made to establish intensions, motivations, emotions, or traits important for maintaining coherence (Hall & Barnes, 2017). Given the wide variety of inferences readers make, inference assessments that more broadly assess these different types could serve to isolate the different points of inference breakdown. Such data could then be used to create more comprehensive inference interventions.

Second, instruction in the current study was only 160 minutes long and designed to promote knowledge acquisition and provide practice in forming knowledge-based inferences. The short duration is consonant with other studies of knowledge-based

inference-making (Cain et al., 2001; Elbro & Buch-Iverson, 2013). However, reading research has established that less skilled adolescent readers are often resistant to intervention and require increased intensity and duration to influence learning (Denton et al., 2013). Future research and practice, therefore, should increase the duration, frequency, and intensity of inference instruction.

Third, future studies should explore multiple intervention components. In the present study, the graphic organizer did cue knowledge retrieval, but the knowledge test demonstrated insufficient knowledge development at posttest, and vocabulary knowledge specific to the curriculum was not assessed. Including explicit instruction in each of the variables that diminish inference-making capacity should be included and evaluated in future multi-component reading interventions.

Fourth, our sample size was small. Although students were randomized to condition, future research should increase the sample size such that it is fully powered to test the effects of intervention. Given the small sample size, for example, it was not possible to observe developmental differences across grades or differences perhaps attributable to demographic variables such as free and reduced-price lunch status.

Fifth, the sample lacked diversity. Future research should diversify the demographics of the sample to better understand how interventions designed to promote knowledge acquisition and teach knowledge-based inferencing generalize across the various subgroups of less skilled readers in the middle grades.

Sixth, research should further explore the hierarchy of error types proposed by this study, as well as develop competing models for consideration. Finally, an adapted version of the cognitive load measure (TLX) was utilized in this study. Future studies might consider using the original measure, with a larger sample size, that is fully powered to detect differences between the intervention and BAU groups.

References

Ackerman, B. P., Silver, D., & Glickman, I. (1990). Concept availability in the causal inference of children and adults. *Child Development, 61,* 230-246. https://doi.org/10.2307/1131062

ACT. (2019). *The condition of college and career readiness.* www.act.org/condition2019

Ahmed, Y., Francis, D., York, M., Fletcher, J., Barnes, M., & Kulsz, P. (2016). Validation of the direct and inferential mediation model of reading comprehension among students in grades 7 through 12. *Contemporary Educational Psychology,* 44-45, 68-82. https://doi.org/10.1016/j.cedpsych.2016.02.002

Archer, A., & Hughes, C. A. (2011). *Explicit instruction: Effective and efficient teaching*. Guilford Press.

Barnes, M. A., Ahmed, Y., Barth, A., & Francis, D. J. (2015). The relation of knowledge-text integration processes and reading comprehension in 7th- to 12th-grade students. *Scientific Studies of Reading, 19*(4), 253-272. https://doi.org/10.1080/10888438.2015.1022650

Barnes, M. A., Dennis, M., & Haefele-Kalvaitis, J. (1996). The effects of knowledge availability and knowledge accessibility on coherence and elaborative inferencing in children from six to fifteen years of age. *Journal of Experimental Child Psychology, 61*, 216-241. https://doi.org/10.1006/jecp.1996.0015

Barth, A. E., Barnes, M., Francis, D., Vaughn, S., & York, M. (2015). Inferential processing among adequate and struggling adolescent comprehenders and relations to reading comprehension. *Reading and Writing, 28*(5), 587-609. https://doi.org/10.1007/s11145-014-9540-1

Barth, A. E., & Elleman, A. (2017). Evaluating the impact of a multi-strategy inference intervention for middle-grade struggling readers. *Language, Speech, and Hearing Services in Schools*, *48*(1), 31-41. https://doi.org/10.1044/2016_lshss-16-0041

Barth, A. E., Vaughn, S., Capin, P., Cho, E., Stillman-Spisak, S., & Martinez, L. (2016). The effects of a text-processing comprehension intervention on struggling middle school readers. *Topics in Language Disorders, 36,* 368-389. ttps://doi:10.1097/TLD.0000000000000101.

Billingsley, G., Thomas, C. N., & Webber, J. (2018). Effects of student choice of instructional method on the learning outcomes of students with comorbid learning and emotional/behavioral disabilities. *Learning Disability Quarterly, 41*(4). https://doi.org/10.1177/0731948718768512

Bowers, L., Huisingh, R., & LoGiudice, C. (2009). *The Listening Comprehension Test—Adolescence*. LinguiSystems, Inc.

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, *29*(6), 850-859. https://doi.org/10.3758/BF03196414

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*(1), 31-42. https://doi.org/10.1037/0022-0663.96.1.31

Caldwell, S. (2013). *Egypt world: Discover the wonders of the ancient land of Tutankhamun and Cleopatra.* Scholastic.

Casteel, M. A. (1993). Effects of inference necessity and reading goal on children's inferential generation. *Developmental Psychology*, *29*(2), 346. https://doi.org/10.1037/0012-1649.29.2.346

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*(4), 293-332. doi:10.1207/s1532690xci0804_2

Cromley J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99,* 311-325. doi:10.1037/0022-0663.99.2.311.

Currie, N. K., & Cain, K. (2015). Children's inference generation: The role of vocabulary and working memory. *Journal of Experimental Child Psychology, 137*, 57-75. doi:https://doi.org/10.1016/j.jecp.2015.03.005

Daugaard, H. T., Cain, K., & Elbro, C. (2017). From words to text: Inference making mediates the role of vocabulary in children's reading comprehension. *Read and Writing, 30*(8), 1773-1788. https://doi.org/10.1007/s11145-017-9752-2

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science, 38*, 105-134. doi:10.1007/s11251-009-9110-0

Denton, C. A., Tolar, T., Fletcher, J. M., Barth, A. E., Vaughn, S. R., & Francis, D. J. (2013). The effects of individualized Tier 3 reading intervention for students with persistent reading difficulties. *Journal of Educational Psychology*, *105*(3), 633-648. https://doi:10.1037/a0032581.

Dexter, D. D., & Hughes, C. A. (2011). Graphic organizers and students with learning disabilities: A meta-analysis. *Learning Disability Quarterly, 34*(1), 51-72. https://doi.org/10.1177/073194871103400104

Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology*, *34*(9), 917-928. https://doi.org/10.1093/jpepsy/jsp004

Elbro, C., & Buch-Iversen, I. (2013). Activation of background knowledge for inference making: Effects on reading comprehension. *Scientific Studies of Reading, 17*, 435-452. https://doi.org/10.1080/10888438.2013.774005

Fritschmann, N. S., Deshler, D. D., & Schumaker, J. B. (2007). The effects of instruction in an inference strategy on the reading comprehension skills of adolescents with disabilities. *Learning Disability Quarterly*, *30*(4), 245-262. https://doi.org/10.2307/25474637

Galy, E., Cariou, M., & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types? *International Journal of Psychophysiology*, *83*(3), 269-275. https://doi.org/10.1016/j.ijpsycho.2011.09.023

Guthrie, J. T., & Davis, M. H. (2003) Motivating struggling readers in middle school through an engagement model of classroom practice. *Reading & Writing Quarterly*, *19*(1), 59-85, doi:10.1080/10573560308203

Hall, C., & Barnes, M. A. (2017). Inference instruction to support reading comprehension for elementary students with learning disabilities. *Intervention in School and Clinic*, *52*(5), 279-286. https://doi.org/10.1177/1053451216676799

Hart, S. G. (2006). NASA-Task Load Index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(9), 90-908. https://doi.org/10.1177/154193120605000909

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press. doi:10.1016/b978-0-08-057065-5.50010-5

Hogan, T. P., Bridges, M. S., Justice, L. M., & Cain, K. (2011). Increasing higher level language skills to improve reading comprehension. *Focus on Exceptional Children*, *44*(3), 1-19. https://doi.org/10.17161/fec.v44i3.6688

Hsieh, H-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277-1288. doi:10.1177/1049732305276687

Kaufman, A. S., & Kaufman, N.L. (2004). *Kaufman Brief Intelligence Test, Second Edition* (K-BIT-2). Pearson Assessment.

Kim, W., Linan-Thompson, S., & Misquitta, R. (2012). Critical factors in reading comprehension instruction for students with learning disabilities: A research synthesis. *Learning Disabilities Research & Practice, 27*(2), 66-78. https://doi.org/10.1111/j.1540-5826.2012.00352.x

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163. https://doi.org/10.1037/0033-295X.95.2.163

Kucan, L., & Beck, I. (1997). Thinking aloud and reading comprehension research: Inquiry, instruction, and social interaction. *Review of Educational Research, 67*(3), 271-299. doi:10.2307/1170566

Laurie-Rose, C., Frey, M., Ennis, A., & Zamary, A. (2014). Measuring perceived mental workload in children. *The American Journal of Psychology, 127*(1), 107. doi:10.5406/amerjpsyc.127.1.0107

Leslie, L., & Caldwell, J.S. (2011). *Qualitative Reading Inventory-5*. Allyn & Bacon.

MacGinitie, W. H. (2000). *Gates-MacGinitie Reading Tests*. Riverside Publishing.

McLeskey, J., Barringer, M-D., Billingsley, B., Brownell, M., Jackson, D., Kennedy, M., Lewis, T., Maheady, L., Rodriguez, J., Scheeler, M. C., Winn, J., & Ziegler, D. (2017). *High-leverage practices in special education*. Council for Exceptional Children & CEEDAR Center. https://ceedar.education.ufl.edu/wp-content/uploads/2017/07/CEC-HLP-Web.pdf

Missouri Department of Education and Secondary Education. (2013). *MAP: LEA guide to the Missouri Assessment Program. Reading.* http://dese.mo.gov/college-career-readiness/asessment/assessment-technical-support-materials

Nachar, N. (2008). The Mann Whitney *U*: A test for assessing whether two independent samples come from the same

distribution. *Tutorials in Quantitative Methods for Psychology, 4*(1), 13-20. https://doi.org/10.20982/tqmp.04.1.p013

O'Brien, E. J., Cook, A. E., & Lorca Jr., R. F. (2015). *Inferences during reading*. Cambridge University Press. https://doi.org/10.1017/CBO9781107279186

Paas, F., & van Merriënboer, J.J.G. (2020). Cognitive-load theory: Methods to manage working memory load in the learning of complex tasks. *Current Directions in Psychological Science*, *29*(4), 394-398. https://doi.org/10.1177/0963721420922183

Pressley, M. (2002). Metacognition and self-regulated comprehension. In A. E. Farstrup & S. J. Samuel (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 291-309). International Reading Association. doi:10.1598/0872071774.13

Scammacca, N. K., Roberts, G. J., Cho, E., Williams, K. J., Roberts, G., Vaughn, S. R., & Carroll, M. (2016). A century of progress: Reading interventions for students in grades 4–12, 1914-2014. *Review of Educational Research*, *86*(3), 756-800. https://doi.org/10.3102/0034654316652942

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4-12: 1980-2011. *Journal of Learning Disabilities*, *48*(4), 369-390. https://doi.org/10.1177/0022219413504995

Scammacca, N., Roberts, G., Vaughn. S., Edmonds, M., Wexler, J., Reutebuch, C. K., & Torgesen, J. K. (2007). Interventions for adolescent struggling readers: A meta-analysis with implications for practice. RMC Research Corporation, Center on Instruction. https://www.centeroninstruction.org/files/Meta-analysis%20Struggling%20Readers1.pdf

Schrank, F. A., McGrew, K. S., & Woodcock, R. W. (2001). *Technical abstract* (Woodcock-Johnson III Assessment Service Bulletin No. 2). Riverside Publishing.

Stevenson, N. A., & Mussalow, P. R. (2018). The effects of planning, goal setting, and performance feedback on avoidance behaviors for struggling readers. *Journal of Positive Behavior Interventions, 21*(3), 171-180. doi:10.1177/1098300718804566

Swanson, E., Barnes, M., Fall, A.-M., & Roberts, G. (2017). Predictors of reading comprehension among struggling readers who exhibit differing levels of inattention and hyperactivity. *Reading & Writing Quarterly, 34*(2), 132-146. doi:10.1080/10573569.2017.1359712

Torgesen, J. K., Wagner, R., & Rashotte, C. (2012). *TOWRE–2 Test of Word Reading Efficiency*. Pro-Ed.

Whitehead, A. (1929). *The aims of education.* Macmillan.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*, 203-214. https://doi.org/10.1901/jaba.1978.11-203

Acknowledgments

The authors wish to acknowledge the invaluable contributions of the middle school staff and students in Laura Speed Elliot Middle School, Boonville School District, and the assistance of Principal Fred Smith.

Parent and Teacher (Dis)Agreement on the Conners Rating Scale: Revised-Long Format

Karen E. Waldie,[1] Christine M. Cornforth,[2] Edwin A. Mitchell,[1] Rebecca F. Slykerman,[1] Claire Wall,[1] and John M. D. Thompson[1]
[1] The University of Auckland, New Zealand
[2] The University of Liverpool, United Kingdom

Abstract

The Conners Rating Scale: Revised-Long version (CRS: R-L) has been used widely in clinic and research settings to measure child behavior and symptoms of attention deficit hyperactivity disorder (ADHD). Despite evidence of poor agreement between parent and teacher reports of child internalizing and externalizing problems, no study to date has examined the levels of agreement between parent and teacher reports of child behavior using the CRS: R-L. Our aim was to measure the level of agreement between parent and teacher perception of child emotion and behavior problems using the CRS: R-L. A total of 591 New Zealand European mothers and their 7-year-old children participated in Phase 4 of the Auckland Birthweight Collaborative (ABC) study. Child behavior was measured when the children were 7 years old using the parent and teacher versions of the CRS: R-L. Parent and teacher agreement was analyzed using weighted kappa coefficient, Pearson correlation coefficient, and Bland-Altman plots. Agreement between parent and teacher reports of child behavior was found to range between poor and low. Agreement decreased with reports of increasing behavioral problems and was generally higher for externalizing subscales and lower for internalizing subscales.

Keywords: Child behavior, inter-rater reliability, parent, teacher, longitudinal, small for gestational age (SGA), Conners Rating Scale: Revised-Long version (CRS: R-L)

It is well understood that behavior problems in children can vary from one situational context to another (Achenbach & McConaughy, 1997; Nass, 2005). Therefore, proving the "pervasiveness" of disorders such as attention deficit hyperactivity disorder (ADHD) requires the involvement of multiple informants in the assessment of children's behavior (World Health Organization, 2004).

When assessing a child's behavior, it is important that the chosen informants are familiar with the child across a range of different settings (Miller et al., 2001). As well as standardized clinical interviews, researchers and clinicians often use the parent and teacher report of a particular behavior rating scale as part of a diagnostic and/or treatment assessment. For example, in an Australian study of children with ADHD, parent and teacher ratings of behavior were used in the diagnostic process in 85%-88% of cases (Efron et al., 2016). While it is well understood that the correspondence between direct observation and behavior rating scale diagnoses supports the validity of rating scales for making a diagnosis, behavior rating scales should not be used as a diagnostic tool in isolation (Kazdin, 1995; Nass, 2005).

Parents and teachers are usually the two groups of respondents who are asked to provide information about a particular child's behaviour. Parents provide valuable information as they are considered to be most familiar with their child across time and in different environments (Cohen-Kettenis & Van-Goozen, 2002). For school-aged children, teachers can also be important second informants. They observe children learning, playing, and interacting with their peers in the school setting, which allows them a unique per-

Waldie, K. E., Cornforth, C. M., Mitchell, E. A., Slykerman, R. F., Wall, C., & Thompson, J. M. D. (2022). Parent and Teacher (Dis)Agreement on the Conners Rating Scale: Revised-Long Format. *International Journal for Research in Learning Disabilities, 5*(2), 36-48. https://doi.org/10.28987/ijrld.5.2.36

spective of a child's social and cognitive functioning (Nottelmann, 2002). Furthermore, teachers also have other children of similar ages to compare with, making them better placed at judging whether a child's behaviours are outside the normal range.

It has been widely observed that the joint use of multiple informants has led to problems of cross-informant effects (e.g., Lohaus et al., 2019). Not surprisingly, studies examining agreement levels between different raters of a child's behaviour have found that agreement is generally higher between the parents of a child than between those who observe the child in different environments; for example, teacher vs. football coach (Achenbach et al., 1987) and preschool teacher vs. parents (Penninx et al., 2008).

Two meta-analyses have examined the cross-informant correlations and situational specificity of child behaviour problems (Achenbach et al., 1987; Lee et al., 1994). Both studies reported low levels of parent and teacher agreement. Specifically, Achenbach and colleagues (1987) included 41 studies in their analysis, of which 4 reported a correlation of 0.5 and above; the remaining studies reported correlations ranging from 0.2 to 0.5. Other studies have revealed similar findings (Gagnon et al., 1992; Lee et al., 1994; Mitsis, 2000; Youngstrom et al., 2000). For example, Youngstrom and colleagues (2000) reported a correlation of 0.3 between parent and teacher reports of child externalising problems and 0.2 for internalising problems.

Further, in a meta-analysis of 49 studies on the agreement between informants on externalising, internalising, and social problems in children with autism spectrum disorder (ASD) or intellectual disability (ID), Stratis and Lecavalier (201) found that, consistent with previous studies, the correlation between informants was higher for externalising problems (r = .42) than for internalising problems (r = .35) or social problems (r = 0.30). Level of agreement for each of these domains differed by the child's age, diagnosis, and IQ.

Mitsis and colleagues (2000) examined parent and teacher agreement for 74 children who were clinically referred with ADHD. The results of their study revealed correlations of 0.3 for ADHD Inattentive Type, 0.4 for ADHD Hyperactive-Impulsive Type, and 0.4 for ADHD Combined Type. The authors also reported that agreement was lower for categorical measures using intra-correlations (Mitsis, 2000). In another study of agreement between parent and teacher reports of disruptive behaviours of children diagnosed with ADHD, Angtrop et al. (2002) found no significant associations between parent and teacher report of ADHD symptoms (r = 0.13 for inattention and r = 0.09 for hyperactivity/impulsivity), moderate agreement for problems relating to conduct disorder (r = 0.36), and good agreement for problems relating to oppositional defiant disorder (r = 0.56).

Correlation between paediatrician- and parent-reported internalising and externalising problems in children with ADHD has also been shown to be poor in an Australian study despite most parents agreeing with the diagnosis of ADHD (Efron et al., 2016). For example, in a study of parent and teacher ratings of preschool children, Orylska and colleagues (2016) found that the agreement between parents and teachers on the Conners Early Childhood Behaviour Scale was low to moderate. Furthermore, they noted that parent and teacher ratings of working memory ability mediated the relationship between inattentive/hyperactive behavior on the Conners Early Childhood Behaviour Scale CEC BEH [S] and fluid intelligence (Orylska et al., 2016).

For adolescents, self-reported symptoms of ADHD can be obtained in addition to observer reports. A study of 80 adolescent survivors of childhood cancer found moderate agreement between parent and teacher reports, with lower agreement between self-reports and observer reports of problems on the Conners 3 Rating Scale. Across all scales, parents reported more problems than either teachers or the young persons themselves (Willard al., 2016). Further, comparing agreement on the Conners 3 Rating Scale between teachers, parents, and youth between 6 to 18 years of age, Izzo et al. (2019) found that correspondence between the three groups of informants was moderate only. Finally, in adults, the concordance between self-reported symptoms of ADHD and informant report (partner or parents) has been shown to range between small to moderate, with the authors concluding that multi-informant information is important in the diagnosis of adult ADHD (Abt-Mörstedt et al., 2015).

The Conners Rating Scale: Revised (CRS-R)

The Conners Rating Scale: Revised (CRS-R) is among the most widely used behavior rating scales in the history of research on children with ADHD (Barkley, 1998). The CRS-R Long Format (CRS: R-L) includes parallel parent and teacher questionnaires, with the aim of measuring the various aspects of problems (including those associated with ADHD) in different situational contexts. Cohen and colleagues (1990) compared parent and teacher ratings on the earlier versions of the Conners and found correlations between parent and teacher ratings to be 0.40 for the ADHD Index scale and 0.42 for the Oppositional scale.

Purpose of the Present Study

The aim of the current study was to examine the level of agreement between parent and teacher ratings of behaviour in a community-based sample of 7-year-old New Zealand European children using the CRS: R-L. Some lack of agreement between the different measures is inevitable, but a focus of this study was the extent by which methods disagree (Bland & Altman, 1986). Despite its widespread use both clinically and in research, to our knowledge no studies to date have reported the levels of agreement between parent and teacher reports of child behaviour problems using the CRS: R-L.

It is noteworthy that the CRS: R-L (Conners, 1997) is no longer published or sold and was replaced by the Conners 3 (Conners, 2008). Despite using an earlier version of the Conners, we believe that the current study makes an important contribution to the wider discussion about agreement between teachers and parents when assessing phenomena as complex as behavior and attention.

Symptoms of ADHD need to be present in at least two different settings to indicate a diagnosis according to the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5; American Psychiatric Association [APA], 2013), so multi-informant assessment approaches are typically adopted. Such information and decisions about the child by specialists (psychologists, social workers, special educators, etc.) are largely based on the information provided by parents and teachers. As such, the current research about rater-agreement remains relevant for practice and has implications for the accurate identification of children with clinically significant attention problems.

Method

Participants and General Procedure

The Auckland Birthweight Collaborative (ABC) Study includes a longitudinal cohort with disproportionate sampling. The methodology of previous phases of the ABC study has been described in detail previously (Thompson al., 2001). In brief, participants are mothers and their singleton infants born full term (≥ 37 completed weeks of gestation) in the Auckland and Waitamata Healthcare regions between 1995 and 1997. Approximately half of the infants were small for gestational age (SGA) with birthweights equal to or below the sex-specific tenth percentile for gestation; the remainder were a random sample of infants born appropriate for gestational age (AGA) (Thompson et al., 1994). Infants were excluded from participating in the study if they did not reside in the study region or if they had any congenital abnormalities likely to affect subsequent growth or development.

Perinatal information was collected on 1,714 participants at birth (Phase 1), of which 863 (50.4%) were male, 870 (50.8%) were born AGA, and 871 (50.8%) had mothers of European ethnicity. The cohort was followed up at one year of age (via a postal questionnaire), and face-to-face assessments were undertaken at 3½ years and beyond. At the 3½-year follow-up, the response rates for other ethnicities was substantially lower than for European participants, so in consultation with our M ori co-investigators and advisors, it was concluded that results for these groups were unlikely to be generalizable, and further follow-up attempts with non-European participants ceased.

Data collection for Phase 4 of the study took place when the children were 7 years old. As noted, due to the differential response rates amongst ethnic groups at previous phases of the ABC study, the present study was restricted to New Zealand European mothers and their children. A total of 871 New Zealand European mothers and infants enrolled at birth were eligible for follow-up at 7 years of age; of these, 591 (67.7%) participated.

Of the 591 New Zealand European mothers and their 7-year-old children, we collected complete sets of parent and teacher reports of child behavior for 553 children. The sample consisted of 292 (52.8%) females and 261 (47.2%) males. Compared with all mothers who were initially recruited into the study, respondents at Phase 4 were more likely to be older at the birth of their child ($p<0.0001$), married ($p<0.0001$), of high socioeconomic status ($p = 0.0002$), have tertiary education ($p<0.0001$), be non-smokers during pregnancy ($p<0.0001$) and not use alcohol during this time ($p = 0.02$), have good social support in pregnancy ($p = 0.03$), and have AGA infants ($p<0.0001$). Respondents and non-respondents did not differ on obstetric factors, including gestational age, infant sex, parity, and levels of maternal stress ($p>0.05$).

Measures

Child behavior was measured when study children were age 7 years old using the parent and teacher versions of the Conners Rating Scale: Revised-Long Format (CRS: R-L; Conners, 1997). Parents and teachers indicate on a 4-point Likert scale how much each item applies to a particular child (0 = "not at all," 1 = "just a little," 2 = "pretty much," and 3 = "very much"). Relevant items are summed to provide total

scores for the following behavioural factors that are not specific to ADHD: Oppositional Problems, Cognition, Hyperactivity, Emotion, Perfectionism, and Anxiety. In addition, several subscales specific to ADHD are also provided, including (a) *Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition* (DSM-R-IV; APA, 1994) symptom subscales (Inattentive, Hyperactive/Impulsive, and Total); (c) Conners Global Indices (GGI; Conners, 1997; Restless-Impulsive, Emotional Liability, and Total); and (c) an ADHD Index scale.

The DSM-IV symptom subscales include the ADHD symptom criteria from DSM-IV (APA, 1994). The Conners Global Index includes 10 items (7 relating to Restless-Impulsive symptoms and 3 relating to Emotional Liability). The CGI is recommended for treatment monitoring and characterization of patterns of behavior (Conners, 1997). The ADHD Index includes 12 items relating to behaviors typical of children with ADHD.

The CRS: R-L parent version was given to one parent (usually the mother) of participating children at the 7-year follow-up, with standardized instructions for completion. The CRS: R-L teacher form was posted to each child's current teacher together with standardized instructions and a prepaid envelope. Raw scores were converted to T-scores, which have a mean of 50 and a standard deviation of 10. All subscales were investigated except Psychosomatic, as this is only available for parent report.

Ethical approval for this research was obtained from the North Health Research Ethics Committee (AKX/02/00/319).

Statistical Analyses

Agreement between parent and teacher report of behavior problems was calculated using several statistical methods. Specifically, the relationships between parent and teacher scores were examined categorically using weighted kappa correlation coefficient (K) and continuously using Pearson correlation coefficient (r). The kappa statistic was interpreted according to Landis and Koch (1977): almost perfect 0.81-1.00, substantial agreement 0.61-0.80, moderate agreement 0.41-0.60, fair agreement 0.21-0.40, slight agreement 0.00-0.20, and poor agreement <0. The strength of the correlation coefficients was interpreted according to published guidelines: excellent >0.90; good 0.80-0.90; moderate 0.50-0.70; low 0.30-0.50; poor <0.30 (Andrews et al., 1993).

Agreement was also examined using the statistical procedure outlined by Bland and Altman (Bland, 1986). Bland-Altman plots were used to examine the agreement between parent and teacher report of behavior problems for all CRS: R-L subscales. The Bland-Altman plot includes the within-pair mean (in this study, this is the mean of the parent and the teacher report of behavior problems) to estimate the severity of a child's behavior problems on the abscissa. The difference between parent and teacher reports was used to represent the inter-rater agreement level on the ordinate.

A large and positive value of the difference indicated that parents reported more problems than teachers, whereas a large and negative value indicated that teachers reported more problems than parents. The smaller the absolute value, the higher the agreement between the parents and teachers report of behavioral symptoms.

Statistical analyses were calculated in SAS version 9.4 (SAS Institute).

Results

Means and standard deviations are presented for parent and teacher report of child behavior for all scales (see Table 1), including the Conners Core Factors, Conners ADHD Index, and CGI subscales. All parental report scores were found to be slightly above the mean (T-score >50). Parents' scores were consistently higher than teacher scores except for the Anxious/Shy subscale; however, these differences were not statistically significant (T>0.05).

Atypical behavior problems were defined by scores that were equal to or above 1 *SD* (≥ T-score of 60). Scores that were below this threshold were considered normal. Parents identified a higher prevalence of all problems than teachers, except for the subscale relating to symptoms of anxiety (teacher prevalence = 21.3%, parent prevalence = 17.0%). Kappa statistics for parent and teacher report of behavior for all CRS: R-L subscales ranged from poor to low. The kappa coefficients were comparable to the Pearson correlation coefficients (see Table 2). However, it is possible that the categorization of subscales resulted in a loss of power, which is reflected in the kappa coefficients.

Pearson and kappa correlation statistics were used to assess the relationship between parent and teacher reports of child behavior for all subscales. As shown in Table 2, all Pearson correlations were positive and significant (p<0.0001), ranging from r = 0.15 or Perfectionism to r = 0.50 for Cognitive problems, indicating a significant but poor-to-moderate relationship between parent and teacher reports. Agreement was higher for the externalizing subscales; for example, Conduct problems (r = 0.28) and Hyperactivity (r = 0.43). Subscales related to internalizing problems were found to have lower levels of agreement (e.g., Anxious/Shy problems; r = 0.19). Levels of parent and teacher agreement for the ADHD subscales ranged from r = 0.24 for CGI:

Emotional Liability to $r = 0.47$ for the Conners ADHD Index. A similar pattern was observed for parent and teacher agreement on the DSM-IV ADHD subscales (Inattentive $r = 0.45$, Hyperactive/Impulsive $r = 0.39$, and Total $r = 0.46$).

Bland-Altman plots were produced for all subscales. Figures 1a-2f show the plots for the Conners Core Factors, and Figures 2a-2g show the Conners DSM-IV Hyperactive-Impulsive, Inattentive, and Combined Type ADHD subscales. All plots reveal very similar patterns with great variation in the differences between parent and teacher reports of child behavior. Furthermore, they reveal that the differences between parent and teacher reports for all subscales increased as the mean of the subscale score increased, as indicted by the funnel shape formed by the points.

Table 1

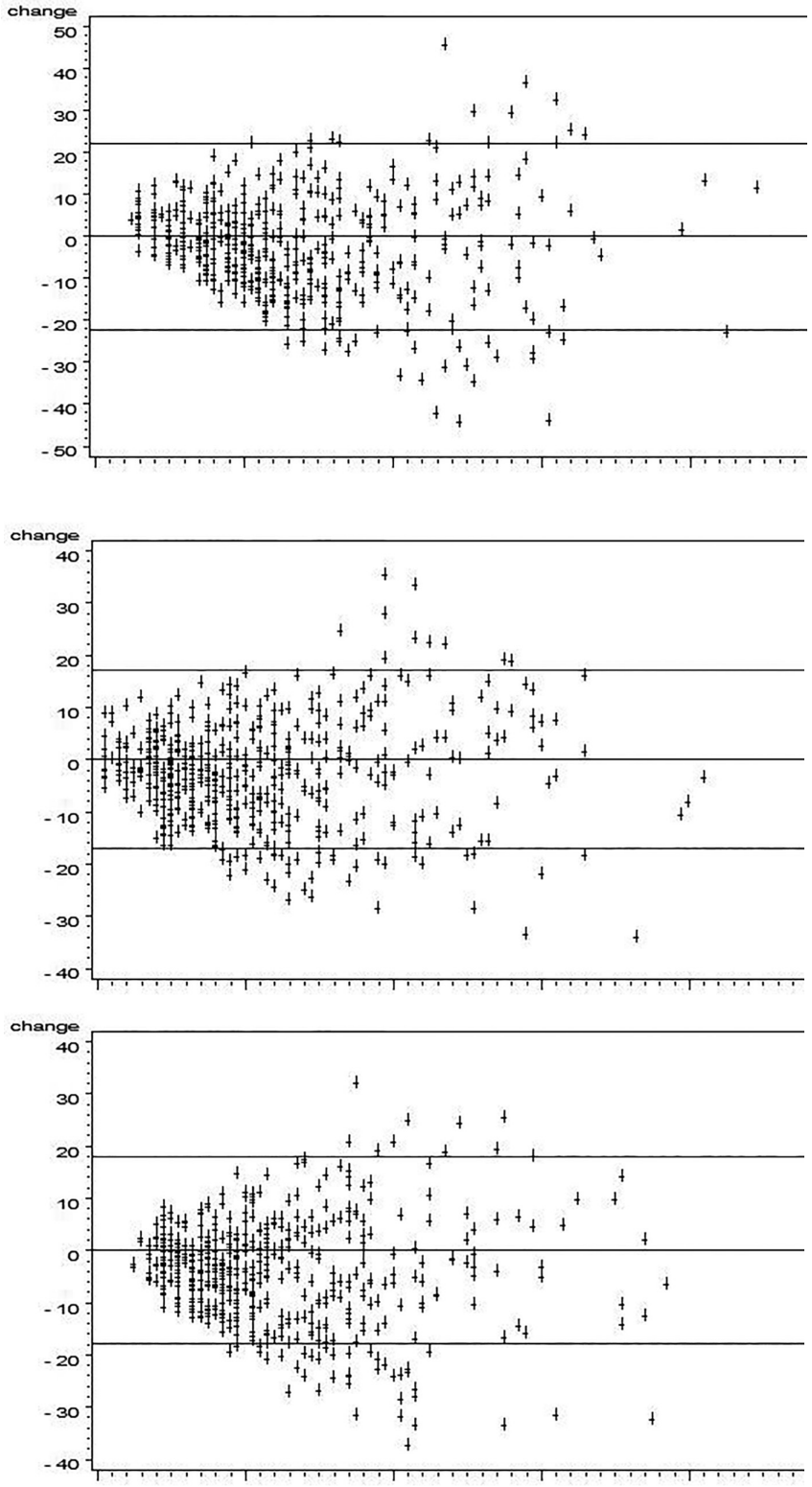
Mean T-Scores and Standard Deviation (SD) for Parent and Teacher Report of Child Behavior by Sex for All CRS: R-L Subscales

	Parent				Teacher			
	Male		Female		Male		Female	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Core Factors								
Conduct	54.62	9.64	54.54	9.52	52.15	8.12	52.56	9.42
Cognitive	51.64	9.34	52.02	7.82	49.32	9.60	49.72	7.47
Hyperactivity	54.14	9.11	53.24	8.24	49.70	8.08	50.68	8.21
Anxious/Shy	50.81	8.98	51.88	9.86	53.63	10.18	52.97	9.53
Perfectionism	52.55	9.17	51.97	9.53	52.30	9.00	50.22	8.36
Social Problems	51.06	8.73	51.04	9.34	49.91	7.80	50.48	8.05
ADHD Subscales								
Conners ADHD Index	51.77	8.73	51.73	7.37	49.87	8.00	49.99	8.49
CGI: Restless Impulsive	54.03	9.25	53.11	8.41	50.91	9.21	51.01	9.20
CGI: Emotional Liability	54.03	9.25	53.11	8.40	50.91	9.20	51.01	9.20
CGI: Total	53.33	8.97	52.53	8.41	50.26	9.22	50.55	9.40
DSM-IV: Inattentive	51.36	8.92	51.00	7.07	48.29	8.41	49.54	7.63
DSM-IV: Hyperactive Impulsive	54.14	9.06	54.12	8.47	49.78	8.40	50.98	9.97
DSM-IV: Total	52.90	9.01	52.52	7.50	48.63	8.02	49.91	7.87

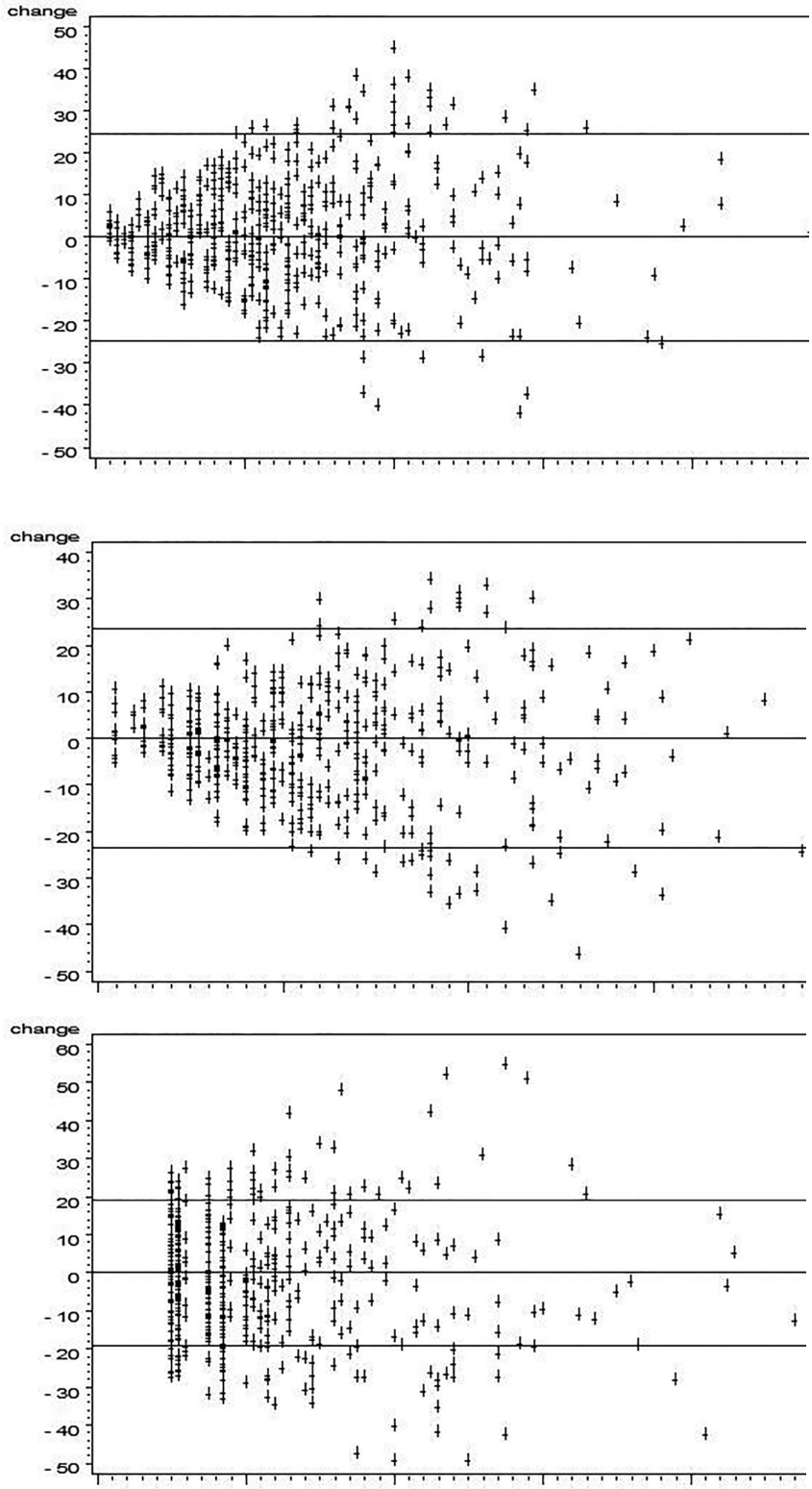
Table 2

Weighted Kappa Coefficient and 95% Confidence Interval (CI) for Parent and Teacher Report of Child Behavior According to All CRS: R-L Subscales

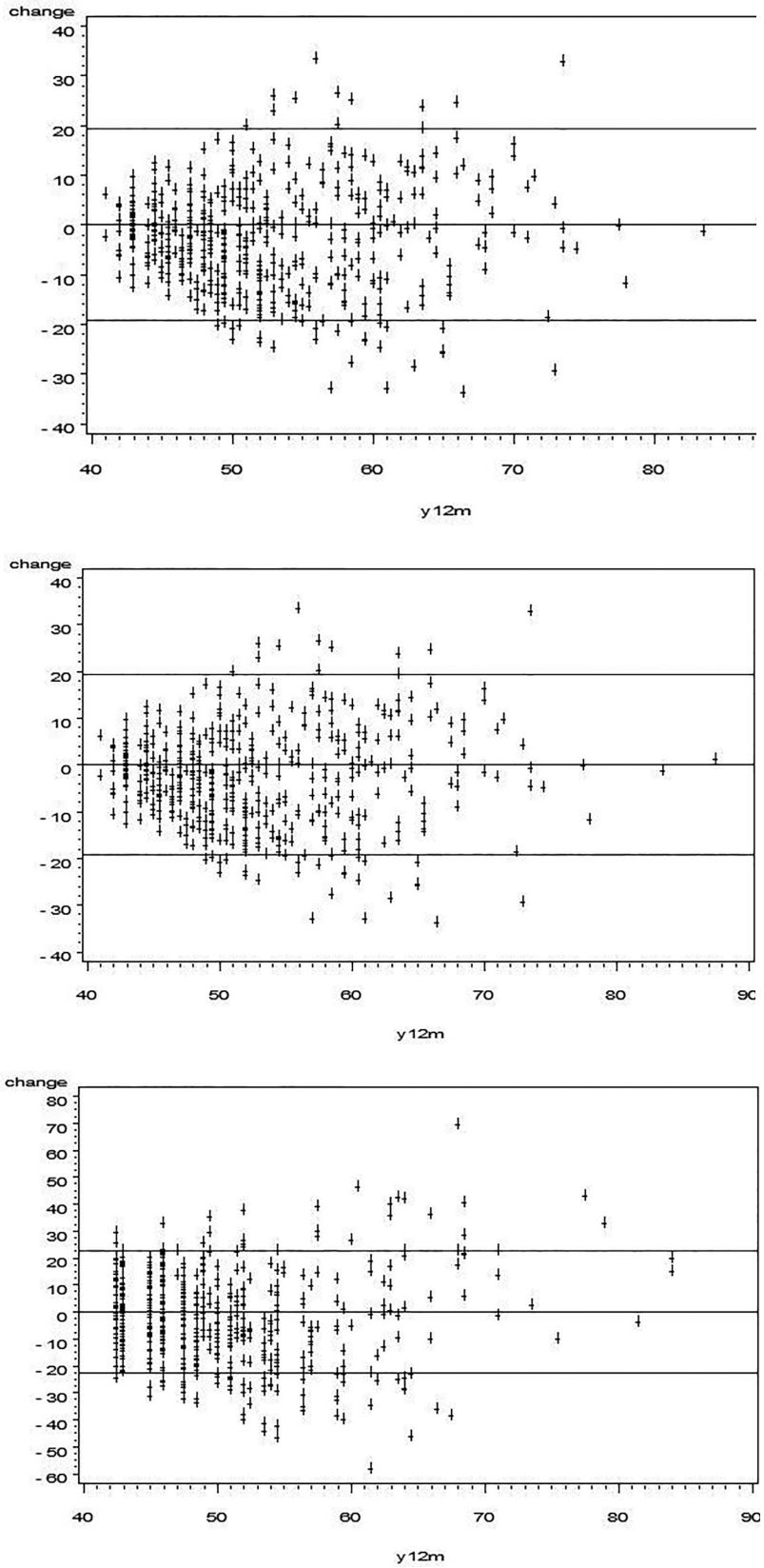
	Weighted Kappa Coefficient		Pearson Correlation Coefficient	
	K	95% CI	*r*	*p*-value
Core Factors				
Conduct	0.15	0.07 – 0.24	0.28	<0.0001
Cognitive	0.30	0.20 – 0.39	0.50	<0.0001
Hyperactivity	0.28	0.18 – 0.38	0.43	<0.0001
Anxious/Shy	0.15	0.06 – 0.23	0.19	<0.0001
Perfectionism	0.13	0.04 – 0.21	0.15	<0.0001
Social Problems	0.27	0.16 – 0.38	0.37	<0.0001
ADHD Subscales				
Conners ADHD Index	0.27	0.18 – 0.36	0.47	<0.0001
CGI: Restless Impulsive	0.26	0.17 – 0.34	0.43	<0.0001
CGI: Emotional Liability	0.14	0.05 – 0.24	0.24	<0.0001
CGI: Total	0.22	0.13 – 0.31	0.40	<0.0001
DSM-IV: Inattentive	0.26	0.16 – 0.37	0.45	<0.0001
DSM-IV: Hyperactive Impulsive	0.27	0.18 – 0.36	0.39	<0.0001
DSM-IV: Total	0.24	0.14 – 0.33	0.46	<0.0001



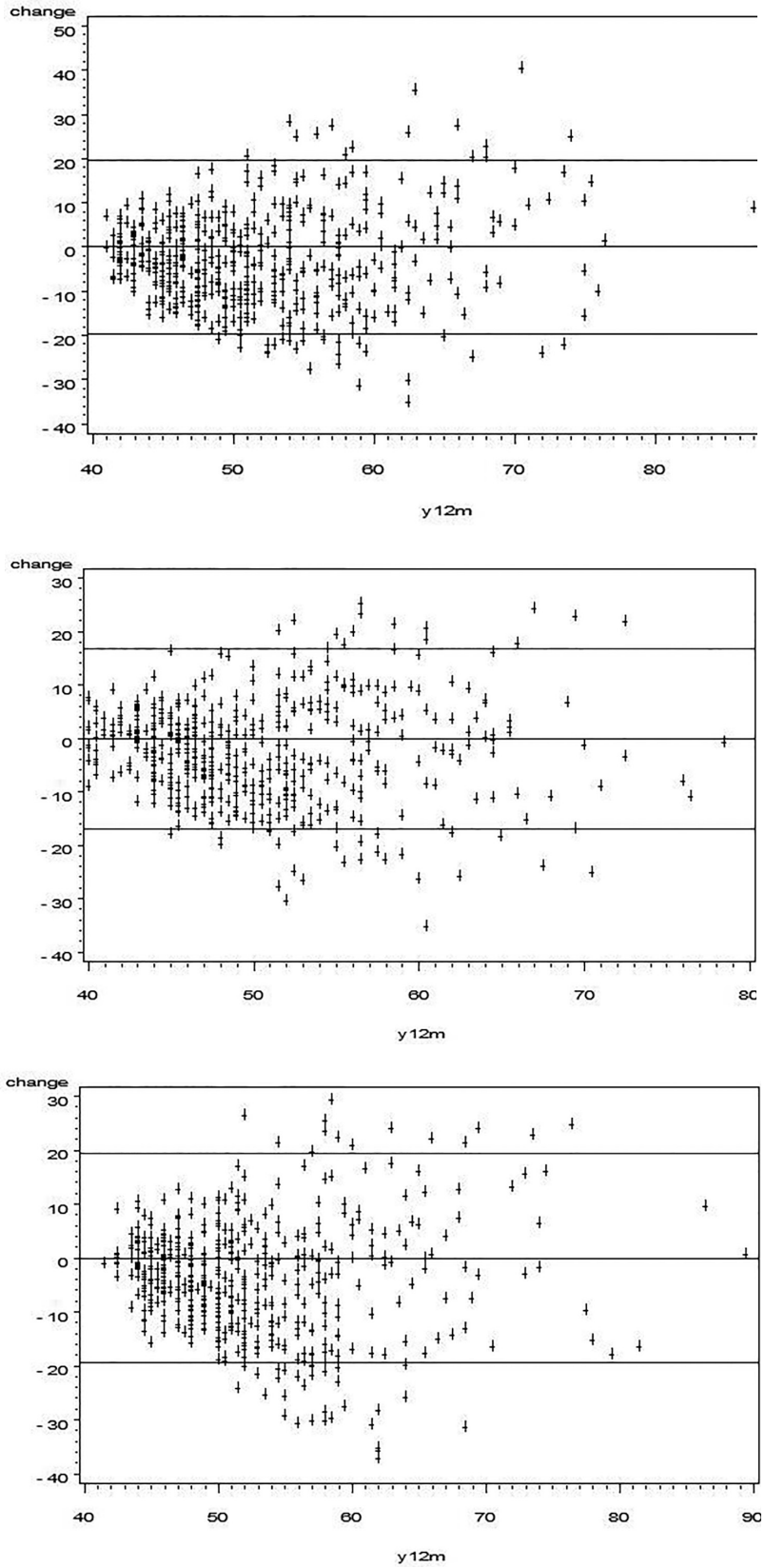
Figures 1a-1c
Bland-Altman plots for CRS: R-L Core Factors



Figures 1d-1f
Bland-Altman plots for CRS: R-L Core Factors



Figures 2a-2c
Bland-Altman plots for CRS: R-L ADHD Subscales



Figures 2d-2f
Bland-Altman plots for CRS: R-L ADHD Subscales

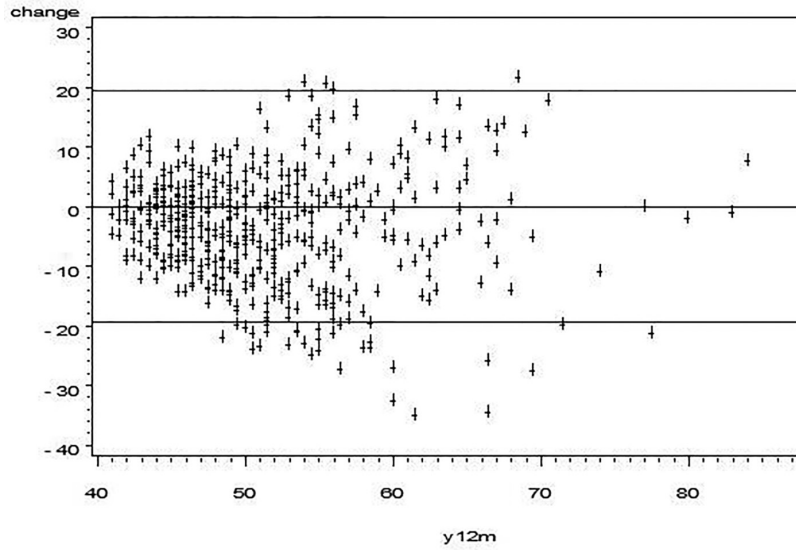


Figure 2g
Bland-Altman plot for CRS: R-L ADHD Subscales

Discussion

To our knowledge, this study is the first to examine parent and teacher agreement using the CRS: R-L. Our main findings were as follows: (a) parent and teacher agreement for all Conners behaviour subscales was poor to low; (b) agreement was highest for the Cognitive problems and DSM-IV Total subscales, and the lowest for the Perfectionism subscale; and (c) parent and teacher agreement declined as behavioural problem scores increased.

The finding that parent ratings were generally higher than teacher ratings is not surprising and is consistent with other studies (Angtrop et al., 2002; Deng et al., 2004; Mitsis, 2000). In a recent study using the latest Conners version (Colomer et al., 2020), compared to adolescents without ADHD, adolescents with ADHD underreported their symptoms and impairment when parent reports, but not teacher reports, were the indicator of performance.

Our study confirms previous findings that parent and teacher reports of child behaviour are not as consistent as one would expect. There are several possible reasons for this. First, children arguably behave differently at school than at home, so the differences in parent and teacher perceptions of child behaviour are due to the different observational contexts. Second, it is also possible that parents and teachers under- or overreport symptoms of behaviour problems depending on the family and/or school climate. For example, parents who are depressed, anxious, or stressed may overreport their children's behaviour problems, and teachers who fear the repercussions of classifying a child as having behaviour problems may also underreport a child's symptoms.

Further, the present study found that agreement between parent and teacher report of ADHD symptoms is low. This is consistent with earlier research (Angtrop et al., 2002; Biederman et al., 1993; Cohen et al., 1990; Loeber et al., 1990; Mitsis 2000). Parent and teacher ratings of oppositional defiant disorder and conduct disorder are generally lower for a normal sample than for a clinically referred group of children (Angtrop et al., 2002).

The current study supports the idea that low levels of agreement do not mean that one informant is providing invalid or unreliable information (Achenbach et al., 1987) and that teachers and parents differ in their perception of problem behaviour (Reid & Maag, 1994). Low levels of parent and teacher agreement may be explained in terms of the situation specificity of the child's behaviour. Thus, it is well known that symptoms of ADHD fluctuate across environments (Barkley, 1998). For example, a child with ADHD may benefit from structured activities, resulting in lower levels of problem behaviour at school. The low levels of agreement nonetheless complicate assessment of ADHD. Some authors (e.g., Collett et al., 2003) have suggested that the term *pervasiveness* does not mean that ratings of both informants need to reach a certain cut-off score in order to meet the diagnostic criteria for ADHD. The DSM-V (APA, 2013) defines pervasiveness as "some impairment from symptoms required in two or more settings" but

does not indicate the degree of impairment required in each setting (p. 62). Findings from this and other studies suggest that assessment protocols are needed in order to clearly define what is meant by pervasiveness in order to overcome the problems of contradictory parent and teacher reports of symptoms.

Similar to other studies in this area, the present study found that agreement was higher for externalising than internalising problems (Achenbach et al., 1987; Loeber et al., 1990). Deng and colleagues (2004) suggested that internalising problems are difficult for teachers to detect in school. If this were true, this study would not find that teachers reported significantly more Anxious/Shy problem behaviours than parents (teacher = 21.3%, parent = 17.0%). One possible explanation for the finding that teachers reported significantly more Anxious/Shy problem behaviours than parents is that children are exposed to new experiences, challenges, or even confrontations in the classroom, which in turn provide more opportunities for anxious, withdrawn behaviours. In contrast, the child's home environment may provide fewer opportunities for such behaviours to be observed.

Consistent with other studies, the present study also found that parent and teacher agreement was highest for the Conners Cognitive Problems subscale (Deng et al., 2004). Willard et al. (2016) reported highest levels of agreement on the Conners 3 Learning Problems subscale in adolescents. As this subscale consists of items that relate to inattention, a possible reason for this finding is that cognitive problems (particularly inattention) are expected to be more stable across environments (Kazdin, 1995).

The present study found that parent and teacher agreement declined as mean scores increased. Although this finding is consistent with two other population-based studies in this area (Kolko & Kazdin, 1993; Deng et al., 2004), it is still surprising. Anecdotally, one would expect that agreement would increase as behaviour problems became more severe. It is important to take into consideration the context specificity of a child's behaviour and how this can impact the level of agreement between different raters. Others have highlighted the importance of this for both clinical and research practice (Achenbach et al., 1987; Deng et al., 2004; Kolko & Kazdin, 1993). It firstly stresses the importance of using multiple informants in measuring children's behavioural problems, and secondly it provides guidance for the clinical evaluation of children's behavioural problems (Deng et al., 2004; Kolko & Kazdin, 1993). If both a child's parent and teacher report relatively few behavioural problems, the child is unlikely to have a disorder. If both a child's parent and teacher report higher levels of behaviour problems, further investigation is required. Moreover, if the reports for parent and teacher differ, further investigation is needed to determine the origin of the difference.

Limitations and Conclusions

The present study has several limitations. First, the CRS: R-L (Conners, 1997) is no longer published or sold and was replaced by the Conners 3 (Conners, 2008) Rating Scales. However, even though our study used an earlier version of the Conners questionnaire, our results reinforce the clinical importance of considering multi-informant perspectives when assessing children. As such, they provide a foundation for future studies on agreement between teacher and parent ratings using later versions of the Conners Rating Scales. It is notable that there are few psychometric studies using the more recent version of the Conners. Only one study has investigated the psychometric properties of the Conners 3 Parent and Teacher scales (Christiansen et al., 2016), and these researchers confirmed the original factor structure (Conners, 2008).

Second, although there are six major ethnic groups in New Zealand (European, Māori, Pacific Peoples, Asian, Middle Eastern/Latin American/African, Other), the sample was restricted to New Zealand European mothers and their children. Because of a poor response rate from children of non-European mothers at ages 1 (Phase 2) and 3.5 years (Phase 3), only participants born to European mothers were included in subsequent phases due to concerns regarding sample representativeness. The response rate for these participants was 67.9% for data used in the present study (Phase 4). Replication with a non-European sample may be useful in establishing the generalizability of our results.

This study has a number of strengths. It was empirically based, using information from multiple informants to assess each child's behaviour. It also adopted empirically based assessment procedures when using psychometrics. These include standardised procedures, the use of multiple items to measure particular aspects of behaviour, aggregated items to provide a quantitative scale for measuring each aspect of behaviour, normative scales to help comparisons of children with relevant reference groups; besides, the measure (Conners) has been tested for reliability and validity to ensure psychometric soundness (Achenbach et al., 1987).

In conclusion, the results of the current study revealed that parent and teacher agreement about behaviour problems using the Conners were poor to low. Agreement was higher for the ADHD subscales,

particularly the DSM-IV subscales, and the Cognitive problems subscale (high scores reflect children who are likely to be inattentive, have organizational problems, have difficulty completing tasks and concentration problems). The study also found that agreement decreases as mean scores increase. It is likely that multiple informants contribute additional variation, which may limit the agreement found between reports by different informants. Children with behavior disorders such as ADHD often exhibit behaviors that are situation-specific; therefore, measures such as the CRS: R-L are likely to provide results that are inconsistent across situations. The American Academy of Paediatrics guidelines (2004) suggest that these discrepancies may be due to differences across settings in terms of expectations, levels of structure, behavioural management strategies, and/or environmental circumstances. In conclusion, the findings from our study indicate that assessment protocols are needed to overcome or identify the reasons for the contradictory parent and teacher reports of symptoms.

References

Abt-Mörstedt, B., Corbisiero, S., Bitto, H., & Stieglitz, R.-D. (2015). Attention-deficit/hyperactivity disorder (ADHD) in adulthood: Concordance and differences between self- and informant perspectives on symptoms and functional impairment. *PLoS ONE, 10*(11). e0141342. 10.1371/journal.pone.0141342

Achenbach, S., & McConaughy, S. M. (1997). *Empirically based assessment of child and adolescent psychopathology – Practical applications* (2nd ed.). Sage.

Achenbach, S., McConaughy, S. M., & Howell, C. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213-232. https://doi.org/10.1037/0033-2909.101.2.213

American Academy of Pediatrics, Steering Committee on Quality Improvement and Management. (2004). Classifying recommendations for clinical practice guidelines. *Pediatrics, 114*(3), 874-877. doi:10.1542/peds.2004-1260

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders, fourth edition* (DSM-IV).

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, 5th edition* (DSM-V).

Andrews, J., Lewinson, P., Hops, H., & Roberts, R. (1993). Psychometric properties of scales for the measurement of psychosocial variables associated with depression in adolescence. *Psychological Reports, 73*, 1019-1046. doi:10.1177/00332941930733pt146

Angtrop, I., Roeyers, H., Oosterlaan, J., & Van Oost, P. (2002). Agreement between parent and teacher ratings of disruptive behavior disorders in children with clinically diagnosed ADHD. *Journal of Psychopathology and Behavioral Assessmen, 24*(1), 67-73. https://doi.org/10.1023/A:1014057325752

Barkley, R. A. (1998). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment*. Guilford.

Biederman, J., Faraone, S. V., Milberger, S., & Doyle, A. (1993). Diagnoses of attention-deficit hyperactivity disorder from parent reports predict diagnoses based on teacher reports. *Journal of the American Academy of Child & Adolescent Psychiatry, 32*(2), 315-317. https://doi.org/10.1097/00004583-199303000-00011

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 1*(8476), 307-310. PMID: 2868172

Christiansen, H., Hirsch, O., Drechsler, R., Wanderer, S., Knospe, E.-L., Gunther, T., et al. (2016). German validation of the Conners 3 Rating Scales for parents, teachers, and children. *Zeitschrift fur Kinder- und Jugendpsychiatrie und Psychotherapie, 44*, 139-147. https://doi.org/10.1024/1422-4917/a000408

Cohen-Kettenis, P., & Van-Goozen, S. (2002). Adolescents who are eligible for sex reassignment surgery: Parental reports of emotional and behavioural problems. *Clinical Child Psychology and Psychiatry, 7*, 412-422. doi:10.1177/1359104502007003008

Cohen, M., Becker, M., & Campbell, R. (1990). Relationships among four methods of assessment of children with attention deficit hyperactivity disorder. *Journal of School Psychology, 28*(3), 189-202. https://doi.org/10.1016/0022-4405(90)90011-U

Collett, B. R., Ohan, J. L., & Myers, K. M. (2003). Ten-year review of rating scales.V: Scales assessing attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*(9), 1015-1037. https://doi.org/10.1097/00004583-200310000-00006

Conners, K. (1997). *Conners Rating Scales – Revised technical manual.* Multi-Health Systems.

Conners, C. K. (2008). *Conners 3rd edition: Manual*. Multi-Health Systems.

Colomer, C., Wiener, J., & Varma, A. (2020). Do adolescents with ADHD have a self-perception bias for their ADHD symptoms and impairment. *Canadian Journal of School Psychology, 35*(4), 238-251. https://doi.org/10.1177/0829573520936457

Deng, S., Liu, X., & Roosa, M. W. (2004). Agreement between parent and teacher reports on behavioral

problems among chinese children. *Journal of Developmental & Behavioral Pediatrics, 25*(6), 407-414. doi:10.1097/00004703-200412000-00004

Efron, D., Bryson, H., Lycett, K., & Sciberras, E. (2016). Children referred for evaluation for ADHD: Comorbidity profiles and characteristics associated with a positive diagnosis. *Child: Care, Health and Development*, *42*(5), 718-724. doi:10.1111/cch.12364

Gagnon, C., Vitaro, F., & Tremblay, R. E. (1992). Parent-teacher agreement on kindergartners' behavior problems. *Journal of Child Psychology and Psychiatry, 33*(7), 1255-1261. https://doi.org/10.1111/j.1469-7610.1992.tb00944.x

Hardy, K. K., Willard, V. W., Gioia, A., Sharkey, C., & Walsh, K. S. (2018). Attention-mediated neurocognitive profiles in survivors of pediatric brain tumors: Comparison to children with neurodevelopmental ADHD. *Neuro Oncology*, *20*(5), 705-715. doi:10.1093/neuonc/nox174

Izzo, V. A., Donati, M. A., & Primi, C. (2019). Assessing ADHD through the multi-informant approach: The contribution of the Conners' 3 Scales. *Journal of Attention Disorders, 23*(6), 641-650. doi:10.1177/1087054718815581

Kazdin, A. (1995). *Conduct disorders in childhood and adolescence.* Sage.

Kolko, D., & Kazdin, A. (1993). Emotional/behavioural problems in clinic and non-clinic children: Correspondence among child parent and teacher reports. *Journal of Child Psychology and Psychiatry, 41*, 253-260. doi:10.1111/j.1469-7610.1993.tb01103.x

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174. https://doi.org/10.2307/2529310

Lee, S. W., Elliott, J., & Barbour, B. (1994). A comparison of cross-informant behavior ratings in school-based diagnosis. *Behavioral Disorders, 19*(2), 87-97. doi:10.1177/019874299401900202

Lohaus, A., Rueth, J.-E., & Vierhaus, M. (2019). Cross-informant discrepancies and their association with maternal depression, maternal parenting stress, and mother-child relationship. *Journal of Child and Family Studies, 29*. https://doi.org/10.1007/s10826-019-01625-z

Loeber, R., Green, S. M., & Lahey, B. B. (1990). Mental health professionals' perception of the utility of children, mothers, and teachers as informants on childhood psychopathology. *Journal of Clinical Child Psychology, 19*(2), 136-143. https://doi.org/10.1207/s15374424jccp1902_5

Miller, M., Bowen, J. R., Gibson, F. L., Hand, P. J., & Ungerer, J. A. (2001). Behaviour problems in extremely low birthweight children at 5 and 8 years of age. *Child: Care, Health & Development, 27*(6), 569-581. doi:10.1046/j.1365-2214.2001.00223.x

Mitsis, E. M., McKay, K. E., Schulz , K. P., Newcorn, J. H., & Halperin, J. M. (2000). Parent-teacher concordance for DSM-IV attention-deficit/hyperactivity disorder in a clinic-referred sample. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(3), 308-313. doi:10.1097/00004583-200003000-00012

Nass, R. D. (2005). Evaluation and assessment issues in the diagnosis of attention deficit hyperactivity disorder. *Seminars in Pediatric Neurology, 12*(4), 200-216. doi:10.1016/j.spen.2005.12.002

Nottelmann, E. (2002). Competence and self-esteem during transition from childhood to adolescence. *Developmental Psychology, 23*, 441-450. https://doi.org/10.1037/0012-1649.23.3.441

Orylska, A., Brzezicka, A., Racicka-Pawlukiewicz, E., Albinski, R., & Sedek, G. (2016). Parent-teacher concordance in rating preschooler difficulties in behavioural and cognitive functioning and their dyadic predicting of fluid intelligence. *Polish Psychological Bulletin*, *47*(1), 81-91. doi:10.1515/ppb-2016-0009

Penninx, B.W.J.H., Beekman, A.T.F., Smit, J. H., Zitman, F. G., Nolen, W. A., Spinhoven, P., et al. (2008). The Netherlands study of depression and anxiety (NESDA): Rationale, objectives and methods. *International Journal of Methods in Psychiatric Research, 17*(3), 121-140. https://doi.org/10.1002/mpr.256

Reid, R., & Maag, J. W. (1994). How many fidgets in a pretty much: A critique of behavior rating scales for identifying students with ADHD. *Journal of School Psychology, 32*(4), 339-354. https://doi.org/10.1016/0022-4405(94)90032-9

Stratis, E. A., & Lecavalier, L. (2017). Predictors of parent-teacher agreement in youth with autism spectrum disorder and their typically developing siblings. *Journal of Autism and Developmental Disorders*, *47*(8), 2575-2585. doi:10.1007/s10803-017-3173-6

Thompson, J. M., Clark, P. M., Robinson, E., Becroft, D. M., Pattison, N. S., Glavish, N., Pryor J. E., Wild, C. J., Rees, K., & Mitchell, E. A. (2001). Risk factors for small-for-gestational-age babies: The Auckland Birthweight Collaborative Study. *Journal of Paediatrics & Child Health, 37*(4), 369-375. doi:10.1046/j.1440-1754.2001.00684.x

Thompson, J. M., Mitchell, E. A., & Borman, B. (1994). Sex specific birthweight percentiles by gestational age for New Zealand. *New Zealand Medical Journal, 107*(970), 1-3. PMID:8295745

World Health Organization. (2004). *World health report 2004: Changing history*.

Youngstrom, E., Loeber, R., & Stouthamer-Loeber, M. (2000). Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems. *Journal of Consulting & Clinical Psychology, 68*(6), 1038-1050. doi:10.1037//0022-006x.68.6.103

The Demise of the Discrepancy Definition of Dyslexia: Commentary on Snowling, Hulme, and Nation

Linda S. Siegel,[1] David P. Hurford,[2] Jamie Metsala,[3] and Timothy N. Odegard[4]
[1]The University of British Columbia
[2]Pittsburg State University
[3]Mount Saint Vincent University
[4]Middle Tennessee State University

Abstract

In this article, we review the proposal of Snowling, Hulme, and Nation (2020) to reinstate the discrepancy definition of dyslexia. Presenting logical reasons and empirical evidence that this definition is not valid, we suggest that there is no need to distinguish between individuals with dyslexia as identified through an IQ-achievement discrepancy framework and nondiscrepant poor readers.

Keywords: Dyslexia, IQ, assessment, dyslexia definition, discrepancy definition

The difficulty in acquiring reading skills continues to be a serious challenge for some. Much is known regarding how to teach the mechanics of reading so that children can become competent readers. In addition, the scientific literature has effectively shown that children who struggle to learn to read can be provided with interventions to help improve their reading skills. The earlier the at-risk child can be identified, the more likely reading failure can be prevented, but for the latter to occur, it is essential that highly predictive and accurate tools be used to identify these children.

The core deficits of dyslexia consist of problems with accuracy and fluency of word recognition and decoding. Specifically, individuals with dyslexia experience difficulties manipulating the sounds of language, mapping sounds onto their representative letters, and using these skills to decode and synthesize words, which, in turn, produces barriers to fluent reading, spelling, comprehension, and writing. Moreover, dyslexia refers to deficiencies in reading skills that are genetically based (Carrion Castillo et al., 2015; Eden et al., 2015; Shao et al., 2016), neurobiologically expressed (Norton et al., 2015), and typically are manifested in phonological processing deficiencies (Vellutino et al., 2004) and poor fluency (Cassar et al., 2005; Lefly & Pennington, 1991; Lyon et al., 2003; Sumner et al., 2014), and lasts a lifetime (Bruck, 1990).

In a recent article, Snowling, Hulme, and Nation (2020) argue that it is challenging to identify children who have dyslexia, and further suggest that it is important to differentiate between children experiencing reading disabilities into (a) those who have dyslexia and (b) those who are nondiscrepant poor readers. That is, according to these authors, it is important to distinguish between poor readers whose reading is significantly lower than their IQ scores (and thus deserve to be identified as having dyslexia) and poor readers who do not show this discrepancy (and who cannot be labeled as having dyslexia). To make this distinction, Snowling et al. argue that the discrepancy definition of dyslexia should be the definition that is used. This requires that to be labeled with dyslexia, a child must demonstrate a discrepancy between their reading abilities and their intellectual functioning as measured by an IQ score. Children with dyslexia, the authors suggest, stand in contrast to children who are poor readers but do not demonstrate discrepancies between their reading skills and intellectual functioning. The latter would be considered "nondiscrepant poor readers" rather than struggling to read due to dyslexia. Nondiscrepant poor readers are those whose reading and IQ scores are similar.

Operationalizing the degree of discrepancy has typically resulted in a difference in which standardized

Siegel, L. S., Hurford, D. P., Metsala, J., & Odegard, T. N. (2022). The Demise of the Discrepancy Definition of Dyslexia: Commentary on Snowling, Hulme, and Nation. *International Journal for Research in Learning Disabilities, 5*(2), 49-54. https://doi.org/10.28987/ijrld.5.2.49

reading scores must be at least one standard deviation below the standardized intelligence score (Farris et al., 2020; Fletcher et al., 2018).

In support of their argument for employing a discrepancy approach to the identification of dyslexia, Snowling and colleagues outline several consequences that followed the removal of the ability-achievement discrepancy criterion from the definition of dyslexia (Snowling et al., 2020). However, eliminating the discrepancy requirement from the definition occurred due to overwhelming evidence that intelligence did not appropriately or accurately differentiate between types of poor readers (described ahead).

The logic of the canonical model of IQ and IQ discrepancy is built on the assumption that IQ measures denote an individual's inherent potential. However, research has made it clear that measures on an IQ test are just as apt to provide a metric of educational opportunities (Ritchie & Tucker-Drob, 2018). As such, the motivation and justification for adopting the canonical model of IQ-achievement discrepancy is based on a flawed assumption. Moreover, educational opportunities are not equally distributed across the population, and reverting to an IQ-discrepancy approach to identifying dyslexia would serve to disadvantage and further disenfranchise certain members of society (Odegard et al., 2022). These include individuals who are not privileged with quality educational opportunities or financial resources to overcome inferior reading instruction by obtaining additional educational services outside of a public-school setting.

Eloranta et al. (2019) provided data germane to conversations surrounding the ability of IQ to provide a stable measure of inherent potential, especially for individuals with reading disabilities. These researchers reported findings from a longitudinal study of Finnish individuals identified with a reading-fluency-based reading disability in childhood. They noted that the verbal IQ scores of the individuals who had been identified with a reading disability had dropped more than 10 standard score points when they were retested as adults. As a result of this drop in scores, individuals with IQ scores in the average to above-average range in childhood had IQ scores in the low-average to below-average range as adults. The same declines were not observed in typically developing individuals sampled from the general population.

To make their argument, Snowling et al. assumed that children with dyslexia are different from nondiscrepant poor readers. Their premise stands on the notion of etiology. That is, they appear to be arguing that children with dyslexia, compared to those with nondiscrepant reading difficulties, have a different set of causes that give rise to their characteristic behavioral deficits in reading and spelling. However, evidence indicates that there is no meaningful distinction between individuals identified with dyslexia using an IQ-achievement discrepancy approach and nondiscrepant poor readers. As a result, there is no utility in differentiating etiology into distinct classes based on intelligence as those classifications will not inform treatment or prognosis. When contextualized in this manner, the assertion that there is utility in differentiating children with dyslexia as defined by discrepancy from those who present with the same behavioral manifestations of reading and spelling deficits who are nondiscrepant is not supported.

Snowling et al. wrote, "While defensible on statistical grounds, to regard dyslexia 'just' as poor reading (and spelling) fails to capture the fact that it has an early onset (in preschool) and is persistent over time" (p. 502). They appear to be arguing that only individuals classified with dyslexia based on an IQ-discrepancy approach exhibit difficulties in preschool and that nondiscrepant poor readers do not. Again, there is no evidence to show that this is the case. Their argument also implicitly suggests that difficulties in preschool are a prerequisite for being identified with reading difficulties that result from dyslexia. As most preschool children are not assessed on skills that are relevant to the acquisition of reading skills, such as phonological awareness, the authors appear to be advocating that unless they demonstrate language-related difficulties in the preschool period, most children who have reading difficulties, even the discrepant ones, cannot be considered to have dyslexia. The literature suggests that poor phonological processing abilities can be recognized during preschool. However, until universal screening occurs in preschool, they cannot be considered a fundamental component of the definition of dyslexia, as evidence of poor preschool phonological processing skills would not be available for a diagnosis of dyslexia for the great majority of children.

Studies that have examined good readers, nondiscrepant poor readers, and children with dyslexia have discovered that both the nondiscrepant poor readers and children with dyslexia performed similarly on reading-related measures and were significantly different from good readers. The two poor reading groups, children with dyslexia and nondiscrepant poor readers, were differentiated based on their intelligence test scores. In studies comparing children with dyslexia and nondiscrepant poor readers, both groups of children performed similarly and very poorly and significantly worse than nondisabled readers (e.g., Hurford, Johnston et al., 1994; Hurford, Schauf et al., 1994; Siegel, 1992; Stanovich, 2005; Vellutino et al., 2000). Further, these studies found no differences between children

with dyslexia and nondiscrepant poor readers in skills related to reading, such as word recognition, regular and exception word reading, pseudoword reading, word and pseudoword spelling, and the recognition of the visual form of sounds.

It should be noted in the Siegel (1992) study that the individuals with dyslexia had significantly higher scores than the nondiscrepant poor readers on full-scale intelligence as well as the verbal and performance composite scores and every subtest of the IQ test. The scores on the language-oriented subtests, Vocabulary, Similarities, and Comprehension, for the group with dyslexia were significantly higher than those of the nondiscrepant poor readers but similar to those of nondisabled readers. These age-appropriate language skills of the group with dyslexia were not sufficient to assist them in performing well on decoding and reading comprehension tasks. In fact, there were no significant differences between individuals with dyslexia and nondiscrepant poor readers in reading comprehension scores; both groups had low word recognition and decoding scores. So, the enhanced language skills of the individuals with dyslexia, as determined by the subtests from IQ tests, did not enable them to overcome their decoding difficulties.

With regard to language skills, Siegel (1992) compared individuals with dyslexia and nondiscrepant poor readers on a variety of language and memory tasks. The group with dyslexia had significantly higher scores than the nondiscrepant poor readers on the simpler syntactic awareness tasks, but the two groups did not differ significantly on the more difficult syntactic tasks. Both groups of children with reading difficulties had significantly lower scores than the nondisabled readers. The individuals with dyslexia had significantly higher scores on short-term and working-memory tasks than the nondiscrepant poor readers. Thus, despite some better scores on language and memory tasks, the children with dyslexia did not perform at a higher level on reading comprehension tasks than the nondiscrepant poor readers. It is clear, therefore, that the core difficulty, located in the word-reading module for both groups, limits text comprehension for both groups.

Even though individuals with dyslexia had better language skills than the nondiscrepant poor readers, there was no indication that it helped their reading performance, nor was their reading performance any different from that of the nondiscrepant poor readers. In addition, there is ample evidence to suggest that there are no significant differences between individuals with dyslexia and nondiscrepant poor readers in the ability to benefit from remediation (e.g., Hurford, Johnston et al., 1994; Stage et al., 2003; Weber et al., 2002).

Snowling et al. write, "An inevitable consequence of removing the discrepancy definition is that more children with poor reading in the context of broader and more serious language difficulties will be labeled 'dyslexic'" (p. 505). They appear to be arguing that children identified with dyslexia, when based on a discrepancy definition, are less likely to have serious language difficulties. There is no evidence for this assertion. Language functioning should be evaluated for all people with reading difficulties, and when deficiencies are found, interventions should be provided. However, the absence or presence of language difficulties is irrelevant for the definition of dyslexia. Snowling et al. noted,

> Since reading for meaning draws on language skills, it follows that many non-discrepant poor readers also have poor reading comprehension skills. Thus, whereas in classic discrepancy-defined dyslexia, reading comprehension is only an issue insofar as poor decoding presents a bottleneck to the construction of meaning, this is not the case for children with dyslexia who have co-occurring language problems; these children have poor reading comprehension too. (Bishop & Snowling, 2004, p. 505)

And, "None of these comorbidities should be viewed as 'core' features of dyslexia, but they can complicate its presentation and response to intervention" (Rose, 2009, p. 506).

The authors appear to be arguing that if individuals with dyslexia have reading comprehension problems, it is a result of decoding difficulties, but if nondiscrepant poor readers have reading comprehension problems, it is a result of decoding and language difficulties. This argument holds no merit and appears circular. That is, if someone has reading comprehension problems and has been discrepancy-defined as having dyslexia, their comprehension problems result from decoding issues, but if it is a nondiscrepant poor reader, then the reading comprehension problems are a result of language problems and possibly decoding. Although Snowling et al. argue that "… following relaxation of the discrepancy definition and hence the IQ cut-off, the number of 'symptoms' co-occurring with dyslexia has increased" (p. 506), there is no evidence to support this claim.

How is dyslexia to be differentiated from poor reading? This is the central question, and Snowing et al. do not provide a clear answer. Snowling et al. wrote,

> Although intellectual disability precludes a diagnosis of specific learning disorder, once the practice of restricting the diagnosis of dyslexia to those principally with above average IQ is abandoned, the kinds of learning difficulties to which the label

> 'dyslexia' applies widen and now include children with a broader range of learning problems. Such children have a range of problems with reading which are not best characterized as affecting only accuracy and fluency. (p. 503)

Again, it is important to note that the individuals with dyslexia and the nondiscrepant poor readers did not differ significantly on various reading and spelling tasks at the word level and did not differ on reading comprehension (Siegel, 1992).

Snowling et al. further state, "An inevitable consequence of removing the discrepancy definition is that more children with poor reading in the context of broader and more serious language difficulties will be labeled 'dyslexic'" (p. 505). Language difficulties and reading difficulties are not necessarily orthogonal dimensions. They may overlap, but each issue may require different types of evaluation and remediation. The presence of other difficulties is irrelevant to the definition of dyslexia, although, of course, these other difficulties, when they exist, should be recognized and treated.

Snowling et al. appear to be arguing that if an individual has difficulties with mathematics, attention, or has social-emotional difficulties, they should not be labeled as having dyslexia. Furthermore, they appear to be arguing that children with dyslexia, vs. nondiscrepant poor readers, have no other presenting difficulties. It is not clear why the authors would suggest that if you have attention and/or language problems, you should not be considered as having dyslexia. The available evidence is that whatever the IQ score, individuals who are poor readers may have other difficulties, and these difficulties should be evaluated and remediated along with the reading difficulty (Catts & Petscher, 2021; Fletcher et al., 2018; Odegard et al., 2022).

These authors further state: " In summary, we argue that cases of 'specific dyslexia' exist, and they are most apparent when a strict discrepancy definition (reading poorer than expected for a child's age and IQ) is adopted. However, when it is dropped, a wider range of difficulties are observed among children with reading disorders" (p. 506). What Snowling and colleagues seem to ignore is that many children with reading difficulties have language, attention, and mathematical difficulties, for example, without regard to their IQ score. The important point is that these difficulties should not be ignored and should be assessed. Rather than an IQ test, specific assessments of language, mathematical skills, and attention will be more useful.

Snowling et al. propose that "the term dyslexia should not be used as a shorthand for 'reading disorder' but should be used to refer to a difficulty with decoding and spelling fluency which is evident from the early school years and persistent over time" (p. 507). We agree but do not think that IQ scores have any place in this conceptualization. However, early efforts to remediate identified deficits may mask the reading problems in some children with dyslexia until the complexity of the word structure overwhelms basic skills acquired in the early grades through remediation. These children would present with late-emerging reading deficits (e.g., Catts et al., 2002; Scarborough, 1990). This is one reason why risk models can be advantageous. Students who are at risk for reading failure could be identified. This would then be noted as a preexisting factor upon later identification if reading deficits should emerge. Efforts should not be delayed with regard to providing remediation to help support children who are at risk for reading failure. In their paper, Snowling et al. note,

> Another issue that has concerned those who do not support the use of the term 'dyslexia' is the fact that the types of intervention that are known to be helpful do not differ from the interventions that are useful for other non-discrepant poor readers. This is, however, a simplistic view. Dyslexia does equate with poor decoding and word reading, and therefore to say it requires similar treatment to poor reading is a tautology. (p. 507)

That is exactly our point; poor word reading is the defining feature of dyslexia regardless of a measured IQ score. We conceptualize dyslexia as difficulty with the accuracy and/or fluency of word recognition.

We are not arguing against the term *dyslexia*. The available evidence indicates that when poor readers have been differentiated into "dyslexia" and "nondiscrepant poor reader" groups, not only do they appear very similar in their reading skills, but they also benefit equally from the same interventions. Although future work should continue to examine the potential of other measures beyond reading to benefit the identification of children at risk for reading failure and to determine their neurological characteristics, the state of the field presently does not support the contention made by Snowling et al. – that the discrepancy definition should be affirmed. In fact, use of the discrepancy definition leads to inherent harm and unintended consequences. Specifically, when the discrepancy formula has been applied, access to special education and related services for children who are nondiscrepant poor readers has been denied (Miciak et al., 2016; Odegard et al., 2022).

At a time when a large percentage of children who are learning to read English have not been exposed to science-based curricula that help them to master such an opaque writing system, a more inclusive definition is more appropriate. Denying struggling readers an appropriate research-based curriculum and then pre-

venting them access to special education and related services to assist them in becoming competent readers is cruel and increases the likelihood that these children will face anxiety and depression (Alexander-Passe, 2015a), poor self-esteem (Terras et al., 2009), parental abuse (Fuller-Thomson & Hooper, 2014), suicidal ideation (Alexander-Passe, 2015b; Dahle et al., 2011), suicide attempts (Fuller-Thomson et al., 2018; Livingston et al., 2018; McBride & Siegel, 1997), school drop-out (Cortiella & Horowitz, 2014), and poor economic earning potential (Cortiella & Horowitz, 2014). Although science is obligated to determine the nature of reading failure, practitioners should be utilizing the fruits of science to assist all poor readers, regardless of their intellectual functioning, in becoming competent readers. There is ample evidence that reading interventions are successful for all struggling readers. Children with reading levels commensurate with their intellectual functioning should not be excluded from interventions that can assist them in becoming competent readers.

If we agree that all children who are struggling readers should be provided with appropriate interventions to help them become competent readers, then we must seriously consider the ramifications of the discrepancy definition and abandon the strategy of differentiating children who will receive interventions based on their IQ score. In summary, both empirically and practically, the use of discrepancy criteria for dyslexia identification is not supported.

References

Alexander-Passe, N. (2015a). Investigating post traumatic stress disorder (PTSD) triggered by the experience of dyslexia in mainstream school education. *Journal of Psychology & Psychotherapy*, *5*(6), 1-10. doi:10.4172/2161-0487.1000215

Alexander-Passe, N. (2015b). Dyslexia: Investigating self-harm and suicidal thoughts/attempts as a coping strategy. *Journal of Psychology & Psychotherapy*, *5*(6), 1-11. doi:10.4172/2161-0487.1000224

Bruck, M. (1990). Word-recognition skills of adults with childhood diagnoses of dyslexia. *Developmental Psychology*, *26*(3), 439-454. doi:10.1037/0012-1649.26.3.439

Carrion Castillo, A., Francks, C., & Fisher, S. E. (2015). *Evaluating the genetic risk for dyslexia in multi-generation families*. Presentation at The Individual Differences in Language Processing Across the Adult Life Span Workshop, Nijmegen, Netherlands.

Cassar, M., Treiman, R., Moats, L., Pollo, T. C., & Kessler, B. (2005). How do the spellings of children with dyslexia compare with those of nondyslexic children? *Reading and Writing*, *18*(1), 27-49.

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech Lang Hear Res, 45*(6), 1142-1157. doi:10.1044/1092-4388 (2002/093)

Catts, H. W., & Petscher, Y. (2021). A cumulative risk and resilience model of dyslexia. *Journal of Learning Disabilities.* doi:10.1177/00222194211037062

Cortiella, C., & Horowitz, S. H. (2014). *The state of learning disabilities: Facts, trends and emerging issues*. National Center for Learning Disabilities.

Dahle, A. E., Knivsberg, A. M., & Andreassen, A. B. (2011). Coexisting problem behaviour in severe dyslexia. *Journal of Research in Special Educational Needs, 11*(3), 162-170.

Eden, G. F., Olulade, O. A., Evans, T. M., Krafnick, A. J., & Alkire, D. R. (2016). Developmental dyslexia. In G. Hickok & S. L. Small (Eds.), *Neurobiology of language* (pp. 815-826). Academic Press.

Eloranta, A., Närhi, V., Ahonen, T., & Aro, T. (2019) Does childhood reading disability or its continuance into adulthood underlie problems in adult-age psychosocial well-being? A follow-up study. *Scientific Studies of Reading*, *23*(4), 273-286.

Farris, E. A., Alexander, E., & Odegard, T. N. (2020). Assessment and identification of learning disabilities. In M. M. Martel (Ed.), *The clinical guide to assessment and treatment of childhood learning and attention problems* (pp. 3-32). Elsevier. https://doi.org/10.1016/B978-0-12-815755-8.00001-0

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2018). *Learning disabilities: From identification to intervention* (2nd ed.). Guilford Press.

Fuller-Thomson, E., Carroll, S. Z., & Yang, W. (2018). Suicide attempts among individuals with specific learning disorders: An underrecognized issue. *Journal of Learning Disabilities*, *51*(3), 283-292.

Hurford, D. P., Johnston, M., Nepote, P., Hampton, S., Moore, S., Neal, J., & Huffman, D. (1994). Early identification and remediation of phonological-processing deficits in first-grade children at risk for reading disabilities. *Journal of Learning Disabilities*, *27*(10), 647-659.

Hurford, D. P., Schauf, J. D., Blaich, T., Moore, K., & Bunce, L. (1994). Early identification of children at risk for reading disabilities. *Journal of Learning Disabilities*, *27*(6), 371-382.

Lefly, D. L., & Pennington, B. F. (1991). Spelling errors and reading fluency in compensated adult dyslexics. *Annals of Dyslexia*, *41*(1), 141-162.

Livingston, E. M., Siegel, L. S., & Ribary, U. (2018). Developmental dyslexia: Emotional impact and consequences. *Australian Journal of Learning Difficulties*, *23*(2), 107-135.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, *53*(1), 1-14.

McBride, H., & Siegel, L. S. (1997). Learning disabilities and adolescent suicide. *Journal of Learning Disabilities, 30*, 652-659.

Miciak, J., Fletcher, J. M., & Stuebing, K. K. (2016). Accuracy and validity of methods for identifying learning disabilities in a response-to-intervention service delivery framework. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of multi-tiered systems of support* (pp. 421-440). Springer Science + Business Media. https://doi.org/10.1007/978-1-4899-7568-3_25

Norton, E. S., Beach, S. D., & Gabrieli, J. D. (2015). Neurobiology of dyslexia. *Current Opinion in Neurobiology*, *30*, 73-78.

Odegard, T. N., Farris, E. A., & Washington, J. A. (2022). Exploring boundary conditions of the listening comprehension-reading comprehension discrepancy index. *Annals of Dyslexia*. https://doi.org/10.1007/s11881-021-00250-0

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological Science, 29*(8), 1358-1369.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development*, *61*(6), 1728-1743.

Shao, S., Kong, R., Zou, L., Zhong, R., Lou, J., Zhou, J., ... & Song, R. (2016). The roles of genes in the neuronal migration and neurite outgrowth network in developmental dyslexia: Single- and multiple-risk genetic variants. *Molecular Neurobiology*, *53*(6), 3967-3975.

Siegel, L. S. (1992). An evaluation of the discrepancy definition of dyslexia. *Journal of Learning Disabilities*, *25*(10), 618-629.

Snowling, M. J., Hulme, C., & Nation, K. (2020). Defining and understanding dyslexia: past, present and future. *Oxford Review of Education*, *46*(4), 501-513.

Stage, S. A., Abbott, R. D., Jenkins, J. R., & Berninger, V. W. (2003). Predicting response to early reading intervention from verbal IQ, reading-related language abilities, attention ratings, and verbal IQ-word reading discrepancy: Failure to validate discrepancy method. *Journal of Learning Disabilities*, *36*(1), 24-33.

Stanovich, K. E. (2005). The future of a mistake: Will discrepancy measurement continue to make the learning disabilities field a pseudoscience? *Learning Disability Quarterly*, *28*(2), 103-106.

Sumner, E., Connelly, V., & Barnett, A. L. (2014). The influence of spelling ability on handwriting production: Children with and without dyslexia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1441.

Terras, M. M., Thompson, L. C., & Minnis, H. (2009). Dyslexia and psycho-social functioning: an exploratory study of the role of self-esteem and understanding. *Dyslexia, 15*, 304-327. doi:10.1002/dys.386

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, *45*(1), 2-40. doi:10.1046/j.0021-9630.2003.00305

Vellutino, F. R., Scanlon, D. M., & Lyon, R. G. (2000). Differentiating between difficult-to- remediate and readily remediated non-discrepant poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities*, *33*(3), 223-238.

Weber, J. M., Marx, P., & Schneider, W. (2002). Profitieren Legastheniker und allgemein lese-rechtschreibschwache Kinder in unterschiedlichem Ausmaß von einem Rechtschreibtraining? [Do dyslexic children and children with general dyslexia benefit to different extents from spelling training?]. *Psychologie in Erziehung und Unterricht*, *49*(1), 56-70.

Transitioning to Postsecondary Education With a Learning Disability: Identifying Similarities and Gaps Between First-Hand Experiences and University Transition Programs

Emma Peddigrew,[1] John McNamara,[1] Naomi C. Z. Andrews,[1] and Samantha Sendzik-Sinclair[2]
[1]Brock University
[2]Learning Disabilities Association of Niagara Region (LDANR)

Abstract

The transition to postsecondary education is accompanied by several social, emotional, and academic challenges, but poses a unique challenge for students with learning disabilities (LD). There is a dearth of research examining the mechanisms that support a positive transition from the perspective of those with LD. Eight first-year university students diagnosed with an LD participated in individual interviews on several key themes: (a) LD and self-esteem/image; (b) accommodations; (c) support networks; and (d) mental health vulnerabilities. In addition, 10 university transition programs across Ontario, Canada, were analyzed. The programs' main objectives included access to resources, social connections, and time and stress management. Results indicated several overlaps between student perspectives and those of universities. However, gaps remain regarding the importance of self-esteem, stigmatization of disability, and mental health. By understanding the mechanisms that contribute to a positive transition, university programs can adopt useful strategies to help support students with LD.

Keywords: Learning disabilities, transition, postsecondary education, university

The transition to postsecondary education is a critical stage in an one's academic, social, and emotional trajectories, and can be especially challenging for students with learning disabilities (LD). Emerging from the 1960s, intensive research on LD has generated immense bodies of knowledge regarding its manifestation within the body, brain, genome, and social structure (Denhart, 2008), resulting in increased awareness and identification of individuals with LD.

The number of university students with LD has grown each year from 5,620 in 2008-09 to 7,064 in 2014-15 (Learning Disability Association of Ontario, 2018). The primary objective of university transition efforts is to support students to be successful in their higher education journey (Learning Disability Association of Niagara Region, 2021). Positive outcomes regarding grades and retention have been reported (Reed et al., 2011), which is important for success, yet psychosocial benefits are of equal importance. Students with LD are a unique population that may be particularly disadvantaged in higher education, leading to low popularity rates (Cavioni et al., 2017), hesitation to self-advocate (Peters, 2011), and being less likely to recognize when faced with an issue that requires support and assistance from others (Martinez et al., 2011).

To address some of their academic, social, physical, and emotional barriers, students with LD should be integrated into university preparation programs. While students with LD experience many of the same types of academic/non-academic stressors as their peers without LD, there is limited research comparing the objectives of current university transition efforts with the actual first-hand experiences of students with LD.

The purpose of this study was to begin to fill this gap in the literature by exploring the transition experience of first-year students with LD in Ontario, Canada, as they move from secondary to postsecondary education. Further research is needed that encompasses the ability for young people to reveal, challenge, and redefine the preconceived notions of their disability during their transition process. Overall, the

Peddigrew, E., McNamara, J., Andrews, N.C.Z, & Sendzik-Sinclair, S. (2022). Transitioning to Postsecondary Education With a Learning Disability: Identifying Similarities and Gaps Between First-Hand Experiences and University Transition Programs. *International Journal for Research in Learning Disabilities, 5*(2), 55-69. https://doi.org/10.28987/ijrld.5.2.55

objective of the current study was to foreground the underrepresented experiences of students with LD to provide valuable knowledge to educators, parents, case workers, and other support staff to inform a smoother transition from high school to university. The narratives within this study will be used to first, understand the experiences of students with LD, and second, delineate 10 Ontario university transition programs. By doing so, it will be revealed how these transition programs work to support the students' needs as well as identify any gaps between students' opinions and the university transition efforts.

The Transition to Postsecondary Education

In *Ensuring a Smooth Transition From Elementary to Secondary School* (2012), the Ministry of Education Quebec (MOEQ) defines transition as a "time during which children gradually adapt their new physical, social, and human environment" (p. 1404). Periods of transition are often stressful and impact one's psychological adjustment and well-being (Rice et al., 2011). The move from high school to postsecondary education is widely considered a significant life transition (Connor, 2012) – an intense period of adaption, in which young people are faced with several psychological, social, environmental, and physical challenges due to the simultaneous changes in school environments, social interactions, and academic expectations (MOEQ, 2012). Nevertheless, research suggests that transition supports can support students' success and foster a more positive move from one school to the next (Morgan & Hertzog, 2001).

Although a successful transition is difficult to define, it is generally measured across multiple dimensions, including social relationships, academic performance, eased parental concerns, increased interest in school-related work, and improved self-esteem (MOEQ, 2012). Gill (2007) found that positive self-concept, supportive peer relationships, and positive thinking significantly impact students' ability to cope with the postsecondary education transition.

Intervention efforts and research focusing on the high school-to-higher-education transition are necessary as this transition period encompasses anxiety and stress, which is correlated with concurrent psychological issues, future adjustment concerns, and attainment issues (Rice et al., 2011). The risks of a poor transition pose a special threat to those who are most vulnerable to academic underachievement, poor self-esteem, anxiety, and school-dropout, such as students with LD (Connor, 2012).

Learning Disabilities and the Transition to Postsecondary Education

Learning disabilities are typically defined as lifelong neurological dysfunctions that interfere with one's ability to store, process, or produce information (Learning Disabilities Association of Canada, 2007). According to the *Learning Disability Mental Health Handbook* (Integra, 2016), learning disabilities affect 5-10% of all Canadians, with 43% of students in Ontario having been diagnosed with educational exceptionalities. Throughout the past 70 years, the field of LD has aimed to support these individuals by generating definitions, understanding neurological contributions, and creating policies and practices.

The often-low popularity rates among students with LD means these adolescents may lack social connections amongst peers at a crucial developmental period (Cavioni et al., 2017). When attending postsecondary education, a student's self-image is often complicated due to the influence of new academic knowledge and exposure to diverse peer groups (Connor, 2012). Therefore, students with LD are faced with the choice of self-disclosure about their disability to new peer groups, and many opt to not self-disclose (Connor, 2012). This social exclusion makes the transition process even more concerning for those with LD.

Students with disabilities make up a unique group of individuals, as they face different methods for receiving academic services during their transition from secondary schools to higher education (Burdge, 2012). Peters (2011) proposed that students with LD may not possess necessary self-advocacy skills, hindering how they receive services and accommodations. When transitioning to postsecondary education, students with LD often avoid seeking academic accommodations for various reasons. For example, they may (a) feel that they are no longer impacted by their LD, (b) be trying to prevent others from knowing about their disability, or (c) simply be unaware of university services or how to access them (Burdge, 2012). Developing self-advocacy skills will allow students with LD to feel confident and comfortable disclosing their disability, thereby helping to alleviate feelings of stress or anxiety (Herridge, 2017). Kimberlin (2009) also found that by improving their self-advocacy skills, students with LD are more likely to seek out reasonable academic accommodations and services needed for success.

While all students encounter different kinds of stressors when beginning postsecondary education, the literature suggests that students with LD report experiencing lower levels of social support, higher peer victimization, lower self-esteem, and higher internalizing and externalizing, putting them in one of the most at-risk groups for mental health concerns (Burdge, 2012; Her-

ridge, 2017; Peters, 2011). "Students with LD often see postsecondary as a new beginning, an opportunity to redefine themselves, including reframing their disability and its relation to their identity" (Connor, 2012, p. 1007). Considering the concerning correlation between LD and mental health, the importance of positive transitional efforts that target not only academic concerns but also social and emotional issues is imperative.

The Current Study

Personal accounts from eight participants guided the current study. Students were asked about their transition in terms of academic concerns, as well as social and emotional concerns, and their impact on mental health, and the major findings were analyzed in an effort to understand how they are reflected within current university transition programs. To that end, 10 transition programs in Ontario, each tailored specifically to students with LD, were analyzed. The transition programs' major objectives were analyzed as direct examples of areas where the programs address students' needs and where any identifiable gaps may be. The following research questions guided the study:

1. What are the academic and socioemotional needs of students with LD that must be acknowledged to ensure they experience a smooth transition to postsecondary education?
2. What are strategies and efforts that benefit incoming postsecondary students with LD during their transition?
3. What areas of the transition experience from firsthand accounts of students with LD coincide with those of current university transition programs in Ontario?
4. What are the gaps or missing objectives between the students with LD and current transitional efforts in Ontario universities?

Methods

Primary Author Positionality

As an academic, my research experience has focused on a wide range of exceptionality-driven, sociocultural, and developmental understandings of young people. I have realized that while I respect the value quantitative research adds, I have more of an interest in the lived experiences that shape the world we live in. Being able to speak to others, physically see their world views, and understand their perspectives on a given topic is something I strive towards. However, this involves much more than simply interviewing others in the hopes of completing a task or research project. Researchers must challenge the feeling of wanting to ignore or dissociate themselves from the research, and instead, acknowledge their positionality, emotional responses to participants, and personal artifacts, that might enhance the research process (Mosselson, 2010).

I uphold privilege, whether I want to or not, by conducting research with others on a topic of my choosing. "As researchers we belong to a moral community. Doing interviews is a privilege granted us, not a right we have. Interviews transform information into shared experience" (Denzin, 2001, p. 24). Although I may think I am studying one topic, for instance, the transition from high school to university for students with LD, it would be naïve to ignore how even amongst shared experiences, these are still uniquely different. Although two individuals may both have a diagnosed LD, their family histories, social backgrounds, cultural beliefs, gender, class, and/or race make their experiences exceptionally distinct. Therefore, when conducting research involving sensitive topics with vulnerable populations, it is crucial to first establish rapport and trust to minimize anxiety and distrust (Pillow, 2003). As a qualitative researcher in an area built upon power, status, and historical understandings of disability, there is analytic work that needs to be done. I have a responsibility to identify shared themes amongst participants, while also highlighting and prioritizing each interviewee's unique perspective. I truly believe that being reflexive involves a methodological alertness to how knowledge is coproduced (Court & Abbas, 2013).

Participants

The Research Ethics Board at Brock University granted clearance for the study and the participant recruitment strategies (File # 19-171). Participants were recruited in January 2020 from a mandatory first-year introductory child and youth studies course with approximately 900 students at Brock University, a moderately sized public research university. At the beginning of the in-person lecture, a verbal invitation was given that informed students of the purpose, methodology, and procedures of the study. Students were also told that they must have a diagnosed LD to participate. Research invitations were also posted on the students' university online portal. Interested participants were asked to email the primary researcher.

A total of eight participants, two males and six females aged 17-18, volunteered. Participants' race and ethnic background was not reported. The participants had no direct relationship to the researcher, other than attending the same university. Participants were informed of their right to skip any questions that made them feel uncomfortable and were told that their answers would

remain confidential. No participant expressed concerns over any of the questions or withdrew from the study.

Procedure

Student Participant Procedures

The qualitative approach provides rich, in-depth information that enabled exploring both the similarities and differences in the transitional experience of students with LD. Data collection strategies were focused on gaining information about the personal experiences from the perspectives of those with LD, as the voices of this marginalized group are not typically considered in this area of research (Connor, 2012). Such interviews reflect more of a conversation, whose purpose is to gather detailed descriptions of one's reality with respect to an interpretation of the meanings of a"described phenomenon."However, the value of interviewing is that it empowers interviewees to express their own thoughts while speaking in their own voice (Alshenqeeti, 2014). For people with LD, who are often silenced by an ableist society, interviewing can enable opportunity to voice personal experiences and perceptions. Due to the sensitivity of some questions (i.e., some participants discussed traumatic experiences), in-depth interviews created trust between the interviewee and interviewer. The individual in-depth interviews were semi-structured in nature and conducted by the primary researcher at Brock University in the researcher's private lab. The interview questions (see Appendix A) were developed to obtain information about the perspectives of students with LD regarding the transition experience. Interviews lasted approximately 35-45 minutes. Interviews were audio-recorded on a passcode-locked computer, with the participants' consent.

University Program Procedures

Ten university programs were selected from the Ontario Universities' Application Centre (OUAC; Ontario Universities' Application Centre, 2021) list of university student transition resources. Of the 27 universities listed, 10 offered programs that had a specific focus on incoming students with LD. Several other postsecondary institutions had universal transition programs that focused on the general population of incoming students, but did not include a specific focus on student with LD. Therefore, the former 10 were well suited for this study, due to their primary focus on LD.

Data Analysis

Student Participant Data Analysis

All eight interviews were transcribed verbatim by the primary researcher. Pseudonyms were used and randomly assigned to protect participants' identity and privacy. Each transcription was read several times for accuracy and consistency and to also become familiar with the content. Due to the COVID-19 pandemic, it was not possible to member-check the transcriptions. However, to ensure trustworthiness of data, all interviews were audio recorded by two separate devices and cross-checked following the interviews. In addition, throughout the interviews, the primary researcher continuously repeated what the participant stated to ensure that each member was heard and understood correctly.

An inductive approach was used for three main reasons: (a) to condense extensive raw data into a summary format; (b) to establish clear links between the research objectives and the summary findings derived from the raw data; and (c) to develop theory regarding the underlying structure of experiences which are evident in the raw data (Thomas, 2003). Transcripts were transferred into NVivo to assist with analysis. A rigorous and systematic reading of the transcripts allowed major themes to emerge. These themes were created from phrases used in specific text segments. Phrases or texts were coded into major categories based on similarities, with the research questions in mind. Although no other members examined the raw data for emergent themes, the selected inductive approach to data analysis helped eliminate researcher bias.

University Program Data Review

The OUAC website list of transition resources (2021) highlights the major objectives of each postsecondary institution's transition program, while also providing links to the programs' information sites. Gaining information from students who are enrolled in each program was expected to generate further themes, but for the purpose of the study, information was gathered from what is made available to the public via the postsecondary education's websites. The main themes and components of each transition program were organized into a table to highlight the similarities and differences amongst programs. Only themes that were consistent amongst at least two programs were included for analysis.

Results

The results will be categorized into two sections. First, we will discuss the major themes that emerged from the personal narratives of the students with LD. Specifically, how these students work to resist the preconceived notions of what LD entails to thrive mentally, physically, academically, and socially during their transition. Although each participant shared their own unique stories, similar experiences emerged for all eight. Four major themes were identified: (a) LD and Self-Esteem/Image, (b) Accommodations, (c) Support Networks, and

(d) Mental Health Vulnerabilities.

Second, the 10 university transition programs will be described and categorized based on their main objectives (see Table 1). Objectives that were discussed by at least two universities are listed by the number of times they were discussed amongst the 10 programs (see Table 2). This information was used to demonstrate the overlaps and gaps between the student data and the university transition objectives (see Table 3).

LD and Self-Image/Esteem

Societal Views

Six of the eight participants spoke about feeling that others would solely define their image based on their diagnosis. This can affect one's self-image and embody the negative connotations many people have about an LD. Participants felt discouraged about how others would view them.

> *I believe it's a way of categorizing people that struggle. So, people that learn through different means, and the minorities, are seen as incompetent in our society.* Rebecca

Some students expressed how their LD quickly became the sole image of them once others around them knew about their diagnosis.

> *I just feel like people thought I was weird in high school because of it [LD]… that's how people define you, or that's how people knew me by.* Brittany

Participants revealed a feeling of not being "good enough" or expected to fail, because of the expectation's society uphold around LD.

> *I've been told a lot that I'm not smart because of my disability. It put me in the mindset of not feeling smart enough to do university and think I was going to waste my family's money on tuition, fail, then be upset with myself for failing and embarrassed.* Rebecca

Consequences for Self-Esteem. Several students expressed how their LD complicated their self-confidence, ultimately raising feelings of self-hatred and remorse. Devin expressed this when stating:

> *I've had horrible self-confidence for a while. I think it's related to having an LD ... it is a bit harder to maintain that self-esteem because there is always that self-loathing.*

Brittany connected their self-esteem struggles with their disability by saying:

> *I can't even count the times I've cried over school. Doing work you don't understand right away is so difficult, especially when everyone else gets it … It puts a damper on self-esteem because you feel like you aren't doing as well and start to internalize that.*

Most participants had similar experiences of not being accepted by others and embodying feelings of shame regarding their diagnosis.

Accommodations

Benefits of Accommodations

Responsibility is placed upon students to disclose their disability prior to entering university. Many students responded positively regarding their current accommodations.

> *In university, accommodations are even more important. I have noticed I have gone over time on every exam. Having that ability has saved my academic life.* Thomas

Accommodations vary depending on the student's needs as well as the university's access to certain resources, support networks, and assistive technologies.

> *I have use of computer, extra time, quiet rooms, and then I have the use of the software that can read to you and note-taking.* Alanna

Most participants expressed that accommodations provided a space for them to feel more at ease regarding their learning. Note-taking, extra exam/assignment time, access to certain technologies, and a quiet workspace, were among some of the accommodations students benefitted from.

Stigma Associated With Accommodations

Although accommodations can provide immense support, participants expressed concern that others see them as a privilege, making students with LD feel guilty, self-conscious, and unworthy of the necessary resources and services provided to them.

> *People think "is she so stupid she needs help with spelling?" Even if you need help with something everyone does, where does that put you on a level of intelligence?* Brittany

> *When people think I'm lucky, I feel guilty… I think other people think that because I have an LD, I am stupid, or not worthy or that I'm taking advantage of a system.* Rebecca

Although some knew their accommodations would benefit their learning, the stigma associated with this assistance was not worth the social consequences to them.

> *The computer for me is a symbol of stress s... as a symbol of my failure and shame.* Devin

Elizabeth voiced the same concern, refusing to take extra time on assignments or exams out of fear of the negative perceptions of others, despite truly needing the accommodation for support:

> *I was worried that when they [professors] are grading, because I got extra time, that it would need to be so much better than everyone else's.*

The balance between accommodations providing necessary support and acting as a marker for disability was a constant juggle for students.

Support Networks

Family and Friends

The role of family and friends as a protective factor was apparent for all participants, indicating that these support networks hold an essential role in aiding a positive transition.

> *My parents are good about not letting me label myself with having an LD, they don't let me use that as an excuse.* Samantha

> *Finding a friend group that I knew I could trust and confide in was super important for me.* Elizabeth

However, without these crucial support networks, individuals may be at risk for a poor transition.

> *A good friend group, which is what I was lacking in and that contributes a lot to the anxiety I felt, but if I had someone to talk to it may have been processed better.* Elizabeth

Support/Case Workers

The appropriate support services available for each student was identified as key factors in helping them navigate the academic, social, and personal demands of university.

> *Knowing your services at university because they are so helpful.* Alanna

This includes having a specific individual, such as a case worker, from the university that can help communicate the student's needs and wants to ensure a smooth transition.

> *My case manager also helped a lot.* Claire

Lack of LD Knowledge

Every participant expressed that there was a clear lack of knowledge regarding what an LD is. Elizabeth described this feeling by stating how:

> *Some people aren't necessarily defining it, because they don't know, but they will say that someone with an LD is dumb, or has a low IQ, a lot of negative stuff.*

> *If somebody doesn't know someone with an LD, they don't really know what it is. It's almost your job to educate them. There is this unknown bubble around LD unless somebody tries to learn more about it.* Claire

The lack of LD awareness for those diagnosed and the general population raises apprehension.

> *We are constantly trying to blend in ... Sometimes people can't handle being different. I feel like you notice your own flaws more… and then you start to internalize that.* Rebecca

Self-Advocacy

University often reflects an independent environment. Therefore, students may be required to reach out to the necessary support systems on their own.

> *I did suffer with anxiety and depression knowing that I was alone and had to do all this work by myself or reach out myself, which scared me.* Claire

Preparation for postsecondary education should go beyond curriculum-based needs. The necessary skills may include self-advocacy to communicate to support staff what accommodations are necessary.

> *You need to advocate for yourself and know what works and what doesn't.* Samantha

However, self-advocacy does not only provide benefits for individuals with LD; others can relate to them as well.

> *By sharing my own experiences, it educates other people and helps people who also struggle with the same things I do. You might as well be your own self advocate.* Rebecca

All participants agreed that support is an important factor to their overall well-being – whether from family, friends, university resources, self-advocacy, or a combination.

Mental Health Vulnerabilities

Mental Health and High School

All participants expressed having faced some form of mental health struggles because of their disability. Elizabeth noted her anxiety-related symptoms when stating:

> *I had them [anxiety symptoms] in high school, that's where it started … again in university, it just hit me, and I hadn't felt it for months and suddenly I was waking up in the middle of the night with a panic attack … university definitely set it off.*

Similar thoughts about mental health were present when Samantha shared:

> *I have my own fair share of mental health … I was diagnosed with anorexia, and after, anxiety and depression … my therapists and doctors said I need control, I didn't feel like I had control in my academics … I think because an LD takes that control away from you.*

These comorbidities were quite common amongst participants, which could indicate that individuals with LD are more vulnerable to mental health concerns.

> *I struggle with anxiety and panic attacks. It could be co-occurring because I was more vulnerable to having worries.* Elizabeth

Mental Health and the Transition to University

The transition may be identified as a period of increased physical, academic, social, and emotional challenges for students with LD. Devin described the transition as an experience of reliving deep-rooted struggles:

> *I've never been good at transitions … a lot of self-aggression and harm started to come back, it's all little things I haven't seen in years came back because of my transition.*

Specifically, the transition itself brought feelings of hardship for seven of the eight participants, including the change of pace, social networks, and increased academic pressure.

> *I have been struggling this first year … I was struggling finding a good friend group … and obviously academically … I wasn't enjoying it and having lots of anxiety.* Elizabeth

Impact of LD on Mental Health

When asked if they thought that having an LD put someone at a greater risk for mental health issues, most of the participants acknowledged a susceptibility.

> *There is a lot more strain on you to do better and you work harder … you get stressed and that leads to anxiety and depression, more so than people who don't have an LD.* Claire

A common theme amongst responses was the pressure felt from societal standards that seemed to conflict with the characteristics of LD.

> *You wish you could be "better" by societal standards. As a result, you dig yourself deeper and deeper. That's what I attribute to leading to my depression.* Devin

This powerful statement seemed to mirror similar feelings from others as well.

> *I think it plays a role; it can cause bullying which can cause depression. I think majority of kids with LD also struggle with mental health.* Samantha

This constant internal and external battle can become detrimental for one's future trajectories, mental health, and overall well-being. Devin was passionate and shared this idea within his interview, which provides a fitting conclusion to this presentation of the testimonies.

> *The problem is because there is something different about you, society tells you you're broken. That's probably one of the main reasons I became depressed … I wanted to do everything perfectly because I knew I couldn't be.*

All participants provided incredibly passionate and meaningful responses. These findings point to the complexity of the transition from high school to postsecondary education for those with LD.

Main Objectives of University Transition Programs

An informal review of universities across Ontario demonstrated that several hold objectives that correlate well with the results of the student interviews. The public transition documents and information from 10 universities were organized into major themes (see Table 1). The main objectives were referenced from the OUAC (2021) University Student Transition Resource guide. Many transition programs included core objectives of getting students connected to others on campus, exploring university services and accommodations, and gaining valuable strategies such as study skills and note-taking. However, programs rarely addressed issues surrounding self-advocacy, mental health, and societal perceptions and stereotypes of disability, which are known to make the transition from high school to postsecondary education difficult.

After analyzing the transition programs, we assessed the main objectives that were consistent amongst at least two of the universities. Table 2 presents each objective as well as how many universities out of the 10 that touched on the given topic. By visualizing the objectives in this way, it becomes clear what universities value most when thinking about students with LD transitioning to postsecondary education. By analyzing both student data and university transition programs, areas of overlap and improvement emerge.

As illustrated in Table 3, none of the 10 transition programs included main objectives that addressed uncovering societal views and stereotypes of LD, which was shown in the student data to complicate students' self-esteem and/or self-image. However, all programs included the benefits of accommodations and how such services and programs can aid in a positive transition for students with LD. Further, all universities except two discussed meeting other peers or family support as being a main objective during the move to postsecondary education. What raises the most concern is that only three programs included discussions around self-advocacy and only three programs addressing the role of mental health such as anxiety or depression.

Discussion

After analyzing both the first-hand experiences in narratives by first-year university students with LD and existing university transition programs, it is apparent that while there are areas in which transition efforts are indeed providing support, there are also gaps within these transitional supports. The findings from this study are situated within the university in which the student participants were enrolled, but also unmask a larger concern for students with LD as they make the complex transition to postsecondary education.

The following discussion is arranged around the four major themes of the study.

Self-Esteem/Image

According to the Transition Planning Resource by the Learning Disabilities Association of Ontario (LDAO; Nichols, 2003), to achieve a successful transition from secondary to postsecondary education, students must be able to understand their LD, present a positive self-image, develop positive personal strategies (such as self-motivation), and develop positive social and prosocial skills. Based on the student interviews, support in these areas must include a general aware-

Table 1
University Transition Programs for Students With Learning Disabilities

Program	University	Main Objectives
Students Taking Academic Responsibility (STAR)	Algoma University	Become familiar with assistive technology, learn different strategies, adopt concrete tools and resources, understand professor expectations
STARTonTrack	University of Guelp	Goal setting, planning, and motivation. Getting connected to people, a community, and groups on campus. Access resources and campus supports. Improve problem-solving
Easy Start Transition Program	University of Guelph-Humber	Explore services, learn about academic accommodations and meet returning students registered with Accessible Learning Services. Learn the differences between the expectations for students in high school compared to college and university. Learn about note-taking, time management, coping with anxiety. Meet other students starting their postsecondary journey and discover the common questions you all have. Learn about what academic accommodations are and how to prepare for university
Laurentian Imitative for Transition (LIFT)	Laurentian	Test-taking strategies, note-taking, study skill strategies, time management tools, understand what services are available and how to access them, time management skills, accessible technologies, how to manage your learning challenges successfully, meet other students
MyStart	Ontario Tech University	Understand differences between high school and postsecondary, self-advocacy, how to use your strengths, how to use accessible software, meet key staff for academic and learning support
On-Line to Success (OLTS)	Queens University	Strategies to deal with increased demands of workload, self-advocacy, and stress management. Identify strengths and challenges, understand personal learning profile, time management, stress management, self-advocacy, and how to navigate accessibility services
Student Heading into Full Time (SHIFT)	Ryerson University	How to register with Academic Accommodation Support, talk to current students with disabilities, understand learning strategies for academic success. Guidance on assistive technology, online lectures, learning strategies, and self-advocacy. Parent information sessions also available
Summer Academic Skills Institute (S.A.S.I)	University of Toronto – Mississauga Campus	Factors of resiliency and why it is important, wellness tools and strategies to manage stress, how to read a course syllabus, how to build a personalized study plan and a weekly schedule, how to take effective notes in lectures and from readings, test-taking strategies for math, essay, and multiple-choice tests, how to write an effective essay, Q&A with volunteer peer mentors
Bridge to University for Students with Learning Disabilities (BUILD)	University of Windsor	Explore the unique opportunities and challenges faced by students with learning disabilities, learn about the variety of services and supports available on campus and through Student Accessibility Services, discover valuable techniques, strategies, and technology to maximize learning potential, and meet fellow students, colleagues, and university staff and faculty
Student Community & Leadership Development (SCLD)	York University	Peer-to-peer programs, support healthy living and well-being, opportunities to develop or enhance academic and leadership skills, community building and making meaningful connections. Student engagement (helping students jump into university life with resources, tips, and services), health education & promotion (offering services that help support your health and well-being), student leadership (focuses on increasing student's leadership capacity, community engagement and social impact), and learning skills services (help develop new academic skills or refresh existing ones, from time management and effective studying skills to school-and work-life balance)

From *"Policy/Program Memorandum No. 8: Identification of and Program Planning for Students With Learning Disabilities,"* by Ontario Ministry of Education, 2014. Retrieved from http://www.edu.gov.on.ca/extra/eng/ppm/ppm8.pdf

Table 2
Transition Programs' Main Objectives

Transition Program Objective	Number of Times Mentioned
Accessible Technology/Software	5
Access to Resources/Services	7
Building Connection to Others	6
Problem Solving/Goal Setting	2
Understanding University Expectation	3
Note-/Test-Taking, Study Skills	5
Time/Stress Management	7
Coping With Mental Health/Well-Being	2
Self-Advocacy/Recognizing Strengths	3

ness of the stigma associated with LD and the association between LD and mental health concerns many students tend to experience. Such awareness can include specific difficulties students with LD encounter and how to become more sensitive and understanding toward them. Yet, this cannot be done without a recognition that these students are able to achieve their full potential. That is, the focus must resonate with their competencies and abilities ahead of the barriers represented by the label of being diagnosed with an LD (Nichols et al., 2003).

Several universities did include some of these key features within their transition programs. For instance, York University discussed the importance of providing student leadership and community engagement opportunities for a successful transition. Similarly, Ontario Tech University discussed the idea of teaching students how to recognize their personal strengths and challenges for an optimal university experience (see Table 1). Overall, transition programs should reflect a culture of participation and awareness that enables those students with LD to combat the stigmatizing stereotypes that view disability and "difference" as a natural and integral part of identity (Connor, 2012).

Lack of LD Awareness

One common theme expressed by participants was a lack of knowledge surrounding LD. Although LDs are common, there is a concerning lack of consistency and knowledge in the general population regarding what LDs are, what they entail, and how individuals are impacted by them. All the subjects in this study expressed a concern over the lack of education surrounding the nature of LD further perpetuating negative connotations and stigma. This lack of knowledge and awareness acts as a barrier to a smooth transition as many feel it creates negative connotations. Therefore, one of the most influential and moving takeaways from this study should be the need for further awareness in an effort to help individuals with LD to feel at peace with their diagnosis.

Several universities recognized this, such as Queens, University of Toronto, University of Windsor, and York University, by including key features in their transition programs that involved students understanding their learning profiles and being given opportunities to meet with other students with LD (see Table 2). However, ways to address the lack of knowledge surrounding LD should go beyond just educating the student with the LD to include more awareness among their university staff and students. Indeed, a lack of understanding of LD among the general public was one of the largest issues during the entire transition experience, as related by the participants.

Accommodations

All 10 institutions discussed the use of accommodations – whether providing information about access to resources, services, or accessible technology or just information about the accommodations available. For example, the Student Heading Into Full Time (SHIFT) Program at Ryerson University focuses on how to register with academic accommodation support services and provides guidance on assistive technology. Similarly, the Easy Start Transition Program at University of Guelph-Humber ensures that students are given an opportunity to explore services and learn about note-taking and how to access academic accommodations.

Accommodation policies and disability programs can make postsecondary education more accessible (Waterfield & Whelan, 2017). However, although accommodations were a key feature identified by the participants for a successful transition, they also pointed out that use

of accommodations often leads to negative connotations and stigma. That is, although being identified with an LD enables access to accommodations, negative social and political implications that are attached to such labeling remain (Waterfield & Whelan, 2017).

Banks (2014) found that, because of the stigma associated with disability culture, many students choose not to disclose their disability or seek the necessary accommodations until after they have experienced academic failure. This was illustrated within the current study when participants expressed how they proceeded to take a full course load, knowing that they had the option not to, but wanted to "fit in" with their peers and undergo the "typical" five-course university semester. While many expressed experiencing academic difficulties, they resisted seeking disability services or utilizing the appropriate accommodations due to the disability stigma they had previously experienced in high school.

Knowing that the current ableist society is not always understanding of those with a disability, some students feel shame, embarrassment, and regret surrounding their accommodations (Ryan, 2007). This notion was supported within the current study as every participant expressed feeling guilty or as if accommodations were "cheating" when using their accommodations. Some participants acknowledged that they knew this was only creating a level playing field and equity for them, but they still struggled to feel equally as "successful" as their peers with additional help, such as extra time on assignments or exams.

These findings are supported within the literature when Hall and Webster (2008), for example, note how students often feel uncomfortable within a merit-based system while relying on accommodations. It could be argued that these underlying feelings of guilt or "cheating" a merit-based system may be due to the underlying construction of disability culture. To combat this, transition programs should ensure that the discourse of dependence and incompetence that surrounds students with LD is deconstructed or, at a minimum, acknowledged in some way.

Support Networks

Family and friends were reported as one of the largest support networks during students' postsecondary education transition. Support from those closest to them provided encouragement that helped students develop and maintain self-confidence and perseverance. The young people within this study greatly valued the role of peer relationships and friendships during this transition period. Social networks of people with LD have been found to be smaller, with proportionately fewer friends, when compared to those without an LD (Hughes et al., 2013). Several students expressed knowing the importance of social networks but had a difficult time creating and maintaining such connections during their transition.

The current findings support previous literature showing that stability in social, emotional, and physical support networks acts as a key factor during the transition process (Martínez et al., 2011). Several programs also reflected the importance of meeting other students and being able to share experiences and concerns with others. For example, University of Guelph (and Guelph-Humber), Ryerson University, University of Windsor, and York University all included aspects of peer support programs and opportunities to meet other students within their transition programs. Hughes et al. (2013) recognized peer support as a protective factor during the transition process as these relationships help students cope with stress and protect one's emotional well-being. More transition programs should focus on the importance of peer relations, especially during adolescence where relationships are often prioritized.

Self-Advocacy

Many individuals with LD lack the self-determination necessary to effectively self-advocate to obtain the reasonable accommodations needed in college (Herridge, 2017). Nearly every university mentioned time and stress management as key features of their transition program (see Table 3). Students also discussed the need for help with organization and time management. Although some knew that it would take them longer to complete certain tasks compared to their peers, they often did not acknowledge it when starting on work that had to be finished on time. Thus, being able to successfully plan assignments and course workloads and understand their personal learning style acted as an academic protective factor (MOEQ, 2012). Developing self-advocacy skills prior to entering postsecondary education is imperative for a positive transition (Geller & Greenberg, 2009) since these allow students to feel more comfortable disclosing their disability and requesting the appropriate supports and accommodations (Herridge, 2017).

Since university is an increasingly more independent environment than high school, it is the responsibility of the individual student to take charge of their overall learning and well-being. Most participants in this study discussed previously suffering from low self-esteem and how, although sometimes it is still present, they have learned about their personal strengths and needs to become the best version of themselves. Yet, these experiences can lead to negative feelings, including anger, sadness, worry, frustration, and lower self-esteem (Integra, 2016). Self-esteem was not explicitly stated as a key feature by any of the 10 university transition programs. Universi-

ties must recognize the unique challenges faced by individuals with LD, which include societal views and their consequences for self-esteem, and work with students to provide them with tools to challenge the definition and stereotypes associated with a "learning disability."

Although self-advocacy was identified as a key feature by participants, only three programs specifically reflected on this within their transition programs (Ontario Tech, Queens, and Ryerson). Findings of the transition research on students with LD generally emphasize the need for external support networks rather than self-initiated efforts to ensure success (Connor, 2012). Preparation for higher education often focuses solely on curriculum-based needs, while failing to prepare students with LD to develop the skills needed to be successful (Burdge, 2012; Herridge, 2017). Therefore, it is the role of all stakeholders involved in the transition process to recognize that there are far more challenges than the commonly acknowledged academic realm. Without the proper coping and self-confidence skills, students with LD find it more difficult to live a positive and fulfilling life. Through obtaining a more comprehensive understanding of students labeled as LD and identifying ways to alleviate their concerns, postsecondary institutions would be better able to assist students in need of academic, social, and emotional support (Herridge, 2017).

Mental Health-Related Difficulties

According to previous research, students with LD are more susceptible to experiencing mental health-related difficulties (Piers & Duquette, 2016; Wilson et al., 2009). This was supported by the current study's findings, which suggest that students with LD are at a greater

Table 3

Overlaps Between Student-Identified Themes and University Transition Program Aims

	University									
Student-Identified Themes	Algoma U	U of Guelph	U of Guelph-Humber	Laur. U	Ontario Tech U	Queens U	Ryerson U	U of Toronto (Miss.)	Windsor U	York U
LD and Self-Esteem/Image										
Societal Views										
Consequences for Self-Esteem										
Accommodations										
Benefits	X	X	X	X	X	X	X	X	X	sX
Stigma										
Support Networks										
Family and Friends		X	X	X			X	X	X	X
Support/Case Workers		X	X	X	X	X			X	X
Lack of LD Education										
Self-Advocacy					X	X	X			
Mental Health Vulnerabilities										
Mental Health and High School										
Mental Health and the Transition			X					X		X

Note. U = University. Laur. = Laurentian. Miss. = Mississauga Campus. X indicates that the theme identified through student narratives was identified in the university program aims.

risk for mental health concerns following the transition to postsecondary education. It is common for LD to be comorbid with other conditions, including emotional, behavioral or attentional disorders, sensory impairments, or other medical conditions (Integra, 2016). All participants expressed struggling with issues such as anxiety, depression, and social skills. Specifically, the construct of their LD was felt by the participants to include constantly feeling "different." Further, aggression, bullying, peer victimization, self-injury, self-hatred, and other antisocial behaviors have been statistically linked to LD (Piers & Duquette, 2016), and were all mentioned by the participants within this study.

Only three transition programs, Easy Start Transition Program at Guelph-Humber, Student Community & Leadership Development at York University, and the Summer Academic Skills Institute at University of Toronto-Mississauga, referred to mental health or mental health-related areas as being a topic of focus. Anxiety, depression, low self-esteem, and poor social relationships all coincide with the overall transitional experience. York University identified health education and promotion, including offering services that help support the student's health and well-being, as a key method for a successful transition, while Guelph-Humber included topics such as coping and anxiety. The University of Toronto-Mississauga mentioned resiliency and wellness tools to manage stress.

Individuals with LD often feel as if they are not meeting others' expectations, that they are letting down their loved ones, and not working hard enough, despite trying their best (Integra, 2016). The findings of the current study reveal a concerning gap in this area. Every single university should be focusing on at least one area of mental health related difficulties for students with LD. Transitional efforts must work to include the impact of mental health more directly, especially when focused on vulnerable populations, such as those with LD , who will otherwise continue to struggle in silence.

Limitations

The present research provides insight into the experiences of eight first-year students with LD. The students chose to participate as they wanted to create awareness of the LD experience and reflect on their personal perceptions, thus encouraging more emancipatory research. While the findings provide in-depth insight into the transition experience for university students with an LD, several limitations must be considered.

While the small number of participants provided for an in-depth analysis of these students' experiences, they were recruited from one university course at the same educational institution and, therefore, the findings cannot be generalized. Further, the interview data represent an understanding of a small sample of students who have disclosed their disability to student accessibility services on campus. Research has suggested that 70% of students with accommodations in high school do not disclose their disability upon graduation (Newman et al., 2011). Additionally, many students with suspected LD are not formally diagnosed for several reasons (e.g., receiving a private assessment being too costly for some families or the waitlists to receive a psycho-educational assessment being months or even years long). Therefore, we acknowledged that the experiences vary depending on the individual and that a greater number of student participants may have provided other perspectives. That is, this study does not "speak to" or represent the voices of every student with an LD or other disabilities. In addition, the interpretations of others involved in the transition experience such as parents, peers, educators, and other professional would have offered more diverse findings.

The outcomes and analysis of the identified transition programs also offer some limitations. That is, since the participating institutions are all located within Ontario, Canada, the results and accompanying interpretations cannot be generalized to other North American or worldwide transitional efforts. Finally, the reliability of results could have been made stronger by having an additional researcher engage in the student data analysis, including data cleaning, to help prevent any inherent bias. Nevertheless, the study provides a rich understanding of the way students with LD navigate the first-year university experience.

Implications for Practice

The transition to postsecondary education for students with LD is a complicated process because of the confluence of individual, peer, family, institutional, and societal-related factors that directly and indirectly impact the experience. The objective of this study was to provide a space for students to share their stories and experiences, including their accomplishments and challenges faced while transitioning from high school to postsecondary education. The findings provide a unique and significant perspective on the high school-to-postsecondary-education transition experience – that of first-year students with LD. This type of research redefines how young people with LD experience and adjust to university, and thereby assist educators, case workers, support staff, and accessibility services in improving transition efforts and informing interventions. In addition, not only do the findings provide insight into the postsecondary education experience for incoming students, they can also offer information for current secondary students regarding

what is needed to support a positive move. In addition, findings can benefit family members by demonstrating the critical role support systems play during the transition and, therefore, enable them to provide adequate support at all stages. Finally, the findings may also offer clinical relevance for identifying types of support that can promote positive resilience and psychosocial adjustment for students with LD.

Transition efforts should be implemented as early as elementary school to provide comprehensive support. Along the way, such support must include a general awareness of the physicality of the transition experience, how to access accommodations and reduce stigma, as well as strategies for maintaining a positive mental health. Reducing barriers is crucial for students with LD transitioning to postsecondary education. From this study it is evident that some students managed to overcome certain barriers with the support of significant stakeholders (i.e., family, friends, case workers), which enabled a positive transitional experience. However, despite accommodations, students with LD continue to experience mental health-related struggles, which are intensified due to the stigmatization of disability. Therefore, services and resources should include input from students with LD. Such input is necessary for productive agency so those with disabilities can communicate their needs not only for university culture but for the larger disability culture as well.

Conclusion

The goal of the present study was to give voice to the experiences and perspectives of students with LD while they transition to postsecondary education. As a result, secondary and postsecondary education personnel can gain insight into how to best support students with disabilities to represent themselves during their transition, instead of being the objects of others' representation (Ashby, 2011). What makes this study unique is the link drawn to current Ontario university transition programs for students with LD. Many programs address the transition experience as a combination of physical, academic, and socio-emotional challenges, yet the larger mental health aspect cannot be ignored. If individuals are unable to maintain a positive self-esteem, create and maintain peer relations, and receive adequate support services for anxiety and depression, we are ultimately failing to ensure the basic needs of those with LD are being met. Although accommodations, meeting other current students, and self-advocacy are key features of a successful transition, other areas require more attention. A focus on reducing disability stigma and a greater emphasis on social supports and mental health advocacy is needed. That is, the ideal transition program would focus not only on the acquisition of basic academic skills but also self-efficacy, relationship skills, self-confidence, and mental health-related factors known to hinder academic performance and overall well-being.

The empowering stories presented here act as a testimony to the individualized experience of students with LD while they transition to university. As a result of having facilitated agency for individuals with LD, this research has led to findings that will help transition programs to more accurately reflect the direct needs of first-year university students with LD.

References

Alshenqeeti, H. (2014). Interviewing as a data collection method: A critical review. *English Linguistics Research, 3*(1), 39-45. http://dx.doi.org/10.5430/elr.v3n1p39

Ashby, C. (2011). Whose "voice" is it anyway?: Giving voice and qualitative research involving individuals that type to communicate. *Disability Studies Quarterly, 31*(4), 55-66. https://doi.org/10.18061/dsq.v31i4.1723

Banks, J. (2014). Barriers and supports to post-secondary transition: Case studies of African American students with disabilities. *Remedial and Special Education, 35*(1), 28-39. https://doi.org/10.1177/0741932513512209

Burdge, M. C. (2012). *Transition from secondary to post-secondary education for students with mild disabilities* [Doctoral dissertation]. ProQuest Dissertations and Theses Global. (UMI No. 1038834743)

Cavioni, V., Grazzani, I., & Ornaghi, V. (2017). Social and emotional learning for children with Learning Disability: Implications for inclusion. *International Journal of Emotional Education, 9*(2), 100-109.

Connor, D. J. (2012). Actively navigating the transition into college: Narratives of students with learning disabilities. *International Journal of Qualitative Studies in Education (QSE), 25*(8), 1005-1036. https://doi.org/10.1080/09518398.2011.590158

Cortiella, C., & Horowitz, S. H. (2014). *The state of learning disabilities: Facts, trends and emerging issues.* National Center for Learning Disabilities.

Court, D., & Abbas, R. (2013). Whose interview is it, anyway? Methodological and ethical challenges of insider-outsider research, multiple languages, and dual-researcher cooperation. *Qualitative Inquiry, 19*(6), 480-488.

Denhart, H. (2008). Deconstructing barriers: Perceptions of students labeled with learning disabilities in higher education. Journal of Learning Disabilities, 41(6), 483-497. https://doi.org/10.1177/0022219408321151

Denzin, N. (2001). The reflexive interview and a performative social science. *Qualitative Research, 1*(1), 23-46.

Geller, L. L., & Greenberg, M. (2009). Managing the transition process from high school to college and beyond: Challenges for individuals, families, and society. *Social Work in Mental Health, 8*(1), 92-116. https://doi.org/10.1080/15332980902932466

Gill, L. A. (2007). Bridging the transition gap from high school to college: Preparing students with disabilities for a successful post-secondary experience. *Teaching Exceptional Children, 40*(2), 12-15. https://doi.org/10.1177/004005990704000202

Hall, C. W., & Webster, R. (2008). Metacognitive and affective factors of college students with and without learning disabilities. *Journal of Post-secondary Education and Disability, 21*(1), 31-34. https://doi.org/10.2307/1511275

Herridge, A. S. (2017). First-year performance: Students with disabilities transitioning to college from high school. *College Student Affairs Leadership, 4*(1), 1-11.

Hughes, L., Banks, P., & Terras, M. (2013). Secondary school transition for children with special educational needs: A literature review. *Support for Learning, 28*(1), 24-34. https://doi.org/10.1111/1467-9604.12012

Integra. (2016). *LDMH: A handbook on learning disabilities and mental health.* Retrieved from https://www.childdevelop.ca/sites/default/files/files/Sept%2022%20Integra%20LDMH%20Handbook%202016.pdf

Kimberlin, S. E. (2009). Political science theory and disability. *Journal of Human Behavior in the Social Environment, 19*(1), 26-43. https://doi.org/10.1080/10911350802619870

Learning Disabilities Association of Canada. (2017). Retrieved from http://www.ldac-acta.ca

Learning Disabilities Association of Niagara Region. (2021). Retrieved from https://ldaniagara.org/resources/post-secondary-supports/

Learning Disabilities Association of Ontario. (2018). Retrieved from https://www.ldao.ca/introduction-to-ldsadhd/articles/about-lds/learning-disabilities-statistics/

Martínez, R., Aricak, O., Graves, M., Peters-Myszak, J., & Nellis, L. (2011). Changes in perceived social support and socioemotional adjustment across the elementary to junior high school transition. *Journal of Youth & Adolescence, 40*(5), 519–530. https://doi.org/10.1007/s10964-010-9572-z

Ministry of Education, Quebec. (2012). *Ensuring a smooth transition from elementary to secondary school.* Retrieved from http://www.education.gouv.qc.ca/fileadmin/site_web/documents/dpse/adaptation_serv_compl/Guide_SoutenirTransitionScolaireQualiteVersSec_a.pdfhttp://www.education.gouv.qc.ca/fileadmin/site_web/documents/dpse/adaptation_serv_compl/Guide_SoutenirTransitionScolaireQualiteVersSec_a.pdf

Morgan, P. L., & Hertzog, C. J. (2001). Designing comprehensive transitions. *Principal Leadership, 1*(7), 10-18.

Mosselson, J. (2010). Subjectivity and reflexivity: Locating the self in research on dislocation. *International Journal of Qualitative Studies in Education, 23*(4), 479-494.

Newman, L., Wagner, M., Knokey, A. M., Marder, C., Nagle, K., Shaver, D., & Wei, X. (2011). *The post-high school outcomes of young adults with disabilities up to 8 years after high school: A report from the national longitudinal transition study-2 (NLTS2).* NCSER 2011-3005. National Center for Special Education Research.

Nichols, E., Paul, C., Vandenbossche, D., Yaworski C., & Ziraldo, L. (2003). *Transition planning resource guide for students with learning disabilities.* (Revised from a document developed for LDAO in May 1999 by Learning Disabilities Association of Ontario.) Retrieved from http://www.access.resources.ldao.ca/main/documents/TransitionPlanningResourceGuide4.pdf

Ontario Ministry of Education. (2014). *Policy/program memorandum No. 8: Identification of and program planning for students with learning disabilities.* Retrieved from http://www.edu.gov.on.ca/extra/eng/ppm/ppm8.pdf

Ontario Universities' Application Centre. (2021, August 10). *University student transition resources.* https://www.ouac.on.ca/university-student-transition-resources/

Peters, J. (2011). *Transition skills of first-year college students with learning disabilities* [Doctoral dissertation]. ProQuest Dissertations and Theses Global. (UMI No. 915692476)

Piers, L., & Duquette, C. A. (2016). Facilitating academic and mental health resilience in students with a learning disability. *Exceptionality Education International, 26*(2), 21-41.

Pillow, P. (2003). Confession, catharsis, or cure? Rethinking the uses of reflexivity as methodological power in qualitative research. *International Journal of Qualitative Studies in Education, 16*(2), 175-196.

Reed, M. J., Kennett, D. J., Lewis, T., & Lund-Lucas, E. (2011). The relative benefits found for students with and without learning disabilities taking a first-year university preparation course. *Active Learning in Higher Education, 12*(2), 133-142. https://doi.org/10.1177/146978741140248

Rice, F., Frederickson, N., & Seymour, J. (2011). Assessing pupil concerns about transition to secondary school. *British Journal of Educational Psychology, 81*(2), 244-263. https://doi.org/10.1348/000709910X519333

Ryan, J. (2007). Learning disabilities in Australian universities: Hidden, ignored, and unwelcome. *Journal of Learning Disabilities, 40*(5), 436-442. https://doi.org/10.1177/00222194070400050701

Thomas, D. R. (2003, August). *A general inductive approach for qualitative data analysis.* Paper presented at School of Population Health, University of Auckland, New Zealand.

Waterfield, B., & Whelan, E. (2017). Learning disabled students and access to accommodations: socioeconomic status, capital, and stigma. *Disability & Society, 32*(7), 986-1006. https://doi.org/10.1080/09687599.2017.1331838

Wilson, A. M., Armstrong, C., Furrie, A., & Walcot, E. (2009). The mental health of Canadians with self-reported learning disabilities. *Journal of Learning Disabilities, 42*(1), 24-40. https://doi.org/10.1177/0022219408326216

Appendix A. Semi-Structured Interview Guide

1. Tell me about your LD (diagnosis, history, etc.). When were you diagnosed, and can you tell me about that process?
2. What made you want to participate in this study?
3. How do you define learning disabilities?
4. Do you like university? Why/why not?
5. How is university different from high school?
6. What were some good things about coming to university?
7. What were some challenging things about coming to university?
8. What expectations did you have from university?
   a. Where did these expectations come from?
9. What was your experience like transitioning from grade 12 to university?
   a. Do you think the transition was easy/hard/both? Why?
10. Do you feel that your high school prepared you for university?
11. you have accommodations in high school? If so, what were they?
12. Do you currently have accommodations in university? If so, what are they?
13. Did your accommodations in high school transfer to university?
   a. Were there barriers to getting the support you needed in university? If so, can you talk about those?
   b. Were you told by anyone to seek out accommodations? If so, who?
14. Were your accommodations in high school effective?
15. Are your accommodations in university effective?
16. Were there support networks or specific individuals who helped your transition?
17. Were you concerned about coming to university because of your learning disability? Why or why not?
18. Can you talk about any potential barriers that you faced specifically, because of your learning disability?
19. Do you think university does a good job with people who have LDs?
   a. Can you give me an example of what went right/wrong with your accommodations?
20. Do you feel as if having a learning disability made it harder for you to transition to university?
21. Did you feel there was a stigma associated with your LD/having accommodations?
22. Do people ever see your LD as the sole definition of yourself?
   a. Do people have a hard time seeing the strengths/positive attributes of your LD?
23. Are you open to telling people about your learning disability? Why or why not?
   a. What about your Professors/TAs?
   b. What do you tell people?
   c. Do you change the story depending on who you are talking to?
24. Have you ever regretted disclosing to someone about your LD? Why/who?
25. When has talking about your LD been useful?
   a. When has it been hurtful/harmful?
26. How have other people defined your learning disability?
27. Would you say that other people's understandings of learning disabilities are similar to your own? Why or why not?
28. Does having an LD ever cause individuals to perceive you in a certain way? If so, how?
29. Did you experience any social or emotional stress or hardships during your transition as a result of your learning disability?
30. Does having an LD make it any harder for you to maintain a positive self-esteem?
31. Has your LD ever made it difficult to maintain social relationships or friendships? If so, why or why not?
32. Do you think that having a learning disability puts an individual at a greater risk for mental health concerns? If so, why do you think?
33. What do you wish people knew about learning disabilities?
34. Would you say that having an LD makes you any more vulnerable to issues like peer victimization, and a negative self-esteem and self-confidence?
   a. In what ways?
35. Have you ever experienced feelings of anxiety or depression during your transition as a result of your LD?
36. Has having an LD created a stigmatization of academic underachievement? How?
37. What do you wish you knew prior to your transition to university that you know now?
38. What are some things that would make your transition process more successful or positive?
39. What would you say to someone who is struggling with their learning disability?
40. Is there anything you wish to share to the field disability studies that you wish everyone would know?

The Effects of Peer Racetracks on the Reading Performance of German L2 Students With and Without Behavioral Difficulties

Anne Barwasser, Susanne Hisgen, and Matthias Grünke
University of Cologne, Germany

Abstract

Reading difficulties in German students are steadily increasing, while reading motivation is decreasing, yet reading is one of the most important aspects of literacy. Complicating matters, reading instruction is challenged by an enormous heterogeneity among the student body, which includes students learning German as a second language and students with behavioral problems. Thus, many teachers are faced with trying to narrow the large gap between strong and weak readers by providing reading support, which often has to start with basic skills, and at the same time meeting the individual students' needs. The present single-case study assessed the effects of a simple peer-tutored motivational reading racetrack intervention on word fluency within a short period of time for students with and without behavioral problems (N = 9) for whom German is an L2. The results are promising with high overlap indices and the followup data showing stable improvements. Limitations and implications are discussed.

Keywords: German second language, literacy intervention, peer tutoring, reading racetracks, word recognition, behavioral difficulties

The German education system is increasingly faced with the challenge of dealing with linguistically heterogeneous students due to increased migration and a growing number of bilingual students acquiring German as an L2 (GL2) (Busse, 2013). In particular, many students have difficulty accessing the German language due to the language systems of their countries of origin, including languages that are transparent or semitransparent and languages that use alphabetic or non-alphabetic written language. In transparent languages, each letter corresponds to one sound, whereas in semitransparent languages (e.g., English) each letter represents different sounds depending on the combination of letters (Buetler et al., 2014; Rodriguez et al., 2016). Alphabetic script is a phonographic script, whose characters primarily represent phonetic segments of spoken language (Gnanadesikan, 2017). The German language is a transparent alphabetic writing system (Perfetti & Dunlap, 2008).

The 2018 Program for International Student Assessment (PISA) showed that students in Germany without an immigrant background scored 63 points higher in reading literacy than students with an immigrant background (Kaya & Elster, 2018), placing Germany among the countries with the largest gap in literacy between L1 and L2 speakers among all Organisation for Economic Cooperation and Development (OECD) countries (OECD, 2010). Furthermore, the Progress in International Reading Literacy Study Survey (PIRLS) reported a significant upward trend in the number of struggling fourth-grade readers in Germany, including GL2. In addition, overall reading motivation decreased across both native and non-native younger and older students (Lafontaine et al., 2018).

Generally speaking, students whose native language is not the language of instruction face a higher risk of school problems than their L1 peers (Golinkoff et al., 2019; Kieffer et al., 2008; Kieffer & Vukovic, 2013; Lesaux & Kieffer, 2010). It is important to learn the German language appropriately in order for these students to have equal educational opportunities in German society (Han, 2012) and, therefore, it is urgent to promote the literacy skills of children learning German.

The first important aspect in second-language reading is word recognition. Without this basal competence, L2 readers have greater difficulty in rapid

Barwasser, A., Hisgen, S., & Grünke, M. (2022). The Effects of Peer Racetracks on the Reading Performance of German L2 Students With and Without Behavioral Difficulties. *International Journal for Research in Learning Disabilities, 5*(2), 70-85. https://doi.org/10.28987/ijrld.5.2.70

reading than L1 readers, and these difficulties often lead to problems in L2 comprehension (Barwasser et al., 2021c; Barwasser et al., 2021d; Dijkstra & Van Heuven, 2002; Schroeter & Schroeder, 2018; Qu et al., 2018;). For example, Cirino et al. (2013) demonstrated the challenges of students with reading difficulties related to basic reading skills (e.g., word recognition). Thus, word recognition also plays a fundamental role in reading comprehension (Perfetti & Stafura, 2014). According to the Dual Route Cascaded Model (DRC; Coltheart et al., 2001), word recognition can occur via two different routes: a phonological route (indirect), in which each word must be recoded letter by letter based on grapheme-phoneme correspondence rules, and an orthographic route (direct), in which written words can be mapped directly onto mental representations of word forms. For the German language, Zari and Nagler (2021) found that orthographic knowledge is a better predictor of comprehension than phonological knowledge, underscoring the importance of strengthening whole-word reading.

Being able to read sight words plays a fundamental role (Balass et al., 2010; Kendeou & O'Brien, 2018) and is crucial in terms of reading fluency and reading comprehension (Barwasser et al., 2021a; Musti-Rao et al., 2015). Sight words can be defined as words that are read within one second of their appearance without phonological recoding (Ehri, 2005). McArthur et al. (2015) showed positive effects of sight word training on practiced and unpracticed words and on fluency in word and sentence reading. Automation in word reading can be particularly challenging for L2 learners, who often require more working memory capacity to read a text in the non-native language (Fraser, 2007). When working memory becomes overloaded, learning is often impossible.

Students with behavioral difficulties, in particular, need special support when faced with such a challenge in a new written language task because they often have serious deficits in achieving reading proficiency (Wanzek et al., 2014). Furthermore, these students are less likely to engage in literacy activities (Becherer et al., 2020; Roberts et al., 2020). Further, students with lower reading skills are more likely to exhibit behavioral challenges than students without (Lin et al., 2013). In terms of L2, language difficulties have also been linked to behavior issues (Jansen et al., 2020), as poorer language skills are a risk factor for developing behavior difficulties (Chow & Wehby, 2018; Hollo et al., 2014). Petersen and LeBeau (2021) reported that social skills, language skills, and behavior difficulties are interrelated and that language ability plays a fundamental role in the development of externalizing behavior difficulties. In sum, the combination of academic and behavioral difficulties/disorders can make it difficult for professionals to provide effective support (Kauffman, 2005), including, in this context, helping struggling students in order to achieve a higher language proficiency level (Grigorenko et al., 2020; Harn et al., 2014).

How to Foster Sight Word Reading

Interventions that can be implemented economically and beneficially are particularly useful (Solis et al., 2017). To address the aforementioned challenges, there are a number of evidence-based interventions that support less proficient readers in inclusive settings (Grigorenko et al., 2020; Harn et al., 2014). Sperling et al. (2019) suggested that a mix of two or more components (e.g., peer tutoring and repeated reading) appears to be most effective in supporting heterogeneous readers equally.

Repeated Reading Through Reading Racetracks

Repeated reading (RR) is necessary to achieve greater reading fluency (Chard et al., 2009), as borne out in a number of studies (e.g., Chard et al., 2002; Ring et al., 2013; Zimmermann et al., 2021). Further, several studies have shown positive effects of repeated word reading on both word reading outcome (e.g., Martens & De Jong, 2008) and general word decoding skills – also in L2 (Shimono, 2018; Van Gorp et al., 2017).

Gamified educational components have been found effective in engaging students in literacy interventions and bringing about positive outcome in reading (Lämsä et al., 2018). One example of integrating RR into an intervention and making it game-like is Reading Racetracks (RT) – a circular board with squares on which cards can be placed with inputs whose content can be trained automatically, such as sight words. RT have been shown to be effective for lower-performing readers in second-language (e.g., Barwasser et al., 2021c; Grünke & Barwasser, 2019; Sperling et al., 2019) and first-language reading (e.g., Barwasser et al., 2021a; Barwasser et al., 2021b; Davenport et al., 2019), as well as for students with learning disorders (LD) and emotional and behavioral disabilities (EBD) (e.g., Barwasser et al., 2021b; Barwasser et al., 2021d).

Peer Tutoring as an Inclusive Tool

To turn an intervention into an inclusion tool where children with different characteristics can participate successfully, a peer-tutorial (PT) element may be added. PT is defined as a method where students work together on a specific exercise (Dufrene et al., 2010). Adding PT to interventions has beneficial effects (Mercer et al., 2011), and several studies have demonstrated its effectiveness across multiple age groups in terms of both academic gains and

social-emotional well-being (e.g., Bowman-Perrott et al., 2014; Moeyaert et al., 2021). With regard to academic achievement, PT has been found successful for reading (Dufrene et al., 2006; Hattie, 2008; Moeyaert et al., 2021). Positive effects have also been demonstrated for students learning an L2 with and without LD (Cole, 2014; Klingner et al., 2014). Moreover, a literature synthesis by Okilwa and Shelby (2010) revealed that PT had a positive impact on academic achievement in a wide range of subject areas for 6- to 12-year-olds, regardless of the type of impairment.

Self-Graphing as a Motivational Boost

Motivation plays a leading role in learning (Marinak & Gambrell, 2008). Thus, it is crucial to make an intervention as engaging as possible, especially for struggling students. Despite its game-like character, RT benefits from self-management components such as self-graphing, whereby students can track their progress and get individual feedback (Sutherland & Snyder, 2007). Beyond motivation, being able to monitor their own progress has positive impacts on students' behavior (Amato-Zech et al., 2006; Legge et al., 2010) as well as academic performance (Sutherland & Snyder, 2007) because they can compare themselves to themselves rather than their peers (Menzies et al., 2009). For example, Guzman et al.'s (2018) meta-analysis of K-12 students' self-monitoring of reading achievement showed a large positive effect on reading achievement as a result of the use of self-graphing.

Research Aim

In light of the fact that the number of less proficient L2 readers in Germany is increasing along with additional difficulties such as behavioral problems, there is an urgent need for interventions that meet the individual needs of the heterogeneous student body. The current study used a peer-tutorial racetrack intervention that is adapted to individual needs through its use of multiple components while promoting one of the most important areas of reading: reading of sight words. Specifically, we looked at the effects of the intervention on third graders with GL2 with and without behavioral problems who showed severe difficulties in literacy, including the students' and teachers' views of the intervention in the context of social validity. Our research questions were as follows:

1. Does a combined RT intervention have positive effects on the sight word reading of struggling GL2 readers with and without behavior difficulties?
2. Are the positive effects maintained after six weeks without intervention?
3. How is the intervention rated by the participants and their teachers?

Methods

Participants and Setting

The study was conducted at an inclusive elementary school in North Rhine-Westphalia, Germany, with students from two third-grade classes. Before the study started, consent forms were distributed to the legal guardians of potential participants. Screenings were conducted to assess students' level of proficiency to choose the final participants and as a basis for dividing the students into tutors and tutees.

The first instrument was the Salzburg Reading Screening Test (SLS 2-9; Mayringer & Wimmer, 2014), which was administered to all students of both classes. In addition, a German vocabulary test, the Integrated Teacher Report Form (Weiß, 2006), was administered – a German version for externalizing behavior problems (Integrated System Teacher Report Form [ITRF-G]; Volpe et al., 2018). Finally, a word pretest was used to crystallize the final words for the intervention. Student characteristics (e.g., date of birth, gender, first and second language) were collected through a teacher questionnaire.

Assessment

Salzburg Reading Screening (SLS 2-9)

The SLS 2-9 was administered as a group test to assess reading fluency. Students are presented with a list of meaningful and nonsensical sentences, with a prompt to check off the meaningful sentences (e.g., *Lemons taste sour*. a) true, b) not true). The evaluation is based on the number of correctly selected sentences, which is converted into a reading quotient (RQ). The RQ is the extent to which the measured reading ability deviates from the average of a norming sample where 100 stands for the mean value with a standard deviation of 15. Norms (reading quotients) for 11,900 students are available for the SLS 2-9. An RQ of 80-89 is considered below average, 70-79 weak, and less than or equal to 69 very weak. With regard to test validity, the test scores are closely related to speed in reading word lists aloud (r = .80 to .90) (Wimmer & Mayringer, 2014). The reliability of the SLS ranges between .87-.95. Participants scoring below RQ 79 were chosen for the study.

ITRF-G

The teachers were asked to complete the German short version of the Integrated System Teacher Report Form (ITRF-G; Volpe et al., 2018) about the potential tutees. The ITRF-G is a universal screening procedure using a 4-point- scale to identify behavior problems in (a) learning-related behavior (APD) with 8 items and

Table 1
Characteristics of Student Participants

Student	Gender	Age/Grade	SN	OPP	APD	Reading RQ	Words (PR)	German L2	L1
Lava	f	8/3	EBD	10	1	69	42	Yes	Serbian
Emyl	m	8/3	/	0	2	67.5	58	Yes	Arabic
Gülcan	f	9/3	/	1	6	66	10	Yes	Kurdic
Ferhat	m	8/3	/	15	13.5	65.5	12	Yes	Turkish
Brav	m	10/3	/	8.5	10.5	65	10	Yes	Romanian
Tugce	f	8/3	/	0	1	64	18	Yes	Turkish
Gül	f	9/3	EBD	14	19.5	63.5	21	Yes	Turkish
Vaneza	f	9/3	/	0	1.5	71	15	Yes	Kurdic
Metül	m	8/3	EBD	12	5	69	42	Yes	Arabic

Note. SN = special needs; f = female; m = male; LD = learning disabilities; EBD = emotional behavioral disorder; L1 - first language; L2 = second language; OPP = disruptive behavior according to the ITRF(OPP); APD = learning-related behavior according to the ITRF(APD); RQ = reading quotient; Words PR = percentile on German Vocabulary Test.

(b) oppositional/disruptive behavior (OPP) with another 8 items. The cutoffs are as follows: 10 for APD, 5 for OPP 5, and 13 for the overall problem value. The ITRF has a high internal consistency (APD: $\alpha = .91$; OD: $\alpha = .87$; total: $\alpha = .91$) (Volpe et al., 2018). We implemented the ITRF because problem behavior can negatively impact reading and writing as well as overall school performance (Campbell et al., 2018).

Word Pretesting

The researcher-developed word pretest was administered to identify words for the intervention that the majority of the participants were not already storing as sight words. It was composed of two PowerPoint presentations with 70 words each and was conducted on two consecutive days. Each word was displayed on a separate slide for one second followed by three slides with hashtags to slow down the speed. The word list was generated on the basis of the German childLex (Schroeder et al., 2015), which displays word frequency. The students' task was to read the words correctly aloud within 1 sec of their appearance (Ehri, 2005). A total of 30 words were selected that could not be read across the majority of participants.

German Vocabulary Test

The German version of the Culture Fair Intelligence Test and the subtest for German vocabulary (CFT-20-R; Weiß, 2006) (cutoff percentile rank <15) was administered as a group screening containing 30 items with words from semantic domains and abstract concepts based on a representative sample of $N = 2724$; it has good reliability ($r = .87$) and a high internal consistency, .86 to .96 (Weiß, 2006). The students are given a series of words, each accompanied by a choice of other words and have to decide which of the words most closely reflect the main words. This task was chosen because vocabulary knowledge is a well-known predictor of success in L2 literacy (Schmitt et al., 2011).

Final Participants

The final participants assumed the role of tutees, each of whom was assigned a tutor, who was identified as having a high reading level on the SLS 2-9 (>120). Students were assigned to pairs by first sorting students according to their RQ. No data was collected on the tutors, except that they had to be able to read all the training words correctly. The final participants were nine tutees ($N = 9$), all were GL2 and had started learning German with the entry of kindergarten at age 3-4. The exception was Brav, who had entered school at age 7.5. According to the ITRF screening, four children showed problem behavior, and three were diagnosed with EBD. Behavioral difficulties in this context means problem behavior independent of an official EBD diagnosis. However, all tutees showed severe difficulties in reading.

We also surveyed the social validity of the intervention as viewed by the two teachers, because it is important to find out whether the method would also be used in schools by teachers. The two teachers were the class teachers of the children, with teacher 1 (female, age 35) teaching five of the children and the other teacher (female, age 39) teaching the other four.

Design

A multiple-baseline design was used across individuals, with the start of the intervention occurring at different times for individuals (Morgan & Morgan, 2009). The participants were randomly assigned to three groups with different phase lengths to fulfil the requirement for a multiple-baseline design of at least three tiers (Tate et al., 2016). The intervention took place three days per week for 10 minutes over a period of seven weeks. Before the start of the intervention, baseline data was collected from the tutees. Then, in the following sessions, the RT was applied followed by data collection. Since the study was conducted during regular class time, the students were pulled out of their classes by the interventionists for the intervention period. Six master's-level students in special education served as test leaders and interventionists. In two configurations, four students were always together in pairs in one group, rotating through the groups in order to prevent an assessor effect. The remaining two students assessed the data. Group 1 had five baseline measurements, Group 2 had six, and Group 3 had seven, which were randomly assigned. Three followup measurements (Phase E) using the same measuring instrument were administered six weeks after the intervention, ended including two weeks of Christmas holidays.

Dependent Variables and Measurement

The dependent variable was the number of correctly read training words. The total was used to determine the extent of the students' word recognition regarding the words of training. The word test was a PowerPoint presentation composed of the 30 training words selected during the pretest. It was constructed the same way as the word pretest and was identical to it in terms of structure and mode of operation. The order of the individual words in the presentation was varied for each test in order to avoid a practice effect. The 30 words each appeared on a separate slide in an automatic 1-sec rhythm (Ehri, 2005), and two interventionists guided the tests. Each word that was correctly read within one second was counted as correct.

Materials

At the beginning of the intervention phase, the student pairs (tutor + tutee) were presented with the RT field and the flashcards. The self-made game field was 30x60cm and contained 30 empty rectangular cells, which were distributed at even intervals on the racetrack (see Figure 1). Each pair of students received a playing field and a game piece. For each of the 30 cells, there was a flashcard showing one of the 30 words to be trained.

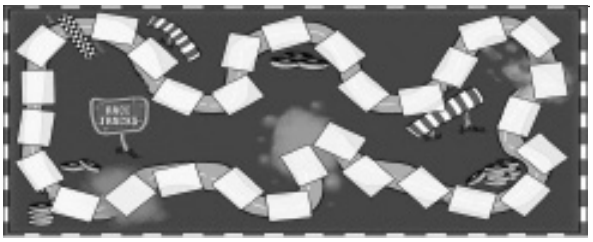


Figure 1
Racetrack Gameboard

For the self-graphing procedure, a sheet was designed that included 12-14 rows (depending on the number of intervention sessions) with 30 small squares for the maximum number of words to be read correctly (see Figure 2). The rows were one below the other so that students could easily track their learning visually. Words for the intervention were of high frequency (mean: 2878) (childLex; Schroeder et al., 2015), meaning that they appear at a frequency of 100 per million words in a corpus (Brysbaert et al., 2018).

Figure 2
Self-Graphing Sheet

Note. Leserennstrecke = Reading Racetracks; Trainingsbogen = Training Sheet.

Procedures

Baseline

Before the actual intervention (Phase B) began, a baseline phase (Phase A) was conducted to determine the current status of the dependent variable and thus be able to estimate the effectiveness of the intervention. However, instead of determining the dependent variable only after each baseline session, children were engaged for 10 minutes beforehand to counter the argument that any improvement during the intervention was due only to a Hawthorne effect – change in behavior by the subjects of a study due to their awareness of being observed. Keeping students engaged and paying attention during the baseline period made the Hawthorne effect less likely. During each baseline session, pairs of students were engaged for 10 minutes in a dif-

ferent program that consisted of cognitive tasks, including logical continuation of a sequence and determining which object did not match the others in the sequence. After each baseline session, students were assessed on the dependent variable.

Intervention

Before the intervention phase began, the tutors were trained for 1.5 hours on how to give feedback in order to fulfill their role as tutors. For this purpose, they were given sample sentences and sample situations while working in pairs.

In the intervention phase, the RT game was played, and prior to its start, it was thoroughly explained to all participants. The pairs remained constant throughout the intervention phase. The RT game was played for 10 min always with the 30 training words, which were printed on individual flashcards and spread across the game board. The tutees moved along the flashcards with their game piece figure one after the other, and the respective card was revealed by the tutor. The tutee read the corresponding word aloud. If it was incorrect, the tutor corrected it after 3 sec to allow for self-correction. If a correction was made, the tutee repeated the correct form of the word aloud. The revealed cards remained face up.

If the tutee read all the cards within the 10 min, they were reshuffled and distributed face down on the field and a new game started for the remainder of the time. After each measurement from the intervention phase on, the tutees recorded the number of correctly read words on their self-graphing sheet. The number of correctly read words was reported to them by the interventionists. The tutees were then allowed to color in the number of boxes in the appropriate rows for the current session on their self-graphing sheet.

Treatment Fidelity

Treatment fidelity aims to improve the accuracy and consistency of an intervention to ensure correct delivery (Smith et al., 2007). The implementation was monitored by means of a treatment fidelity questionnaire, which was completed by the interventionists after each unit and by an external person after one third of the intervention time. To ensure faithful implementation of the training, the treatment sheet listed various components, including the environment where the support was given, the provision of materials, the process of support and feedback, and the recording of the handling of student behavior. Example questions were as follows: "Did the support take place without external disturbances?," "Were the materials ready for support?," and "Did the tutor have 3 seconds for self-correction?." Overall interrater reliability was 98% across the interventionists as well as between the interventionists and the external raters.

Social Validity

In order to increase the social validity of the study, two questionnaires were distributed to class teachers and the participants after the end of the intervention, asking them to record their acceptance of the intervention. The statement items both questionnaires are listed in Table 2. The items were rated on a 5 point-Likert scale from 0 = "totally not agree" to 4 = "totally agree." For the student survey, the interventionists were not present to avoid biased results. Further, to prevent students' low reading level from influencing the results, the teachers went over the statements individually with each student.

Table 2
Social Validity Items

Students	Racetrack helped me to be able to read words correctly.
	I think the intervention also helps other students with difficulties in reading.
	I understood well the meaning of the intervention.
	I learned a lot during the intervention.
	I was happy to come to the intervention.
	I enjoyed the intervention.
	I would participate again.
Teachers	Automation is especially important in the context of reading.
	The intervention is a good way to improve students' reading fluency.
	The intervention is an appropriate way to train reading fluency of sight words.
	I would use the intervention in my classroom.
	The total time required to complete the intervention was manageable.
	The material resources required for this intervention are adequate.

Results

For the data analysis, the statistical program R and the SCAN package for single-case analysis (Wilbert & Lueke, 2021) were used. First, visual inspection was applied followed by overlap measures and regression analysis at Level 2, once for each group and once across all groups, with a particular focus on possible Phase A trends, the increase from Phase A to Phase B (slope), and the direct increase at the onset of the intervention (level). The Non-Overlap of All Pairs (NAP; Parker et al., 2011a), the Percentage Exceeding the Median Trend (PEM; Ma, 2006), the Percentage of All Non-Overlapping Data (PAND; Parker et al., 2007), and Tau-U (Parker et al., 2011b; A vs. B + TrendB – TrendA) were selected as overlap measures.

The NAP is the percentage of all pairwise comparisons between baseline and the intervention phase and displays improvement across phases (Parker et al., 2011a). Thus, it can be used when data shows some variation in the phases. The PEM is the percentage of data points exceeding the baseline phase median, which is a good overlap index when there are floor or ceiling effects in the baseline phase. PAND refers to the total number of data points that do not overlap between phases while using all data from both phases, which makes it more concrete and reliable (Parker et al., 2007). Tau-U is a combination of Kendall's rank correlation and Mann-Whitney U while being able to control for possible baseline trends (Parker et al., 2011b). To directly test treatment effects, a regression analysis on level 2 (across all students) was performed to estimate level and slope effect from Phase A to Phase B (see Tate et al., 2016). First an analysis was performed for each group and afterwards across all groups. Finally, the SCAN package (Wilbert & Lueke, 2021) can control for autocorrelation single-case analysis.

Visual Analysis and Descriptive Statistics

The baselines could be characterized as not having much variation across participants, which can also be seen when looking at the standard deviation (*SD*) in Phase A, which ranged from 0.5 to 2.19 with mean values ranging from 0.40 to 12.00. Nevertheless, the data for Lava ($M = 2.60$, $SD = 2.19$), Tugce ($M = 3.50$, $SD = 2.17$), and Gül ($M = 1.71$, $SD = 1.11$) went slightly upward towards the end of the baseline while Brav's seemed to go a bit down. Emyl started with higher values ($M = 12.00$, $SD = 1.00$) than the other participants, and quickly reached the maximum number of words to be read in the intervention phase. Lava also reached the maximum number very rapidly, although her baseline data was significantly lower. Brav ($M = 0.40$, $SD = 0.55$), Ferhat ($M = 1.33$, $SD = 0.82$), and Gül had the lowest mean values at baseline ($M = 1.71$, $SD = 1.11$). Metül ($M = 5.57$, $SD = 0.79$) and Vaneza ($M = 5.42$, $SD = 0.79$) also had flat baselines, which seemed to go down for a short time, but then stabilized again towards the end of the baseline.

Regarding Phase B, all children improved. Brav showed a rather low mean value, due to the fact that he needed a certain amount of time for the data to increase. Although he was the one with the least increase, he still reached a maximum value of 20 in Phase B. Lava was very strong in Phase B and, like Emyl, achieved the maximum possible value of 30. Tugce also displayed a vast increase and reached a value of 29.00. Gülcan (*M* Phase A = 5.00; *M* Phase B = 18.71), Metül (*M* Phase A = 5.57; *M* Phase B = 20.50), and Vaneza (*M* Phase A = 5.43; *M* Phase B = 19.25) had a similar increase from Phase A to Phase B in mean values. Gül (*M* Phase A = 1.71; *M* Phase B = 20.09) and Ferhat (*M* Phase A = 1.33; *M* Phase B = 14.92) also had similar Phase A mean values but Gül showed a stronger increase than Ferhat.

Overall, all participants showed a clear increase without a followup drop with visible trends in Phase B and level effects from Phase A to Phase B. Thus, the followup measurements proved to be stable, showing that the students could still read the words after six weeks without intervention. Only Vaneza (23, 26, 24), Lava (29, 30, 28), and Gül (24, 26, 24) displayd a slight drop at the third followup measurement (see Figure 3 and Table 3).

Overlap Indices

According to the NAP, all subjects achieved a strong effect (99.00-100.00; $p<.001$) except for Brav, who displayed a moderate effect (92.00; $p<.01$). The PEM showed a highly effective treatment for all. The PAND displayed strong effects (94.74-100.00) for all subjects except Brav, who had a moderate effect (77.78). In terms of the Tau-U values (considering Phase A trends), all participants presented large changes (>.60, $p<.001$) except Gülcan, who exhibited a very large change (0.81, $p<.001$).

With regard to the regression analysis, first per group, a significant level effect from Phase A to Phase B ($p<.01$) could be seen for Group 1 but no significant slope effect. Overall, there was no Phase A trend. In Group 2, there was a significant slope effect ($p<.01$) from Phase A to Phase B with an average improvement of 1.139 words per session. There was also no Phase A trend. Group 3 displayed a significant level effect from Phase A to Phase B ($p<.001$) and a significant slope effect ($p<.01$), with a beta coefficient of 1.078. There was also a significant level effect from Phase B to Phase E ($p<.01$). The regression analysis across all groups showed a significant level effect ($p<.001$) as well as slope effect ($p<.001$) from A to B and a level effect from B to E ($p<.001$). The students were able to read an average of 1.056 words per session.

Social Validity

Students

The results of the social validity survey for the students were as follows. All items were rated between "partly agree" and "strongly agree." First, all children found that the RT helped them read words correctly and that the intervention can help other children as well. Lava and Gülcan were the only ones to "partly agree" with this statement. Second, the students understood the intervention well and felt that they have learned a lot. Overall, the intervention was perceived as very pleasant. Again, Lava and Gülcan were the only

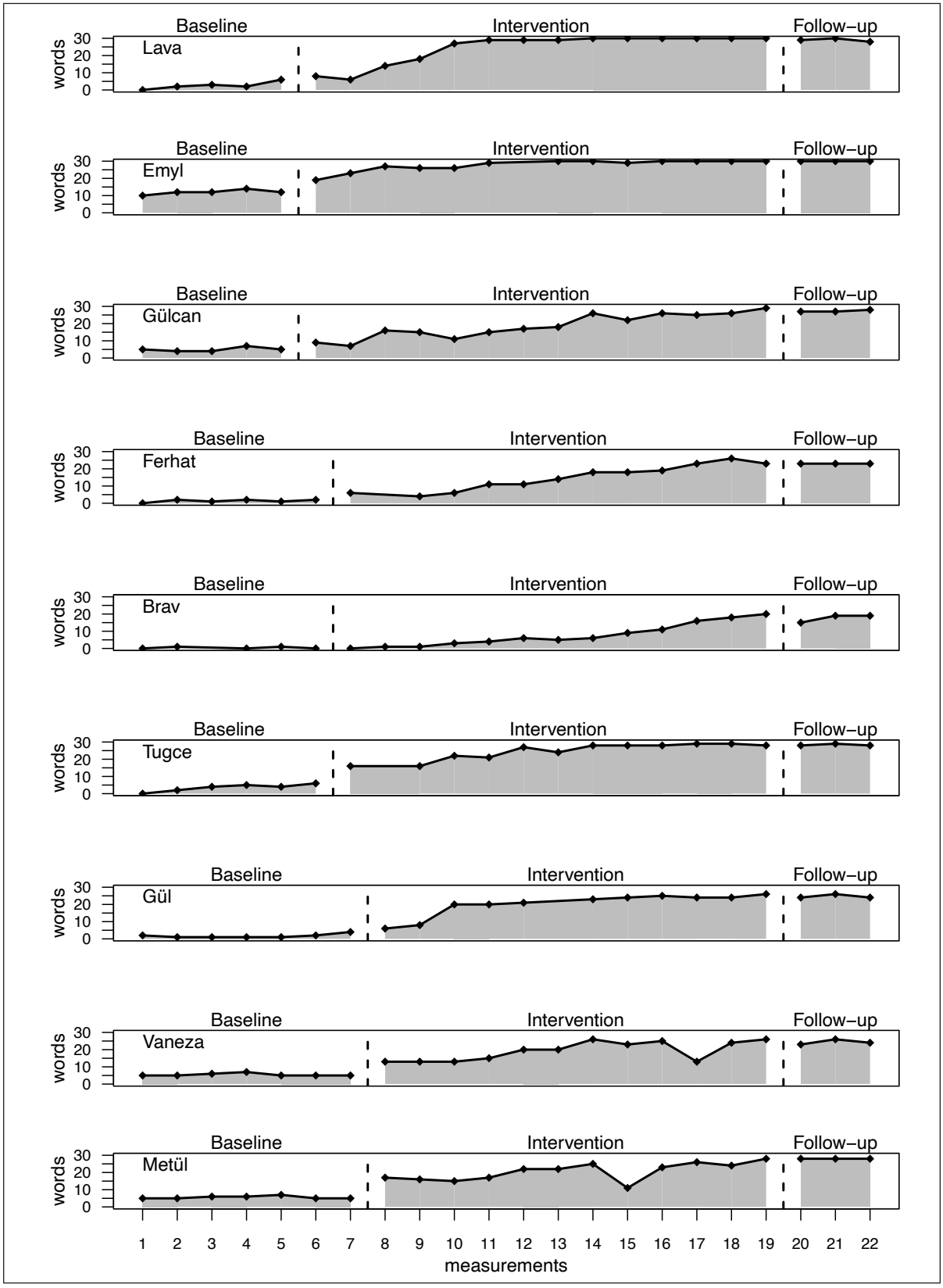


Figure 3
Words Read Correctly

Table 3
Descriptive Data and Overlap Indices for Each Participant for Words Read Correctly

	Lava	Emyl	Gülcan	Ferhat	Brav	Tugce	Gül	Vaneza	Metül
N(A)	5	5	5	6	6	6	7	7	7
N(B)	14	14	14	13	13	13	12	12	12
N(E)	3	3	3	3	3	3	3	3	3
M(A) (SD)	2.60 (2.19)	12.00 (1.00)	5.00 (1.22)	1.33 (0.82)	0.40 (0.55)	3.50 (2.17)	1.71 (1.11)	5.43 (0.79)	5.57 (0.79)
M(B) (SD)	24.29 (8.84)	27.62 (3.40)	18.71 (7.04)	14.92 (7.38)	7.69 (6.69)	24.67 (4.85)	20.09 (6.77)	19.25 (5.53)	20.50 (5.18)
Max(B)	30.00	30.00	29.00	26.00	20.00	29.00	26.00	26.00	28.00
M(E) (SD)	29.00 (1.00)	30.00 (0.00)	27.33 (0.58)	23.00 (0.00)	17.67 (2.31)	28.33 (0.58)	24.67 (1.15)	24.33 (1.53)	28.00 (0.00)
Max(E)	30.00	30.00	28.00	23.00	19.00	29.00	26.00	26.00	28.00
NAP (p)	99.00 (<.001)	100.00 (<.001)	99.00 (<.001)	100.00 (<.001)	92.00 (<.001)	100.00 (<.001)	100.00 (<.001)	100.00 (<.001)	100.00 (<.001)
PEM	100.00	100.00	100.00	100.00	92.31	100.00	100.00	100.00	100.00
PAND	94.74	100.00	94.74	100.00	77.78	100.00	100.00	100.00	100.00
Tau-U (p)	0.78 (<.001)	0.68 (<.001)	0.81 (<.001)	0.73 (<.001)	0.75 (<.001)	0.64 (<.001)	0.69 (<.001)	0.72 (<.001)	0.70 (<.001)

Note. N = measurements; *M* = mean; *SD* = standard deviation; Max = maximum value; A = Phase A; B = Phase B; E = follow-up; *p* = *p* value; PEM = Percentage Exceeding the Median Trend; PAND = Percentage of All Non-Overlapping Data; NAP = Non-Overlap of All Pairs.

ones to state "partly agree." All would like to participate again; Lava was unsure, however.

Teachers

The teachers also rated the intervention as positive. All were of the opinion that the intervention is a good way to train reading fluency and sight words. In addition, teachers responded "agree" and "totally agree" to the statement that their students had become better at reading overall. All teachers responded that they would use the intervention in their own classes. The materials seemed appropriate as did the timing across the board. Only one teacher stated "partly agree" when asked about time management. However, this did not refer to the intervention itself, but to the individual measurements after each session.

Discussion

Main Findings

The aim of the present study was to evaluate the effect of a peer-tutorial racetrack game on automating sight words in primary school students with and without behavioral difficulties who have reading difficulties in their GL2. The intervention was specifically designed to meet individual student needs within a heterogeneous study body such as is found in an inclusion classroom (meaning that different students can learn together with one method). The results show a clear overall increase for each child with flat baselines in a very short time and stable data at followup.

Looking at the results in more detail, the fact that there is no significant slope effect for Group 1 is partly because Emyl already had higher values in the baseline and that he, as well as Lava, rapidly reached the maximum possible number of words. Otherwise, all showed an increase and also stable followup data in Group 1. Gülcan improved significantly more slowly than those of other group members. In the German vocabulary screening, she only had a percentile of 10, which is low, and might have had an impact on her reading performance (e.g., Schroeter & Schroeder, 2017). Lava indicated "partly agree" on the social validity survey in response to whether she enjoyed the intervention and would participate. Since she reached the maximum number of words quickly, boredom might explain her response. The ITRF screening showed that she has severe problems with disruptive behavior and had been diagnosed with EBD. However, problem behavior did not seem to have an impact on her results.

Group 2 displayed a steady increase. Tugce even reached the upper range quite quickly and reacted positively to the onset of the intervention even though she was one of the weakest readers. Compared to the others, Brav needed more time until he could correctly read several words. He had a percentile in German vocabulary of 10 and was the only students with Romanian as L1. His low vocabulary knowledge and language background might have played a role for the slower increase (Lesaux & Kieffer; 2010; Schroeter & Schroeder, 2017). Brav also exhibited problems in OPP and APD, which relates to concentration and, thus, could have been a factor. All three students rated the intervention positively. Ferhat also had problems in the areas of OPP and APD. However, this did not seem to affect the success of the intervention much for him. The followup data was stable overall; Brav was the only one who dropped, but only minimally. However, since he was the student who seemed to have the most difficulty in memorizing the words, it was possible that he needed more automation.

Across the board, Group 3 also showed significant improvement. Vaneza and Metül responded positively to the intervention while Gül made a huge jump at the third intervention session. Followup data was stable for each child. Gül was the weakest reader according to the prior reading screening, had an EBD diagnosis, and problems in OPP as well as APD. Perhaps she had to settle in before she could focus on the intervention and/or perhaps memorizing sight words was still difficult for her at first because she was trying to read via phoneme-grapheme correspondence, which is typically for German readers (Landerl & Wimmer, 2008).

Overall, the data does not show that the students with behavioral difficulties had more problems than their peers benefiting from the intervention. Looking specifically at the three students with an official EBD diagnosis, it is not clear that EBD had a negative impact on the intervention, as all three benefited from the intervention in a very short period of time, even though they started with very low baseline scores. Interesting, these findings differ from other researchers arguing that students with behavioral difficulties have greater problems benefiting from literacy interventions and that teachers have difficulty finding appropriate reading interventions for them (Becherer et al., 2020; Forlin & Chambers, 2011; Nelson et al., 2003; Roberts et al., 2020). Motivation plays a major role in behavior, so perhaps the playful and motivating nature of the racetracks helped ensure that students with behavioral difficulties benefit to the same extent as those without behavioral difficulties. Likewise, we cannot fully determine whether students' different L1s had an impact, because the progress of the students cannot be differentiated by L1. In terms of gender, Brav and Ferhat were the weakest overall, but this does not apply to Emyl and Metül. Therefore, it cannot be said with certainty that there was a difference between the subjects in terms of gender.

Based on the fairly clear results, we can conclude that PT RT, which was only conducted for 10-minute sessions, after a very short time, had a positive effect on the learning of sight words of less proficient GL2 primary school readers with Kurdish, Turkish, Arabic, Serbian, and Romanian language backgrounds, with and without behavioral difficulties. Followup data also showed that the children were able to recall the words after six weeks. In addition, there was no difference between children with and without behavioral problems, and no difference by age.

The finding that problem behavior did not seem to play a major role is consistent with Kempe et al. (2011), who argued that reading problems and externalizing behavior should be seen as independent of each other. The findings are also consistent with findings from PT studies (e.g., Moeyaert et al., 2021; Okilwa & Shelby; 2010) and previous studies using reading racetracks (Barwasser et al., 2021a; Barwasser et al., 2021b; Grünke & Barwasser, 2019; Sperling et al., 2019). Further, the results also support studies with a focus on self-graphing (e.g., Legge et al., 2010; Sutherland & Snyder, 2007).

Limitations

Despite the promising results, several limitations of the present study warrant mentioning. As with all individual case studies, the sample size was small, making it difficult to draw generalized conclusions. However, the advantage of such studies lies in the opportunity they offer to directly see individual responses to an intervention and, therefore, be better able to assess and adapt it to specific characteristics. In addition, they allow researchers to see through the learning process how long it takes for an intervention to elicit responses in the students. Furthermore, it is unclear what influence the children's languages of origin played. According to Zigler and Goswami (2005), learning to read is a process of understanding orthography-phonology mapping, which may explain differences in reading in different languages. Overall, While there was no clear difference in this respect between the children overall, the influence of the various L1s on GL2 is unknown. Further, we did not measure reading in children's L1 – thus, they might have problems in L2 reading but not in L1 – maybe as a result of undergoing a language learning process.

Additionally, at this point we can only draw conclusions about children with the language backgrounds

mentioned above and, of course, only about German as L2. Thus, our findings cannot automatically be transferred to other languages and sight word reading, even if a transfer could be expected due to prior studies focusing on sight word training in, for example, English (McArthur et al., 2015).

For Emyl, the already quite high values in the baseline were of interest, with a ceiling effect expected to follow. Nevertheless, his direct response to the intervention and the initially strong increase show that the intervention seemed to be effective for him as well. Plus, we cannot be sure which effects the RT with PT had vs. the self-graphing, since self-monitoring is also considered to promote reading (Menzies et al., 2009). However, since most instruction in Germany classrooms uses reinforcers and research has shown the importance of motivation, we decided to also reinforce while combining the RT with self-graphing.

As a last limitation, it should be mentioned that the very positive effects may also be due to the small group size and, therefore, the more intensive teaching. And yet, in a previous study, Barwasser et al. (2021b) also found strong effects in a classroom setting of a racetrack intervention. However, the latter study was conducted with children without special needs support.

Implications

Based on the results and limitations, our study has several implications for practice and research. As a next step, one could look at other language backgrounds as well as other languages in which sight vocabulary is promoted in PT. In addition, since an L1 influences the L2 (Shum et al., 2016), it would be interesting to measure L1 reading beforehand to see connections. Reading motivation could also be surveyed before and after using questionnaires to see correlations since reading and motivation are strongly related (Marinak & Gambrell, 2008). Further, many children fail in reading because they have not yet fully internalized the L2 alphabetic principle. On the basis of a previous screening, more information could be obtained in order to assess whether a given intervention might be too difficult.

Regarding the intervention itself, one could choose a larger pool of training words simply to minimize a training effect. To avoid a training effect in the current study, the words in the measurements were shown in randomized order each time, but this does not completely rule out a training effect. Moreover, use of a larger pool of training words would minimize the likelihood of ceiling effects, besides offering the children an opportunity to store a larger sight vocabulary. However, since not all of the children in the study were able to memorize all of the words at the end, the time of the intervention and possibly the number of training weeks would have to be increased in order to create automation. Especially children with weaker skills need a high degree of automation and to have words presented to them with high frequency (Chard et al., 2009).

In future studies, we would also like to focus on the tutors, both in terms of how they liked the intervention and to determine if their skills improved. For this purpose, we could use another questionnaire and general reading screenings before and after the intervention – and also with the tutees – in order to detect influences of the intervention on their general reading fluency. Plus, one could embed the training words in texts and, by measuring before and after, determine if the text can be read faster as a result. In addition, one could separate the RT from the self-graphing within a single-case design with A B BC phases or an experimental group design with the control group receiving only the Racetracks. Moreover, one could contrast one intervention with another and generate a larger sample to make more generalized statements about the effectiveness of PT RT in GL2. Referring back to the limitation of the small group size, a further study could be implemented in a classroom with students with special needs to see if these students need to be in a small group to improve or if a classroom setting would lead to the same positive effects.

Conclusion

Overall, this study showed that it is possible, in a short time and in a simple way, to improve the sight vocabulary of poorly reading primary students in the second language, some of whom also show behavioral difficulties. Training sight vocabulary is essential for reading (Musti-Rao et al., 2015) and sets the trajectory for reading fluency and comprehension. Our intervention adds another piece to the puzzle of reading research and through peer tutoring makes Reading Racetracks a tool that can be used in an inclusive way, allowing all children to participate. Use of the intervention can contribute to narrowing the large gap between students with and without German L2 in reading and give everyone the same educational opportunities while enhancing their language skills.

References

Barwasser, A., Urton, K., & Grünke, M. (2021a). Effects of peer-tutorial reading racetracks on sight word reading of secondary school students with learning disabilities and emotional behavioral disorder. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2021.671385

Barwasser, A., Urton, K., Grünke, M., Sperling, M., & Coker, D. L. (2021b). Fostering word fluency of struggling third graders from Germany through motivational peer-tutorial reading racetracks. *Reading & Writing, 35*, 29–53.

Barwasser, A., Schmitz, S., & Grünke, M. (2021c). A sub-lexical patterns training with Racetracks on trained and untrained words in low-literacy German students with behavior problems with and without learning disabilities. *Learning Disabilities: A Contemporary Journal.*

Barwasser, A., Nobel, K., & Grünke, M. (2021d). Peer-reading racetracks for word reading of low-achieving graduating students with learning disabilities and behavioural problems. *British Journal of Special Education.*

Amato-Zech, N. A., Hoff, K. E., & Doepke, K. J. (2006). Increasing on-task behavior in the classroom: Extension of self-monitoring strategies. *Psychology in the Schools*, *43*(2), 211–221. https://doi.org/10.1002/pits.20137

Balass, M., Nelson, J. R., & Perfetti, C. A. (2010). Word learning: An ERP investigation of word experience effects on recognition and word processing. *Contemporary Educational Psychology*, *35*(2), 126–140. https://doi.org/10.1016/j.cedpsych.2010.04.001

Becherer, J., Koeller, O., & Zimmermann, F. (2020). Externalizing behaviour, task-focused behaviour, and academic achievement: An indirect relation? *British Journal of Educational Psychology, 91*(1), 27–45.

Bowman-Perrott, L., Burke, M. D., Zhang, N., & Zaini, S. (2014). Direct and collateral effects of peer tutoring on social and behavioral outcomes: A meta-analysis of single-case research. *School Psychology Review*, *43*(3), 260–285. https://doi.org/10.1080/02796015.2014.12087427

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, *27*(1), 45–50. https://doi.org/10.1177/0963721417727521

Buetler, K. A., de León Rodríguez, D., Laganaro, M., Müri, R., Spierer, L., & Annoni, J. M. (2014). Language context modulates reading route: an electrical neuroimaging study. *Frontiers in Human Neuroscience, 8*, 83.

Busse, V. (2013). An exploration of motivation and self-beliefs of first year students of German. *System*, *41*(2), 379–398. https://doi.org/10.1016/j.system.2013.03.007

Campbell, A. R., Bowman-Perrott, L., Burke, M. D., & Sallese, M. R. (2018). Reading, writing, math, and content-area interventions for improving behavioral and academic outcomes of students with emotional and behavioral disorders. *Learning Disabilities: A Contemporary Journal, 16,* 119–138.

Chard, D., Vaughn, S., & Tyler, B. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities*, *35*(5), 386–406. https://doi.org/10.1177/00222194020350050101

Chard, D. J., Ketterlin-Geller, L. R., Baker, S. K., Doabler, C., & Apichatabutra, C. (2009). Repeated reading interventions for students with learning disabilities: Status of the evidence. *Exceptional Children*, *75*(3), 263–281. https://doi.org/10.1177/001440290907500301

Chow, J. C., & Wehby, J. H. (2018). Associations between language and problem behavior: A systematic review and correlational meta-analysis. *Educational Psychology Review*, *30*(1), 61–82. https://doi.org/10.1007/s10648-016-9385-z

Cirino, P. T., Romain, M. A., Barth, A. E., Tolar, T. D., Fletcher, J. M., & Vaughn, S. (2013). Reading skill components and impairments in middle school struggling readers. *Reading & Writing*, *26*(7), 1059–1086. https://doi.org/10.1007/s11145-012-9406-3

Cole, M. W. (2014). Speaking to read: Meta-analysis of peer-mediated learning for English language learners. *Journal of Literacy Research*, *46*(3), 358–382.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Davenport, C. A., Konrad, M., & Alber-Morgan, S. R. (2019). Effects of reading racetracks on sight word acquisition for deaf kindergarteners. *The Journal of Deaf Studies and Deaf Education*, *24*(2), 173–185. doi:10.1093/deafed/eny038

Dijkstra, T., & Van Heuven, W. J. (2002). Modeling bilingual word recognition: Past, present and future. *Bilingualism: Language and Cognition*, *5*(3), 219–224.

Dufrene, B. A., Reisener, C. D., Olmi, D. J., Zoder-Martell, K., McNutt, M. R., & Horn, D. R. (2010). Peer tutoring for reading fluency as a feasible and effective alternative in response to intervention systems. *Journal of Behavioral Education*, *19*(3), 239–256. https://doi.org/10.1007/s10864-010-9111-8

Ehri, L. C. (2005). Development of sight word reading: Phases and findings. In M. J. Snowling & C. Hulme (Eds.), *Blackwell handbooks of developmental psychology. The science of reading: A handbook* (pp. 135–154). Blackwell Publishing. https://doi.org/10.1002/9780470757642.ch8

Forlin, C., & Chambers, D. (2011). Teacher preparation for inclusive education: Increasing knowledge but raising concerns. *Asia-Pacific Journal of Teacher Education, 39*(1), 1732.

Fraser, C. A. (2007). Reading rate in L1 Mandarin and L2 English across five reading tasks. *The Modern Language Journal*, *91*(3), 372–394. https://doi.org/10.1111/j.1540-4781.2007.00587.x

Gnanadesikan, A. E. (2017). Towards a typology of phonemic scripts. *Writing Systems Research*, *9*(1), 14-35.

Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019). Language matters: Denying the existence of the 30-million-word gap has serious consequences. *Child Development*, *90*(3), 985–992. DOI: 10.1111/cdev.13128

Grigorenko, E. L., Compton, D. L., Fuchs, L. S., Wagner, R. K., Willcutt, E. G., & Fletcher, J. M. (2020). Understanding, educating, and supporting children with specific learning disabilities: 50 years of science and practice. *American Psychologist*, *75*(1), 37. https://doi.org/10.1037/amp0000452

Grünke, M., & Barwasser, A. (2019). Enhancing sight word fluency of second-language elementary students through reading racetracks. *International Journal of Technology and Inclusive Education, 8*(1), 1373–1378.

Guzman, G., Goldberg, T. S., & Swanson, H. L. (2018). A meta-analysis of self-monitoring on reading performance of K–12 students. *School Psychology Quarterly*, *33*(1), 160.

Han, W. J. (2012). Bilingualism and academic achievement. *Child Development*, *83*(1), 300–321. https://doi.org/10.1111/j.1467-8624.2011.01686.x

Harn, B., Fritz, R., & Berg, T. (2014). Effective literacy instruction in inclusive schools. In J. McLeskey, N. L. Waldron, F. Spooner, & B. Algozzine (Eds.), *Handbook of effective inclusive schools: Research and practice* (pp. 229–246). Routledge. https://doi.org/10.4324/9780203102930

Hollo, A., Wehby, J. H., & Oliver, R. M. (2014). Unidentified language deficits in children with emotional and behavioral disorders: A meta-analysis. *Exceptional Children*, *80*(2), 169–186. https://doi.org/10.1177/001440291408000203

Jansen, R., Maljaars, J., Verhappen, A., Zink, I., Steyaert, J., & Noens, I. (2020). Problem behavior in young children referred with language difficulties: Relations to language and intentional communication. *Autism & Developmental Language Impairments*, *5*. https://doi.org/10.1177/2396941519900076

Kauffman, J. M. (2005). *Characteristics of children's behavior disorders* (7th ed.). Merrill.

Kaya, V. H., & Elster, D. (2018). German students' environmental literacy in science education based on PISA data. *Science Education International*, *29*(2), 75–87. https://doi.org/10.33828/sei.v29.i2.2

Kempe, C., Gustafson, S., & Samuelsson, S. (2011). A longitudinal study of early reading difficulties and subsequent problem behaviors. *Scandinavian Journal of Psychology*, *52*(3), 242–250. https://doi.org/10.1111/j.1467-9450.2011.00870.x

Kendeou, P., & O'Brien, E. J. (2018). *Reading comprehension theories: A view from the top down.* Routledge. doi:10.4324/9781315687384

Kieffer, M. J., Lesaux, N. K., & Snow, C. E. (2008). Promises and pitfalls: Implications of No Child Left Behind for identifying, assessing, and educating English language learners. In G. Sunderman (Ed.), *Holding NCLB accountable: Achieving accountability, equity, and school reform* (pp. 57–74). Corwin Press.

Kieffer, M. J., & Vukovic, R. K. (2013). Growth in reading-related skills of language minority learners and their classmates: More evidence for early identification and intervention. *Reading and Writing, 26*(7), 1159–1194. https://doi. org/10.1007/s11145-012-9410-7

Klingner, J. K., Boele, A., Linan-Thompson, S., & Rodriguez, D. (2014). Essential components of special education for English language learners with learning disabilities: Position statement of the Division for Learning Disabilities of the Council for Exceptional Children. *Learning Disabilities Research & Practice*, *29*(3), 93–96.

Knoepke, J., Richter, T., Isberner, M. B., Naumann, J., & Neeb, Y. (2014). Phonological recoding, orthographic decoding, and comprehension skills during reading acquisition. *Zeitschrift für Erziehungswissenschaft*, *17*, 447–471. https://doi.org/10.1007/s11618-014-0560-z

Lafontaine, D., Dupont, V., & Schillings, P. (2018). Does reading literacy instruction vary according to language or culture? Similarities and differences in English-, German- and French-speaking education systems in PIRLS 2016. *IEA Compass: Briefs in Education*, *4*, 1–11.

Lämsä, J., Hämäläinen, R., Aro, M., Koskimaa, R., & *Äyrämö*, S.-M. (2018). Games for enhancing basic reading and maths skills: A systematic review of educational game design in supporting learning by people with learning disabilities. *British Journal of Educational Technology*, *49*(4), 596–607. https://doi.org/10.1111/bjet.12639

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology*, *100*(1), 150–161. https://doi.org/10.1037/0022-0663.100.1.150

Legge, D. B., DeBar, R., & Alber-Morgan, S. R. (2010). The effects of self-monitoring with a MotivAider(R) on the on-task-behaviour of fifth and sixth graders with autism and other disabilities. *Journal of Behaviour Assessment and Intervention in Children*, *1*(1), 43–52. https://doi.org/10.1037/h0100359

Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal*, *47*(3), 596–632. https://doi.org/10.3102/0002831209355469

Lin, Y. C., Morgan, P. L., Hillemeier, M., Cook, M., Maczuga, S., & Farkas, G. (2013). Reading, mathematics, and behavioral difficulties interrelate: Evidence from a cross-lagged panel design and population-based sample of US upper elementary students. *Behavioral Disorders*, *38*(4), 212–227. https://doi.org/10.1177/019874291303800404

Ma, H. H. (2006). An alternative method for quantitative synthesis of single subject researches: Percentage of data points exceeding the median. *Behavior Modification*, *30*(5), 598–617. https://doi.org/10.1177/0145445504272974

Mahdavi, J. N., & Tensfeldt, L. (2013). Untangling reading comprehension strategy instruction: Assisting struggling readers in the primary grades. *Preventing School Failure: Alternative Education for Children and Youth*, *57*(2), 77–92. https://doi.org/10.1080/1045988X.2012.668576

Marinak, B. A., & Gambrell, L. B. (2008). Intrinsic motivation and rewards: What sustains young children's engagement with text? *Literacy Research and Instruction*, *47*(1), 9–26. https://doi.org/10.1080/19388070701749546

Martens, V.E.G., & de Jong, P. F. (2008). Effects of repeated reading on the length effect in word and pseudoword reading. *Journal of Research in Reading*, *31*(1), 40–54. https://doi.org/10.1111/j.1467-9817.2007.00360.x

Mayringer, H., & Wimmer, H. (2014). *SLS 2-9. Salzburger Lese-Screening für die Schulstufen 2–9* [*Salzburg Reading Screening for Grades 2–9*]. Hogrefe.

McArthur, G., Castles, A., Kohnen, S., Larsen, L., Jones, K., Anandakumar, T., & Banales, E. (2015). Sight word and phonics training in children with dyslexia. *Journal of Learning Disabilities*, *48*(4), 391–407. https://doi.org/10.1177/0022219413504996

Menzies, H., Lane, K., & Lee, J. (2009). Self-monitoring strategies for use in the classroom: A promising practice to support productive behavior for students with emotional or behavioral disorders. *Beyond Behavior*, *18*(2), 27–35.

Mercer, C. D., Mercer, A. R., & Pullen, P. C. (2011). *Teaching students with learning problems*. Pearson.

Moeyaert, M., Klingbeil, D. A., Rodabaugh, E., & Turan, M. (2021). Three-level meta-analysis of single-case data regarding the effects of peer tutoring on academic and social-behavioral outcomes for at-risk students and students with disabilities. *Remedial and Special Education*, *42*(2), 94–106. https://doi.org/10.1177/0741932519855079

Morgan, D. L., & Morgan, R. K. (2009). *Single-case research methods for the behavioral and health sciences*. SAGE Publications. https://doi.org/10.4135/9781483329697

Musti-Rao, S., Lo, Y., & Plati, E. (2015). Using an iPad® app to improve sight word reading fluency for at-risk first graders. *Remedial and Special Education*, *36*(3), 154–166. https://doi.org/10.1177/0741932514541485

Nelson, J. R., Benner, G. J., & Gonzalez, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy Interventions: A meta analytic review. *Learning Disabilities Research & Practice, 18*, 255–267.

Okilwa, N. S., & Shelby, L. (2010). The effects of peer tutoring on academic performance of students with disabilities in grades 6 through 12: A synthesis of the literature. *Remedial and Special Education*, *31*(6), 450–463. https://doi.org/10.1177/0741932509355991

Organisation for Economic Cooperation and Development (OECD). (2010). *PISA 2009. Assessment Framework: Key competencies in reading, mathematics and science*. https://doi.org/10.1787/9789264062658-en

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percent of all nonoverlapping data (PAND): An alternative to PND. *The Journal of Special Education*, *40*(4), 194–204. https://doi.org/10.1177/00224669070400040101

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011a). Effect size in single case research: A review of nine non-overlap techniques. *Behavior Modification*, *35*(4), 302–322. https://doi.org/10.1177/0145445511399147

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011b). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, *42*(2), 284–299. https://doi.org/10.1016/j.beth.2010.08.006

Perfetti, C. A., & Dunlap, S. (2008). Learning to read: General principles and writing system variations. In *Learning to read across languages* (pp. 25–50). Routledge.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37. https://doi.org/10.1080/10888438.2013.827687

Petersen, I. T., & LeBeau, B. (2021). Language ability in the development of externalizing behavior problems in childhood. *Journal of Educational Psychology*, *113*(1), 68–85. https://doi.org/10.1037/edu0000461

Qu, Q., Cui, Z., & Damian, M. F. (2018). Orthographic effects in second-language spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(8), 1325.

Ring, J. J., Barefoot, L. C., Avrit, K. J., Brown, S. A., & Black, J. L. (2013). Reading fluency instruction for students at risk for reading failure. *Remedial and Special Education*, *34*(2), 102–112. https://doi.org/10.1177/0741932511435175

Roberts, G. J., Cho, E., Garwood, J. D., Goble, G. H., Robertson, T., & Hodges, A. (2020). Reading interventions for students with reading and behavioral difficulties: A meta-analysis and evaluation of co-occurring difficulties. *Educational Psychology Review*, *32*(1), 17–47. https://doi.org/10.1007/s10648-019-09485-1

Rodriguez, D. D. L., Buetler, K. A., Eggenberger, N., Preisig, B. C., Schumacher, R., Laganaro, M., ... & Müri, R. M. (2016). The modulation of reading strategies by language opacity in early bilinguals: an eye movement study. *Bilingualism: Language and Cognition*, *19*(3), 567–577.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *Modern Language Journal*, *95*, 26–43. https://doi.org/10.1111/j.1540-4781.2011.01146.x

Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., & Kliegl, R. (2015). ChildLex–Eine lexikalische Datenbank zur Schriftsprache für Kinder im Deutschen [childLex-A lexical database of written language for children in German]. *Psychologische Rundschau*, *66*(3), 155–165. https://doi.org/10.1026/0033-3042/a000275

Schroeter, P., & Schroeder, S. (2017). The impact of L2 German on component processes of reading. *Journal of Research in Reading*, *40*(1), 107–124. https://doi.org/10.1111/1467-9817.12078

Schroeter, P., & Schroeder, S. (2018). Differences in visual word recognition between L1 and L2 speakers: The impact of length, frequency, and orthographic neighborhood size in German children. *Studies in Second Language Acquisition*, *40*(2), 319–339.

Shimono, T. R. (2018). L2 reading fluency progression using timed reading and repeated oral reading. *Reading in a Foreign Language*, *30*(1), 152–179.

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

Smith, S. W., Daunic, A. P., & Taylor, G. G. (2007). Treatment fidelity in applied educational research: Expanding the adoption and application of measures to ensure evidence-based practice. *Education and Treatment of Children*, *30*(4), 121–134. https://doi.org/10.1353/etc.2007.0033

Solis, M., Scammacca, N., Barth, A. E., & Roberts, G. J. (2017). Text-based vocabulary intervention training study: Supporting fourth graders with low reading comprehension and learning disabilities. *Learning Disabilities: A Contemporary Journal*, *15*(1), 103–115.

Sperling, M., Barwasser, A., & Grünke, M. (2019). The effects of a reading racetrack intervention on the sight word fluency of learning-disabled elementary school students with German as second language. *Insights into Learning Disabilities, 16*(1), 79–90.

Sutherland, K. S., & Snyder, A. (2007). Effects of reciprocal peer tutoring and self-graphing on reading fluency and classroom behavior of middle school students with emotional or behavioral disorders. *Journal of Emotional and Behavioral Disorders*, *15*(2), 103–118.

Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., et al. (2016). The single-case reporting guideline in behavioural interventions (SCRIBE) 2016 statement. *Journal of School Psychology 56,* 133–142. https://doi: 10.2522/ptj.2016.96.7.e1

Van Gorp, K., Segers, E., & Verhoeven, L. (2017). The role of feedback and differences between good and poor decoders in a repeated word reading paradigm in first grade. *Annals of Dyslexia*, *67*, 1–25. https://doi.org/10.1007/s11881-016-0129-z

Volpe, R. J., Casale, G., Mohiyeddini, C., Grosche, M., Hennemann, T., Briesch, A. M., & Daniels, B. (2018). A universal behavioral screener linked to personalized classroom interventions: Psychometric characteristics in a large sample of German schoolchildren. *Journal of School Psychology*, *66*, 25–40. https://doi.org/10.1016/j.jsp.2017.11.003

Wanzek, J., Al Otaiba, S., & Petscher, Y. (2014). Oral reading fluency development for children with emotional disturbance or learning disabilities. *Exceptional Children*, *80*(2), 187–204. https://doi.org/10.1177/001440291408000204

Weiß, R. H. (2006). *Grundintelligenztest Skala 2 - Revision (CFT 20-R) mit Wortschatztest und Zahlenfolgentest.* [*Basic Intelligence Test Scale 2 - Revision (CFT 20-R) with vocabulary test and number sequence test.*] Hogrefe.

Wilbert, J., & Lueke, T. (2021). *Analyzing single-case data with R and SCAN.* Available at https://jazznbass.github.io/scan-Book/

Zarić, J., & Nagler, T. (2021). Reading comprehension on word-and sentence-level can be predicted by orthographic knowledge for German children with poor reading proficiency. *Reading and Writing, 34*(8), 2031–2057.

Zigler, J., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Zimmermann, L. M., Reed, D. K., & Aloe, A. M. (2021). A meta-analysis of non-repetitive reading fluency interventions for students with reading difficulties. *Remedial and Special Education, 42*(2), 78–93.

The International Academy for Research in Learning Disabilities

44th Annual IARLD Conference

Preconference Activities – October 3, 2022
Conference – October 4 and 5, 2022

***University of Oviedo* (Asturias), Spain**

Conference Chairs:

Professors **Celestino Rodríguez** and
Paloma Gonzalez Castro
Department of Psychology
University of Oviedo

Visit www.iarld.com for details

Conference Information:

Faculty of Teacher Training and Education
University of Oviedo
Email: rodriguezcelestino@uniovi.es
mgcastro@uniovi.es

Phone: +34 (985) 103402

Manuscript Submission Guidelines

English is used for submissions to the journal, correspondence and publication. All submissions must be formatted consistent with the 7th edition of the *Publication Manual of the American Psychological Association* (APA). Manuscripts must include a 100- to 150-word abstract summarizing the contents.

A critical concern in the learning disabilities field is the definition of the population. Therefore, authors are expected to operationally define the study participants in accordance with professional standards (see CLD Research Committee: Rosenberg et al., 1993. Minimum standards for the description of participants in learning disabilities research. *Learning Disability Quarterly, 26*(4), 210-213). In addition, parameters of the setting in which the research took place are to be clearly delineated. Such descriptions facilitate replication and application of results. Manuscripts that fail to specify participant and setting variables will be rejected or returned to the authors for clarification. Authors of research manuscripts are encouraged to include brief (e.g., one to two sentences) explanations of why specific procedures and analysis methods were employed. Peer reviewers will evaluate the appropriateness of this and all aspects of the reported study.

Manuscripts are not to exceed 35 double-spaced pages, consistently employing a 12-point font (including references, tables, figures, and appendices). Please limit tables and figures to those essential to conveying your content. Please present figures and tables in portrait format and use grey scale rather than colored images.

The manuscript submission and review process will be conducted electronically. To submit your manuscript digitally, you must use one of the following formats: Microsoft Word (.doc or .docx), RTF, or PDF. Manuscripts must be saved as"letter" ("U.S. letter") page length or paper size. Submissions in Word or RTF format will be converted into a PDF before being sent for blind peer review. If you submit a PDF file, it is your responsibility to ensure it can be read and printed by others. To avoid delays, please embed all fonts and use Adobe PDF Distiller instead of PDF Writer to ensure that others can view the article exactly as intended. No email attachment should exceed 15 MB.

Each manuscript must be accompanied by a cover letter that communicates (a) that the manuscript is an original work, (b) that the manuscript is not under consideration by any other journal, and (c) any other disclosures as required by the APA (e.g., ethical treatment of participants, financial relationship disclosures). A cover page that provides the name, affiliation, mailing address, phone number, fax, and email address of each author must also accompany each manuscript. All communications will be with the lead author. No author identifying information should be included directly in manuscripts submitted for review (e.g., explicit reference to previous publications, acknowledgments, author biographies). Please send the required cover letter, separate cover page with author identifying information, and manuscript as three separate attachments to a single email to IJRLDeditors@gmail.com (please do not send tables, figures, appendices or other supporting materials in separate files). Manuscripts may not be submitted by fax or in paper form.

For further information, please refer to:http://www.iarld.com/wp-content/uploads/2022/03/IJRLD-Call-for-Manuscripts_031122dkr.pdf

International Journal for Research in Learning Disabilities
Lynch School of Education and Human Development
140 Commonwealth Avenue
Chestnut Hill, MA 02467
USA

The IARLD (International Academy for Research in Learning Disabilities) is an international professional organization dedicated to conducting and sharing research about individuals who have learning disabilities.

The IARLD is an elected group of premier scientists, educators and clinicians in the field of learning disabilities throughout the world. The Academy was formed in 1976 by Dr. William Cruickshank (United States of America) and Dr. Jacob Valk (The Netherlands), meeting in Canada with the intention of providing a forum for the exchange of information and the advancement of knowledge regarding learning disabilities.

Since its inception, the Academy has realized its mission of being a professional, international, interdisciplinary consortium of scientists. The Academy currently has a membership of nearly 200 distinguished scholars, representing 26 different countries and thirty disciplines.

IARLD members represent:

- distinguished researchers,
- distinguished practitioner/clinicians,
- young researchers, and
- promising doctoral students.

www.iarld.com

ISSN: 2325-565X (print)
ISSN: 2329-3764 (online)