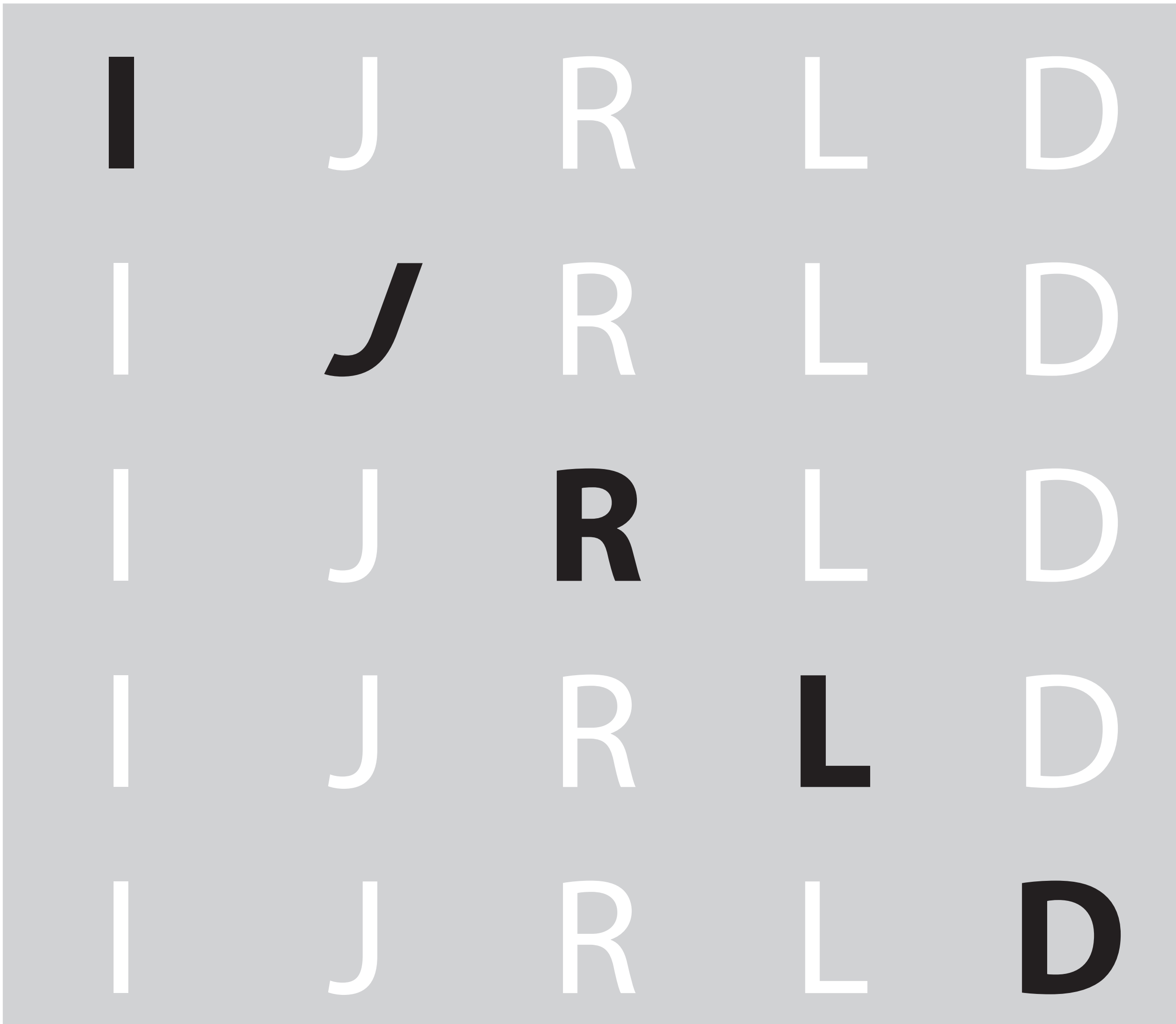
International Journal for Research in Learning Disabilities

volume 4, issue 2, 2020

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 Journal for Research in Learning Disabilities

formerly *Thalamus*

journal of the *International Academy for Research in Learning Disabilities*

David Scanlon, Editor, Boston College, USA

Kirsten McBride, Copy Editor, Kansas City, Missouri, USA

IJRLD

International Journal for Research in Learning Disabilities

Volume 4, No. 2, 2020

Table of Contents

Guest Reviewer

Daniela Lucangeli, *Università de Padova, Italy*

William M. Cruickshank Memorial Lecture Delivered at the 2019 Conference of the International Academy for Research in Learning Disabilities

Extricating Reading Science From Entrenched Anglocentricism, Eurocentricism, and Alphabetism and Embracing Global Diversity: A Personal Journey

David L. Share
Department of Learning Disabilities and Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, Faculty of Education, University of Haifa, Israel

Abstract

The following semi-autobiographical essay tells a cautionary tale about the entrenched Anglocentrism, Eurocentrism, and Alphabetism in reading and reading disabilities (dyslexia) research. Having been born, raised, and educated in an entirely monolingual English-speaking environment, I later migrated to a country where non-European languages (Hebrew and Arabic) were the linguae francae and, furthermore, written in a non-alphabetic script. Over the period of a decade or so, I gradually, and sometimes painfully, came to understand that I needed to revise or discard many of my Anglophone insights into reading and reading disabilities. The culmination of this "awakening" came with the writing of my Anglocentricities critique (Share, 2008), where I argued that the extreme ambiguity of English spelling-sound correspondence had confined reading science to an insular, Anglocentric research agenda addressing theoretical and applied issues with limited relevance for the majority of the world's literacy learners. I subsequently extended this argument to Eurocentricism and Alphabetism – the misguided belief that alphabets are inherently superior to non-alphabetic writing systems (Share, 2014). More recently, together with Peter T. Daniels, a linguist specializing in writing systems, I have been exploring writing system diversity and its implications for learning to read around the world (Daniels & Share, 2018). I hope my story will help raise awareness of the need to move our field a step closer toward embracing global diversity.

Keywords: reading, reading disabilities, learning disabilities, Anglocentrism, Eurocentrism, Alphabetism, writing systems

Share, D. L. (2020). Extricating Reading Science From Entrenched Anglocentricism, Eurocentricism, and Alphabetism and Embracing Global Diversity: A Personal Journey. *International Journal for Research in Learning Disabilities, 4*(2), 3-14. https://doi.org/10.28987/ijrld.4.2.3

Growing up in an Anglocentric Bubble

I was born, raised, and educated in an entirely monolingual English-speaking environment. Although my parents spoke no fewer than 11 languages between them, English was the only common tongue in my family, and in London in the late 1950s and early '60s that was the only language I recall hearing. We moved to Sydney, Australia, when I was six years old, and although the local children made fun of my accent, English, once again, was the only language I heard at home, on the television and radio, around the neighborhood, and at school. There were sizeable Italian- and Greek-speaking communities in Sydney, but that was a distant part of a sprawling metropolis. I have no memories of exposure to any other (spoken) language.

In (Jewish heritage) Sunday school, I learned to decode Hebrew in preparation for my Bar-Mitzvah at age 13, but recitation, not comprehension, was paramount, so I had negligible understanding of what I was decoding. In fact, no one actually spoke Hebrew at the Sunday school. In high school, we briefly studied some "foreign" languages, including Latin, French, and German. I elected to study French and German for matriculation, but the status of "foreign" languages was so low, and the teaching so poor (I don't think our German teacher ever spoke German with us – perhaps he couldn't), that even after some six years of study, I still couldn't speak, read, or write either language at a basic functional level.

Academic Socialization in the Anglocentric Bubble

During my four-year undergraduate studies in psychology at the University of New South Wales, I don't recall ever hearing anything about L2 (second language), English as a Foreign Language (EFL), or bilingualism. My very first "experiment" (looking at the role of phonological recoding in young children's word reading) – my BSc Honors thesis – was carried out in a private Catholic school not far from the university. It never occurred to me to ask if the children (grades 3 and 6) were native English speakers/readers or bilingual, although it's quite possible that the teachers avoided sending me children who were not native English speakers. My PhD research, a large-scale longitudinal study of reading acquisition from kindergarten to grade 2, was undertaken in Geelong, in the state of Victoria – a smallish town of just over 100,000 residents. It never occurred to us that some of our sample of 543 children (Share, Jorm, Maclean, & Matthews, 1984) might not be native-English speakers, perhaps bilingual, or perhaps had learned to read in a script other than English. It is quite possible there were no such outliers in our sample because Geelong, at least in the early 1980s, was considered to be a rural or "country" town, whereas "minorities" speaking "foreign" languages lived in the big cities like Melbourne and Sydney.

The first two years of my three-year post-doc were spent at Otago University in Dunedin, New Zealand, studying the now world-famous Dunedin longitudinal sample. This sample was a birth cohort – all the babies born in the local hospital during 1972-1973 – so all were native-born English-speaking New Zealanders. Again, I don't ever recall the question of bilingualism or additional languages ever arising in our research unit. Unlike the northern parts of New Zealand, I doubt there were any indigenous Maoris in this sample either. I was still in an Anglophone "bubble" studying monolingual English speakers learning to read their native English.

Unsurprisingly, my entire academic socialization/orientation, and all my own research in reading and reading disabilities, was entirely about monolingual English-speakers learning to read their native English. What's more, all our theoretical frameworks were exclusively focused on English. Max Coltheart's dual-route masterpiece (Coltheart, 1978) had burst onto the scene in 1978; the questions that empirical research into reading needed to address were: Were there two routes or one? Was phonology an essential or non-essential component of skilled word recognition? Was phonics or whole language the optimal method for teaching reading? And, back then (in the 1980s), phonemic awareness was fast becoming the hot topic in reading acquisition, culminating in Marilyn Adams' blockbuster, *Beginning to Read* (Adams, 1990).

First Steps Outside the Anglophone Bubble

After completing my post-doc, I migrated to Israel in 1988, taking an academic post at the University of Haifa. This was the Middle East (although a partly Westernized and Americanized version); Israelis spoke non-European Semitic languages – Hebrew and Arabic – and the writing wasn't the familiar Roman alphabet. As an aside, it's worth mentioning that almost all reading research that has addressed bilingualism, bidialectism, biscriptism, and multilingualism has been written by (primarily) monolingual English speakers tackling the challenges of migrants from non-Anglophone countries arriving in English-speaking countries learning to speak and to read English.

I was different, here was an English-speaker arriving in a non-English-speaking country learning to be literate in a non-European language and script. I was embarking on an academic career unaware of the ramifications of these cultural, linguistic, and graphonomic differences (*graphonomy* is the latest name for the study of writing systems; Daniels, 2018). Such was my hubris that I made a bet with my friend and co-author Linda Siegel (the prominent Canadian LD researcher) that, within one year, my spoken Hebrew would be indistinguishably native. I not only lost the bet with Linda, but even now, 30 years later, my children enjoy poking fun at my faulty Hebrew. (It's grammatical gender that poses insurmountable obstacles for English speakers, at least late-learning ones like me!)

A reading researcher needs a solid grounding in the language and the writing system that children are learning to read. I expected this would take a year, perhaps two. I had allowed myself a full six months of intensive Hebrew studies before I commenced teaching (in Hebrew) at Haifa University. Looking back, I realize it took me more than a decade to understand the language and the writing system. From the very outset, colleagues at Haifa University in the reading field as well as Ministry of Education officials were gently reminding me that Hebrew is not quite like English.

One of my first research projects was to determine if phonemic awareness was relevant to Hebrew. I drew up a proposal for an experimental training study and submitted my request for ethical approval to the national supervisor of kindergartens. My proposal was promptly rejected, but the supervisor was considerate enough to invite me to discuss the proposal in person. She did her best to explain that phonological awareness in Hebrew is not the same as in English. I don't recall the specifics of her argument, but I came away confused and frustrated (although she was very obliging to this naïve young newcomer and granted the official approval I needed to carry out my Anglocentric study). I now understand that neither Hebrew nor Arabic phonological awareness (PA) can be regarded as a linear "string" of phonemes in the same way that the letters in the Roman alphabet you are now reading line up neatly in a single row.

Other colleagues, such as Zvia Breznitz (2006) and, later, Tami Katzir were talking about the construct of fluency the "neglected" aspect of reading (Allington, 1983) that was "rediscovered" in the National Reading Panel Report (2000). All our reading measures in both Australia and New Zealand were accuracy measures such as the Burt Word Reading Test. There was one speed/rate measure in the Neale Reading Analysis (Neale, 1999) that assessed text reading errors, passage reading time, and comprehension. But for single-word reading, all our measures were accuracy alone, which made perfect sense for a language in which a list of words ranging from short, high-frequency words to long low-frequency irregular words discriminates all levels of reading ability – from novices through to experts – and where no one reaches ceiling!

In the Deakin longitudinal study (Share et al., 1984), we developed a list of "sight words" – simple high-frequency words – and a list of pseudowords, but recorded only reading accuracy, not rate/time. And, in the course of numerous publications submitted to British, Australian, and U.S. journals (every one, of course, Anglophone), no reviewer questioned our reliance on word reading accuracy alone or asked to see what predicted word reading *rate*. Studies in more regular orthographies, on the other hand, typically rely on measures of reading rate or fluency rather than accuracy for the simple reason that after grade 1, children reach near-ceiling levels of accuracy. Generations of English-speaking dyslexics have been diagnosed with the Wide-Ranging Achievement Test (WRAT) (Jastak & Wilkinson, 1984) and/or the Woodcock Word Attack measures (Woodcock, McGrew, & Mathers, 2000) and the reading disability/dyslexia status determined solely on the basis of word reading accuracy. This is finally changing with the availability of the TOWRE (Torgesen, Rashotte, & Wagner, 1999) and the DIBELS (Kaminsky & Good, 1998).

First Misgivings

At a symposium chaired by the prominent Australian LD researcher Brian Byrne on the topic of universal and language-specific aspects of developmental dyslexia held at the International Congress of Psychology in Stockholm in 2000, I presented some of these "misgivings" in a talk entitled *Confessions of an ex-Anglocentric Reading Researcher*. My presentation was not well received. But undeterred, I eventually wrote up these ideas in my Anglocentricities critique in 2008, arguing that much of the then-current theoretical and applied research in our field was driven by an insular Anglocentric research agenda preoccupied with the challenges posed by the extreme ambiguity of English spelling-sound correspondence. As a consequence, I suggested that the Anglophone literature may have only limited relevance for other languages and orthographies. In the following, I return to the arguments raised a decade ago, updating and refining a number of the points.

The Coltheart/Baron version of dual-route theory. In my 2008 critique, I began by noting that the classic Coltheart/Baron dual-route model (Baron & Strawson, 1976; Coltheart, 1978) was originally a response to the challenges of correctly pronouncing irregular words such as *was, some, yacht,* and *choir* because applying spelling-sound rules (GPCs) (in English) will result in an incorrect ("regularized") pronunciation. Hence the need for a second mechanism or "lexical" route that circumvents reliance on rules.

Many authors have re-interpreted (indeed *mis*-interpreted) dual-route theory in a somewhat less Anglocentric way, proposing (erroneously) that the two routes represent phonological vs. orthographic mechanisms. However, as I clarify in the 2008 paper, this was not Coltheart's concern. Nor was it his intention to draw a distinction between familiar and unfamiliar word strings. Indeed, he clearly and knowingly eschewed the older familiar/unfamiliar distinction, "this is not quite the right distinction, at least for English" (Coltheart, 2005, p. 205).

As noted by Ziegler and Goswami (2005), many, indeed most, alphabetic writing systems are highly regular (at least from spelling to sound [e.g., German or Spanish]), and have few irregular spellings, so who needs a second route? The focus on regular spellings (pronounceable via rules) as opposed to irregular spellings (inaccessible to the rules) diverted attention away from the original familiar vs. unfamiliar distinction which, I maintain, is a crucial and overarching dualism relevant to all words in all possible orthographies. This is the developmental, item-based unfamiliar-to-familiar dualism that was briefly outlined in Share 2008 and further elaborated in Share [2019]; it draws a distinction between a slower, effortful, serial, letter-by-letter reading process in the case of unfamiliar words and a rapid, effortless, holistic, one-step process for reading familiar words. Furthermore, there is evidence that this dualism is a universal property of learning to read (Share, 2019).

Preoccupation with accuracy and neglect of rate and fluency. The fact that accurate pronunciation of "irregular" spellings permeates all levels of the English lexicon may explain the preoccupation with (oral) reading accuracy and the relative neglect by researchers and educators alike of the question of reading rate and fluency (at least until the National Reading Panel Report, 2000). Even Wolf and Bowers' (1999) double-deficit model of dyslexia (which foregrounded reading fluency and not just reading accuracy) was probably on the sidelines of dyslexia research owing to the Anglocentric preoccupation with accuracy. As discussed below, the focus on accuracy profoundly influenced definitions of reading disability/dyslexia (Lyon et al., 2005). As I noted above, the only word reading tests I was aware of in my academic socialization were accuracy measures, such as the Burt, WRAT, and Woodcock. It was only the recent Anglicization of European fluency measures such as the Dutch One-Minute Test (Brus & Voeten, 1973), reincarnated as the TOWRE, that finally brought reading fluency into the classroom and clinic, and still our field debates the nature of fluency and automaticity. So it appears that the overwhelming problems of reading accuracy in English helped shackle the study of fluency and its operationalization.

It is now common knowledge among reading researchers that in regular orthographies, accuracy reaches ceiling performance early, and reading speed/fluency becomes the critical measure for assessing individual and developmental differences. The case of adult English-speaking dyslexics reinforces this picture. Like young German-speaking dyslexics, English adult dyslexics make relatively few accuracy errors, but a slow reading *rate* is the essential sign of dyslexia, and perhaps also late-acquired L2 English. My own late-acquired L2 Hebrew reading accuracy is fine, but I read a single page of Hebrew in about the same time it takes a native speaker to read three pages.

The special problems of pronouncing irregular English spellings may also have biased our field toward oral reading measures at the expense of (rapid) silent reading, which, of course, is the true benchmark of skilled word and text reading. I must admit that there are other (historical and pedagogical) reasons for the bias toward oral as opposed to silent reading, but it seems fair to say that English spelling irregularity did not help to remedy the oral reading bias.

Phonological awareness: Rethinking the linear Anglocentric/alphabetic "string-of-beads" view. It is now universally accepted that awareness of the phonological units in speech that are represented in the written symbols (whether phonemes, sub-syllabic units, or whole syllables) is crucial for learning to read, although the strength and timing of the reading-PA connection may vary across orthographies. However, the Anglophone literature has promoted a very particular notion of PA, one that I call a linear, "alphabetic" "string-of-beads" view. Much like the letters in English spellings, neatly sitting in a row like a string of beads, the Anglocentric view of phonological awareness depicts each of the sounds in a spoken word (phonemes) as strung along a single (acoustic) axis, thereby promoting the view of PA as awareness of a (linear) string of (discrete) phonemes. But many writing systems exhibit substantial non-linearity (Daniels & Share, 2018), or rather multilinearity with multiple axes. Even in many European alphabets (e.g., French and German) and non-European alphabets such as Vietnamese, extra-lineal diacritics are common.

Beyond Europe, non-linearity is a fundamental feature of most non-alphabetic scripts such as Semitic abjads – Arabic and Hebrew, Brahmi-derived Indic abugidas (also called *akshara* scripts or alphasyllabaries), and Chinese characters, which are found not only in Chinese, but also in Korean and Japanese. Phonological awareness in these scripts both reflects and reinforces a non-linear conceptualization of PA. For example, Hebrew and Arabic have a very different phonological structure (primarily "core [CV] syllables" plus consonantal phonemes) that calls for a different non-alphabetic (and non-linear) approach to the question of PA and instruction.

The late Professor Dina Feitelson, a distinguished Israeli reading teacher and researcher in my department at Haifa, vehemently asserted that these core syllables (called *tserufim* in Hebrew) should be taught to beginning readers in first grade. According to Feitelson, the combinations of consonant and appended vowel signs, which form an integral unit in the *vertical* (not horizontal) dimension, were an *indivisible* unit, hence teaching children to divide the word *bag* into three units, /b/ /a/ /g/, made no sense pedagogically in Hebrew and (by extension) Arabic.

Similarly, Buckley (2018) makes a strong case that not only Hebrew and Arabic but almost all so-called syllabic or moraic scripts represent core syllables. In the Brahmi-derived Indic abugidas of South and Southeast Asia or askhara scripts, syllabic awareness, not phoneme awareness, is crucial in the initial years of learning to read; only later, around grade 5 does awareness of individual phonemes become a necessity (Nag, 2017; Share & Daniels, 2014). As for Chinese, the initial/final (C/V, C/VV/, CVC) subdivision of the Chinese syllable (not to be confused with the controversial onset-rime notion debated in the Anglophone literature) (Goswami & Mead, 1992; Nation & Hulme, 1997) is also at odds with the linear string-of-beads view of PA.

The definition and diagnosis of reading disability and the wait-to-fail model. Definitions of dyslexia in the English-speaking world have traditionally been operationalized with measures of accuracy, specifically oral reading accuracy (see, e.g., Siegel, 1999; Stanovich, 1999). Writing on this topic, Lyon and colleagues (2005) noted,

> [R]eading fluency is rarely assessed in current identification procedures … Slow-reading students who score within the average range on both the untimed reading measures and the IQ/test will typically be denied services because there is no discrepancy – even though they also have a disability that requires specialized services and/or accommodations. (p. 267)

It was only around the turn of the 21st century that learning disabilities (LD) researchers recognized that reading fluency and not just accuracy was a crucial aspect of reading disability (RD) and began to incorporate fluency into definitions of RD/dyslexia (American Psychiatric Association, 2013; British Psychological Association, 1999; Lyon, Shaywitz, & Shaywitz, 2003).

It has long been considered common knowledge (perhaps myth) that it takes three years to learn to read English before children can read to learn (Chall, 1983; Singer, 1978). We now know that in "normal" alphabetic orthographies, it takes only a single year to master the basics of decoding (Seymour, Aro, & Erskine, 2003). The same applies to the highly regular pointed Hebrew and *mashkul* Arabic (Saiegh-Haddad, 2018; Share, 2017).

The unusually prolonged phase of learning to read in English was institutionalized in the British and Australian systems in the form of the infants school – kindergarten to grade 2. If the child was deemed "ready," s/he graduated to the primary school – grades 3 to 6. Accordingly, grade 3 was considered the appropriate time to establish remedial reading centers on the premise that until that point a child was still learning to read, so earlier identification of a reading difficulty would be premature. I have wondered whether the traditional reluctance to "intervene" earlier than grade 3 might be attributable to the well-entrenched belief in the three-year learning-to-read phase. The IQ-reading discrepancies and two-year gaps required to diagnose an RD all seem to match this model, hence the wait-to-fail model (Lyon et al., 2003) may be yet another casualty of the complex spelling-sound system of English.

Reading development: Stages and phases. Models of word reading acquisition developed by English language researchers almost invariably include one or more phases, or stages, in which the novice reader is unable to exploit all the grapheme-phoneme information available in a printed word, relying instead either on partial letter-sound cues (e.g., the initial letter or initial and final letters, often in conjunction with contextual cues) or on purely global visual information, such as word length and envelope, or salient visual (non-phonological) features of selected letters. The best-known terms for these stages or strategies are *partial alphabetic* (Ehri, 1995) and *logographic* (Frith, 1985). These phenomena appear to be a product of an unusually protracted period of early reading development jointly attributable to encouragement of early literacy during the preschool years followed by a prolonged period of code learning. In many more regular orthographies, these phenomena appear to be far less prevalent (see, e.g., Landerl, 2000; Wimmer & Hummer, 1990).

Reading instruction. Much of the debate on methods of reading instruction has centered on the "villain" of spelling-sound irregularity (Snow & Juel, 2005). The complexities of English spelling and the difficulties of decoding have been prominent in whole language theories of reading (Smith, 1978). In regular alphabets, phonics instruction is typically the dominant approach with less or no emphasis on "preparatory," "early," or "emergent" literacy activities such as letter knowledge (see Share, 2008, pp. 601-603).

Has the Field Moved Forward?

The Anglocentricities paper (Share, 2008) was conceived two decades ago, and the published version is now over a decade old. Is our field less Anglocentric today? Yes and no. There has unquestionably been important progress toward developing a more wide-angled perspective on reading and RD. A growing proportion of theoretical and empirical work is addressing basic issues in languages and orthographies other than English or Western European languages (e.g., Joshi & McBride, 2019; Saiegh-Haddad & Joshi, 2014; Verhoeven & Perfetti, 2017; Verhoeven, Pugh, & Perfetti, 2019).

However, the theoretical and applied frameworks developed for English are still all too often generalized to other languages and writing systems without due consideration for linguistic and writing system diversity. Almost all publications by English-language researchers continue to omit any "... in English" qualification in the titles of their papers – *A New Whiz-Bang+++ Model of Learning to Read ... in English*? – as if the results of studies conducted in English alone enjoy the privileged status of universal applicability, unlike researchers investigating other languages, who are obliged to qualify their findings by adding the ... *in Chinese/Arabic/Korean,* etc., disclaimer that automatically demarcates the findings as language-specific and hence not necessarily universally applicable. But Anglocentrism is not the only form of ethnocentrism in our field. There is another obstacle to progress, which I call Alphabetism, itself an expression of Eurocentrism.

Eurocentrism and Alphabetism

When reading researchers today seek enlightenment on the subject of writing systems, they look to Gelb (1952), the founding father of the field of grammatology, now relabeled *graphonomy*. Like Taylor (1883) before him, Gelb (1952) was a proponent of an evolutionary view of writing systems history, from "primitive" pre-alphabetic systems to the venerated alphabet. Consistent with the "ontogeny recapitulates phylogeny" idea, Gelb's inexorable "three great steps [logographic-to-syllabic-to-alphabetic] by which writing evolved from the primitive stages to a full alphabet" (p. 203) was embraced by almost all reading researchers, despite its repudiation by recent scholarship in the field of writing systems research (Coulmas, 2009; Daniels, 1992, 2018; Mattingly, 1985; Olson, 1994; Rogers, 2005).

It needs to be pointed out, however, that the "culture" of Alphabetism, like culture in general, is often invisible; that is, its presence is more often discernible in acts of omission than commission. Nonetheless, this alphabetic bias is ubiquitous, and is manifest in the following:

1. Unqualified generalizations about reading across languages and/or across orthographies in papers that refer almost exclusively to English or to European alphabets (see, e.g., Caravolas et al., 2013; Ziegler et al., 2010; Ziegler & Goswami, 2005).

2. Implicit or explicit acceptance of Gelb's evolutionary theory in leading texts on reading development aimed at educators,

   > Taking the final step toward the creation of a true alphabetic writing system, the Greeks assigned a symbol to each consonant and vowel of their language … In many ways, the individual development of the children who are discovering the alphabetic principle in English writing recapitulates human history. (Moats, 2000, pp. 82-83)

3. Continued reiteration of Gelb's views in even the most up-to-date and authoritative texts on the psychology of reading (e.g., Dehaene, 2009; Rayner, Pollatsek, Ashby, & Clifton, 2012).

> [I]n an evolutionary sense, the alphabet is the"fittest ..." (Rayner et al, 2012, p. 37)
> The history of writing suggests a clear evolutionary trend ...These systems evolved to a logographic system, which in turn evolved to syllabic systems and finally to alphabetic systems ... Such an evolutionary argument suggests that alphabets are fitter (in the Darwinian sense). (Rayner et al., 2012, pp. 46-47)

4. Reference to non-alphabetic systems as imperfect or defective (e.g., Hannas, 2003; Rayner et al., 2012).

   > The Semitic writing systems ... and the languages of India still incompletely represent vowels (Rayner et al., 2012, p. 36) ... In this sense, many of these scripts are not fully alphabetic. (Rayner et al., 2012, p. 37
   > The Phoenician system, however, was not perfect. It failed to represent all vowels ... It was the Greeks who finally created the alphabet as we know it … For the first time in the history of mankind, the alphabet allowed the Greeks to have a complete graphic inventory of their language sounds. (Dehaene, 2009, p. 193)

   Moving slightly further afield, consider what one eminent sociolinguist wrote about the marvels of alphabetic writing.

   > The basic difference between Western alphabetic and East Asian syllabic writing acts on several levels to promote or inhibit creativity, particularly that associated with breakthroughs in science … syllabic literacy entails a diminished propensity for abstract and analytical thought … Certain Asian characteristics credited with blocking creativity, such as conservative political and social institutions and group-oriented behavior, derive in part from effects that the orthography has had on the minds of individuals. (Hannas, 2003, p. 203)

5. The use of alphabetic terminology (e.g., *letters, graphemes*) to describe and label the functional architecture (and even the anatomical brain structures) of reading (letter detectors, letterbox area, universal letter shapes; Dehaene, 2009) purported to be universal in reading. Whereas the concept of a letter (or grapheme) is widely used (but not entirely unproblematic) in European alphabets, it has questionable applicability to many writing systems, including Chinese characters, Japanese Kanji, Brahmi-derived Indic aksharas or Mayan glyphs. Even the notion of the phoneme as the fundamental unit of analysis of speech may be an artifact of West European alphabetic literacy (Daniels, 2018).

As a postscript to this section on Eurocentrism and Alphabetism, it is worth keeping in mind that the alphabet, a uniquely European creation, was first disseminated throughout Europe with the spread of Christianity, then across the globe by European colonizers, traders, and, above all, missionaries, who never thought to question whether their own writing systems would be optimal for non-European languages. They took it for granted that the ideal orthography was alphabetic, operating on the principle of one letter for one sound (phoneme) for both consonants and vowels under the motto "consonants as in English, vowels as in Italian" (Gleason, 1996).

But are alphabets optimal? The evidence is still too scarce to make a call, but there are several lines of counter-evidence (reviewed in Share, 2014) converging on the conclusion that syllable-based writing systems are, in many cases, superior to alphabets. And we can hardly ignore the recent 2018 PISA (Programme for International Student Assessment; Organisation of Economic Co-operation and Development [OECD], 2019) results showing that the Chinese topped the international league tables in reading comprehension.

Embracing Global Diversity

Some years ago, one of my PhD students brought to my attention a volume on *The World's Writing Systems* (Daniels & Bright, 1996). It was a revelation for me: I couldn't put it down – all 900 pages! It remains the world's most comprehensive and authoritative exposition of the world's writing systems – a work that is surprisingly un-encyclopedic in the way it weaves a coherent narrative throughout a 900-page mosaic. Mid-way through reading this volume, I got an email from none other than Peter Daniels asking if I would be interested in writing a chapter on literacy across orthographies for a revision of *The World's Writing Systems.* Daniels opened the door to an entire world of writing systems – ancient, modern, exotic. I came to understand that any theory about how children learn to read and write requires a deep understanding not only of the language, but of the particular writing system.

Learning about the amazing diversity of writing systems (compare Ogham, Mayan, and Arabic), I came to appreciate that a majority of children around the globe do not learn to read in European languages and alphabetic scripts. And, furthermore, only a small minority are monolingual speakers learning to read their native tongue.

Dissatisfied with the one-dimensional approaches to orthographic variety (*Orthographic Depth*, Katz & Feldman, 1981; Katz, & Frost, 1992; Frost, 2005; and *Psycholinguistic Grain-Size Theory,* Ziegler & Goswami, 2005), which dwell almost exclusively on spelling-sound irregularity and ignore many other important dimensions of orthographic complexity, Daniels and I began exploring the implications of writing system variation for theories of learning to read and dyslexia. Building on a set of five dimensions originally presented by Daniels at a workshop in Haifa in 2012, we expanded this set to 10 dimensions of orthographic complexity, summarized below, each of which is liable to create obstacles for the learner.

1. *Linguistic distance*

   Possibly the most potent factor in learning to read across the globe, *linguistic distance,* refers to the situation in which the language or dialect spoken by the literacy learner is not the same as the language of written texts. The dimension of linguistic distance may be regarded as a continuum of varying degrees of spoken/written divergence. This ranges from the extreme case of speakers of one language learning to read in a totally different language (such as Arabic speakers learning to read in English or French), through diglossic situations, in which two distinct varieties of the same language exist side by side in a single speech community but are used for different purposes. Alongside a low-prestige everyday conversational dialect (or dialects), there exists a high-prestige, typically grammatically more complex variety used in formal spoken settings and, in the case of literate communities, written settings. Dialect variation (such as African-American Vernacular English [AAVE]) represents yet another case of linguistic distance.

   The evidence is overwhelming that linguistic distance has a profoundly detrimental impact on learning to read (August, Shanahan, & Escamilla, 2009; Gatlin & Wanzek, 2015; Myhill, 2009; Saiegh-Haddad & Schiff, 2016).

2. *Non-linearity*

   In many scripts, symbols are not arrayed along a single axis. Extra-lineal diacritics are found in many European alphabets, but in non-alphabetic scripts, symbols commonly appear above, below, within, or even surrounding, other symbols. We know very little about the implications of different forms of non-linearity. In Devanagari, for example, a noninitial /i/ is written *before* the consonant after which it is pronounced, and this non-linearity appears to create problems for the learner (Kandahai & Sproat, 2010). In some scripts, characters are nested in syllabic units such as Indic aksharas and Korean syllable blocks, which may facilitate reading acquisition (at least initially) by obviating the need to access phonemes.

   Chinese semantic-phonetic character compounding typically positions the semantic component to the left of the phonetic, but this component can also appear to the right, above, below, or surrounding the phonetic. Learning the many positional regularities (and exceptions) of a script would be expected to tax visual-spatial skills.

3. *Visual confusability*

   The shapes of individual letters and characters in scripts – their ductus – represent another neglected aspect of orthography that has been shown to impact symbol discriminability, learnability, and processing speed. Visual complexity has been shown to contribute significantly to variation in symbol learning among beginning readers (Nag, Snowling, Quinlan, & Hulme, 2014) and to processing speed among skilled adult readers (Chang et al., 2018; Pelli et al., 2014).

4. *Historical change: Retention of historical spellings despite pronunciation change*

   Any living language is constantly changing, but orthographies are altered or reformed only occasionally, if at all. If not realigned, pronunciation will drift further and further away from spelling over time. (This is a prime cause of what Liberman et al. [1980] labeled *orthographic depth*, as well as of Ziegler and Goswami's *inconsistency* and Coltheart's dual-route *irregularity*.) This category includes two subtypes: words spelled the same but pronounced differently (*bough/cough/dough/through/tough*, which once all rhymed), as well as words spelled differently but pronounced the same (*meet/meat/mete*). Even many "shallow" scripts often have multiple options for spelling sounds (e.g., Spanish). In most orthographies, sound-to-spelling consistency is typically more complex than spelling-to-sound consistency, with the result that spelling is typically more challenging than reading.

5. *Spelling constancy despite morphophonemic alternation*

   In many writing systems, another source of spelling-sound and sound-spelling inconsistency is morphophonemic: namely, the orthography does not change when either morphemes or phonemes undergo conditioned alternations: /haws/ *(house)* becomes /hawz/ when the plural suffix is added, but the spelling does not change *(houses)*.

6. *Omission of phonological elements*
   In most writing systems in which phoneme-level information is explicit or potentially extractible, consonant representation often takes precedence over vowel representation (consider English abbreviations and text messaging), probably because consonants are less predictable and convey more information about lexical identity (see Adams, 1990; Gnanadesikan, 2017; Shimron, 1993). In Brahmi-derived Indic abugidas, the most common (default) vowel is not explicitly marked in post-initial positions but is inherent in all simple consonant aksharas. In Hebrew and Arabic, many or most vowels are normally not written, creating extensive homography. And in many languages, stress is not marked *contráct* (v.)/*cóntract* (n.). For example, the four accent marks that designate the four tones of Mandarin Chinese in pinyin romanization are hardly ever written, yet tonal awareness has been shown to differentiate good and poor readers in mainland China (Ding, Liu, McBride, & Zhang, 2015). In many African tone languages, especially those in which tone carries a high functional load, good and poor readers alike struggle with the lack of tone marking in the Roman orthographies imposed by the Europeans (Coulmas, 1989; Roberts, Walter, & Snider, 2016).

7. *Allography*
   Many upper- and lowercase letter pairs in English are identical (Ss, or near-identical Ff), but others (Ee, Bb, Gg) seem unrelated. In Indic scripts, initial forms of vowels can differ considerably from post-initial forms (*matras*). Similarly, many Indic consonant ligatures (see Dimension 9) may not all be readily recognizable as allographs of the full, unreduced radical. For example, in Chinese, many of the some 200+ semantic radicals assume reduced forms when juxtaposed in semantic-phonetic compounds. In Arabic, it is commonly assumed that many letters alter their form depending on their position in a word (initial, medial, final, and separate). For many sets of allographs, the common letter form is easily discerned, for others the similarity is harder to discern, and for some, the letters seem entirely unrelated.

8. *Dual-purpose letters*
   In Semitic scripts some letters serve multiple purposes, representing either a vowel or consonant. For example, the letter H in English serves as both a consonant and a diacritic in the digraphs (th, sh, ph, wh, ch).

9. *Ligaturing*
   Ligaturing, the joining of letters, is often considered a source of difficulty in a number of scripts (e.g., Arabic and Indic scripts). Whereas in Arabic, ligaturing usually involves letter shapes augmented by the addition of connecting strokes, ligaturing may also involve diminished or reduced forms, as in Indic scripts. Consonants can be ligatured horizontally, or vertically, sometimes with identifiable components but sometimes with unpredictable modifications.

10. *Inventory size*
    The difference between scripts in the total number of letters or characters has been termed by Nag (2007) the *contained-extensive* dimension of writing system variation (also *orthographic breadth*). The number of characters in phonemic writing systems (alphabets and abjads) are typically in the region of a few dozen, but if Indic consonantal compounds or the Korean syllable blocks are taught and learned (at least initially) as undivided wholes, then they number in the hundreds. The many thousands of Chinese characters reduce to some 200 radicals/semantic components and 700–800 phonetic components.

Summary

This multidimensional framework is no more than a starting point for discussion of writing system diversity. Other dimensions may need to be added, or perhaps some can be coalesced. It might be possible to reduce all 10 dimensions to a three-dimensional framework based on (a) visual-orthographic factors, (b) phonology, and (c) morphology/meaning. This scheme also suggests that any one script's complexity may be a unique combination of some or all of the 10 dimensions above. This explains why a script (e.g., Arabic) might have near-perfect spelling-sound consistency – a presumably "shallow" or "transparent" script – yet be complex for very different reasons. At the very least, the proposed 10 dimensions make it clear that a one-dimensional (one-size-fits-all) approach is untenable once we move beyond European alphabets.

Concluding Remarks

In the age of globalization and the internet, falling borders and barriers have revealed a remarkable panoply of cultural, linguistic, and script diversity. Western monolingual Anglophone reading science, which has dominated reading research until recently, is now waking up to the fact that most of the world does not speak English or any other European language, is not monolingual, and, furthermore, does not become literate in an alphabetic script. Nor is there compelling evidence that this uniquely European creation – the alphabet – is inherently superior to other non-alphabetic writing systems. Western reading science has made undeniably significant advances, but we cannot indiscriminately generalize our theories and findings to other languages and writing systems. Current one-dimensional frameworks for conceptualizing the challenges of learning to read across languages and orthographies (e.g., *orthographic depth* and *psycholinguistic grain-size theory*) need rethinking. Consideration of the full spectrum of the world's writing systems reveals multiple dimensions of orthographic complexity, each liable to create obstacles for children learning to read and write. The time has come to extricate reading science from entrenched Anglocentricism, Eurocentricism and Alphabetism and embrace global diversity.

References

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

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5).* Arlington, VA: Author.

August, D., Shanahan, T., & Escamilla, K. (2009). English language learners: Developing literacy in second-language learners – Report of the National Literacy Panel on Language-Minority Children and Youth. *Journal of Literacy Research, 41*(4), 432-452.

Baron, J., & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 2*(3), 386.

British Psychological Society. Division of Educational and Child Psychology. (1999). *Dyslexia, literacy and psychological assessment.* Leicester, UK: Author.

Buckley, E. (2018). Core syllables vs. moraic writing. *Written Language & Literacy, 21*(1), 26-51.

Brus, B. T., & Voeten, M. J. M. (1973). *Een-minuut-test: vorm A en B; schoolvorderingentest voor de technische leesvaardigheid, bestemd voor het tweede tot en met het zesde leerjaar van het basisonderwijs; verantwoording en handleiding*. Amsterdam, The Netherlands: Berkhout.

Caravolas, M. (2013). Learning to spell in different languages: How orthographic variables might affect early literacy. In P. G. Aaron (Ed.), *Handbook of orthography and literacy* (pp. 511-526). London, UK: Routledge.

Chall, J. S. (1983). *Stages of reading development.* New York, NY: McGraw-Hill.

Chang, L.-Y., Plaut, D. C., & Perfetti, C. A. (2016). Visual complexity in orthographic learning: Modeling learning across writing system variations. *Scientific Studies of Reading, 20*, 64-85.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). London, UK: Academic Press.

Coltheart, M. (2005). Modeling reading: The dual-route approach. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 6-23). Oxford, UK: Blackwell.

Coulmas, F. (1989). *The writing systems of the world*. Oxford, UK: Blackwell.

Daniels, P. T. (1992). The syllabic origin of writing and the segmental origin of the alphabet. In P. A. Downing, S. D. Lima, M. Noonan (Eds.), *The linguistics of literacy* (pp. 83-110). Amsterdam, The Netherlands: John Benjamins Publishing.

Daniels, P. T. (2018). *An exploration of writing*. London, UK: Equinox Publishing Limited.

Daniels, P. T., & Bright, W. (Eds.). (1996). *The world's writing systems.* New York, NY: Oxford University Press.

Daniels, P. T., & Share, D. L. (2018). Writing system variation and its consequences for reading and dyslexia. *Scientific Studies of Reading, 22*(1), 101-116.

Dehaene, S. (2009). *Reading in the brain.* London, UK: Penguin.

Ding, Y., Liu, R.-D., McBride, C., & Zhang, D. (2015). Pinyin invented spelling in Mandarin Chinese-speaking children with and without reading difficulties. *Journal of Learning Disabilities, 48,* 635-645.

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18,* 116-125.

Frith, U. (1985). Beneath the surface of developmental dyslexia. *Surface Dyslexia, 32,* 301-330.

Frost, R. (2005). Orthographic systems and skilled word recognition processes in reading. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 272–295). Oxford, UK: Blackwell.

Gatlin, B., & Wanzek, J. (2015). Relations among children's use of dialect and literacy skills: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 58,* 1306-1318.

Gelb, I. J. (1952). *A study of writing: The foundations of grammatology.* New York, NY: Taylor & Francis.

Gleason, A. (1996). Christian missionary activities. In P. T. Daniels & W. Bright (Eds.), *The world's writing systems* (pp. 777-780). New York, NY: Oxford University Press).

Gnanadesikan, A. E. (2017). Towards a typology of phonemic scripts. *Writing Systems Research, 9,* 1-22.

Goswami, U., & Mead, F. (1992). Onset and rime awareness and analogies in reading. *Reading Research Quarterly,* 153-162.

Hannas, W. C. (2003). *The writing on the wall: How Asian orthography curbs creativity.* Philadelphia, PA: University of Pennsylvania Press.

Jastak, S., & Wilkinson, G. S. (1984). *The wide range achievement test-revised.* Wilmington, DE: Jastak Associates.

Joshi, R. M., & McBride, C. (2019). *Handbook of literacy in Akshara orthography* (pp. 3-9). New York, NY: Springer, Cham.

Kaminsky, R. A., & Good III, R. H. (1998). Assessing early literacy skills in a problem-solving model: Dynamic Indicators of Basic Early Literacy Skills. In M. R. Shinn (Ed.), *The Guilford school practitioner series. Advanced applications of Curriculum-Based Measurement* (p. 113–142). New York, NY: The Guilford Press.

Kandhadai, P., & Sproat, R. (2010). Impact of spatial ordering of graphemes in alphasyllabic scripts on phonemic awareness in Indic languages. *Writing Systems Research, 2,* 105-116.

Katz, L., & Feldman, L. B. (1981). Linguistic coding in word recognition: Comparisons between a deep and a shallow orthography. In A. M. Lesgold & C. A. Perfetti (Eds.), *Linguistic coding in word recognition, interactive processes in reading* (pp. 85-106). Hillsdale, NJ: Erlbaum.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. *Advances in Psychology, 94,* 67-84.

Landerl, K. (2000). Influences of orthographic consistency and reading instruction on the development of nonword reading skills. *European Journal of Psychology of Education, 15*(3), 239.

Liberman, I.Y., Liberman, A. M., Mattingly, I. G., & Shankweiler, D. L. (1980). Orthography and the beginning reader. In J. F. Kavanagh & R. L. Venezky (Eds.), *Orthography, reading, and dyslexia* (pp. 137-153). Baltimore, MD: University Park Press.

Lyon, G. R., Fletcher, J. M., Shaywitz, S. E., Shaywitz, B. A., Torgesen, J. K., Wood, F. B., et al. (2005). Rethinking learning disabilities. In C. E. Finn, Jr., A. J. Rotherham, & C. R. Hokanson, Jr. (Eds.), *Rethinking special education for a new century* (pp. 259–287). Washington, DC: Thomas B. Fordham Foundation, Progressive Policy Institute.

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

Mattingly, I. G. (1985). Did orthographies evolve? *Remedial and Special Education*, *6*(6),18-23.

Moats, L. C. (2000). *Speech to print: Language essentials for teachers*. Baltimore, MD: Paul H. Brookes Publishing Company.

Myhill, J. (2009). *Towards an understanding of the relationship between diglossia and literacy: A survey commissioned by the language and literacy committee*. Haifa, Israel: University of Haifa.

Nag, S. (2017). Learning to read alphasyllabaries. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of reading development* (pp. 5-98). Amsterdam, The Netherlands: John Benjamins Publishing.

Nag, S., Snowling, M., Quinlan, P., & Hulme, C. (2014). Child and symbol factors in learning to read a visually complex writing system. *Scientific Studies of Reading, 18*(5), 309-324.

Nation, K., & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. *Reading Research Quarterly, 32*(2), 154-167.

Neale, M. D. (1999). *Neale analysis of reading ability: Reader.* Melbourne, Australia: ACER Press, Australian Council for Educational Research Limited.

Olson, D. R. (1994). *The world on paper.* Cambridge, UK: Cambridge University Press.

Organisation of Economic Co-operation and Development (OECD). (2019). *PISA 2018 results (Volume I). What students know and can do.* Geneva, Switzerland: Author.

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research, 46*(28), 4646-4674.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr, C. (2012). *Psychology of reading.* New York, NY: Psychology Press.

Roberts, D., Walter, S. L., & Snider, K. (2016). Neither deep nor shallow: a classroom experiment testing the orthographic depth of tone marking in Kabiye (Togo). *Language and Speech, 59*(1), 113-138.

Rogers, H. (2005). *Writing systems: A linguistic approach.* Oxford, UK: Blackwell.

Saiegh-/Haddad, E. (2018). MAWRID: A model of Arabic word reading in development. *Journal of Learning Disabilities, 51*(5), 454-462.

Saiegh-Haddad, E., & Joshi, R. M. (Eds.). (2014). *Handbook of Arabic literacy: Insights and perspectives* (Vol. 9). New York, NY: Springer Science & Business Media.

Saiegh-Haddad, E. (2017). Learning to read Arabic. In L. Verhoeven & C. A. Perfetti (Eds.), *Reading acquisition: Cross-linguistic and cross-script perspectives* (pp. 127-154). Cambridge, UK: Cambridge University Press.

Saiegh-Haddad, E., & Joshi, R. M. (Eds.). (2014). *Handbook of Arabic literacy: Insights and perspectives* (Vol. 9). New York, NY: Springer Science & Business Media.

Saiegh-Haddad, E., & Schiff, R. (2016). The impact of diglossia on voweled and unvoweled word reading in Arabic: A developmental study from childhood to adolescence. *Scientific Studies of Reading, 20*, 311–324.

Seymour, P. H., Aro, M., Erskine, J. M., & Collaboration with COST Action A8 Network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143-174.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of over-reliance on an "outlier" orthography. *Psychological Bulletin, 134,* 584-616.

Share, D. L. (2014). Alphabetism in reading science. *Frontiers of Psychology, 5*, 1-3.

Share, D. L. (2017). Learning to read Hebrew. In L. Verhoeven & C. A. Perfetti (Eds.), *Reading acquisition: Cross-linguistic and cross-script perspectives* (pp. 155-180). Cambridge, UK: Cambridge University Press.

Share, D. L. (2019). *Blueprint for a universal model of learning to read; The Combinatorial model*. Manuscript submitted for publication.

Share, D. L., & Daniels, P. T. (2016). Aksharas, alphasyllabaries, abugidas, alphabets and orthographic depth: Reflections on Rimzhim, Katz and Fowler (2014). *Writing Systems Research*, *8*(1), 17-31.

Share, D. L., Jorm, A. F., Maclean, R., & Matthews, R. (1984). Sources of individual differences in reading acquisition. *Journal of Educational Psychology*, *76*(6), 1309.

Shimron, J. (1993). The role of vowels in reading: A review of studies of English and Hebrew. *Psychological Bulletin, 114*, 52–67

Siegel, L. S. (1999). Issues in the definition and diagnosis of learning disabilities: A perspective on Guckenberger v. Boston University. *Journal of Learning Disabilities*, *32*(4), 304-319.

Singer, H. (1978). Developmental changes in reading instruction: From learning to read to learning from text. *The Florida Reading Quarterly*, *14*(2), 10-13.

Smith, F. (1978). *Understanding reading: A psycholinguistic analysis of reading and learning to read* (2nd ed.). New York, NY: Holt, Rinehart, and Winston.

Snow, C. E., & Juel, C. (2005). Teaching children to read: What do we know about how to do it? In M. J. Snowling & C. Hulme (Eds.), *Blackwell handbooks of developmental psychology. The science of reading: A handbook* (p. 501–520). Oxford, UK: Blackwell Publishing.

Stanovich, K. E. (1999). The sociopsychometrics of learning disabilities. *Journal of Learning Disabilities*, *32*(4), 350-361.

Taylor, I. (1883). *The alphabet: an account of the origin and development of letters* (Vol. 2). London, UK: K. Paul, Trench & Company.

Torgesen, J. K., Rashotte, C. A., & Wagner, R. K. (1999). *TOWRE: Test of word reading efficiency*. Toronto, ONT, Canada: Psychological Corporation.

Verhoeven, L., & Perfetti, C. (Eds.). (2017). *Learning to read across languages and writing systems*. Cambridge, UK: Cambridge University Press.

Verhoeven, L., Perfetti, C., & Pugh, K. (Eds.). (2019). *Developmental dyslexia across languages and writing systems*. Cambridge, UK: Cambridge University Press.

Wimmer, H., & Hummer, P. (1990). How German-speaking first graders read and spell: Doubts on the importance of the logographic stage. *Applied Psycholinguistics*, *11*(4), 349-368.

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, *91*(3), 415.

Woodcock, R. W., McGrew, K., & Mathers, N. (2000). *Woodcock–Johnson–III. Tests of Achievement: Norm tables*. New York, NY: Riverside Publishing.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3.

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., ... & Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science*, *21*(4), 551-559.

Academic and Cognitive Remediation for Students With Learning Disabilities: A Comparison Between Orton Gillingham and NILD Educational Therapy

Josephine A. Stebbings and Erin Kline
Fraser Academy

Abstract

This study compared the effectiveness of Orton Gillingham (OG) tutoring and National Institute for Learning Development Educational Therapy (NILD). A randomized controlled trial using 27 participants determined whether academic and cognitive outcomes differed between the groups over a 9-month period. Participants had designated learning disabilities (LD) in reading (RD), written expression (WD), and/or math (MD), and received 64 hours of one-to-one remediation. Pre and post measures were conducted using WJ IV Cognitive and WJ IV Achievement assessments. Results indicated that the standard score median increased in both groups for Fluid Reasoning, Reading Comprehension, and Word Attack with Fluid Reasoning and Reading Comprehension increasing more dramatically for the NILD group and Word Attack increasing more dramatically for the OG group. Outcomes are congruent with the instructional focus of each approach thus supporting the efficacy of both approaches which currently lack robust empirical evidence. This preliminary research provides important contributions for researchers and practitioners considering interdisciplinary remedial options for students with LD.

Keywords: learning disabilities, cognitive remediation, reading remediation, Orton-Gillingham, NILD, neuroplasticity

As a society, we have a collective responsibility to teach our children to learn and to thrive as independent adults. Consequently, we must ensure that our education system is doing its part. It is the job of educators to prepare all children for success in whatever life path they choose (Province of British Columbia, 2018). The goal of educating students with specific learning disabilities (LD) is no exception. Specifically, educators must support students with LD in reading (RD), written expression (WD), and math (MD) to help them develop tools for independent learning in the classroom and in life (National Institute for Learning Development [NILD], n.d.).

One way to ensure learner independence is by providing appropriate remediation in literacy. To date, much of the research on word reading difficulties in students with LD has focused on phonological awareness and rapid naming, but recent research suggests that other cognitive factors such as visual attention span, short-term memory, and working memory also play a significant role in word reading level (Shaul, Katzir, Primor, & Lipka, 2016). In addition, weaknesses in other cognitive domains such as visual-spatial attention are considered to be important risk factors for reading challenges (Franceschini, Gori, Ruffino, Pedrolli, & Facoetti, 2012).

According to de Lima, Azoni, and Ciasca (2013), “children with dyslexia had more difficulties with performing visual-spatial and auditory attention tasks, as well as tasks involving executive functions (EF) such as flexibility, inhibitory control, strategy use, working memory, and verbal fluency” (p. 1). Shaul

Stebbings, J. A., & Kline, E. (2020). Academic and Cognitive Remediation for Students With Learning Disabilities: A Comparison Between Orton Gillingham and NILD Educational Therapy. *International Journal for Research in Learning Disabilities, 4*(2), 15-34. https://doi.org/10.28987/ijrld.4.2.15

et al. (2016) concluded that "word reading is a complex ability which builds on a wide base of cognitive abilities" (p. 51). Consequently, support exists for the notion that reading remediation of students with LD would benefit from a multidisciplinary interventional approach that connects education to neuroscience and cognitive psychology (Gollery, 2018).

Language is the foundation for reading and literacy education (Moats, 2010). To access language systems, literacy instruction must include orthography, phonetics, phonology, morphology, syntax, semantics, discourse, and pragmatics (Moats, 2010). Siegel and Mazabel (2013) identified six processes as being significant in the development of reading skills: phonology, syntax, working memory, semantics, morphology, and orthography. For students with LD, in particular, metalinguistic awareness, which involves an understanding of phoneme-grapheme correspondence, is critical to the development of reading (Ehri, 2005).

In a meta-analysis of 22 randomized controlled studies conducted from 1985 to 2010 on treatment approaches for children and adolescents with RD, Galuschka, Ise, Krick, and Schulte-Körne (2014) looked at reading fluency, phonemic awareness, reading comprehension, phonics, auditory processing, and the use of coloured overlays. The results revealed that phonics instruction was not only the most frequently investigated treatment approach, but also the only approach whose effectiveness on the reading and spelling performance of children and adolescents with reading disabilities has been statistically confirmed. By comparison, the mean effect size of the other treatments did not reach statistical significance (Galuschka et al., 2014).

It is important to recognize, however, that effect size outcomes should not be discounted as statistically relevant in interpretation of the data. As such, a comprehensive multiple-treatment approach for reading is recommended. Moats (2010) agreed, stating that "instruction that builds phonemic awareness, phonic decoding skills, text reading fluency, vocabulary and various aspects of comprehension is the best antidote for reading difficulty" (p. 15).

Orton Gillingham (OG)

The Orton Gillingham (OG) remedial method employs the recommended structured approach to literacy instruction. Originally developed in the 1930s by Samual Orton and Anna Gillingham (Academy of Orton Gillingham Practitioners and Educators, 2018), the OG method uses direct instruction that is phonologically based and employs sequential, multisensory teaching, whereby phonemes and graphemes are connected to keywords that trigger the sound-symbol relationship.

Rose and Zirkel (2007) listed more than 15 commercial OG-based reading programs, including Wilson Reading System (Wilson, 1996), the Spalding Method (Spalding & Spalding, 1990), the Slingerland Approach (Slingerland & Aho, 1996), Alphabetic Phonics (Cox, 1992/2007) and Project Read (Enfield & Greene, 1997). However, in a meta-analysis conducted between 1979 and 2001 using experimental or quasi-experimental design, Ritchey and Goeke (2006) found only 12 published peer-reviewed studies that met the criteria of scientifically based research outlined by the No Child Left Behind Act (U.S. Department of Education, 2001). These authors found that "OG instruction has yet to be comprehensively studied and reported in peer-reviewed journals. The small number of existing studies lack methodological rigor that would be required for publication in current peer-refereed journals" (2006, p. 182).

Furthermore, according to the U.S. Department of Education's Institute of Education Sciences (IES; 2010) What Works Clearinghouse (WWC), no studies of Orton Gillingham-based strategies – unbranded or branded (including Alphabetic Phonics, Barton, Dyslexia Training Program, Herman Method, Voyager Reading Program, Wilson Reading System) – fell within the scope of the evidence standards for the Students with Learning Disabilities protocol (IES, n.d.a). This means that the WWC was unable to draw any conclusions based on the available research about the effectiveness or ineffectiveness of these interventions for students with LD. Only Project Read phonology fell within the scope of the Students with Learning Disabilities WWC evidence standards; however, no discernible effects on general reading achievement for students with LD were found (IES, 2010b).

Nevertheless, it is important to note that the "long term, extensive use of OG has been encouraged by anecdotal evidence and personal experience" (Ritchey & Goeke, 2006, p. 172). This "practice to research gap" (Ritchey & Goeke, 2006, p. 182) exists largely due to the endorsement of OG by parents who observe true progress in their child's reading. Thus, the popularity of OG is reflected in an examination of 27 reading methodology case law decisions in the United States, which revealed that OG was the preferred methodology requested by parents, at 67%, followed by Lindamood-Bell, at 16% (Rose & Zirkel, 2007). Given these findings, the lack of rigorous scientific research highlights the need for increased study on the effectiveness of OG intervention.

In spite of the lack of evidence to support OG-based reading interventions, a meta-analysis of several hundred studies conducted in 2000 by the National Reading Panel (NRP) determined that phonemic awareness training and systematic phonics instruction makes a significant improvement in reading growth in children compared to unsystematic instruction or no phonics instruction (pp. 2–130). Furthermore, an examination by the WWC of effective reading instruction for emergent readers, struggling readers, and/or readers with LD, showed positive effects on alphabetic instruction in three programs, potentially positive effects in three additional programs, and evidence of a small effect in one program. However, no specific outcomes were evaluated by the WWC for students with LD equivalent in age to the participants in the present study (IES, 2010c). Consequently, although OG-based methodology has yet to be legitimized, evidence suggests that the phonemic awareness and phonics instruction used in OG are successful instructional strategies for reading.

National Institute for Learning and Development Educational Therapy (NILD)

The NILD also adheres to a structured, sequential, multisensory keyword remedial approach to literacy instruction based on OG methodology, identified as the Blue Book method (Dwyer, 2000). However, NILD's remedial approach differs from OG in its overarching educational paradigm, which includes an emphasis on neurocognitive and social-emotional development. This paradigm is based on Feurstein's (1980, as cited in Mentis, Dunn-Bernstein, & Mentis, 2008) concept of structural cognitive modifiability, which posits that cognition is plastic and can be structurally changed through a mediated learning approach. The approach is supported by Vygotsky's (1978) theory of the Zone of Proximal Development (ZPD), whereby a mediator or more capable peer scaffolds student learning to achieve greater results than the student could do independently.

As for OG, scientific research on NILD's effectiveness remains inconclusive. Thus, a review of existing NILD studies appears to suggest a lack of rigorous scientific methodology. Stanley's (2008) study of 29 students was a quasi-experimental design, using students in grades 2 to 5 who presented as below-grade-level readers although not necessarily with an LD diagnosis.

Szabo and Balla's (2017) study of 13 students from ages 6 to 12 received pre/post math tests after receiving NILD's math intervention or no special intervention. Although greater gains were seen after seven months in the NILD treatment group than the control group, these results were not statistically confirmed. Furthermore, other studies suffer from a lack control groups (Benson & Scott, 2005; Keafer, 2008), consistency in length of intervention (Gollery, 2018), quantitative measurement (Hutchinson, 1999), randomized selection (Cawthon & Maddox, 2009; Hopkins, 1996; Szabo & Balla, 2017), and controls for IQ (Cawthon & Maddox, 2009; Gollery, 2018).

Despite these shortcomings, the available NILD studies provide some promising preliminary evidence. Specifically, significant outcomes were observed on phonological processing, decoding, and fluency (Stanley, 2008). Increases were also seen in math performance and cognitive performance compared to a control group (Szabo & Balla, 2017), within-group gains over time (Hopkins, 1996), and large effect size results on the Weschler Intelligence Scale for Children (WISC-IV; Wechsler, 2003) pre/post full-scale IQ scores (Gollery, 2018), with the perceptual reasoning subtest showing the largest mean growth. Nevertheless, the lack of rigorous research on the effectiveness of NILD highlights the need for further study.

The NRP's (2000) recommendations for comprehension instruction align with much of NILD's philosophy, stating that "reading comprehension is a cognitive process that integrates complex skills" (pp. 4-1). The NRP has advocated a "transactional strategy instruction" (pp. 4-46), whereby mental processes and cognitive strategies are explicitly discussed to aid comprehension. Other instructional techniques for reading comprehension recommended by the NRP include self-generation of questions, summarization and generalization, visualization, mnemonic strategies, and comprehension self-monitoring through the reading process. These strategies are supported by NILD's mediated instructional approach.

Other Approaches

What is clear, according to Swanson, Harris, and Graham (2014), is that the research is not conclusive as to whether various remedial delivery methods have an influence on student outcomes. According to Swanson et al., intervention efforts for children with language LD are not sufficient for complete remediation of underlying language deficits, suggesting a potential need for other types of remediation, including cognitive interventions. Shaywitz, Modyu, and Shaywitz (2006) claimed, "The utilization of advances in neuroscience to inform educational policy and practices provides an exciting example of translational science being used for the public good" (p. 281).

Based on Swanson et al.'s (2014) assertion that current interventions do not include remediation for the underlying language deficits, there exists a need to determine whether interventions that target the underlying cognitive deficits would yield improvement in cognitive results. Furthermore, if cognitive improvements were noted, further studies could suggest that, over time, greater gains would be seen in academic outcomes.

The current study sought to compare a commonly known intervention for LD (OG) with an alternative intervention that also targets cognition (NILD) to determine if differences in academic and cognitive outcomes between the two types would inform future remedial interventions. The importance of integrating a cognitive approach to academic remediation may be understood by Hale and Fiorello's (2004) cognitive testing hypothesis (CHT). Among the four premises of this theory, the authors stated that a number of complex cognitive and neuropsychological processes have been empirically linked to academic achievement and that academic deficits must be remediated and/or compensated for based on underlying cognitive strengths and weaknesses. NILD (n.d.) has subscribed to this neuroscientific approach to remediation of students with LD.

Evidence suggests alternative methods of developing cognition. Mindfulness meditation, for example, "might cause neuroplastic changes in the structure and function of brain regions involved in regulation of attention, emotion and self-awareness" (Tang, Hözel, & Posner, 2015, as cited in McKay, 2018, Evidence that Mindfulness Meditation Affects the Brain section, para. 2). Just one year of music instruction has been correlated with improved scores in "general intelligence skills such as literacy, verbal memory, visuospatial processing, mathematics and IQ" ("First evidence," 2006, para. 1).

In summary, despite the proliferation of LD resources, there remains a lack of solid evidence to support an interdisciplinary approach to remediation. This study aimed to fill this gap by exploring outcomes of two intervention approaches.

The Current Study

The purpose of this preliminary study was to determine if there was a difference in academic and cognitive outcomes between an OG group receiving a pure literacy-based interventional approach and an NILD group receiving targeted cognitive remediation in the context of literacy instruction. Evidence suggests "a strong convergence between the brain's structural and functional alterations in the left hemisphere of the dyslexic brain" (Linkersdörfer, Lonnemann, Lindberg, Hasselhorn, & Fiebach, 2012, para. 31). This notion creates the possibility that by strengthening the brain's neuropathways through cognitive remediation, academic functions can be improved. This, in turn, suggests a need for a multidisciplinary approach to literacy remediation.

Our primary aim was to determine if there were differences in cognitive outcomes between the OG group and the NILD group, measured by fluid reasoning standard scores. A secondary aim was to determine if different instructional approaches resulted in differences in outcomes on word attack and reading comprehension standard scores post intervention. (Measurement of neural pathways was beyond the scope of the study.)

It is important to examine the two types of interventions because both claim to remediate literacy, but only NILD claims to remediate the specific cognitive processes necessary for growth in literacy. We examined pre and post academic achievement and cognitive results using the Woodcock Johnson IV Tests of Achievement (WJ-Ach) and Woodcock Johnson IV Tests of Cognitive Abilities (WJ-Cog), respectively, for both the OG and NILD groups (Schrank, McGrew, Mather, Wendling, & LaForte, 2014). For the WJ-Cog test, we selected the variable of fluid reasoning to measure general cognitive growth as this measure represents the broad ability to reason by forming novel concepts and solving problems without requiring background knowledge. This composite score consists of Concept Formation and Number Series subtests. For the WJ-Ach test, we selected two reading variables, word attack and reading comprehension. To facilitate the inquiry, we posed the following research questions:

> *Research Question 1:* Does the type of remedial intervention produce different cognitive results between the OG and the NILD groups?
>
> H1: Larger gains will be seen on cognitive measures of fluid reasoning in the NILD group than in the OG group post intervention.
>
> *Research Question 2:* Does the type of remedial intervention produce different academic results between the OG and NILD groups?
>
> H2: Larger gains will be seen on academic measures of reading (word attack) in the OG group than in the NILD group post intervention.

H3: Larger gains will be seen on academic measures of reading (comprehension) in the NILD group than in the OG group post intervention.

Method

Research Participants

The study adhered to the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* and followed the core principles of respect for persons, concern for welfare, and justice (Government of Canada, Panel on Research Ethics, 2014). Free, informed, and ongoing consent was obtained in writing from parents and student participants, who voluntarily agreed to participate.

Participants were 27 students in grades 7, 8, and 9 from an urban independent school in British Columbia serving grade 2-12 students with LD. Participant ages ranged from 143 to 177 months with an average of 159.69 months or 13.3 years (see Table 1). Remediation is provided as part of the regular school schedule, and no pull-outs were required. All participants had one or more diagnosed LDs in either reading (RD), written expression (WD), or math (MD), neurocognitive disorder not otherwise specified (NOS), or LD unspecified (UNSP; see Table 2).

Participants were randomly assigned to the OG group or the NILD group using a stratified process. Names were printed on strips of paper, separated by grade, and then gender. Beginning with grade 9 boys, names were alternately assigned to the NILD and OG groups. This was followed by grade 9 girls, grade 8 boys, grade 8 girls, grade 7 boys, and finally grade 7 girls.

Participants were not stratified by LD type for two reasons. First, the majority of them had multiple diagnoses, and all educators designed individualized lessons based on students' academic and cognitive profiles, reflecting the school's practice of having each educator prescribe the intervention plan for both NILD and OG students based on formal and informal diagnostics. Second, in both OG and NILD interventions, approximately 30% of each lesson is developed from predetermined scope and sequence of options chosen at the discretion of the educator to target the students' particular areas of need. Despite the randomized assignment to an intervention type, both groups had relatively evenly distributed diagnoses between groups. As a result, it was not expected that distribution of LD types between groups would materially impact the study outcomes (see Table 2).

Each participant received 64.2 hours of one-to-one instruction in either OG (N = 15) or NILD (N = 12) for 70 minutes on alternating days with two sessions and three sessions per week, respectively, across 27 weeks. The OG group consisted of 11 males and 4 females whereas the NILD group was comprised of 8 males and 4 females. Three male students from the NILD group withdrew from the study mid-year due to departures from the school.

Selection Criteria

Research participants met the selection criteria based on their percentile scores on the WISC General Ability Index (GAI) and the Cognitive Proficiency Index (CPI). The GAI consists of the Verbal Comprehension Index (VCI) scores and the Perceptual Reasoning Index (PRI) scores; the CPI consists of the Working Memory Index (WMI) scores and Processing Speed Index (PSI) scores (Wechsler, 2003).

To meet the selection criteria, participants needed a VCI score and/or a PRI score of greater than the 25th percentile and a CPI score of less than 25th percentile on the WISC IV assessment (Wechsler, 2003). Thus, all participants had average or above-average general cognitive ability and below-average cognitive efficiency in addition to a designated LD based on psychoeducational assessment reports conducted by registered educational psychologists. These reports were provided by families upon their child's admission to the school. Students were excluded from the study if they were new to the school or had received previous NILD intervention.

Remedial Educators

All educators were full-time learning specialists (LS) and either professionally certified NILD educational therapists (PCET) and/or certified OG associate practitioners by either the Canadian Academy of Therapeutic Tutors or the Academy of Orton Gillingham Practitioners and Educators. The average years of educator experience in one-to-one remedial work for students with LD was seven for NILD and six for OG. NILD educational therapists had recently completed Level III training to achieve certification.

Study Design and Intervention

Pre and post intervention testing was conducted by qualified B level testers for the WJ-Ach and C level testers for the WJ-Cog. Testers used the WJ-Ach B version for the posttest and the WJ-Ach C version for the pretest to ensure test-retest reliability.

Table 1
Participant Demographics

Variable	OG group	NILD group
Number of participants		
Male	11	8
Female	4	4
Age in months		
Mean	160	160
Race/ethnicity		
Caucasian	14	10
African American/African Canadian	0	1
Asian/Asian Indian	1	1
Socioeconomic status		
High	15	12
Grade level		
Mean years in special education placement	2.27	1.33
Level of Placement	7,8,9	7,8,9
Intelligence (WJ-IV Cog.)		
Mean W score[a]	504	504
Range	490–524	488–519
Overall academic achievement[b]		
Mean W score	502	502
SD	12.9	13.9
Range	476–520	478–520
Specific academic achievement[c]		
Mean W score	509	503
SD	9.1	13.8
Range	490–522	472–523

Note. WJ-IV Tests of Cognitive Abilities (Jaffe, 2009). [a]The W score is a foundational metric for all derived scores standard, percentile, and relative proficiency index on an equal interval scale that allows comparison of differences between two sets of scores. [b]Measured using WJ-IV Tests of Achievement, Academic Skills Composite. [c]Measured using WJ-IV Tests of Achievement, Reading Composite.

OG group. Participants in the OG group received the school's regular one-to-one OG remedial intervention whereby the LS determined the instructional focus and lesson design based on students' individual profiles. Student information was gathered from psychoeducational assessments, individualized education programmes (IEP), informal diagnostic lessons, and LS observations.

Lessons in OG comprised eight core elements, visual drill, auditory drill, phonological awareness activities, spelling dictation, blending drills, rapid reading drills, discretionary new concept instruction, and novel reading (Leopold, 2017), using direct instruction in a multisensory, cumulative, and sequential manner. Phonemes/graphemes were introduced individually using a keyword approach with progression occurring once

Table 2
Learning Disability Diagnosis by Student

Identifier	DSM reading	DSM writing	DSM math	LD-NOS	LD Unspecified
		DSM diagnosis per student for OG group			
OG1	1	0	0	0	0
OG2	1	1	1	0	0
OG3	1	1	1	0	0
OG4	1	1	0	0	0
OG5	1	1	0	0	0
OG6	1	1	0	1	0
OG7	1	1	1	0	0
OG8	1	1	1	0	0
OG9	1	0	0	1	0
OG10	1	1	0	0	0
OG11	0	0	0	0	1
OG12	1	1	1	0	0
OG13	0	1	1	0	0
OG14	0	1	1	1	0
OG15	0	0	0	0	1
Total	11	11	7	3	2
		DSM diagnosis per student for NILD group			
NILD1	1	1	0	0	0
NILD2	1	1	1	0	0
NILD3	0	1	1	1	0
NILD4	1	1	1	0	0
NILD5	1	1	1	0	0
NILD6	1	1	1	0	0
NILD7	1	1	1	0	0
NILD8	1	0	1	0	0
NILD9	0	0	0	0	1
NILD10	1	1	1	0	0
NILD11	1	1	1	1	0
NILD12	0	1	0	0	0
Total	9	10	9	2	1

Note. LD-NOS = Learning disorder not otherwise specified; LD Unspecified = British Columbia Ministry designation for learning disability "Q."

Table 3
Pedagogical Similarities and Differences Between OG and NILD Approaches

Aspect	OG	NILD	Both OG and NILD
Diagnostic and Statistical Manual Reading Disorder	Starts phonological instruction at level of student independent success; uses direct instruction.	Starts phonological instruction at beginning of Blue Book; applies ZPD strategy and mediated instruction using Socratic questioning.	Sequential, cumulative, and multisensory; apply OG structured literacy approach.
	Does not move on until mastery of concept is achieved.	Moves on to new linguistic pattern or concept regardless of mastery, and reinforces previous concepts in subsequent lessons to achieve mastery.	Focus on phonology; develop skills in decoding accuracy, fluency, and comprehension.
	Does not assume student knows anything that has not been directly instructed; repetition to automaticity.	Teaches students to apply cognitive strategies such as compare, label, search for patterns, use precise accurate language, and connect to prior knowledge to discover concepts.	Emphasis on student success – student gains confidence in process; individualized instruction.
	Highly structured lessons using flash card system to introduce and reinforce concepts.	Structured literacy lesson that includes explicitly targeting cognitive skills at the input, elaboration, and output phases to develop concepts.	Keyword approach; lessons contain components of phonemic awareness, phonology, blending, syllabication, word reading, reading fluency, and morphology.
Diagnostic and Statistical Manual Disorder of Written Expression	Modeling; direct instruction and cueing; repetition and review; application of learning.	Modeling and mediated instruction; applying specific thinking actions; application of learning.	Sequential, cumulative, multisensory; focus on spelling, sentence type and structure, grammar, paragraph and essay organization and structure, editing strategies.
Diagnostic and Statistical Manual Math Disorder	Not applicable.	Math block: Builds number sense, memory strategies, math vocabulary, abstract reasoning, and real-world application; emphasizes process over product and justification of responses using precise and specific language.	Not applicable.
General strategies	Direct instruction.	Mediated instruction.	Sequential, cumulative; step-by-step, repetition, and review; use of mnemonics, extra time for processing, executive function strategies; cues to focus, breaks as needed, success oriented.

the concept had been mastered. Lessons addressed the following areas: reading (decoding, fluency, comprehension), written expression (spelling, sentence structure, grammar, mechanics, ideation, outlining, structure, editing), and executive functioning (EF) such as planning, prioritizing, and organization (see Table 3).

NILD group. Participants in the NILD group received NILD educational therapy instead of the standard OG. This approach differs from OG in that it incorporates a mediated instructional approach (Feuerstein, Rand, Hoffman, & Miller, 1980) using open-ended questioning based on student responses. NILD prompts students to connect to prior knowledge and apply specific thinking strategies in addition to learning the alphabetic code in order to build literacy. This alphabetic instruction employs The Blue Book, which differs from OG in that students learn multiple spelling options for each phoneme by linguistic pattern recognition, mnemonics, and chunking. Like OG, NILD uses a keyword approach and visual, auditory, and oral rehearsal. Sequential progression occurs regardless of mastery, which is achieved through subsequent questioning and review. A key focus of the NILD approach is targeted cognitive remediation, which underlies literacy and learning (see Table 3).

NILD lessons consist of five core elements: rhythmic writing, dictation and copy, buzzer, Blue Book, and math block. Math block is designed to develop number sense, memory strategies, math vocabulary, abstract reasoning, and real-world application. Emphasis is placed on process over product and justification of responses using precise and specific language.

In addition, NILD includes up to 19 discretionary supplementary elements selected by the LS to target academic and cognitive functions based on individual student need. These supplementary elements target visual and auditory discrimination, visual and auditory working memory, sequential processing, visual-spatial processing, EF, reading fluency and comprehension, written and oral language, and math. Students identify and verbalize their cognitive strategies, which are mediated by the instructor (see Table 3). Both OG and NILD lessons incorporate 5 to 10 minutes of EF strategies known as "EF Time" from the school's executive function curriculum (Huopalainen & Stebbings, 2017).

Study Fidelity

Contributing to study fidelity was the fact that educators used standardized lesson plan templates reflecting either the OG or NILD approach. To ensure study fidelity, a teacher-on-call (TOC) with an MEd degree conducted monthly observations of all OG and NILD lessons using a checklist to verify that the stated elements of the lessons were followed and the appropriate methodology was applied. Educators also kept daily anecdotal observational notes that were reviewed by the TOC to provide further evidence of appropriate strategy implementation. All records were reviewed by the TOC, and no variations from the prescribed instructional delivery model were reported.

Results

Data were analyzed by the University of British Columbia's Applied Statistics and Data Science Group. To test the primary and secondary hypotheses of estimating differences between the groups, a linear model was used with post measurements as outcomes, group as a main effect, and pre measurements as covariates. Additionally, age and gender were added as covariates. Thus, the estimated group differences in post measurement were adjusted for their baseline (pre) measurements, age, and gender. Linear modeling was chosen due to the imbalanced number of observations in each of the groups.

Standard Scores of Pre and Post Fluid Reasoning by Group

With regard to the first outcome of the primary hypothesis, it was expected that NILD would produce greater outcomes on the fluid reasoning variable after intervention. Results indicated that the median of the standard score for fluid reasoning increased over time for both NILD and OG, with NILD increasing more dramatically. The median of the post standard score for fluid reasoning was higher in NILD than in OG. However, variability in the data was large in both groups due to the small sample size, so there were no statistically significant differences between or within groups (see Figure 1).

Furthermore, the results suggested that the variability of standard scores of fluid reasoning within each student differed: from very small, to modest, to very large (see Figure 2). With a 95% confidence interval (CI), the average fluid reasoning scores were higher in the NILD group at both pre and post time points. Again, due to the wide range of CI, differences between NILD and OG were not statistically significant (see Figure 3). See Table 4 for the means and standard deviations for pre and post measurements.

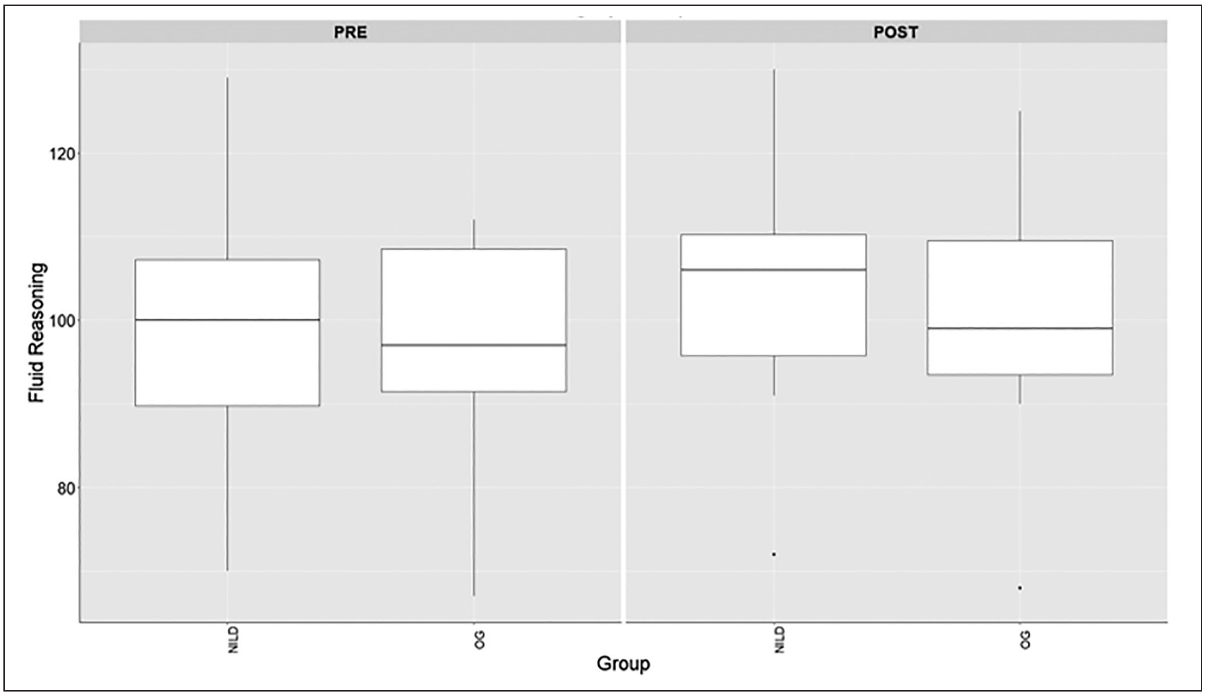


Figure 1. Boxplot of pre and post standard scores for fluid reasoning by group.

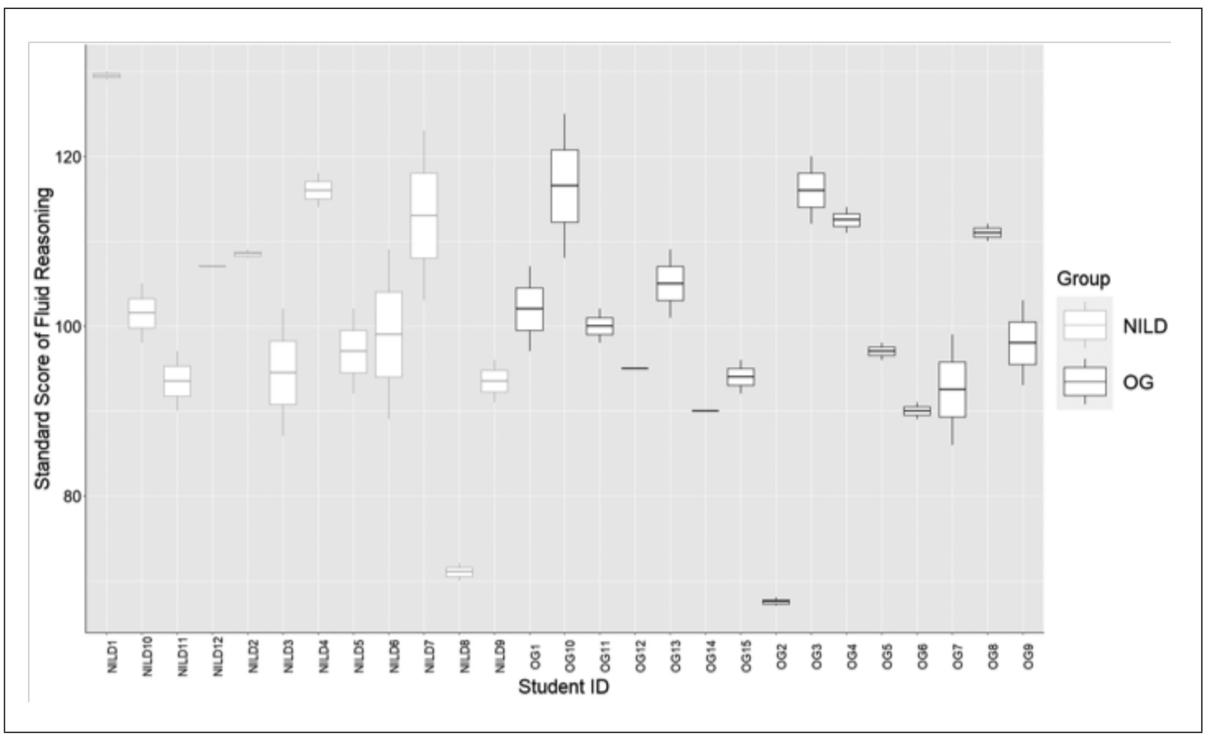


Figure 2. Boxplots of standard scores for fluid reasoning for each student by group.

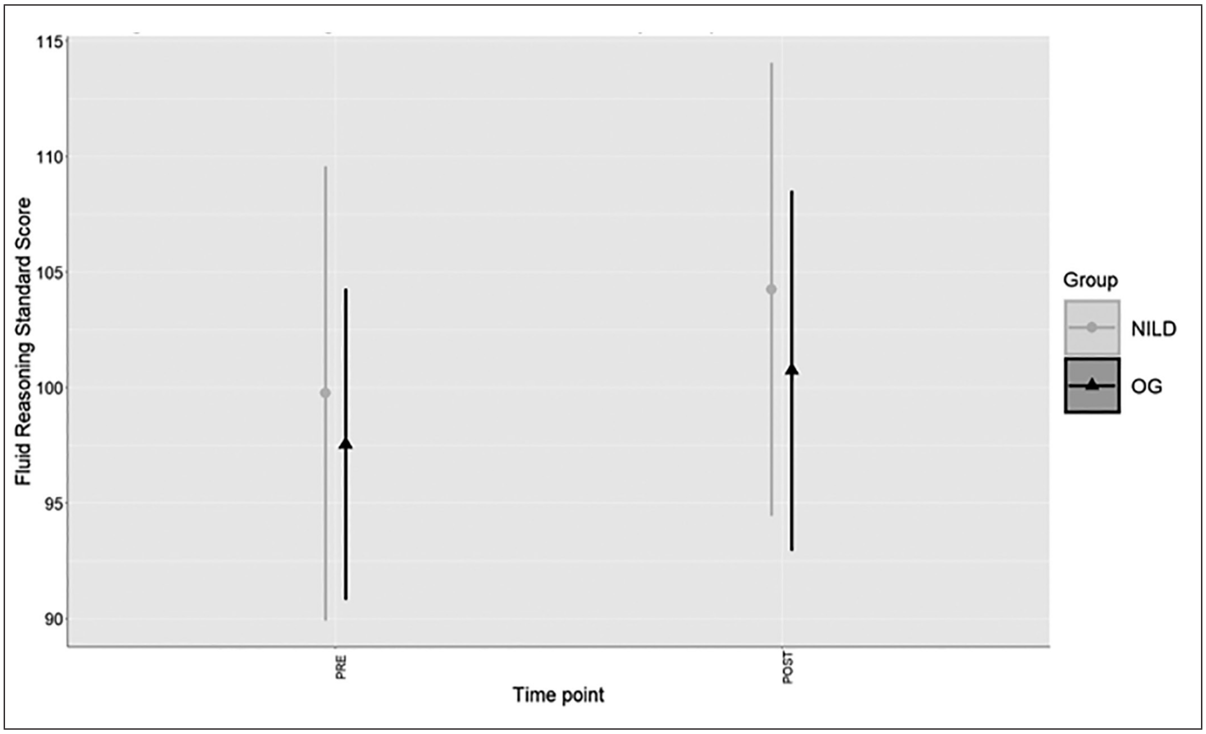


Figure 3. Average fluid reasoning standardized scores at each time point by group.
Note: The vertical lines through the points represent corresponding 95% CI.

Table 4
Three Standard Scores Outcomes

		Pre		Post	
Outcome	Group	*M*	*SD*	*M*	*SD*
FR	NILD	99.75	15.375	104.25	15.33
FR	OG	97.53	12.047	100.73	13.97
WA	NILD	94.00	15.696	102.25	11.78
WA	OG	96.67	12.028	107.60	11.59
RC	NILD	84.42	15.524	93.75	14.32
RC	OG	93.07	9.169	97.13	12.42

Note. FR = Fluid reasoning composite; WA = Word attack; RC = Reading comprehension.

Test of Primary Hypothesis on the Primary Outcome Fluid Reasoning, Linear Model

To test the first hypothesis for fluid reasoning, a linear model was used with post fluid reasoning standard scores specified as an outcome and group as a main effect. The group differences in post fluid reasoning standard scores were adjusted for pre fluid reasoning standard scores, age, and gender. Using this model, no significant difference was found between the groups in post fluid reasoning standard scores (see Table 5). However, the means estimate of post fluid reasoning standard scores for NILD was higher than OG by 1.486; 95% CI (-5.52, 8.49; see Table 6). The CI for fluid reasoning was not adjusted for multiple testing since this was our primary outcome.

Standard Scores of Pre and Post Word Attack by Group

In the secondary outcome of the secondary hypothesis, it was expected that OG would produce

Table 5
Coefficients From the Linear Model Fit for Fluid Reasoning

Coefficient	Estimate	Std. Error	*t*	*p*
(Intercept)	12.4497	25.8666	0.4813	0.6351
SS_FR_PRE	0.8922	0.1281	6.9669	0.0000
GroupNILD	1.4860	3.3794	0.4397	0.6644
Age	0.1425	1.7006	0.0838	0.9340
GenderMale	-0.7964	3.6807	-0.2164	0.8307

Note. SS_FR_POST~SS_FR_PRE+Group+Age+Gender.

greater outcomes on the post word attack standard scores. Results indicated that the median of the standard score for word attack increased over time for both NILD and OG, with OG increasing more dramatically. The median of the post standard score for word attack was higher in OG than in NILD. Again, due to higher variability in the small sample size, statistically significant outcomes were not expected between groups (see Figure 4). Not surprisingly, when looking at the variability of word attack standard scores within each student, results differed from very small, to modest, to very large (see Figure 5).

Average word attack scores in both OG and NILD groups increased over time. Notably, the word attack standard score was higher in OG at both pre and post time points. Again, as expected, based on the small sample size and variability, there were no statistically significant differences between the groups (see Figure 6). Table 4 lists the means and standard deviations for pre and post measurements.

Test of Secondary Hypothesis for the Secondary Outcome Word Attack, Linear Model

To test the secondary hypothesis for word attack standard scores, a linear model was again used, where post word attack standard score was the outcome variable and group was the main effect. The group differences were adjusted for pre word attack standard scores, age, and gender. Using this model, there was no significant difference between groups (see Table 7). However, the means estimate of post word attack standard scores for NILD was lower than for OG by 4.533; that is, the group difference estimate for NILD com-

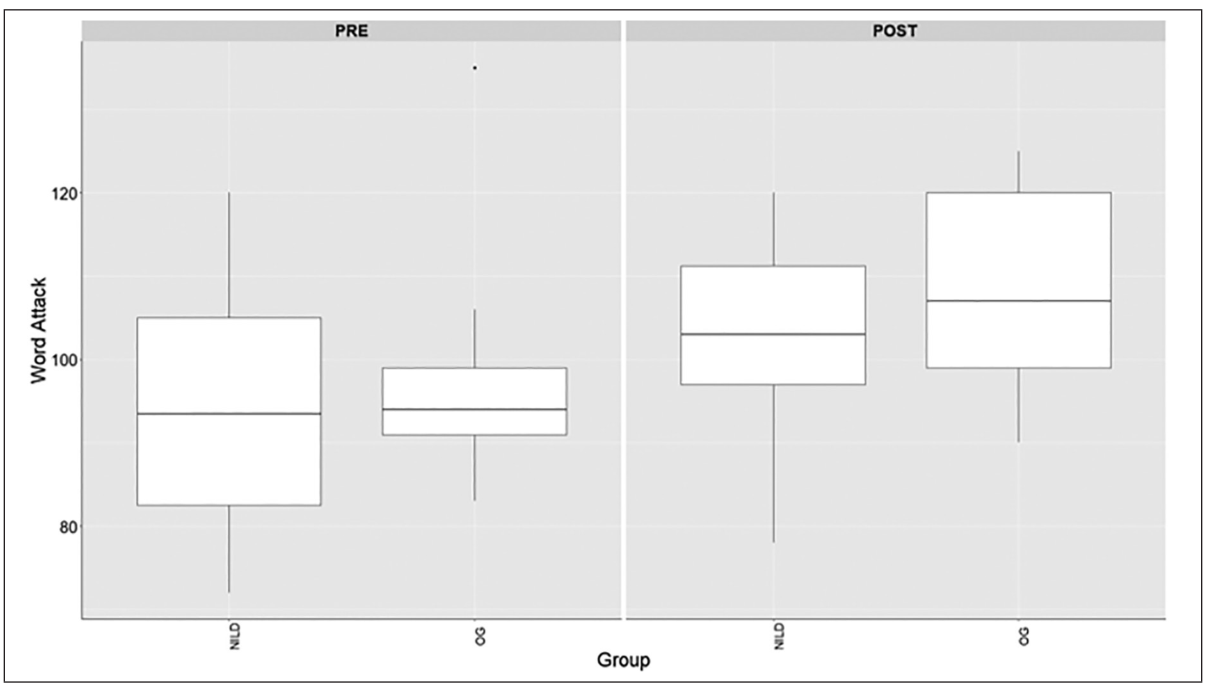


Figure 4. Boxplot of pre and post standard scores for word attack by group.

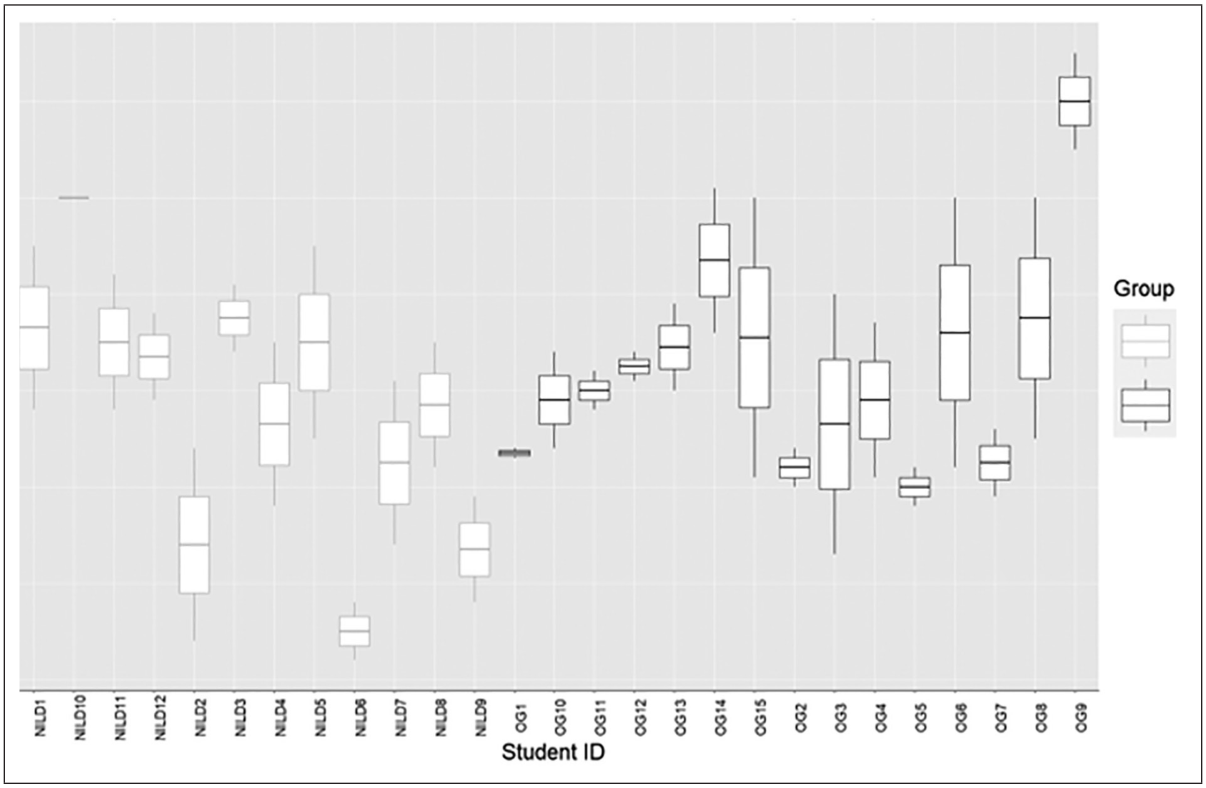


Figure 5. Boxplots of standard scores for word attack for each student by group.

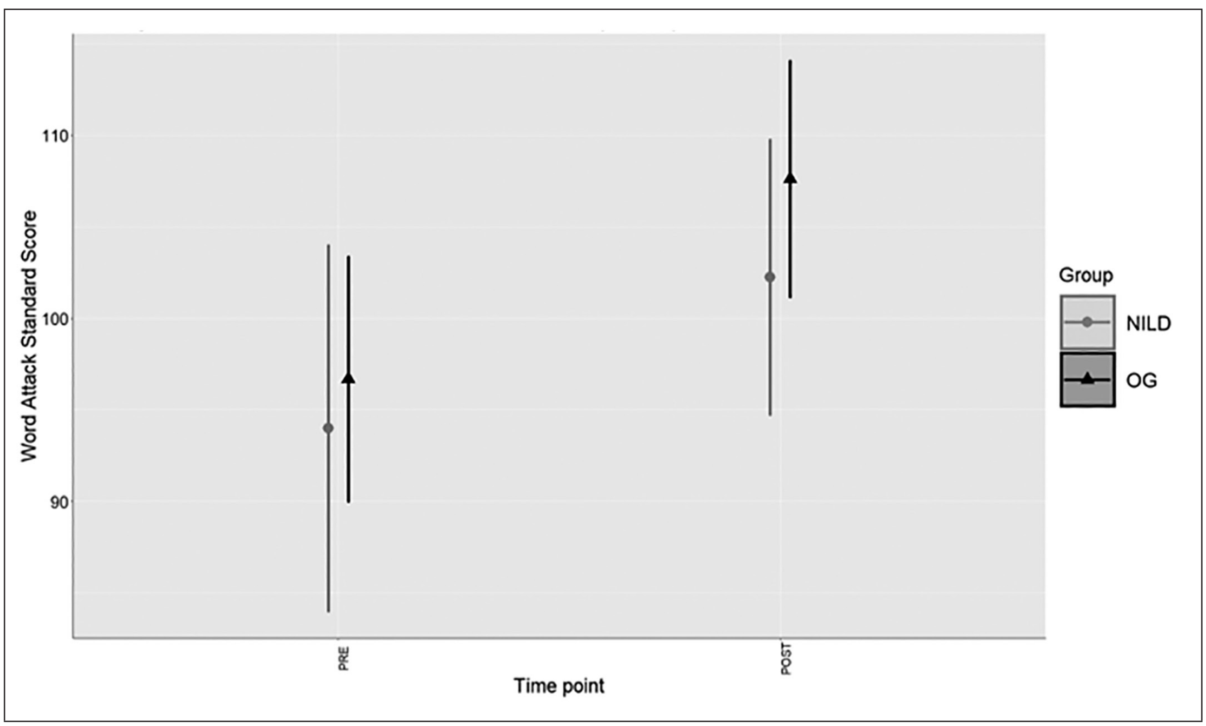


Figure 6. Average word attack standardized scores at each time point by group.
Note: The vertical lines through the points represent corresponding 95% CI.

Table 6
Estimates and Confidence Intervals for All Three Outcomes

Hypothesis	Outcome	NILD.OG.Diff	Lower_CI	Upper_CI	Cohen f^2
1	FR	1.485962	-5.52252	8.494448	0.008788
2	WA	-4.53262	-13.327	4.261793	0.069866
3	RC	4.382788	-3.5928	12.35838	0.079424

Note. FR = Fluid reasoning composite; WA = Word attack; RC = Reading comprehension.

Table 7
Coefficients From the Linear Model Fit for Word Attack

Coefficient	Estimate	Std. Error	*t*	*p*
(Intercept)	70.2563	26.8743	2.6143	0.0158
SS_WA_PRE	0.5075	0.1368	3.7094	0.0012
GroupNILD	-4.5326	3.6560	-1.2398	0.2281
Age	-0.4478	1.8409	-0.2433	0.8101
GenderMale	-8.0409	3.9729	-2.0239	0.0553

Note. SS_WA_POST~SS_WA_PRE+Group+Age+Gender.

pared to that of OG was -4.533 with 97.5% CI (-13.3, 4.26; see Table 6). The CI of the group effect estimate was adjusted for two tests (word attack and reading comprehension) using the Bonferroni correction technique for multiple testing.

Standard Scores of Pre and Post Reading Comprehension by Group

It was expected that NILD would produce greater outcomes than OG on the reading comprehension variable post intervention. Results indicated that the median of the standard score for reading comprehension increased over time for both NILD and OG, with NILD increasing more dramatically. The median of the post standard score for reading comprehension was only slightly higher in OG than in NILD. However, variability in the post reading comprehension scores was greater in the OG group than in the NILD group, so statistically significant outcomes were not expected between groups (see Figure 7). As anticipated, when looking at the variability of standard scores of reading comprehension within each student, results differed from very small, to modest, to very large (see Figure 8).

Average reading comprehension scores in both the OG and NILD groups increased over time. Notably, the reading comprehension standard score was higher in NILD at both pre and post time points. Again, as expected, based on the small sample size and great variability, there were no statistically significant differences between the groups (see Figure 9). Table 4 lists the means and standard deviations for pre and post measurements.

Test of Secondary Hypothesis on Secondary Outcome Reading Comprehension, Linear Model

To test the secondary hypothesis on reading comprehension standard scores, a linear model was used, with post reading comprehension standard scores as an outcome variable and group as a main effect. The group differences were adjusted for pre reading comprehension standard scores, age, and gender. Using this model, there was no significant difference between the groups (see Table 8). However, the means estimate of post reading comprehension standard scores for NILD was higher for OG

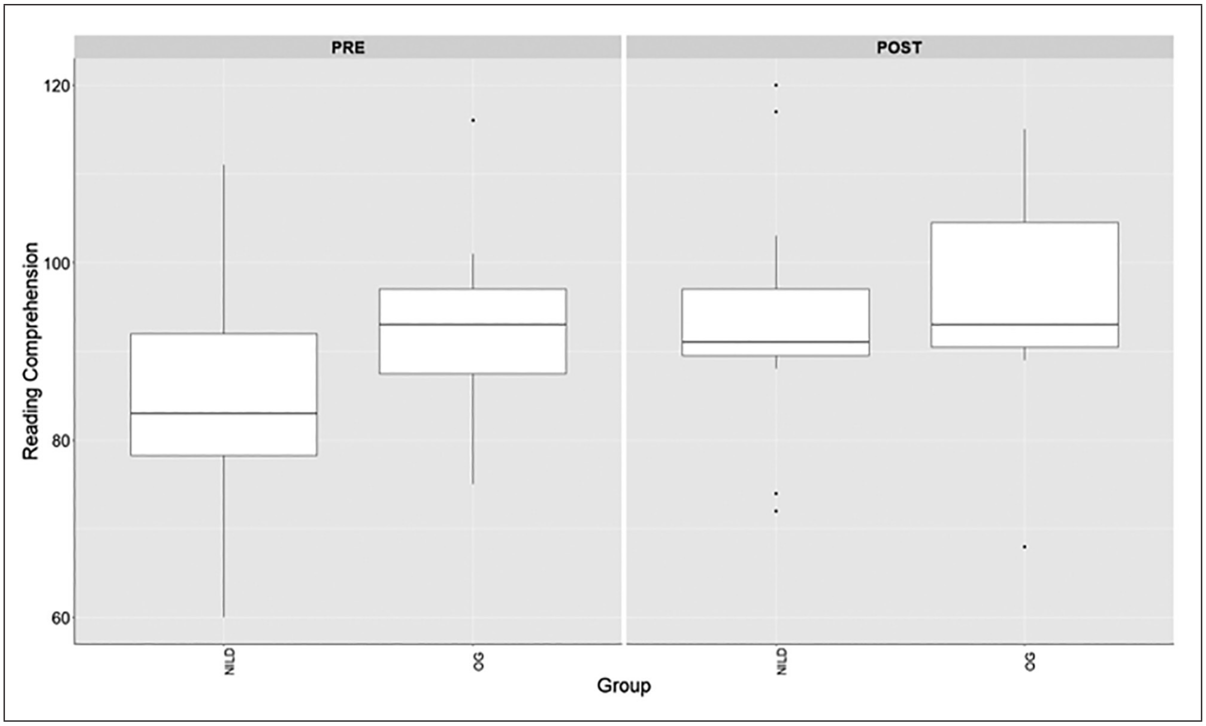


Figure 7. Boxplot of pre and post standard scores for reading comprehension by group.

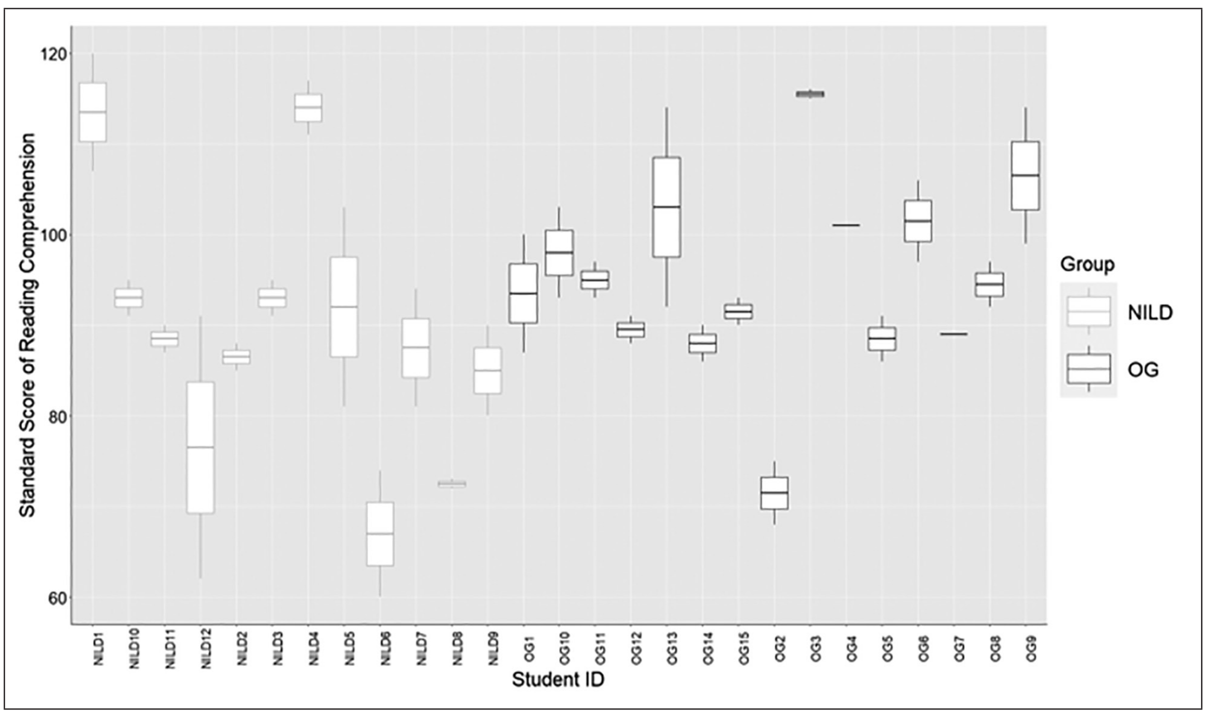


Figure 8. Boxplots of standard scores for reading comprehension for each student by group.

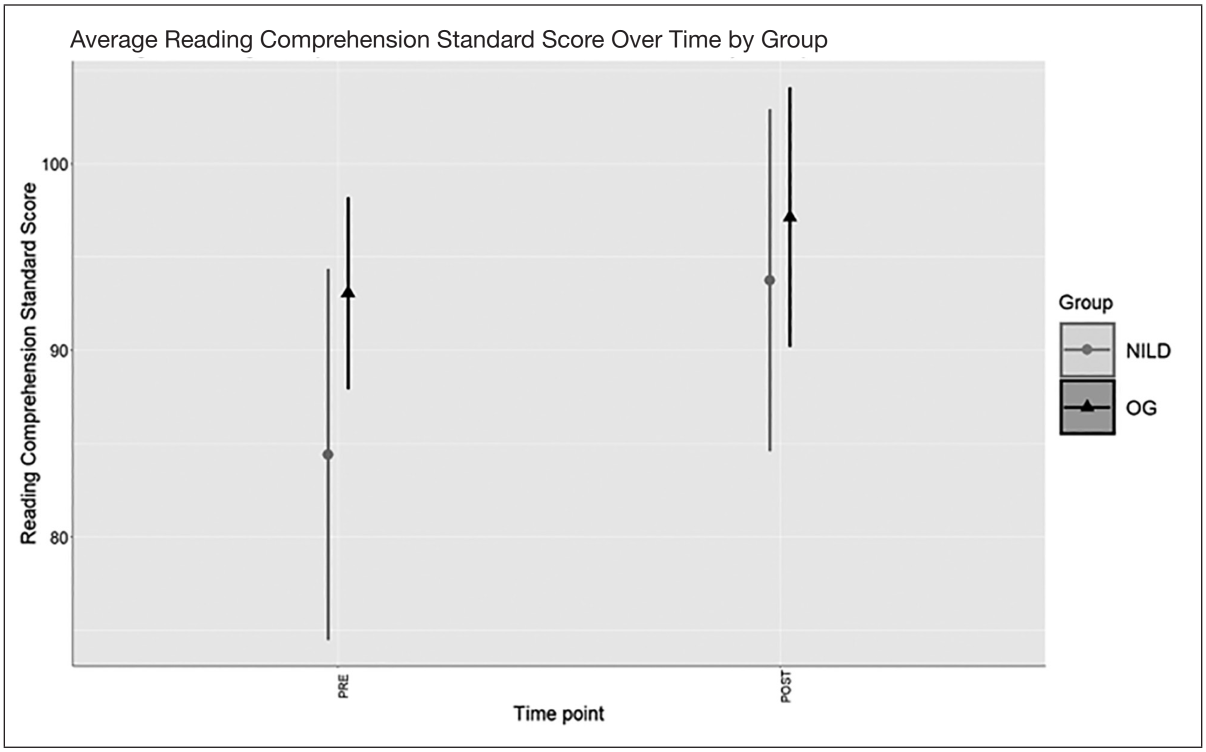


Figure 9. Average reading comprehension standardized scores at each time point by group.
Note: The vertical lines through the points represent corresponding 95% CI.

Table 8
Coefficients From the Linear Model Fit for Reading Comprehension

Coefficient	Estimate	Std. Error	*t*	*p*
(Intercept)	61.6428	24.4229	2.524	0.0193
SS_RC_PRE	0.8647	0.1337	6.469	0.0000
GroupNILD	4.3828	3.3156	1.322	0.1998
Age	-3.7027	1.5567	-2.378	0.0265
GenderMale	4.2937	3.4946	1.229	0.2322

Note. SS_RC_POST~SS_RC_PRE+Group+Age+Gender.

by 4.383; 97.5% CI (-3.60,12.36; see Table 6). The confidence interval was adjusted for two tests (for the two secondary outcomes word attack and reading comprehension) using the Bonferroni correction technique for multiple testing.

Discussion

This study offers preliminary support for the notion that targeted remedial interventions can impact academic outcomes. While it is interesting to see higher cognitive scores on fluid reasoning following cognitive remediation, the effect becomes more relevant when it impacts academic outcomes such as reading comprehension, as demonstrated by the NILD group. Equally meaningful, the intervention primarily designed to build basic reading skills was able to do just that, congruent with the instructional focus of OG. These findings further assert the notion that there is a place for both direct instruction and a mediated learning approach in remedial pedagogy.

One of the most interesting outcomes of the study was that OG showed greater results in reading decoding whereas NILD showed greater results in reading comprehension. These findings can be directly attributed to the instructional approaches used in each group.

Berninger and Richards' (2002) concept of high- and low-level reading skills provides a useful model to illustrate the two approaches. For example, OG primarily emphasizes low-level reading skills such as word recognition; in essence, learning to read. Conversely, NILD, while also remediating phonological skills and phonics, primarily focuses on the higher-level aspects of reading such as comprehension; in essence, reading to learn. These preliminary results clearly support the notion that the type of instructional approach to reading impacts academic outcomes. This idea is further supported by Fiorello, Hale, and Snyder (2006) with reference to the response to intervention (RTI) model (RTI Action Network, n.d.). Fiorello et al. stated,

> Despite the well-established research base for instruction in multiple areas of reading including phonemic awareness, phonics, fluency and comprehension (NRP, 2000), RTI interventions typically focus on one or a few areas of difficulty (e.g., phonemic awareness) to the exclusion of many others such as higher level comprehension skills. There are numerous reasons for reading disability, and focusing on a single determinant cannot effectively identify or serve all children with the disorder. (2006, p. 836)

The aim of both the OG and the NILD remediation is to foster learner independence. To do so, Fiorello et al. (2006) suggested that it is important to "develop individualized interventions based on cognitive processing strengths and weaknesses" (p. 837) and that remediation should include a neurological approach that acknowledges "the cognitive processes that underlie reading performance rather than focusing on visible input or output demands" (p. 837). The preliminary findings of this study support previous assertions that underlying cognitive processes can impact academic achievement, confirming the value of an interdisciplinary interventional approach to students with LD.

This study provides preliminary evidence that targeted cognitive remediation positively impacts reading comprehension, whereas targeted phonics remediation positively impacts reading decoding. Although both intervention types support the development of literacy, they do so using markedly different approaches, suggesting that there is value in identifying the primary area of student need in order to determine the best interventional approach.

Recommendations

Not only does this study provide preliminary evidence of where the two interventions, OG and NILD, are more effective, it also presents practical support for an approach that can address weaknesses in academic and cognitive processes, thus increasing the scope of remedial interventions. Based on the current study, we can provide preliminary recommendations for practice in the following ways.

First, for a nonreader or emergent reader with LD, intervention should include an OG remedial model. The focus for OG intervention is learning to read. Diagnosing this type of LD and prescribing treatment does not necessarily require a full cognitive assessment as the focus is on the visible input or output demands, an example of which is RTI.

Second, for a student with LD who has basic decoding skills but whose cognitive profile is impacting learning, intervention should include a cognition-targeted NILD approach in conjunction with phonological instruction. Diagnosing this type of LD and prescribing treatment would benefit from both an academic and a cognitive assessment to fully address the underlying cognitive deficits and enable appropriate remediation.

Third, both types of remediation yielded increases in posttest scores. These results suggest that both students with low-level and high-level literacy needs can benefit from instruction in OG and NILD. The task is to determine how to target instruction based on each individual student's learning profile.

Directions for Future Research

The findings of this study suggest promising results for students with LD. Future research should consider the impact of these interventions on larger sample sizes, different age groups, and in small-group settings. Further, a longitudinal study could assess the retention of academic and cognitive gains.

Limitations

Some limitations of the study need to be considered. The first pertains to the length of educator experience between groups. Educators in both groups were familiar with the student population and had a similar average number of years' experience (OG, $M = 7$; NILD, $M = 6$); however, the NILD group had only recently achieved professional certification. It was not feasible to access certified NILD educational therapists with the same experience as the OG practitioners due to lack of availability.

A second limitation is that all participants came from high socioeconomic backgrounds, therefore having the means to afford private school education. Students from differing socioeconomic backgrounds would likely experience a range of parental support impacting their academic and cognitive growth. However, the challenge of providing one-to-one intervention to a different demographic would likely be prohibitive due to limited human and financial resources.

A third limitation is that there was no control group due to ethical considerations of withholding intervention students with diagnosed LD. Furthermore, there was no readily available LD population in the public or private system who received no intervention of any kind. In sum, the strength of this student population was that it provided a unique opportunity to study the effects of two types of LD intervention in identical settings. While having a no-intervention control group would have made it possible to assess the efficacy of each of the interventions in general, in future research one could assess the effectiveness of each intervention in general by including a control group with a standard intervention (Singal, Higgins, & Waljee, 2014).

A fourth limitation involves the length of intervention. Although positive results were achieved by both groups over 9 months, it would be beneficial to study performance over a longer period of time to determine further gains and retention. Lastly, a larger sample size in future research would be important in supporting these findings.

Conclusions and Future Implications

The results of this preliminary study suggest that effective reading instruction requires both a low-level phonological approach such as OG and a high-level cognitive approach that emphasizes comprehension, such as NILD. That is, students benefit from an interdisciplinary approach that targets academics and the underlying cognitive processes and is based on individual learning profiles.

The promising results of this study suggest that researchers can feel optimistic in conducting further studies using these remedial approaches. The challenge for psychologists and educators is to appropriately diagnose students with LD in a cost-effective way, so that they can provide the optimum intervention to promote learning independence and maximize potential. It is of the utmost importance that educational policymakers determine a feasible and economically viable way to train faculty and implement OG and NILD or similar remediation in all schools, private and public. In this way, students with LD can think, learn, and thrive as independent adults in society.

References

Academy of Orton Gillingham Practitioners and Educators. (2018). *What is the Orton Gillingham approach?* Retrieved from https://www.ortonacademy.org/resources/what-is-the-orton-gillingham-approach/

Benson, B., & Scott, K. (2005, September). *Data analysis to determine the effectiveness of NILD educational therapy for students with learning disabilities:* A collaborative study by NILD and Descubre. Retrieved from http://nild.org/research.doc

Berninger, V. W., & Richards, T. L. (2002). *Brain literacy for educators and psychologists*. Seattle, WA: Academic Press.

Cawthon, H. C., & Maddox, J. S. (2009). *Small-group versus one-on-one educational therapy for struggling readers and writers*. Retrieved from http://nild.org/AET%20Journal%20-%202011%20summer.pdf

de Lima, R. F., Azoni, C.A.S., & Ciasca, S. M. (2013). Attentional and executive deficits in Brazilian children with developmental dyslexia. *Psychology*, *4*(10), 1–6. https://doi.org/10.4236/psych.2013.410A001

Dwyer, K. M. (2000). *The Blue Book method: An associative key word approach.* Hilton Head, SC: Interactive Educational Systems.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, *9*(2), 167–188. Retrieved from https://miblsi.org/sites/default/files/Documents/Presentations/AnitaArcherWorkshops/January2014/LearningtoReadWords.pdf

Feuerstein, R., Rand, Y., Hoffman, M., & Miller, R. (1980). *Instrumental enrichment: An intervention program for cognitive modifiability.* Baltimore, MD: University Park Press.

Fiorello, C. A., Hale, J. B., & Snyder, L. E. (2006). Cognitive hypothesis testing and response to intervention for children with reading problems. *Psychology in the Schools*, *43*(8), 835– 853. Retrieved from http://www.iapsych.com/wj3ewok/LinkedDocuments/fiorello2006hl.pdf

First evidence that musical training affects brain development in young children. (2006 September 20). *Science Daily*. Retrieved from https://www.sciencedaily.com/releases/2006/09/060920093024.htm

Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K., & Facoetti, A. (2012). A cognitive and linguistic approach to predicting and remediating word reading difficulties in young readers. In R. Schiff & R. M. Joshi (Eds.) *Interventions in learning disabilities: A handbook of systematic training programs for individuals with learning disabilities* (pp. 47–66). Geneva, Switzerland: Springer.

Galuschka, K., Ise, E., Krick, K., & Schulte-Körne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: A meta-analysis of randomized controlled trials. *PLOS One*, *9*(2). https://doi.org/10.1371/journal.pone.0089900

Gollery, T. (2018, January 27). *NILD educational therapy improves thinking and learning*. Paper presented at the NILD NW Regional Conference, Clyde Hill, Washington.

Government of Canada, Panel on Research Ethics. (2014). *Ethics framework: Core principles*. Retrieved from http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/chapter1-chapitre1/

Hale, J. B., & Fiorello, C. A. (2004). *School neuropsychology: A practitioner's handbook.* New York, NY: Guilford Press.

Hopkins, K. R. (1996). *A study of the effect of interactive language in the stimulation of cognitive function for students with learning disabilities* (Unpublished doctoral dissertation). School of Education, The College of William and Mary, Williamsburg, Virginia. Retrieved from https://pdfs.semanticscholar.org/2c53/6965fe56e5e3dec45d752b8f1a8f381cfbd0.pdf

Huopalainen, S., & Stebbings, J. (2017). *Fraser Academy's executive function curriculum.* Manuscript in preparation.

Hutchinson, S. K. (1999). *Voices heard in educational therapy for the remediation of learning disabilities in Christian schools: An analytic description of questions asked and answered* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia, Pennsylvania. Retrieved from https://repository.upenn.edu/dissertations/AAI9934161

Jaffe, L. E. (2009). *Development, interpretation, and application of the W score and the relative proficiency index* (Woodcock-Johnson III Assessment Service Bulletin No. 11). Rolling Meadows, IL: Riverside. Retrieved from https://www.hmhco.com/~/media/sites/home/hmh-assessments/clinical/woodcock-johnson/pdf/wjiii/wj3_asb_11.pdf?la=en

Keafer, K. A. (2008). *Effects of National Institute for Learning Development educational therapy for students with learning disabilities* (Unpublished doctoral dissertation). Regent University, Virginia Beach, Virginia.

Leopold, K. (2017). *Orton-Gillingham training manual: AOGPE Training* (Unpublished manuscript). Amenia, NY: Academy of Orton-Gillingham Practitioners and Educators.

Linkersdörfer, J., Lonnemann, J., Lindberg, S., Hasselhorn, M., & Fiebach, C. J. (2012). Grey matter alterations co-localize with functional abnormalities in developmental dyslexia: An ALE meta-analysis. *PLoS ONE*, *7*(8), e43122. https://doi.org/10.1371/journal.pone.0043122

McKay, S. (2018, October 30). *The neuroscience of mindfulness.* The Chopra Centre. Retrieved from https://chopra.com/articles/the-neuroscience-of-mindfulness-meditation

Mentis, M., Dunn-Bernstein, M., & Mentis, M. (2008). *Mediated learning: Teaching, tasks, and tools to unlock cognitive potential*. Thousand Oaks, CA: Corwin Press.

Miller, D. (2014). *Using the WJ IV cognitive, oral language and achievement tests in research.* Texas Woman's University, Woodcock Institute. Retrieved from https://drive.google.com/drive/folders/1EtMNJMTikZRUhKtK5FyFj9n9QTij76m6

Moats, L. C. (2010). *Speech to print: Language essentials for teachers* (2nd ed.). Baltimore, MD: Paul H. Brookes.

National Institute for Learning and Development. (n.d.). *What is NILD educational therapy?* Retrieved from http://nild.org/nild-educational-therapy-2

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implication for reading instruction.* Washington, DC: National Institute of Child Health and Human Development.

Province of British Columbia. (2018). *BC's new curriculum.* Retrieved from https://curriculum.gov.bc.ca/graduation

Ritchey, K. D., & Goeke, J. L. (2006). Orton-Gillingham and Orton-Gillingham-based reading instruction: A review of the literature. *The Journal of Special Education*, *40*(3), 171–183.

Rose, T. E., & Zirkel, P. (2007). Orton-Gillingham methodology for students with reading disabilities: 30 years of case law. *Journal of Special Education*, *41*(3), 171–185. https://doi.org/10.1177/00224669070410030301

RTI Action Network. (n.d.). *What is RTI?* Retrieved from http://www.rtinetwork.org/learn/what/whatisrti

Schrank, F. A., McGrew, K. S., Mather, N., Wendling, B. J., & LaForte, E. M. (2014). *Woodcock Johnson IV tests of achievement*. Rolling Meadows, IL: Riverside.

Shaul, S., Katzir, T., Primor, L., & Lipka, O. (2016). A cognitive and linguistic approach to predication and remediation of word reading difficulties in young readers. In R. Schiff & R. M. Joshi (Eds.), *Interventions in learning disabilities: A handbook of systematic training programs for individuals with learning disabilities* (pp. 47–66). Retrieved from https://www.amazon.com/Interventions-Learning-Disabilities-Systematic-Individuals/dp/3319312340

Shaywitz, S. E., Modyu, M., & Shaywitz, B. A. (2006). Neural mechanisms in dyslexia. *Association for Psychological Science*, *15*(6), 278–281. https://doi.org/10.1111/j.1467-8721.2006.00452.x

Siegel, L. S., & Mazabel, S. (2013). Basic cognitive processes and reading disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 189–213). New York, NY: Guilford.

Singal, A. G., Higgins, P.D.R., & Waljee, A. K. (2014). A primer on effectiveness and efficacy trials. *Clinical and Translational Gastroenterology*, *5*(1), e45. https://doi.org/10/1038/ctg.2013.13

Stanley, S. K. (2008). An analysis of Rx for Discovery Reading for elementary students below grade level in reading. *Discoveries*, *25*(1), 4–9. Retrieved from http://discoveryprogram-inc.com/learning-center/wp-content/uploads/Analysis-of-Rx_Stanley.pdf

Swanson, H. L., Harris, K. R., & Graham, S. (2014). *Handbook of learning disabilities* (2nd ed.). New York, NY: Guilford.

Szabo, I., & Balla, A. (2017). The efficiency of the NILD educational therapy in the development of mathematical thinking. *Romanian Journal of School Psychology*, *10*(19), 51–61. Retrieved from https://www.ceeol.com/search/viewpdf?id=564018

U.S. Department of Education. (2001). *No child left behind: Elementary and secondary education act* (Public Law PL 107-110). Retrieved from https://www2.ed.gov/nclb/landing.jhtml

U.S. Department of Education, Institute of Education Sciences. (2010a). *Find what works based on the evidence.* Retrieved https://ies.ed.gov/ncee/wwc/FWW/Results?filters=,Children-Youth-with-Disabilities

U.S. Department of Education, Institute of Education Sciences. (2010b). *What Works Clearinghouse intervention reports for Orton Gillingham Based Strategies*. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_ortongill_070110.pdf

U.S. Department of Education, Institute of Education Sciences. (2010). *What Works Clearinghouse intervention evidence snapshot.* Retrieved from https://ies.ed.gov/ncee/wwc/EvidenceSnapshot/397

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wechsler, D. (2003). *WISC IV: Intelligence scale for children.* San Antonio, TX: Pearson. Wechsler, D. (2014). *WISC V: Intelligence scale for children*. San Antonio, TX: Pearson.

Acknowledgments

Special thanks to Fraser Academy learning specialists for providing remedial instruction, to Linda Siegel and Colleen Reid for their consultation services, and to Fraser Academy families for their willingness to participate in this study.

Risk Factors Associated With Language Delay in Preschool Children

Alexandra Galvin[1], Georgina Davis[1], Denise Neumann[1], Lisa Underwood[2], Elizabeth R. Peterson[1,2], Susan M. B. Morton[2,3], Karen E. Waldie[1,2]

Abstract

A variety of antenatal risk factors have been established as being detrimental to a child's developing language ability. Our aim was to examine the relationship between exposure to cumulative risk (CR), including antenatal maternal, perinatal, and maternal health characteristics, and children's preschool language ability. Analyses were comprised of interviews and observational data from 5,721 children and their mothers enrolled in the longitudinal *Growing up in New Zealand* cohort study. Language ability status was measured using the PPVT-III and DIBELS letter naming task (LNF) task at age 4.5 years. Results showed that CR was significantly associated with language status on both measures after controlling for multiple covariates. Improving maternal awareness and support during the perinatal period may reduce the number of risks a fetus is exposed to, which may aid early childhood language development.

Keywords: cohort, longitudinal, cognition, language, child, antenatal, perinatal, risk factors, cumulative risk

[1]School of Psychology, The University of Auckland, Auckland, New Zealand; [2]Center for Longitudinal Research – He Ara ki Mua, The University of Auckland, Auckland, New Zealand; [3]School of Population Health, The University of Auckland, Auckland, New Zealand.

Introduction

An understanding of words and letters lays the foundation for further learning, including reading and writing (Whitehurst & Lonigan, 1998), and has long-term effects on academic achievement, health, and future opportunities (Kaplan, Damphousse, & Kaplan, 1994; Masten et al., 2005; McCaul, Donaldson, Coladarci, & Davis, 1992; Obradović Burt, & Masten, 2009; Welsh, Nix, Blair, Bierman, & Nelson, 2010). The brain structure from birth to 5 years old provides the optimal time and conditions for language acquisition, with the peak period believed to occur around 16-18 months (Fox, Levitt, & Nelson, 2010; Phillips & Shonkoff, 2000; Werker & Tees, 2005). Phonological sensitivity, for example, was reported to have an 83% stability from the approximate ages of 4-6 years and 95% from the approximate ages of 7-9 years (Wagner et al., 1997).

Lower-ability language development is one of the most prevalent developmental challenges among preschool-aged children, and is often accompanied by a lifetime of social, emotional, academic, and economic challenges (Wake et al., 2012). Perhaps most troubling are associations between speech and language impairments and psychiatric disorders (Clegg, Hollis, Mawhood, & Rutter, 2005), including attention deficit problems, behavioural difficulties, and anxiety (Baker & Cantwell, 1987a, 1987b; Beitchman et al., 1986; Clegg et al., 2005; King, Jones, & Lasky,1982). Therefore, not only does lower-ability language have a negative effect on school life, it also has lasting connotations for a child's development into adulthood and future success. For the purpose of this study, language ability was assessed by measuring receptive language and early literacy.

Galvin, A., Davis, G., Neumann, D., Underwood, L., Peterson, E. R., Morton, S.M.B., & Waldie, K. E. (2020). Risk Factors Associated With Language Delay in Preschool Children. *International Journal for Research in Learning Disabilities, 4*(2), 35-52. https://doi.org/10.28987/ijrld.4.2.35

Perinatal Influences on Language Status

The development of language ability is determined by a multitude of biological and environmental factors, including gender (Halpern, 2012). In research on cognition and sex differences, it appears that while, overall, males and females are far more similar than they are different with respect to general intellect, there are nuances between the sexes with regard to the narrow abilities that make up general cognitive capacity (Hyde, 2005). As such, there is widely established evidence of a typical gendered variation throughout cognitive development, with preschool-aged girls developing verbal and language fluency, complex play, and planning skills (Barbu, Cabanes & Le Maner-Idrissi, 2011; Kramer, Delis, Kaplan, O'Donnell, & Prifitera, 1997; Unterrainer, Ruh, Loosli, Heinze, Rahm, & Keller, 2013) at an earlier age than their male counterparts. Thus, it is important to control for sex when investigating the development of language ability.

The effects of maternal alcohol consumption during pregnancy on a child's cognitive abilities, including language development, are well established (Davies et al., 2011; Lairoque, Kaminski, Dehaene, Subtil, Delfosse, & Querleu, 1995; Russell, Czarnecki, Cowan, McPherson, & Mudar, 1991). Thus, consumption of alcohol during pregnancy can result in a host of developmental abnormalities, with fetal alcohol spectrum disorder being the direct result of prenatal alcohol use. One study found that both alcohol consumption and preterm birth were associated with lower-ability language development (Peyre et al., 2014). Indeed, preterm delivery is also associated with lower-ability language status and communication impairments (Boyle et al., 2012; Stene-Larsen, Brandlisstuen, Lang, Landolt, Latal, & Vollrath, 2014; Zambrana, Vollrath, Sengpiel, Jacobsson, & Ystrom, 2015). Similarly, smoking cigarettes during pregnancy can induce a plethora of health problems for a mother and her unborn child (Abel, 1980). This is also the case for offspring brain growth in that prenatal smoking produces neurotoxic effects on the developing fetal brain (Ekblad, Korkeila, & Lehtonen, 2015).

Parity has been considered a predictor factor in a child's development. The firstborn child is overrepresented among Nobel Prize winners, tertiary education populations, and, on average, eldest children have higher academic and intelligence test scores than their later-born siblings (Adams, 1972; Altus, 1965; Arap-Maritim, 2009; Clark & Rice, 1982) even though the effects are small (Ernst & Angst, 2012). Cognitive ability may even be stratified ordinally among siblings, with first-borns outperforming second-born siblings, who outperform third-borns, and so on (Belmont & Marolla, 1973). Perhaps because a mother with multiple children has less time to focus on each child individually, resulting in a lack of necessary stimuli, children with two or more siblings are more likely to have delayed language ability (Halpern, Giugliani,Victora, Barros, & Horta, 2002; Quevedo et al., 2011). However, when factors such as family size and parental education are taken into account, parity no longer predicts cognitive ability (Damian & Roberts, 2015; Kanazawa, 2012; Rodgers, Cleveland, van den Oord, & Rowe, 2000).

Though the subject is widely debated, there is some evidence that children born as a result of fertility treatments, such as in-vitro fertilization (IVF), are at greater risk of cognitive delay (Sandin, Nygren, Iliadou, Hultman, & Reichenberg, 2013). Another form of reproductive technology, intracytoplasmic sperm injection (ICSI), has also been shown to put offspring at a higher risk of lower IQ scores compared to children conceived without assistance (Knoester et al., 2009). However, other studies have not found significant differences in cognition for children born via IVF or ICSI treatments compared to those conceived without assistance (Bay, 2014; Bay, Mortensen, & Kesmodel, 2014; Leslie, Gibson, McMahon, Cohen, Saunders, & Tennant, 2003).

There is a biological component to language learning (Rice, 2013; Tomblin, 2009; Vargha-Khadem, Watkins, Alcock, Fletcher, & Passingham, 1995), and, in addition, biology also interacts with the environment, making the role of parents and caregivers an important aspect of a child's language development (Hoff, 2006; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 199; Rowe, 2008). Because mothers largely constitute infants' social environment, mediating between children and their world to facilitate learning, they play a key role in their child's development (Bradley & Caldwell, 1984; Broen, 1972; Pan, Rowe, Singer, & Snow, 2005). For instance, delayed language and cognitive abilities are correlated with socioeconomic status (SES) in that children born into disadvantaged, low-SES families may not realise and fulfil their true potential (Davies, Crothers, & Hanna, 2010; Hart & Risley, 2003; Schady et al., 2015). The language ability gap between social classes is potentially due to lesser exposure to learning opportunities or materials and limited access to quality educational experiences (Willingham, 2012). Counter to a child from a less affluent background, advantaged children tend to have a relative abundance of resources and stimulating opportunities (Bradley & Corwyn, 2002).

Maternal Health Influences on Language Status

Mothers' mental health is a strong predictor of their children's mental health and cognition (Kim, Bale, & Epperson, 2015). Women are particularly vulnerable to adverse mental health during pregnancy (Waldie et al., 2015) and up to a year afterward (Underwood, Waldie, D'Souza, Peterson, & Morton, 2016). In particular, maternal stress, anxiety, and depression are common mental health issues in women of childbearing age (Lancaster, Gold, Flynn, Yoo, Marcus, & Davis, 2010; Woods, Melville, Guo, Fan, & Galvin, 2010). Laplante et al. (2004) found that toddlers whose mothers had been pregnant with them during a natural disaster displayed lower language development scores than standardized norms. Other studies have found that antenatal stress and anxiety can lead to changes in fetal brain structure and function as a result of reduced blood flow (Fisk & Glover, 1999; Lupien, McEwen, Gunnar, & Heim, 2009; Welberg & Seckl, 2001). For example, irregularities have been found in the orbito-frontal cortex, leading to impairments in working memory and a reduced ability to regulate behaviour and attention, which can result in attention deficit-hyperactivity disorder (ADHD) and externalising problems (Dieter, Emory, Johnson, & Raynor, 2008; Heron, O'Connor, Evans, Golding, Glover, & ALSPAC Study Team, 2004; Kinsella & Monk, 2009; Mennes, Stiers, Lagae, & Van den Bergh, 2006). Further, postnatal depression has been found to be detrimental to children's language learning due to poorer quality communication or fewer interactions between mother and child (Grace, Evindar, & Stewart, 2003; McLearn, Minkovitz, Strobino, Marks, & Hou, 2006; Rothbart, 2011; Saarni, Campos, Camra, & Witherington, 1998).

Other maternal factors that are related to offspring language include the planning of pregnancy (de La Rochebrochard & Joshi, 2013), body mass index scores (BMI), and nutrition status. For example, children of unplanned pregnancies have poorer verbal abilities (Carson, Kelly, Kurinczuk, Sacker, Redshaw, & Quigley, 2011) and perform worse on reading and math tasks than children whose mothers had planned their pregnancies (Joyce, Kaestner, & Korenman, 2000). Women whose pregnancies are unintended are more likely to engage in smoking, alcohol consumption, and illicit drug use, and less likely to take vitamins than mothers whose pregnancies were planned (Than, Honein, Watkins, Yoon, Daniel, & Correa, 2005). For example, Han, Nava-Ocampo, and Koren (2005) found that women with unintentional pregnancies were more likely to be exposed to cigarette smoke, X-rays, alcohol, and potentially harmful medications.

Maternal BMI classified as overweight or obese is associated with poor child cognitive outcomes (Daraki et al., 2017; Pugh et al., 2015). Further, maternal folate or folic acid intake during pregnancy has been strongly linked with irregularities in fetal brain development and greater risk of developing neural tube defects (De Wals et al., 2007; Roza, van Batenburg-Eddes, & Steegers, 2010) and behaviour problems (D'Souza, Waldie, Peterson, Underwood, & Morton, 2019). Understanding the effects of maternal health on child growth and development is essential. Hence, determining the relationship between maternal health, sociodemographic indicators, and early language development is the primary focus of this research.

With a growing number of longitudinal studies on health and development emerging, it has become possible to investigate more rigorously the extent to which vital influences such as prenatal and early-life factors impact developmental outcomes (Maccani & Marsit, 2009; Wang, Walker, Hong, Bartell, & Wang, 2013) and informing theory and practice within both the scientific community and health initiatives around the world (Poulton, Moffitt, & Silva, 2015). While the relationships between environmental conditions and language development require further investigation (Peterson & Pennington, 2015), we do know that unfavourable conditions during the gestational and perinatal periods negatively impact offspring development in multiple spheres (Liu et al., 2016).

Cumulative Risk and Language Ability

Though earlier research on this subject is limited, there has been some investigation into the relationship between cumulative risks (CR) and language ability. For example, Stanton-Chapman, Chapman, Kaiser, and Hancock (2004) looked at the impact of CR on language development in 3-year-old children from low-income families. Their CR model, which consisted of social and environmental risks present when the child was born, was significantly related to language development. Their findings showed that an increase in risks amplified the likelihood of lower-ability language status (Stanton-Chapman et al., 2004). Similarly, Burchinal, Roberts, Hooper, and Zeisel (2000) found an association between cumulative sociocultural and caregiving risks with language development throughout the first four years of African American children's lives. This study, focused on home environment, maternal IQ, and a mother's responsiveness to her child as components of CR, revealed that the probability of lower-ability language development increased alongside CR exposure (Burchinal et al., 2000).

Conversely, Sylvestre and Mérette (2010) found that CR had no significant impact on a child's language development, but that individual risk factors had a greater influence. Their longitudinal sample of children from 2 to 36 months old who had experienced parental neglect was more likely to have a language ability if they also had lower cognitive development, whereas the other psychological and biological risk factors were not significantly associated with language status when examined through a CR model (Sylvestre & Mérette, 2010).

Contradictory findings such as those cited above emphasize the need for further study on the relationship between CR and language development. In the current study, we utilised data from 6,822 *Growing up in New Zealand* study children and their families. We first catalogued risk factors that have been associated with delayed cognitive development: exposure to alcohol and cigarette smoke, maternal depression and/or anxiety during pregnancy, born with low birth weight, not being the first-born child, premature or overdue birth, mother not taking folate during the pregnancy, overweight or underweight maternal BMI, use of any fertility treatments, presence of any birth complications, and unplanned pregnancy.

We expected that children who are exposed to more risks are more likely to be classified as having lower-language ability than individuals exposed to fewer risks. As such, the current research fills the need for a longitudinal study that defines the cumulative contributions of certain gestational, perinatal, and postpartum risk factors of emergent lower-ability language development.

Methods

Participants and General Procedure

Participants consisted of expectant mothers and their children (N = 6,822; 52% male) enrolled in the *Growing up in New Zealand* (GUiNZ) longitudinal cohort study. The study focused on 5,721 children and their mothers who participated in our language ability assessment and provided personal information related to our CR index. This cohort is comprised of a socioeconomically and ethnically diverse sample of children that is broadly generalizable to current NZ births, coming from a geographical area that approximately mirrors one third of the New Zealand (NZ) birth population (Morton et al., 2012). Mothers were invited to participate if they were due to give birth between April 25, 2009, and March 25, 2010, and were residing in three bordering District Health Boards (DHBs) in the North Island of NZ: Auckland, Counties Manukau, and Waikato.

Data collection waves (DCWs) took place at various points in the child's development. Antenatal data collection occurred during the third trimester of pregnancy, with a face-to-face computer-assisted interview. Postnatal data collection was carried out nine months after childbirth, at ages 2 and 4.5 years old, using the same method as the antenatal data collection, as well as child observations at age 2 and 4.5 years.

Measures

DIBELS Letter Naming Fluency task. The Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Letter Naming Fluency (LNF) task was used at the 4.5-year DCW. The DIBELS LNF task is a standardised, individually administered test that indicates a child's early reading fluency (University of Oregon, 2019). It is designed for children entering primary school, who are 4.5–6 years of age.

The child of interest is presented with a page of upper- and lowercase letters, randomly ordered. The child has one minute to name as many letters as possible, providing the name of the letter rather than the sound it produces (Good & Kaminski, 2007). The child must read successively, across the page, pointing to each letter while saying its name. Children who cannot name any of the presented letters are considered at risk for developing literacy issues (Walsh, Price, & Gillingham, 1988). The DIBELS was an ideal measurement for the present study as it has been validated in a New Zealand longitudinal sample across primary school (Suggate, Schaughency, & Reese, 2012).

Peabody Picture Vocabulary Test (PPVT-III). An adapted version of the Peabody Picture Vocabulary Test (PPVT-III), developed by the Longitudinal Study of Australian Children (LSAC), was administered at 4.5 years to measure receptive language in English. The test involves an examiner revealing a series of images (four to a page) and reciting a word that corresponds to one of the images (Taylor et al., 2013). The child is tested on the ability to point to the correct image.

For the current study, PPVT-III scores were calculated by taking the total number of correct responses and submitting it to an item response theory factor analysis. The scores, which ranged from 0-40, were adjusted and converted to a z score. Children who scored one standard deviation below the mean were categorized as having lower-ability receptive language, and all other children were scored as not having lower ability. This is the standardized scoring technique for the task (Dunn & Dunn, 2007).

PPVT and DIBELS language ability outcome groups. In order to identify the children who were most at risk of developing late literacy, the PPVT and DIBELS scores were analysed separately. The DIBELS was completed by 5,409 (79%) children, whereas 5,533 children completed the PPVT task (81%) at 4.5 years of age. If the child was considered to have a lower-language ability on one or both measures, he or she was classified as having lower-language ability.

Because language ability was assessed only in English, the language ability of bilingual/multilingual children may be underestimated. Of those categorized here as lower ability, 52.5% spoke only English (n = 892) and 47.3% were bilingual or multilingual (n = 806). Within the higher-language ability group, 74.7% spoke only English, and 25.3% were bilingual or multilingual.

Though there is a discrepancy in the proportion of bilingual/multilingual children between groups, this is a multifaceted variable that cannot be easily controlled for. Because this is a complex subject, further investigation on the relationship between bilingual/multilingual status and language ability requires a separate study. However, all of the participating children were born in New Zealand, where English is the prominent language. It will be of interest to assess this association in future followup data collections and to directly test how multilingualism impacts cognitive development.

Cumulative risk variables. A cumulative risk (CR) index was developed for antenatal and perinatal predictor variables by (a) dichotomizing all predictor variables into 1 = risk or 0 = not at risk categories; and (b) adding all risk categories into a continuous CR variable. Multiple risk metrics are generally untenable due to low statistical power (Evans, Li, & Whipple, 2013).

Antenatal depression symptoms were assessed using the Edinburgh Postnatal Depression Scale (EPD; Cox, Holden, & Sagovsky, 1987), a 10-item scale designed to identify depressive symptoms during pregnancy and the postnatal period. Although the EPDS was designed for screening depression postnatally, it has been demonstrated to be a valid measure of depression in pregnant women (Gibson, McKenzie-McHarg, Shakespeare, Price, & Gray, 2009; Kozinszky & Dudasm, 2015; Murray & Cox, 1990). A cut-off of 13 was used to indicate clinically significant levels of depression.

Perceived maternal stress was measured using the abbreviated version (10 items) of the Perceived Stress Scale (PSS; Cohen, Karmack, & Mermelstein, 1983), which has been reported to have validity for use during pregnancy (Solivan, Xiong, Harville, & Buekens, 2015). Items on the PSS classify the level of stress to which participants assign daily activities. The maximum score for this measure is 40, with higher scores signifying greater levels of stress. Perceived stress scores at or above the median were considered as indicating a risk (Lamb et al., 2014).

Experiences of antenatal anxiety and panic attacks were assessed using modified questions from the New Zealand Health Survey (Ministry of Health, 2007). Women were asked if they had experienced anxiety or panic attacks before, during, or both before and during their current pregnancy. Those who indicated they had any anxiety and/or panic attacks during their pregnancy were classified as having antenatal anxiety.

Mothers were also asked about their alcohol consumption both before and during their current pregnancy throughout each trimester using questions derived from New Zealand's National Nutrition Survey (Russell et al., 1999); any alcohol consumption during the pregnancy was considered a risk.

Based on maternal self-report, the following were also considered risks: not having taken folate or folic acid as a supplement at any time during or before their pregnancy and presence of maternal smoking during pregnancy or second-hand smoke exposure during pregnancy. Finally, fertility awareness and weight loss, ovulation induction (with clomiphene citrate in the community or in an infertility clinic), tubal surgery, and IVF treatment or other specified fertility treatments were also assessed as risk factors.

Maternal BMI was assessed using self-reported height and weight pre-pregnancy against standard measures from the World Health Organisation (WHO), calculated as (weight (kg)/height (m2) (WHO, 2017). BMI measures were then categorically grouped with a BMI of ≤ 18.5 as underweight, ≥ 18.5 as normal, ≥ 25kg/m2 defined as overweight, and ≥ 30 kg/m2 as obese. Having maternal BMI classified as underweight, overweight, or obese was considered a risk.

Not being a firstborn child, presence of any birth complications, being premature (less than 37 weeks) or overdue (post-term was categorised above 41 weeks' gestation) were all risk factors. Having a high (>4000 grams) or low (<2500 grams) birth weight and if the pregnancy was unplanned were also categorised as risks.

With regard to frequency of risks, the highest number was 12 risks (out of a total of 14 possible risk factors); the lowest was zero risks (2.6% of the sample). Almost a third of participants reported being exposed to two or three risks (31.4%). A total of 61.1% of children were exposed to three or more risks. The risk categories were added for an overall CR variable.

In this context, the term *risk* is not necessarily a direct causative suggestion, but more loosely referred to as situations that have been correlated in research with putting certain individuals at a higher likelihood of poor cognitive outcomes. In order to be included in our analyses, individuals must have at least seven data points if they had a risk factor. If they had no risk factors, they needed to have a minimum of 10 data points in order to reduce bias. Based on these criteria, 62 individuals who did not meet the threshold were omitted from the analyses.

Control Variables

A range of variables previously found to be associated with language development were controlled for in the current study. Maternal variables measured at the antenatal DCW included mother's ethnicity (European, Māori, Pacific, Asian, Other), mother's education (no secondary school, secondary school/diploma/trade certificate, Bachelor's degree or higher), and mother's age when pregnant (≤ 30 years, or ≥ 31 years). Further control variables from the antenatal period included area-level deprivation (high, medium, and low) related to socioeconomic status and based on the New Zealand deprivation index (Salmond et al., 2007), and rurality (urban or rural). The analyses also controlled for child sex and age (in days) when assessed at the 4.5-year DCW.

Data Analysis

Following descriptive statistics, a hierarchical multivariable logistic regression was performed to ascertain the effect of CR on the likelihood of having lower-language ability, controlling for sociodemographic variables. All risk variables were included, with analyses adjusting for the seven sociodemographic covariates.

All statistical analyses were completed using IBM SPSS Statistics Version 25; statistical significance for the analyses was given at a *p* value of 0.05.

Results

Of the 6,822 individuals enrolled in the GUiNZ cohort, our study included the 5,721 for whom complete information was available on risk exposure as a function of language ability. Of this group, 5,533 children completed the PPVT, and 5,409 completed the DIBELS.

As shown in Table 1, mothers tended to be 31 years of age or older (52.3%), were predominantly European (57.5%), and had completed some form of education (93.7% had gained at least secondary school qualifications or higher). Area deprivation groups were similar, with around a third of participants coming from low (26.3%), medium (37.6%), or highly (36.1%) deprived areas; a majority of participants came from urban living areas (92.6%). Most women did not consume any alcohol (71.3%), smoke (90.2%), or expose themselves to second-hand smoke (93.4%) during pregnancy. Folate intake pre-pregnancy and/or during the first trimester was prevalent, with only 17% of women reporting taking no folate during the early stages of pregnancy. Fewer than half, 44.8%, of women in the study had BMIs that put them as being underweight, overweight, or obese. Regarding mental health, 11.5% of the women were identified as having antenatal depression symptoms, and 3.4% had doctor-diagnosed anxiety and panic attacks during their pregnancy. Fertility treatment was used to conceive a child by 10.3% of the women, and 61.9% of the women planned their pregnancy. With regard to birth factors, 34.3% of children were born via an assisted delivery, and 13.5% of children experienced birth complications. Over half the children were not firstborn, and 91.8% of the children were born at term with 78.8% at an appropriate birth weight.

A significant association was found between the PPVT and CR category, $X^2(14) = 77.68$, $p<.001$. The model for the DIBELS associated with CR was also significant, $X^2(14) = 391.41$, $p<.001$. Of the 5,533 children who took the PPVT, 1,202 (21.7%) were considered as having a lower-language ability. The DIBELS reflected that 1,010 (18.7%) of the total 5,409 were considered at lower ability. Of the individuals who partook in both assessments, 68.1% were within the higher-ability group for both tests, whereas 31.9% scored within the lower-ability group for one or both tests.

Logistic Regression

Table 2 shows the associations between covariates and CR variables and the language ability out-

Table 1
Descriptive Sociodemographic Covariates of Pregnant Women and Children and Perinatal Predictors Deriving Cumulative Risk Index

Variable		N	%
Mother Age			
	Younger[1]	2,727	47.7
	Older[2]	2,994	52.3
Child Age During 4.5 Year DCW			
	Younger[3]	2,818	49.3
	Older[4]	2,899	50.7
Ethnicity			
	European	3,240	57.5
	Māori	745	13.2
	Pacifica	715	12.7
	Asian	756	13.4
	MELAA, other, NZer[5]	179	3.2
Deprivation			
	Low	1,504	26.3
	Medium	2,148	37.6
	High	2,068	36.1
Education			
	No Secondary School Qualification	361	6.3
	Secondary School Qualification	1,290	22.6
	Diploma or Trade Certificate	1,743	30.5
	Bachelor's Degree	1,378	24.1
	Higher Degree	941	16.5
Child Gender			
	Boy	2,948	51.5
	Girl	2,773	48.5
Rurality			
	Urban	5,298	92.6
	Rural	423	7.4
Alcohol Consumption During Pregnancy			
	Yes	1,638	28.7
	No	4,077	71.3
Smoking During Pregnancy			
	Yes	512	9.8
	No	4,716	90.2
Prenatal Secondary Hand Smoke Exposure			
	Yes	346	6.6
	No	4,882	93.4
Antenatal Depression			
	Depression	592	11.5
	No Depression	4,570	88.5
Perceived Stress			
	Higher Stress	2,663	50.9
	Lower Stress	2,567	49.1

Table 1 (cont.)

Variable		N	%
Anxiety During Pregnancy			
	Anxiety	193	3.4
	No Anxiety	5,522	96.6
Fertility Treatment			
	Yes	366	10.3
	No	3,161	88.7
Maternal BMI			
	Underweight, Overweight or Obese	2,276	44.8
	Normal BMI	2,805	55.2
Folate During Pregnancy			
	No folate	887	17.0
	Pre-pregnancy and/or Trimester-One Intake	4,337	83.0
Delivery Type			
	Assisted Birth	1,956	34.3
	Spontaneous Vaginal	3,753	65.7
Parity			
	Firstborn	2,392	41.8
	Subsequent Born	3,329	58.2
Birth Complications			
	Yes	772	13.5
	No	4,949	86.5
Term			
	Term	5,247	91.8
	Pre- or Post-Term	471	8.2
Birth Weight			
	Appropriate	4,505	78.8
	Low or High	1,212	21.2
Pregnancy Planned			
	Yes	3,485	61.9
	No	2,144	38.1

[2] Mothers who were 30 years of age and/or younger while pregnant.
[3] Mothers who were 31 years of age and/or older while pregnant.
[4] Children who were younger than 4.5 years of age during this DCW.
[5] Children who were older than 4.5 years of age during this DCW.
[6] MELAA = Middle Eastern, Latin American, or African; NZer = New Zealander.

come at the 4.5 year DCW. Table 3 reflects he associations between covariates, CR, and the PPVT. Finally, Table 4 depicts the associations between covariates, CR, and the DIBELS.

The following covariates were significantly associated with language ability: child's age, mother's age, mother's ethnicity, gender, deprivation, and maternal education. The analysis also revealed that CR was significantly associated with language ability ($p = 0.014$) after controlling for the covariates.

Discussion

The aim of this study was to explore the relationship between mothers' CR factors and their child's language development. At 4.5 years of age, children were tested for English language ability using the PPVT-III and the DIBELS, and were either considered lower-ability in language development or higher-ability based on the two test results. A CR model comprised of 14 gestational, perinatal, and postpartum risk factors was created to assess

Table 2
Associations Between Covariates and Cumulative Risk Variable and Language Ability at Age 4.5 Years

Variable		*B* (SE)	OR (95% CI)
Mother Age*			
	Younger[6]	0.23 (0.07)	1.26 (1.10, 1.43)
	Older[7]		
Child Age During 4.5 Year DCW*			
	Younger[8]	-0.002 (0.001)	1.00 (1.00, 1.00)
	Older[9]		
Ethnicity*			
	European		
	Māori	1.02 (0.10)	2.77 (2.29, 3.37)
	Pacifica	1.46 (0.11)	4.30 (3.46, 5.31)
	Asian	0.88 (0.10)	2.42 (2.01, 2.92)
	MELAA, Other, NZer[10]	0.56 (0.18)	1.75 (1.23, 2.47)
Deprivation*			
	Low		
	Medium	0.42 (0.93)	1.52 (1.26, 1.82)
	High	-0.005 (0.09)	1.00 (0.84, 1.80)
Education*			
	No Secondary School Qualification	0.53 (0.15)	1.70 (1.27, 2.27)
	Secondary School Qualification	0.40 (1.00)	1.50 (1.23, 1.82)
	Diploma or Trade Certificate	0.44 (0.09)	1.56 (1.30, 1.86)
	Bachelor's Degree	-0.02 (0.11)	0.98 (0.79, 1.22)
	Higher Degree		
Child Gender*			
	Boy	0.42 (0.06)	1.53 (1.34, 1.73)
	Girl		
Rurality			
	Urban		
	Rural	0.20 (0.13)	1.22 (0.95, 1.56)
Cumulative Risk*			
(continuous)		0.07 (0.02)*	1.07 (1.03, 1.10)*

*Risk categories = $p < .05$.

[7] Mothers who were 30 years of age and/or younger while pregnant.
[8] Mothers who were 30 years of age and/or younger while pregnant.
[9] Children who were younger than 4.5 years of age during this DCW.
[10] Children who were older than 4.5 years of age during this DCW.
[11] MELAA = Middle Eastern, Latin American, or African; NZer = New Zealander.

Table 3
Associations Between Covariates and Cumulative Risk Variable and PPVT at Age 4.5 Years

Variable		*B* (SE)	OR (95% CI)
Mother Age*			
	Younger[11]	0.26 (0.08)	1.30 (1.12, 1.51)
	Older[12]		
Child Age During 4.5 Year DCW*			
	Younger[13]	-0.002 (0.001)	1.00 (1.00, 1.00)
	Older[14]		
Ethnicity*			
	European		
	Māori	1.07 (0.11)	2.92 (2.36, 3.61)
	Pacifica	1.65 (0.11)	5.21 (4.20, 6.48)
	Asian	1.42 (0.11)	4.15 (3.38, 5.10)
	MELAA, Other, NZer[15]	0.87 (0.20)	2.38 (1.63, 3.50)
Deprivation*			
	Low		
	Medium	0.49 (0.11)	1.64 (1.32, 2.03)
	High	0.10 (0.11)	1.11 (0.90, 1.37)
Education*			
	No Secondary School Qualification	0.59 (0.16)	1.81 (1.33, 2.46)
	Secondary School Qualification	0.50 (0.11)	1.66 (1.33, 2.06)
	Diploma or Trade Certificate	0.39 (0.11)	1.48 (1.20, 1.82)
	Bachelor's Degree	-0.05 (0.14)	0.95 (0.73, 1.25)
	Higher Degree		
Child Gender*			
	Boy	0.34 (0.07)	1.41 (1.22, 1.62)
	Girl		
Rurality			
	Urban		
	Rural	0.20 (0.13)	1.22 (0.95, 1.56)
Cumulative Risk* (continuous)		0.06 (0.02)*	1.06 (1.02, 1.10)*

*Risk categories = $p < .05$.

[12] Mothers who were 30 years of age and/or younger while pregnant.
[13] Mothers who were 30 years of age and/or younger while pregnant.
[14] Children who were younger than 4.5 years of age during this DCW.
[15] Children who were older than 4.5 years of age during this DCW.
[16] MELAA = Middle Eastern, Latin American, or African; NZer = New Zealander.

Table 4
Associations Between Covariates and Cumulative Risk Variable and DIBELS at Age 4.5 Years

Variable		*B* (SE)	OR (95% CI)
Mother Age*			
	Younger[16]	0.06 (0.08)	1.07 (0.91, 1.24)
	Older[17]		
Child Age During 4.5 Year DCW*			
	Younger[18]	-0.002 (0.001)	1.00 (1.00, 1.00)
	Older[19]		
Ethnicity*			
	European		
	Māori	0.84 (0.11)	2.32 (1.88, 2.87)
	Pacifica	0.96 (0.12)	2.60 (2.08, 3.26)
	Asian	-0.25 (0.14)	0.78 (0.59, 1.08)
	MELAA, Other, NZer[20]	0.22 (0.21)	1.24 (0.82, 1.89)
Deprivation*			
	Low		
	Medium	0.32 (0.11)	1.38 (1.11, 1.72)
	High	0.04 (0.10)	1.04 (0.85, 1.27)
Education*			
	No Secondary School Qualification	0.69 (0.16)	1.99 (1.45, 2.72)
	Secondary School Qualification	0.40 (0.12)	1.49 (1.18, 1.88)
	Diploma or Trade Certificate	0.51 (0.11)	1.66 (1.34, 2.06)
	Bachelor's Degree	-0.02 (0.14)	0.98 (0.74, 1.28)
	Higher Degree		
Child Gender*			
	Boy	0.37 (0.07)	1.45 (1.25, 1.68)
	Girl		
Rurality			
	Urban		
	Rural	0.33 (0.14)	1.38 (1.05, 1.82)
Cumulative Risk*			
(continuous)		0.07 (0.02)*	1.07 (1.03, 1.11)*

*Risk categories = $p < .05$.

[17] Mothers who were 30 years of age and/or younger while pregnant.
[18] Mothers who were 30 years of age and/or younger while pregnant.
[19] Children who were younger than 4.5 years of age during this DCW.
[20] Children who were older than 4.5 years of age during this DCW.
[21] MELAA = Middle Eastern, Latin American, or African; NZer = New Zealander.

the degree to which risk exposure affects language. Through logistic regression, the relation between CR and language status was analysed in 5,721 children and their mothers.

Our study revealed a high prevalence rate of lower-ability language status with almost a third (31.9%) of children presenting a lower language ability based on our criteria. Thus, the greater number of maternal risks children are exposed to, the more likely they are to experience delays in language development. All controlled covariates were significant except for rurality.

As expected, the number of maternal risks are strongly linked to a child's language ability. Prior research assessing the relationship between CR and language development with 3-year-olds of low-income families (Stanton-Chapman et al., 2004) found a relationship similar to that of our study. And like our study, the accumulation of risk factors increased the likelihood of delayed language development (Stanton-Chapman et al., 2004). Burchinal and colleagues (2000) also found that an increase in CR exposure led to an increase in lower language ability in children. However, Sylvestre and Mérette (2010) noted that CR had an insignificant impact on a neglected child's language development. Rather, these researchers found that individual risks explain language ability more concisely than a CR model. The children in all three of these studies were neglected by their parents or came from low-SES families. Our research differs as our CR model focuses on gestational, perinatal, and postpartum risk factors while controlling for sociodemographic covariates, such as deprivation, from a large and diverse group of mothers and children.

Despite conflicting results with regard to CR and language ability, past research on individual risks supports our findings. All risks included in our CR model were associated with cognitive development. Our study found that almost a third of mothers had consumed alcohol during pregnancy, which is an established risk factor for language and cognitive delays (Davies et al., 2011; Lairoque et al., 1995; Russell et al., 199). More than half of the children were not first-borns, a status associated with poorer cognitive ability (Belmont & Marolla, 1973). Almost half of mothers were underweight, overweight, or obese according to their BMI, a health concern associated with a child's lower-ability cognitive outcomes (Daraki et al., 2017; Pugh et al., 2015). Further, more than one third of the pregnancies were unplanned, a risk connected with lower verbal ability (Carson et al., 2011).

Though the mechanism behind language ability is relatively unknown, several of the risk factors included in this study are known to alter fetal development. For example, maternal perceived stress, which was experienced by a majority of mothers in this study, can negatively impact the intrauterine environment, leading to bio-behavioural abnormalities in newborns (Dieter et al., 2008). While there are no direct neural connections between mother and fetus (Dieter et al., 2008), the fetal environment is altered by maternal stress as there is a connection between maternal and fetal cortisol levels (Talge, Neal, & Glover, 2007), with an estimated 10-20% of maternal cortisol passing through to the fetus (Gitau, Cameron, Fisk, & Glover, 1998; Glover, Bergman, Sarkar, & O'Connor, 2009). Further, recent research from a longitudinal birth cohort found that adolescents carrying the TT genotype of the rs12193738 polymorphism on the KIAA0319 gene exposed to high maternal stress during pregnancy had significantly poorer reading abilities than offspring exposed to low maternal stress (D'Souza et al., 2016). Therefore, research now focuses on a pregnant mother's hypo-thalamic-pituitary-adrenal (HPA) axis as the underlying mechanism for long-term effects on a child's development (Talge et al., 2007). Nevertheless, more research regarding the mechanisms behind other risk factors is needed to determine the biological origins for language ability.

Strengths and Limitations

There are some important limitations to this study that must be taken into consideration when reviewing the findings. The antenatal data were collected late during the mother's pregnancy, making the results subject to potential bias in recalling information regarding pre- and early pregnancy. Also, as with many longitudinal studies, there are cases of missing data that relay a different sociodemographic distribution, which limits the overall generalisability of the results. Further, the CR approach limits risk intensity, making it impossible to determine the statistical interactions between risk factors (Evans et al., 2013). That is, when using the CR approach, risks are granted equal weight rather than each risk being viewed as having a unique effect, done in traditional multiple regression (Hall et al., 2010). This method is also subject to loss of information as individual measures are dichotomised, potentially creating an erroneous statistical relationship (Hall et al., 2010; MacCallum, Zhang, Preacher, & Rucker, 2002).

Another limitation is that we only measured English-language skills and did not control for multilingual status. The association between bilingualism/multilingualism and language ability is complex, and we acknowledge that some children may have ended up in our lower-ability category because of their non-monolingual status. However, all study members were born in New Zealand, where English is the predominant language. We intend to further explore the topic to better understand the effect multilingualism has on language ability.

Further, we did not account for any additional instruction or learning opportunities children might have been exposed to. However, at age 4.5, unless children are being taught by parents to recognise or write letters, they are unlikely to be learning letters on their own. It is possible, but if so, it would apply to a small minority.

Despite these limitations, there are also several strengths to the study. Previous research has mainly focused on the related risk factors of language ability in isolation, whereas the current study examined the effects of cumulative risk. Therefore, this multivariate analysis expands on our understanding of the relevance of both the antenatal environment and the perinatal events throughout early development. Additionally, the participants came from a large and diverse population. With 5,721 children and their mothers, the sample well represented a range of demographics, including ethnic minority and low-SES groups, making our results generalisable to children across New Zealand. Due to the up-to-date and detailed nature of the GUiNZ data, we were able to pinpoint a large variety of antenatal and perinatal risk factors associated with language development and control for a broad range of sociodemographic confounds.

Identification of controllable risk factors that may delay a child's early language ability can aid in the prevention and intervention of lower language ability. Findings support the improvement of education for new mothers on the dangers or risks, such as second-hand smoke exposure during pregnancy, and the importance of folate intake during the first trimester. The ongoing GUiNZ research has great potential to further investigate the long-term effects of CR on offspring language ability throughout development.

Conclusions

The extent of CR exposure has a significant impact on a child's language development. This study contributes to an understanding of CR's influence on pregnancy and language ability. Specifically, our findings emphasize the need for improving maternal education regarding modifiable factors such as smoking, taking vitamins, and keeping within a healthy weight gain during pregnancy. As a child's likelihood of having lower-ability language status increases after being exposed to more than three risks during and surrounding pregnancy, it is important that a mother understands the cumulative power of risks on a child's future language status.

References

Abel, E. L. (1980). Smoking during pregnancy: A review of effects on growth and development of offspring. *Human Biology*, *54*(4), 593-625. Retrieved from https://www.jstor.org/stable/41464571

Adams, B. (1972). Birth order: A critical review. *Sociometry, 35*(3), 411-439. doi:10.2307/2786503

Altus, W. D. (1965). Birth order and academic primogeniture. *Journal of Personality and Social Psychology, 2*(6), 872. http://dx.doi.org/10.1037/h0022705

Arap-Maritim, E. K. (2009). Birth order and selection for university education in Kenya. *Journal of Perceptual and Motor Skills, 109*(2), 387-394. doi:10.2466/PMS.109.2.387-394

Baker, L., & Cantwell, D. P. (1987a). Factors associated with the development of psychiatric illness in children with early speech/language problems. *Journal of Autism and Developmental Disorders*, *17*(4), 499-510. https://doi.org/10.1007/BF01486966

Baker, L., & Cantwell, D. P. (1987b). A prospective follow up of children with speech/language disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *26*, 546-553. https://doi.org/10.1097/00004583-198707000-00015

Barbu, S., Cabanes, G., & Le Maner-Idrissi, G. (2011). Boys and girls on the playground: Sex differences in social development are not stable across early childhood. *PLoS ONE, 6*(1), 16-40. http://doi.org/10.1371/journal.pone.0016407

Bay, B. (2014). Fertility treatment: Long-term growth and mental development of the children. *Danish Medical Journal, 61*(10), 49-60. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25283630

Bay, B., Mortensen, E. L., & Kesmodel, U. S. (2014). Fertility treatment and child intelligence, attention, and executive functions in 5‐year‐old singletons: A cohort study. *BJOG: An International Journal of Obstetrics and Gynaecology, 121*(13), 1642-1651. doi:10.1111/1471-0528.12907

Beitchman, J. H., Nair, R., Clegg, M., Ferguson, B., & Patel, P.G. (1986). Prevalence of psychiatric disorders in children with speech and language disorders. *Journal of the American Academy of Child Psychiatry, 25*(4), 528–535.

Belmont, L., & Marolla, F. A. (1973). Birth order, family size and intelligence. *Science Magazine, 182*(4117), 1096-1101. doi:10.1126/science.184.4133.114

Broen, P. (1972). The verbal environment of the language-learning child. *ASHA Monogr. 17*. Retrieved from https://eric.ed.gov/?id=ED098768

Boyle, E. M., Poulsen, G., Field, D. J., Kurinczuk, J. J., Wolke, D., Alfirevic, Z., & Quigley, M. A. (2012). Effects of gestational age at birth on health outcomes at 3 and 5 years of age: Population based cohort study. *BMJ: British Medical Journal, 344*(7848), 17. doi:https://doi.org/10.1136/bmj.e896

Bradley, R. H., & Caldwell, B. M. (1984). The HOME Inventory and family demographics. *Developmental Psychology, 20*(2), 315-320. http://dx.doi.org/10.1037/0012-1649.20.2.315

Bradley, R. H., & Corwyn, R. F. (2002). Socio-economic status and child development. *Annual Review of Psychology, 53*(1), 371-399. https://doi.org/10.1146/annurev.psych.53.100901.135233

Burchinal, M. R., Roberts, J. E., Hooper, S., & Zeisel, S. A. (2000). Cumulative risk and early cognitive development: A comparison of statistical risk models. *Developmental Psychology*, *36*(6), 793-807. doi:10.1037//0012-1649.36.6.793

Carson, C., Kelly, Y., Kurinczuk, J. J., Sacker, A., Redshaw, M., & Quigley, M. A. (2011, July). Effect of pregnancy planning and fertility treatment on cognitive outcomes in children at ages 3 and 5: Longitudinal cohort study. *BMJ: British Medical Journal*, 343. doi: https://doi.org/10.1136/bmj.d4473

Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental language disorders – a follow-up in later adult life. Cognitive, language and psychosocial outcomes. *Journal of Child Psychology & Psychiatry*, *46*(2), 128-149. doi:10.1111/j.1469-7610.2004.00342.x

Clark, R. D., & Rice, G. A. (1982). Family constellations and eminence: The birth orders of Nobel Prize winners. *Journal of Psychology*, *110*(2), 281-287. http://dx.doi.org/10.1080/00223980.1982.9915350

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behaviour, 24*(4), 385-396. doi:10.2307/2136404

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *The British Journal of Psychiatry, 150*(6), 782-786. https://doi.org/10.1192/bjp.150.6.782

Damian, R. I., & Roberts, B. W. (2015). The associations of birth order with personality and intelligence in a representative sample of U.S. high school students. *Journal of Research in Personality*, *58*, 96-105. https://doi.org/10.1016/j.jrp.2015.05.005

Daraki, V., Roumeliotaki, T., Koutra, K., Georgiou, V., Kampouri, M., Kyriklaki, A., … Chatzi, L. (2017). Effect of parental obesity and gestational diabetes on child neuropsychological and behavioral development at 4 years of age: The Rhea mother-child cohort, Crete, Greece. *European Child & Adolescent Psychiatry*, *26*(6), 703-714. doi:10.1007/s00787-016-0934-2

Davies, E., Crothers, C., & Hanna, K. (2010). Preventing child poverty: Barriers and solutions. *New Zealand Journal of Psychology, 39*(2), 20-31.

Dunn, L., & Dunn, D. (2007). *PPVT-4: Peabody picture vocabulary test manual*. Minneapolis, MN: Pearson.

Davies, L., Dunn, M., Chersich, M., Urban, M., Chetty, C., Olivier, L., & Viljoen, D. (2011). Developmental delay of infants and young children with and without fetal alcohol spectrum disorder in the Northern Cape Province, South Africa. *African Journal of Psychiatry*, *14*(4), 298-305. http://dx.doi.org/10.4314/ajpsy.v14i4.7.

De Wals, P., Tairou, F., Van Allen, M. I., Uh, S. H., Lowry, R. B., Sibbald, B., Evans, J. A., Van den Hof, M. C., Zimmer, P., Crowley, M., & Fernandez, B. (2007). Reduction in neural-tube defects after folic acid fortification in Canada. *New England Journal of Medicine, 357*(2), 135-142. doi:10.1056/NEJMoa067103

Dieter, J. N., Emory, E. K., Johnson, K. C., & Raynor, B. D. (2008). Maternal depression and anxiety effects on the human fetus: Preliminary findings and clinical implications. *Infant Mental Health Journal: Official Publication of The World Association for Infant Mental Health*, *29*(5), 420-441. doi:10.1002/imhj.20192

D'Souza, S., Backhouse-Smith, A., Thompson, J.M.D., Slykerman, R., Marlow, G., Wall, C., … Waldie, K. E. (2016). Associations between the KIAA0319 dyslexia susceptibility gene variants, antenatal maternal stress, and reading ability in a longitudinal birth cohort. *Dyslexia*, *22*(4), 379. doi:10.1002/dys.1534

D'Souza S., Waldie, K. E., Peterson, E. R., Underwood, L., & Morton, S.M.B. (2019). Antenatal and postnatal determinants of behavioural difficulties in early childhood: Evidence from growing up in New Zealand. *Child Psychiatry & Human Development, 50*(1), 45-60. doi:10.1007/s10578-018-0816-6

Ekblad, M., Korkeila, J., & Lehtonen, L. (2015). Smoking during pregnancy affects fetal brain development. *Acta Paediatrica, 104*(1), 12-18. doi:10.1111/apa.12791

Ernst, C., & Angst, J. (2012). *Birth order: Its influence on personality*. New York, NY: Springer Science and Business Media.

Evans, G. W., Li, D., & Whipple, S. S. (2013). Cumulative risk and child development. *Psychological Bulletin, 139*, 1342-1396. https://doi.org/10.1037/a0031808

Fisk, N. M., & Glover, V. (1999). Association between maternal anxiety in pregnancy and increased uterine artery resistance index: Cohort based study. *BMJ: British Medical Journal, 318*(7177), 153-157. doi:https://doi.org/10.1136/bmj.318.7177.153

Fox, S. E., Levitt, P., & Nelson III, C. A. (2010). How the timing and quality of early experiences influence the development of brain architecture. *Child Development*, *81*(1), 28-40. doi:10.1111/j.1467-8624.2009.01380.x

Gibson, J., McKenzie-McHarg, K., Shakespeare, J., Price, J., & Gray, R. (2009). A systematic review of studies validating the Edinburgh Postnatal Depression Scale in antepartum and postpartum women. *Acta Psychiatrica Scandinavica, 119*(5), 350-364. doi:10.1111/j.1600-0447.2009.01363.x

Gitau, R., Cameron, A., Fisk, N. M., & Glover, V. (1998). Fetal exposure to maternal cortisol. *The Lancet*, *352*(9129), 707-708. doi:10.1016/S0140-6736(05)60824-0

Glover, V., Bergman, K., Sarkar, P., & O'Connor, T. G. (2009). Association between maternal and amniotic fluid cortisol is moderated by maternal anxiety. *Psychoneuroendocrinology, 34*(3), 430-435. doi:10.1016/j.psyneuen.2008.10.005

Good, R. H., & Kaminski, R. A. (Eds.). (2007). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement. Retrieved from http://dibels.uoregon.edu/

Grace, S. L., Evindar, A., & Stewart, D. E. (2003). The effect of postpartum depression on child cognitive development and behavior: A review and critical analysis of the literature. *Archives of Women's Mental Health*, *6*(4), 263-274. doi:10.1007/s00737-003-0024-6

Hall, J. E., Sammons, P., Sylva, K., Melhuish, E., Taggart, B., Siraj-Blatchford, I., & Smees, R. (2010). Measuring the combined risk to young children's cognitive development: An alternative to cumulative indices. *British Journal of Developmental Psychology*, *28*(2), 219-238. http://dx.doi.org/10.1348/026151008X399925

Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4th ed.). New York, NY: Psychology Press.

Halpern, R., Giugliani, E. R., Victora, C. G., Barros, C. B., & Horta, B. L. (2002). Risk factors for suspected delays in neuropsychomotor development at 12 months of life. *Revista Chilena de Pediatría, 73*(5), 529-539. http://dx.doi.org/10.4067/S0370-41062002000500016

Han J. Y., Nava-Ocampo, A. A., & Koren, G. (2005). Unintended pregnancies and exposure to potential human teratogens. *Birth Defects Research Part A: Clinical and Molecular Teratology, 73*(4), 245-248. https://doi.org/10.1002/bdra.20132

Hart, B., & Risley, T. R. (2003). The early catastrophe. The 30 million word gap. *American Educator*, *27*(1), 4-9.

Heron, J., O'Connor, T. G., Evans, J., Golding, J., Glover, V., & ALSPAC Study Team. (2004). The course of anxiety and depression through pregnancy and the postpartum in a community sample. *Journal of Affective Disorders*, *80*(1), 65-73. doi:10.1016/j.jad.2003.08.004

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*, 55-88. https://doi.org/10.1016/j.dr.2005.11.002

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*, 236-248. http://dx.doi.org/10.1037/0012-1649.27.2.236

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60(6), 581-592. doi:10.1037/0003-066X.60.6.581

Joyce, T., Kaestner, R., & Korenman, S. (2000). The effect of pregnancy intention on child development. *Demography, 37*(1), 83-94. Retrieved from http://www.jstor.org.ezproxy.auckland.ac.nz/stable/2648098

Kanazawa, S. (2012). Intelligence, birth order, and family size. *Personality and Social Psychology Bulletin*, *38*(9), 1157-1164. https://doi.org/10.1177/0146167212445911

Kaplan, D. S., Damphousse, K. R., & Kaplan, H. B. (1994). Mental health implications of not graduating from high school. *The Journal of Experimental Education*, *62*(2), 105-123. Retrieved from http://www.jstor.org/stable/20152404

Kim, D. R., Bale, T. L., & Epperson, C. N. (2015). Prenatal programming of mental illness: Current understanding of relationship and mechanisms. *Current Psychiatry Reports, 17*(2), 5. http://doi.org/10.1007/s11920-014-0546-9

King, R. R., Jones, C., & Lasky, E. (1982). In retrospect: A fifteen year follow-up report of speech-language-disordered children. *Language, Speech and Hearing Services in Schools*, *13*(1), 24-32. https://doi.org/10.1044/0161-1461.1301.24

Kinsella, M. T., & Monk, C. (2009). Impact of maternal stress, depression & anxiety on fetal neurobehavioral development. *Clinical Obstetrics and Gynecology*, *52*(3), 425-440. doi:10.1097/GRF.0b013e3181b52df1

Knoester, M., Helmerhorst, F. M., Vandenbroucke, J. P., van der Westerlaken, L. A., Walther, F. J., Veen, S., & Leiden. Artificial Reproductive Techniques Follow-up Project. (2009). Cognitive development of singletons born after intracytoplasmic sperm injection compared with in vitro fertilization and natural conception. *Obstetrical and Gynecological Survey, 64*(1), 18-19. doi:10.1016/j.fertnstert.2007.06.090

Kozinszky, Z., & Dudas, R. B. (2015). Validation studies of the Edinburgh Postnatal Depression Scale for the antenatal period. *Journal of Affective Disorders, 176*, 95-105. doi:10.1016/j.jad.2015.01.044

Kramer, J. H., Delis, D. C., Kaplan, E., O'donnell, L., & Prifitera, A. (1997). Developmental sex differences in verbal learning. *Neuropsychology, 11*(4), 577-584. doi:10.1037/0894-4105.11.4.577

Lairoque, B., Kaminski, M., Dehaene, P., Subtil, D., Delfosse, M.-J., & Querleu, D. (1995). Moderate prenatal alcohol exposure and psychomotor development at preschool age. *American Journal of Public Health*, *85*(12), 1654. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1615719/

Lamb, Y. N., Thompson, J. M., Murphy, R., Wall, C., Kirk, I. J., Morgan, A. R., & ABC Study group. (2014). Perceived stress during pregnancy and the catechol-O-methyltransferase (COMT) rs165599 polymorphism impacts on childhood IQ. *Cognition, 132*(3), 461-470.

Lancaster, C. A., Gold, K. J., Flynn, H. A., Yoo, H., Marcus, S. M., & Davis, M. M. (2010). Risk factors for depressive symptoms during pregnancy: a systematic review. *American Journal of Obstetrics and Gynecology, 202*(1), 5-14. doi:10.1016/j.ajog.2009.09.007

Laplante, D. P., Barr, R. G., Brunet, A., Fort, G. G. D., Meaney, M. L., Saucier, J. F., … King, S. (2004). Stress during pregnancy affects general intellectual and language functioning in human toddlers. *Pediatric Research*, *56*(3), 400. https://doi.org/10.1203/01.PDR.0000136281.34035.44

de La Rochebrochard, E., & Joshi, H. (2013). Children born after unplanned pregnancies and cognitive development at 3 years: Social differentials in the United Kingdom millennium cohort. *American Journal of Epidemiology, 178*(6), 910-920. doi:10.1093/aje/kwt063

Leslie, G. I., Gibson, F. L., McMahon, C., Cohen, J., Saunders, D. M., & Tennant, C. (2003). Children conceived using ICSI do not have an increased risk of lower-ability mental development at 5 years of age. *Human Reproduction, 18*(10), 2067-2072. https://doi.org/10.1093/humrep/deg408

Liu, L., Wang, J., Shanshan, S., Luo, X., Rui, K., Xiaohui, Z., & Song, R. (2016). Descriptive epidemiology of prenatal and perinatal risk factors in a Chinese population with reading disorder. *Scientific Reports, 6*, 36697. doi:10.1038/srep36697

Lupien, S. J., McEwen, B. S., Gunnar, M. R., & Heim, C. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nature Reviews Neuroscience, 10*(6), 434-445. doi:10.1038/nrn2639

MacCallum, R. C, Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*(1), 19-40. doi:10.1037//1082-989X.7.1.19

Maccani, M. A., & Marsit, C. J. (2009). Epigenetics in the placenta. *American Journal of Reproductive Immunology, 62*(2), 78-89. doi:10.1111/j.1600-0897.2009.00716.x

Masten, A. S., Roisman, G. I., Long, J. D., Burt, K. B., Obradović J., Riley, J. R., ... & Tellegen, A. (2005). Developmental cascades: Linking academic achievement and externalizing and internalizing symptoms over 20 years. *Developmental Psychology*, *41*(5), 733. doi:10.1037/0012-1649.41.5.733

McCaul, E. J., Donaldson Jr, G. A., Coladarci, T., & Davis, W. E. (1992). Consequences of dropping out of school: Findings from high school and beyond. *The Journal of Educational Research*, *85*(4), 198-207. Retrieved from https://doi.org/10.1080/00220671.1992.9941117

McLearn, K. T., Minkovitz, C. S., Strobino, D. M., Marks, E., & Hou, W. (2006). Maternal depressive symptoms at 2 to 4 months post partum and early parenting practices. *Archives of Pediatrics & Adolescent Medicine*, *160*(3), 279-284. doi:10.1001/archpedi.160.3.279

Mennes, M., Stiers, P., Lagae, L., & Van den Bergh, B. (2006). Long-term cognitive sequelae of antenatal maternal anxiety: involvement of the orbitofrontal cortex. *Neuroscience & Biobehavioral Reviews*, *30*(8), 1078-1086. doi:10.1016/j.neubiorev.2006.04.003

Morton, S. M., Atatoa Carr, P. E., Grant, C. C., Robinson, E. M., Bandara, D. K., Bird, A., & Perese, L. M. (2012). Cohort profile: Growing up in New Zealand. *International Journal of Epidemiology, 42*(1), 65-75. https://doi.org/10.1093/ije/dyr206

Murray, D., & Cox, J. L. (1990). Screening for depression during pregnancy with the Edinburgh Depression Scale (EDDS). *Journal of Reproductive and Infant Psychology, 8*(2), 99-107. https://doi.org/10.1080/02646839008403615

New Zealand Health Survey Public Health Intelligence. (2007). *New Zealand Health Survey*. Wellington, New Zealand: Ministry of Health.

Obradović J., Burt, K. B., & Masten, A. S. (2009). Testing a dual cascade model linking competence and symptoms over 20 years from childhood to adulthood. *Journal of Clinical Child & Adolescent Psychology*, *39*(1), 90-102. doi:10.1080/15374410903401120

Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development*, *76*(4), 763-82. doi: 10.1111/j.1467-8624.2005.00876.x

Peterson, R., & Pennington, B. (2015). Developmental dyslexia. *Annual Review of Clinical Psychology, 11*, 283-307. https://doi.org/10.1146/annurev-clinpsy-032814-112842

Peyre, H., Bernard, Y. J., Forhan, A., Charles, M., De Agostini, M., Heude, B., & Ramus, F. (2014). Predicting changes in language skills between 2 and 3 years in the EDEN mother-child cohort. *PeerJ, 2*, e335. doi: 10.7717/peerj.335.

Phillips, D. A., & Shonkoff, J. P. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.

Poulton, R., Moffitt, T. E., & Silva, P. A. (2015). The Dunedin multidisciplinary health & development study: overview of the first 40 years, with an eye to the future. *Social Psychiatry and Psychiatric Epidemiology, 50*(5), 679-693. http://doi.org/10.1007/s00127-015-1048-8

Pugh, S. J., Richardson, G. A., Hutcheon, J. A., Himes, K. P., Brooks, M. M., Day, N. L., & Bodnar, L. M. (2015). Maternal obesity and excessive gestational weight gain are associated with components of child cognition. *Journal of Nutrition*, *145*(11), 2562-2569. doi:10.1542/peds.2014-3058

Quevedo, L. A., Silva, R. A., Godoy, R. , Jansen, K. , Matos, M. B., Tavares Pinheiro, K. A., & Pinheiro, R. T. (2012), The impact of maternal post‐partum depression on the language development of children at 12 months. *Child: Care, Health and Development, 38*(3), 420-424. doi:10.1111/j.1365-2214.2011.01251.x

Rice, M. L. (2013). *Toward a genetics of language*. New York, NY: Psychology Press. https://doi.org/10.4324/9780203763476

Rodgers, J. L., Cleveland, H. H., van den Oord, E., & Rowe, D. C. (2000). Resolving the debate over birth order, family size, and intelligence. *American Psychologist, 55*(6), 599-612. http://dx.doi.org/10.1037/0003-066X.55.6.599

Rothbart, M. K. (2011). *Becoming who we are.* New York, NY: Guilford.

Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language, 35*, 185-205. https://doi.org/10.1017/S0305000907008343

Roza, S. J., van Batenburg-Eddes, T., & Steegers, E.A.P. (2010). Maternal folic acid supplement use in early pregnancy and child behavioural problems: The Generation R Study. *British Journal of Nutrition*, *103*(3), 445-452. doi:10.1017/S0007114509991954

Russell, M., Czarnecki, D. M., Cowan, R., McPherson, E., & Mudar, P. J. (1991). Measures of maternal alcohol use as predictors of development in early childhood. *Alcoholism: Clinical and Experimental Research*, *15*(6), 991-1000. doi:https://doi.org/10.1111/j.1530-0277.1991.tb05200.x

Russell, D. G., Parnell, W. R., Wilson, N. C., Faed, J., Ferguson, E., Herbison, P., & Wilson, B. (1999). *NZ food: NZ people. Key results of the 1997 national nutrition survey*. Wellington, New Zealand: Ministry of Health.

Saarni, C., Campos, J. J., Camras, L. A., & Witherington, D. (1998). Emotional development: Action, communication, and understanding. In N. Eisenberg, W. Damon, & R. M. Damon (Eds.), *Handbook of child psychology, volume III.* (pp. 226-299). New York, NY: John Wiley and Sons.

Salmond, C., Crampton, P., & Atkinson, J. (2007). *NZDep2006 index of deprivation*. Wellington, NZ: University of Otago.

Sandin, S., Nygren, K. G., Iliadou, A., Hultman, C. M., & Reichenberg, A. (2013). Autism and mental retardation among offspring born after in vitro fertilization. *JAMA, 310*(1), 75-84. doi:10.1001/jama.2013.7222

Schady, N., Behrman, J., Caridad Araujo, M., Azuero, R., Bernal, R., Bravo, D., … Vakis, R. (2015). Wealth gradients in early childhood cognitive development in five Latin American countries. *Journal of Human Resources, 50*(2), 446-463.

Solivan, A. E., Xiong, X., Harville, E. W., & Buekens, P. (2015). Measurement of perceived stress among pregnant women: a comparison of two different instruments. *Maternal and Child Health Journal, 19*(9), 1910-1915. http://doi.org/10.1007/s10995-015- 1710-5

Stanton-Chapman, T. L., Chapman, D. A., Kaiser, A. P., & Hancock, T. B. (2004). Cumulative risk and low-income children's language development. *Topics in Early Childhood Special Education*, *24*(4), 227-237. https://doi.org/10.1177/02711214040240040401

Stene-Larsen, K., Brandlistuen, R. E., Lang, A. M., Landolt, M. A., Latal, B., & Vollrath, M. E. (2014). Communication impairments in early term and late preterm children: A prospective cohort study following children to age 36 months. *Journal of Pediatrics*, *165*(6), 1123-1128. doi:10.1016/j.jpeds.2014.08.027

Suggate, S. P., Schaughency, E. A., & Reese, E. (2012). Children learning to read later catch up to children reading earlier. *Early Childhood Research Quarterly*, *28*, 33-48.

Sylvestre, A., & Mérette, C. (2010). Language ability in severely neglected children: A cumulative or specific effect of risk factors? *Child Abuse & Neglect*, *34*, 414-428. https://doi.org/10.1016/j.chiabu.2009.10.003

Talge, N. M., Neal, C., & Glover, V. (2007). Antenatal maternal stress and long-term effects on child neurodevelopment: how and why? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *48*(3-4), 245. doi:10.1111/j.1469-7610.2006.01714.x

Taylor, C. L., Christensen, D., Lawrence, D., Mitrou, F., & Zubrick, S. R. (2013). Risk factors for children's receptive vocabulary development from four to eight years in the longitudinal study of Australian children. *PloS One, 8*(9), e73046.

Than, L. C., Honein, M. A., Watkins, M. L., Yoon, P. W., Daniel, K. L., & Correa, A. (2005). Intent to become pregnant as a predictor of exposures during pregnancy: Is there a relation? *Journal of Reproductive Medicine, 50*(6), 389-396. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/16050563

Tomblin, J. B., & Bavin, L. B. (2009). Children with specific language impairment. In E. L.Bavin (Ed.), *The Cambridge handbook of child language* (pp. 417-432). Cambridge, UK: Cambridge University Press.

Underwood, L., Waldie, K., D'Souza, S., Peterson, E. R., & Morton, S. (2016). A review of longitudinal studies on antenatal and postnatal depression. *Archives of Women's Mental Health, 19*(5), 711-720. doi:10.1007/s00737-016-0629-1

University of Oregon. (2019). *DIBELS letter naming fluency.* Retrieved from https://dibels.uoregon.edu/assessment/dibels/measures/lnf.php

Unterrainer, J. M., Ruh, N., Loosli, S.V., Heinze, K., Rahm, B., & Kaller, C. P. (2013) Planning steps forward in development: In girls earlier than in boys. *PLoS One, 8*(11), 80-92. https://doi.org/10.1371/journal.pone.0080772

Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P., & Passingham, R. (1995). Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proceedings of the National Academy of Sciences*, *92*(3), 930-933. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC42734/

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

Wake, M., Levickis, P., Tobin, S., Zens, N., Law, J., Gold, L., … Reilly, S. (2012). Improving outcomes of preschool language ability in the community: Protocol for the Language for Learning randomised controlled trial. *BMC Pediatrics*, *12*, 96. doi:10.1186/1471-2431-12-96

Waldie, K. E., Peterson, E. R., D'Souza, S., Underwood, L., Pryor, J. E., Carr, P. A., ... & Morton, S. M. (2015). Depression symptoms during pregnancy: Evidence from growing up in New Zealand. *Journal of Affective Disorders*, *186*, 66-73. doi:10.1016/j.jad.2015.06.009

Walsh, D., Price, G., & Gillingham, M. (1988). The critical but transitory importance of letter naming, *Reading Research Quarterly, 23*(1), 108-22. http://dx.doi.org/10.2307/747907

Wang, G., Walker, S. O., Hong, X., Bartell, T. R. ,& Wang, X. (2013). Epigenetics and early life origins of chronic noncommunicable diseases. *Journal of Adolescent Health, 52*(2), 14-21. doi:10.1016/j.jadohealth.2012.04.019

Welberg, L. A., & Seckl, J. R. (2001). Prenatal stress, glucocorticoids and the programming of the brain. *Journal of Neuroendocrinology*, *13*(2), 113-128.

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology*, *102*(1), 43. doi:10.1037/a0016738

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *46*(3), 233-251. doi:10.1002/dev.20060

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69*, 848–872.

Willingham, D.T. (2012). Ask the cognitive scientist: Why does family wealth affect learning? *American Educator, 36*(1), 33-39. https://eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=EJ971756

Woods, S. M., Melville, J. L., Guo, Y., Fan, M. Y., & Galvin, A. (2010). Psychosocial stress during pregnancy. *American Journal of Obstetrics and Gynecology, 202*(1), 61-81.

World Health Organisation. (2017). *BMI classification.* Geneva, Switzerland: Author.

Zambrana, I. M., Vollrath, M. E., Sengpiel, V., Jacobsson, B., & Ystrom, E. (2016). Preterm delivery and risk for early language abilitys: a sibling-control cohort study. *International Journal of Epidemiology*, *45*(1), 151-159. doi:10.1093/ije/dyv329

[1] *Growing up in New Zealand* was funded by the New Zealand Ministries of Social Development, Health, Education, Justice and Pacific Island Affairs; the former Ministry of Science Innovation and the former Department of Labour (now both part of the Ministry of Business, Innovation and Employment); the former Ministry of Women's Affairs (now the Ministry for Women); the Department of Corrections; the Families Commission (now known as the Social Policy Evaluation and Research Unit); Te Puni Kokiri; New Zealand Police; Sport New Zealand; the Housing New Zealand Corporation; and the former Mental Health Commission, The University of Auckland and Auckland UniServices Limited. Other support for the study has been provided by the NZ Health Research Council, Statistics New Zealand, the Office of the Children's Commissioner and the Office of Ethnic Affairs.

Acknowledgments

We acknowledge the research by Charlotte Anderson and Sarah Herbert, whose theses inspired the current research. We thank Professor Elaine Reese for her contributions to tool selection and review of early versions of this manuscript.

Private Speech Use in Mathematics Problem Solving: A Review of Studies Comparing Children With and Without Mathematical Difficulties

Snorre A. Ostad
University of Oslo

Abstract

Recent studies have concluded that children's development of private speech (private speech internalization) is related to and important for developing mathematical ability. In this article, we review a project consisting of studies exploring the cognitive factors that may underlie differences between the use of private speech by children with (MD) and without (MN) mathematical difficulties. The main issue of interest was whether private speech internalization is related to children's mathematical achievement, task-specific strategies, phonological awareness, and phonological memory, and whether any such relationships are modulated by age and mathematical achievement. The findings not only confirm that private speech internalization relate to mathematical achievement, they also highlight possible parallels between the contributions of strategies, phonological awareness, and phonological memory to subsequent mathematical achievement. Overall, the results seem to provide evidence for the hypothesis that mathematical achievement is causally related to phonological abilities – which underpin the internalization of private speech – rather than being directly related to the private speech internalization.

Keywords: mathematical difficulties, private speech, private speech internalization, strategy use, phonological awareness, phonological memory

Difficulties in Mathematics

Children with MD have a specific difficulty in mastering calculation despite adequate instruction and the absence of intellectual disability. It is estimated that 5-8% of students have a cognitive or neuropsychological deficit that interferes with calculation or word problem solving (Badian, 1983; Geary, 1993; Ostad, 1998).

To understand and address problems in mathematical cognition, it is essential to first identify the core difficulties that discriminate children with MD from their peers without such challenges. Previous theoretical models in the area of mathematical learning have focused on various sources of MD. Thus, difficulties in visuospatial representation of numerical information (Geary, 1993), impaired working memory (Geary, Hamson, & Hoard, 2000; Passolunghi & Siegel, 2004; Swanson & Sachse-Lee, 2001), use of developmentally immature calculation procedures (Jordan & Montani, 1997; Ostad, 1997; Ostad & Sorensen, 2007), difficulty learning basic arithmetical facts or retrieving them once they are learned (Geary, 1993; Jordan, Hanich, & Kaplan, 2003; Ostad, 1999; Swanson & Sachse-Lee, 2001), and poor conceptual understanding (Geary et al., 2000; Jordan et al., 2003) are all observed in young children with MD.

Ostad, S. A. (2020). Private Speech Use in Mathematics Problem Solving: A Review of Studies Comparing Children With and Without Mathematical Difficulties. *International Journal for Research in Learning Disabilities, 4*(2), 53-65. https://doi.org/10.28987/ijrld.4.2.53

Proposed mechanisms underlying these deficits range from fundamental deficits in potentially innate systems for processing number and magnitude (Butterworth, 2005) to specific deficits or delays in the mechanisms underlying mathematical ability (Hecht, Torgesen, Wagner, & Rashotte, 2001; Ostad, 2013; Swanson & Jerman, 2006); for further details, see Geary, Hoard, and Bailey (2012).

Children's Private Speech

Considerable research has investigated the cognitive abilities underpinning reading. For example, children's decoding reading ability is known to be highly correlated with phonological ability (Dolcos & Albarracin, 2014; Durand, Hulme, Larkin, & Snowling, 2005; Wagner, Torgesen, & Rashotte, 1994). Longitudinal correlation studies have made an important contribution to the emerging consensus that certain kinds of phonological processing ability; that is, the ability to use phonological or sound information in processing written and oral language is causally related to the development of reading skills and phonological impairments are a leading cause of reading disabilities (Ramus & Szenkovits, 2008; Snowling, 2001; Wagner et al., 1994).

More than 30 years of research on dyslexia has resulted in the identification of three main dimensions to the phonological deficit associated with the disorder: deficits in *phonological memory, phonological awareness, and phonological re- or decoding* (Hecht al., 2001; Kulak, 1993; Snowling, 2001). Of interest to the current study, it is possible that the poor mathematical ability of persons with MD is due to deficits in one or more of these dimensions (Hecht et al., 2001; Rasmussen & Bisanz , 2005; Wagner & Torgesen, 1987).

Vygotsky (1934/1986) hypothesised that the phenomenon of private speech (self-talk used by children in various situations that is not addressed to others) reflects children's potential for self-directed planning, guiding, and monitoring of personal goal-directed activity. From the Vygotskian perspective, and that, therefore, children's private speech can be considered an important intrapsychic tool for regulating thought and behaviour (Berk & Winsler, 1995; Winsler & Naglieri, 2003).

Development of private speech follows a developmentally typical course of increasingly sophisticated private speech categories, culminating in what we refer to as *private speech internalization*. Researchers have shown an overall ontogenetic pattern, whereby children's overt private speech is gradually replaced by partially internalized whispers, inaudible mutterings, and silent inner speech as they progress through elementary school (Berk, 1992; Diaz & Berk, 1992; Flavell, Green, Flavell, & Grossman, 1997; Kohlberg, Yaeger, & Hjertholm, 1968; Winsler & Naglieri, 2003).

Other aspects of children's private speech have also been explored, including its relationships with children's task performance and on-task behaviour (Winsler, Diaz, & Montero, 1997), task and setting influences on such speech (Winsler, Carlton, & Barry, 2000), its use among children with behaviour problems, learning impairment, or attention deficits (Berk & Landau, 1993), and task-related utterances to self during problem-solving (Winsler, Feder, Way, & Manfra, 2006).

The majority of recent studies have concluded that internalization of private speech is related to and important for the development of mathematical ability (Berk & Landau, 1993; Ostad, 2015; Winsler & Naglieri, 2003). Underlying this conclusion is the belief that arithmetic knowledge is stored in a sound-based, or phonological, form. Several variants of the phonological storage hypothesis have been proposed (Anderson-Day & Fernyhough, 2015; Cohen & Dehaene, 2000; McCloskey, 1992; Robinson, Menchetti, & Torgesen, 2002). It remains far from clear, however, whether private speech internalization itself plays any causal role in the development of mathematical ability.

The overall purpose of the project reported here was to add to our knowledge of the cognitive abilities underpinning private speech internalization. More specifically, whether private speech internalization is related to children's mathematical achievement, use of task-specific strategies, phonological awareness and phonological memory, and whether any such relationships are modulated by age and mathematical achievement.

The Present Project: Common Methods

Separate laboratory investigations were carried out to examine children's private speech, strategy use, phonological awareness, and phonological memory, respectively. For each child, the four investigations were finished within a seven-day period. The same children participated in all the investigations.

The study used a cross-sectional design. To ensure that the behaviour of interest was recorded under comparable conditions across all subjects and grade levels, the design included research procedures for individual observation developed for use outside the classroom. Order of administration of the investigations was counterbalanced across subjects at each of the grade levels included in the study.

Participants

Control for mathematical ability was based on the arithmetic subtest of the Wechsler Intelligence Scale for Children (WISC-R; Undheim, 1977) and a Standard Mathematics Performance Test (Hammervoll & Ostad, 1999). The latter appears to tap very basic number processing aptitudes, in which children with MD have deficits (Geary et al., 2012).

MD children, defined for the purposes of this series of studies as the less mathematically able children in their grades, were chosen from a population of children who met the following two criteria: (a) attained stanine scores in the range of 1-3 on the standard mathematics test (i.e., a score among the 23% weakest for the grade group) and (b) attained scaled scores in the range of 1-7 on the WISC-R subtest (i.e., a score among the 25% weakest for the relevant group).

To avoid false negatives, MN children, in turn, were defined as the most mathematically able children in all grades. These children were selected by asking general education classroom teachers to nominate the most mathematically able children in their grades they taught, and based on these nominations the researchers assigned a same-sex MN student as a match for each MD child. The final sample of MN children satisfied both of the following criteria: (a) attained stanine scores in the range 7-9 on the standard mathematics test (i.e., among the 23% strongest in the relevant grade group and (b) attained scaled scores in the range 12-19 on the WISC-R (i.e., among the strongest 25% for the relevant grade group).

The sample comprised 134 children (half MN children and half MD children) from five state schools in Norway. All children belonged to one of the following grade groups: Group 1: children in grades 2 and 3 (n = 22 + 22); Group 2: children in grades 4 and 5 (n = 22 + 22); and Group 3: children in grades 6 and 7 (n = 23 + 23). The respective mean ages for MD and MN children in each group were as follows, Group 1: 7.7 and 7.8 years; Group 2: 9.7 and 9.6 years; and Group 3: 11.7 and 11.8 years.

There were no statistically significant age differences between children with and without MD within the grades groups, $F(1,42) = 2.19$, $p > .05$ for Ggr1; $F(1,42) = 0.98$, $p > .05$ for Ggr2; $F(1,43) = 2.07$, $p > .05$ for Ggr3. Inspection of the MD children's records showed that the schools' support services had already identified all 67 children as being in need of a special programme of mathematics teaching. Children in special education classes were not included as members in the sample.

Instruments

Assessment tools of private speech were constructed from the 64 possible pairwise additive combinations of the integers 2 to 9, with tie problems (e.g., 2 + 2) excluded. The remaining problems, 56 single-digit addition problems in the form a + b, were divided into two halves (half a and half b). The two halves were counterbalanced so that all 56 problems were pair-wise matched (e.g., 9 + 8 and 8 + 9). Lots were drawn so that the one problem in each pair was randomly assigned to one half, with the other problem from the pair going to the other half. Only one of the two halves (half a or half b) was used in testing each child. The problems (28 problems in each of the two halves) were placed horizontally at the centre of 21 x 10 cm printed booklets, with one problem per sheet. Other equipment placed on the table where the children were tested included paper, pencil, and 40 cubes (sides measuring about 1.5 cm).

Observational Procedure

According to Fuson (1979), the presence of another person creates the challenge of separating social speech from private speech. To minimize communication with the researcher in the current study, the children were informed ahead of time that they had to answer all the problems without any reference to the researcher. At the start of the study, the children were told to solve problems that would be presented on the (28) pages in the"book." During the time interval when the child turned to the next page, the private speech category for solving the problem was recorded by the researcher using the classification scheme developed for the investigation.

The reliability of this structured data collection format was previously tested. Two master's degree students who had received extensive training in the research procedure developed for the study coded, independently of each other, private speech responses based on 100 randomly chosen addition problems. The overall inter-coder reliability (averaging the three private speech levels) for the broad category classification was 0.96. Inter-coder reliability for each private speech level was, as follows: Audible: 1.00; Inaudible: 0.97; Silence: 0.95.

Relationship Between Private Speech and Mathematical Achievement

There is no consensus on methods or standardised criteria for categorising units of private speech (Girbau, 2002). The categorisation system used in the studies reported here integrated techniques and ideas from earlier research (Berk, 1986; Girbau, 2002; Kohlberg et al., 1968). Specifically, we defined categories of private speech in relation to private speech internalization (Kohlberg et al., 1968; Vygotsky, 1987) as follows: (a) externalised verbal production by means of words or sounds, (b) externalised manifestation of private speech, and (c) silence (Berk & Landau, 1993; Girbau, 2002).

More specifically, we classified private speech as (a) Audible private speech (high, normal, or low) – private speech that was potentially intelligible to a listener and could be transcribed; (b) Inaudible private speech that was detectable by face-to-face observation but unintelligible to even a very nearby listener. This category is used to refer to externalised private speech that is not loud enough for a listener to discern any semantic content to verbalisation, including inaudible muttering, and may be evident from lip and tongue movements (Girbau, 2002). It seems to be widely accepted that private speech according to this classification unit represents an external manifestation of inner speech (Berk & Landau, 1993; Fuson, 1979; Kolberg et al., 1968). (c) Subvocal private speech; that is, private speech that is silent and not reflected in any external verbal production or in lip or tongue movements (Girbau, 2002; Vygotsky, 1987). In this report, subvocal speech was operationalised as a purely implicit, covert, mental process that was undetectable through face-to-face-observation. It seems reasonable to consider subvocal speech as representing the most internalised form of private speech.

These three categories of private speech were coded as audible private speech, inaudible private speech, and silence. The number of answers belonging to each of the three private speech categories formed the basis for a unitary measure of the children's private speech developmental level, also known as the *private speech internalization score* (the PSI score).

The three subcategories' relative contributions to the development of mathematical aptitude (Ostad, 2013; Ostad & Sorensen, 2007) were roughly estimated and scored in the proportion 0:1:2, respectively. (For example, the PSI score = 26 for the child with 11 answers belonging to category audible, 8 to inaudible, and 9 to in silence.) With reference to the private-speech literature (Berk, 1992; Diaz & Berk, 1992; Flavell et al., 1997; Kohlberg et al., 1968; Winsler & Naglieri, 2003), we suggest that the PSI score is at least a theoretically continuous variable.

Results

Table 1 presents descriptive information about the private speech use, expressed as PSI scores, for solving the corresponding number fact problems. The data (mean score and standard deviation), common for the three studies, are broken down initially by grade groups (grades 2-3, grades 4-5, and grades 6-7), and then further by achievement groups (MN children and MD children).

To examine whether or not the private speech development of the MD children showed different developmental patterns from those of the MN children, we performed a univariate ANOVA analysis, identifying achievement groups as between-subjects

Table 1
Means and Standard Deviations for Private Speech Internalisation Scores (PSI Score) by Grade and Mathematical Achievement Group

Mathematical Achievement Group	Ggr 1		Ggr 2		Ggr 3	
	M	*SD*	*M*	*SD*	*M*	*SD*
MD children	16.73	4.95	18.73	2.90	22.78	4.26
MN children	25.14	5.05	32.73	5.80	42.17	8.10

Note. Number of problems = 28. Ggr1= Grade group 1 (Grades 2-3); Ggr2 = Grade group 2 (Grades 4-5); Ggr3 = Grade group 3 (Grades 6-7); MD = Children with mathematical difficulties; MN = Children without mathematical difficulties.

factors and PSI score as the dependent factor. The test indicated that a child's achievement group had a significant effect: F (1, 132) = 115.94, $p < .01$ (R^2 = .468). Furthermore, ANOVAs were performed to determine if the PSI scores of MN and MD children differed in the three grade groups. The analysis revealed significant differences for all three groups: $F_{Gr2\text{-}3}$ (1, 42) = 31.03, $p < .01$ (R^2 = .425); $F_{Gr4\text{-}5}$ (1, 42) = 102.58, $p < .01$ (R^2 = .710); $F_{Gr6\text{-}7}$ (1, 44) = 103.23, $p < .01$ (R^2 = .701). (For further details about the results, see Ostad and Askeland, 2008.)

Relationship Between Private Speech and Strategy Use

Several studies have investigated the strategy use of individuals of various levels of arithmetical experience and ability, and in samples varying in age from very young children to adults (Carpenter & Moser, 1982; De Chambier & Zeiger, 2018; Dowker, 2005; Geary et al., 2000; Ostad, 1998; Siegler, 1988; Thevenot, Barrouillet, Castel, & Uittenhove, 2016). In particular, the developmental level of children's strategy use has been viewed as the basis of theoretical models of the pupils' mathematical ability (Siegler & Jenkins, 1989). For example, researchers have observed inter-individual variation in strategy used to solve a basic problem among children of a given age. There has also been interest in intra-individual variability.

Research over the past three decades has uncovered developmental changes in children's use of problem-solving strategies (Geary, 1993; Ostad, 1998; Siegler & Jenkins, 1989). MN children appear to progress from using overt strategies, to verbal counting, to fact retrieval (Carpenter & Moser, 1982; Siegler, 1988). By contrast, MD children seem to be characterised by (a) use of backup strategies only, (b) use of the most primary backup strategies, (c) use of a limited range of strategies, and (d) limited change in the use of strategies during the school years (Ostad, 1997). It is therefore not surprising that it has been hypothesised that differences in strategy use can be used to distinguish between individuals of different intellectual levels or aptitudes (Dowker, 2005; Ostad, 2000).

Assessment of Task-Specific Strategies

Several systems have been used to classify children's strategy use (Carpenter & Moser, 1982; Geary et al., 2000; Groen & Parkman, 1972; Ostad, 1999; Siegler & Shrager, 1984). In the present study, we used the three main aspects of the strategy use internalization process: backup strategies, decomposition strategies, and direct retrieval strategies (Geary et al., 2012; Ostad, 1997); for further details, see Ostad and Sorensen, 2007.)

As a basis for the data analysis, we introduce the notion of *strategy use internalization,* which refers to the movement through the typical chain of increasingly sophisticated task-specific strategy categories. We suggest that strategy use internalization might load along a continuum (analogue theoretical axis) from primary backup strategies at one extreme to direct retrieval strategies at the other.

The number of strategies belonging to each category (backup, decomposition, retrieval) formed the basis for a unitary measure of children's strategy developmental level, also called the *strategy-use internalization score* (STRAT score). The three subcategories' relative contributions to the development of mathematical aptitude were roughly estimated and scored in the proportion 0:1:2, respectively. For example, the STRAT score was 26 for the child with 11 answers belonging to backup, 8 to decomposition, and 9 to retrieval. With reference to the strategy use researchers (e.g., de Chambrier & Zesiger, 2018; Geary et al., 2012; Jordan et al., 2003; Ostad, 1997) have suggested that the STRAT score is at least a theoretically continuous variable.

The reliability of this structured collection format was previously tested. A preliminary examination of the constancy of children's strategy use related to the three main categories (i.e., backup, decomposition, and retrieval strategies) was conducted as follows: 10 children representing each of the six grade performance groups (a total of 60 children) were given a set of 15 addition problems, selected by drawing lots from the addition problems included in the study. These sets of problems were presented to the children twice with a time limit of 60 min. The frequency of using the same strategy category for solving each of the problems twice was expressed in terms of percentage of occurrence. In all the six groups, the percentage was between 97 and 100.

Table 2 presents descriptive data showing children's strategy use expressed as STRAT scores. The data (mean scores and standard deviations) are broken down initially by grade groups (Gr 2-3, Gr 4-5, and Gr 6-7), and then further by achievement groups (MN and MD children).

To examine whether or not the development of the strategy use of MD children showed different developmental patterns from those of MN children, we performed a univariate ANOVA identifying

Table 2
Means and Standard Deviations for Strategy Use Internalisation Scores (STRAT Score) by Grade and Mathematical Achievement Group

Mathematical Achievement Group	Ggr 1		Ggr 2		Ggr 3	
	M	*SD*	*M*	*SD*	*M*	*SD*
MD children	.91	1.15	2.41	1.65	4.17	1.59
MN children	12.27	5.99	19.41	2.41	25.61	6.43

Note. Number of problems = 28. Ggr1= Grade group 1 (Grades 2-3); Ggr2 = Grade group 2 (Grades 4-5); Ggr3 = Grade group 3 (Grades 6-7); MD = Children with mathematical difficulties; MN = Children without mathematical difficulties.

achievement groups as between-subjects factors and STRAT score as the dependent variable. The test indicated achievement group significant effect, $F(1, 132) = 294.51$, $p < .01$ ($R^2 = .688$). Furthermore, ANOVAs were performed to determine whether the STRAT scores of MN and MD children differed in the three grade groups. The analysis indicated significant differences for all three groups: $F_{Gr2\text{-}3}$ (1, 42) = 76.27, $p < .01$ ($R^2 = .636$); $F_{Gr4\text{-}5}$ (1, 42) = 441.18, $p < .01$ ($R^2 = .911$); $F_{Gr6\text{-}7}$ (1, 44) = 240.48, $p < .01$ ($R^2 = .842$).

Correlation analysis (Pearson) was also performed at three levels: all participants, grade group, and achievement group. Results of the analysis indicated significant correlations between STRAT scores and PSI scores when all children in the study were included, $r(132) = .763$, $p < .01$. Furthermore, significant correlations emerged across all three grade groups: r_{Ggr1} (44) = .500, $p < .01$; r_{Ggr2} (42) = .823, $p < .01$; r_{Ggr3} (44) = .748, $p < .01$. At the achievement group level, the analysis indicated significant correlations for MN children, r (67) = .516, $p < .01$, and significant correlations for MD children $r(67) = .322$, $p <.01$.

More specifically, unlike MN children, the MD children rarely used private speech-decomposition strategy combinations. The silence-decomposition strategy combination was mainly used by MN children. The most striking difference between MN children and MD children, however, is related to the silence private speech-retrieval strategy combination. Thus, we observed that MN children used this combination far more often than the MD children and that the difference between the two achievement groups increased with grade.

Consistent with earlier research, the results indicated that that mathematical development is reflected in use of a rich and varied repertoire of task-specific strategies, which become more efficient over time as a result of internalisation (Carpenter & Moser, 1982; Geary, 1993; Ostad, 1998). In comparison, MD children were less likely to use the more internalised, less audible forms of private speech, suggesting a possible link between strategy development and private speech internalization. In accordance with research in the field of private speech, the results suggest that successful mathematical development might be a function of efficiency in the production of task-relevant private speech (Berk & Landau, 1993; Harris, 1986; Hecht et al., 2001; Winsler & Naglieri, 2003). It could be argued that since MD children make less use of internalised private speech than their mathematical-typical counterparts, their MD can at least partly be attributed to a failure to use self-directed language to guide strategy use. (For further details from the study, see Ostad and Sorensen, 2007.)

Relationship Between Private Speech and Phonological Awareness

In general terms, phonological awareness relates to the auditory and oral manipulation of sounds, and refers to the ability to analyze the sound structure of oral language (Krajewski & Schneider, 2009). Phonological awareness can also be described as one's awareness of, and access to, the sound structure of oral language (Dolcos & Albarracin, 2014; Wagner & Torgesen, 1987).

No area of reading research has gained as much attention over the past three decades as phonological awareness. However, although it is well known that phonological awareness affects literacy development, its relation to mathematical competences is not well investigated (Krajewski & Schneider, 2009). Surprisingly, no published study on cognitive correlates of primary school mathematics performance has investigated the role of phonological awareness in private speech internalization.

Assessment of Phonological Awareness

Earlier research in this area did not use assessment instruments developed specifically for the mathematics domain; however, the relatively well-developed research literature in literacy (e.g., Chard & Dickson, 1999; Hecht et al., 2001) has provided insight into potential methods of assessing phonological awareness within the domain of mathematics.

A wide variety of tasks have been used to measure phonological awareness, including rhyming tasks, phoneme counting tasks, sound comparison tasks, blending tasks, segmentation tasks, and deletion tasks. There seems to be ample evidence that the tasks that require manipulation of phonemes are the most difficult. However, just how these tasks relate to each other is far from clear, and this makes it difficult to compare scores from different phonological awareness tasks (Adams & Hitch, 1997; Antony & Francis, 2005; Chard & Dickson, 1999; Hambleton, Swaminathan, & Roger, 1991; Stahl & Murry, 1994).

The test developed for this study was based on instruments used in the literacy domain (e.g., Anthony & Francis, 2002; Hambleton et al., 1991; Stahl & Murry, 1991). The test consisted of tasks related to two-, three-, and four-digit number words arranged in order of difficulty. Four tasks were associated with each number: blending, analysis, deletion, and substitution, in order of difficulty. All tasks require the subject to hold segmented phonological units in short-term memory and to access those representations consciously. In other words, the phonological awareness tasks required subjects to have a special type of access to phonological representations, namely *conscious access*, which may place special demands on retrieval mechanisms.

The total number of points were worked out in order of item difficulty and scored in the proportion 1: 2: 3 for number words with 2, 3, and 4 phonemes (digits), respectively. The participants gained points from the actual number word only if they answered all four tasks satisfactorily. The testing was ended after two subsequent mistakes. The total number of points (max 32 points) was transformed to scaled scores, named PHON scores. (For further details of the assessment process, see Ostad, 2013.)

Table 3 gives descriptive information (means and standard deviations) about the participants' phonological awareness expressed as PHON scores. The data are organized by grade group and by achievement group.

To examine whether the phonological awareness development of MD children showed different developmental patterns from those of MN children, we performed a univariate ANOVA, with achievement groups as between-subjects factors and PHON score as the dependent variable. The test indicated a significant achievement group effect, $F(1, 132) = 251.34$, $p < .01$ ($R^2 = .656$).

Tests of between-subject effects (ANOVAs) were performed to determine whether phonological awareness of MD and MN children differed in all three age groups. The analysis indicated significant phonological awareness differences in all three grade groups: F_{Ggr1} (1, 42) = 98.44, $p < .01$ ($R^2 = .701$); F_{Ggr2} (1, 42) = 207.66, p < .01 ($R^2 = .832$); F_{Ggr3} (1, 44) = 232.30, < .01 ($R^2 = 0.842$).

Correlation analysis was performed at three levels: all participants, grade group, and achievement group. The analysis indicated significant correlations between PSI scores and PHON scores when all children in the study were included, $r(132) = .767$, $p < .01$. Furthermore, the results indicated significant correlations across all three grade groups: r_{Ggr1} (42) = .551, $p < .01$; r_{Ggr2} (42) = .838, $p < .01$; r_{Ggr3} (44) = .720, $p < .01$. However, at the achievement group level, the analysis indicated significant correlations for MN children, r (65) = .585, $p < .01$, and non-significant correlations for MD children, $r(65) = .124$, $p > .05$.

Table 3

Means and Standard Deviations for Phnological Awarness Scores (PHON Score) by Grade and Mathematical Achievement Group

Mathematical Achievement Group	Ggr 1		Ggr 2		Ggr 3	
	M	*SD*	*M*	*SD*	*M*	*SD*
MD children	5.27	1.38	5.59	1.29	6.09	1.91
MN children	9.50	1.43	13.55	2.24	16.39	2.60

Note. Ggr1= Grade group 1 (Grades 2-3); Ggr2 = Grade group 2 (Grades 4-5); Ggr3 = Grade group 3 (Grades 6-7); MD = Children with mathematical difficulties; MN = Children without mathematical difficulties; DFS = Forward digital recall score; DBS = Backward digital recall score.

Among the MN children, there was a grade-depended shift from the lower to the higher levels of phonological awareness, whereas among MD children phonological awareness did not seem to progress beyond the level associated with the lower grades in MN children. These results are consistent with phonological deficit theory (Ramus & Szenkovits, 2008; Rasmussen & Bisanz, 2005; Wagner & Torgesen, 1987) in showing that young children with low or impaired phonological awareness seem to be at risk of developing problems with mathematics.

One of the most important insights from this study concerns developmental changes in the relationship between private speech and phonological awareness in children with and without MD. Thus, the MN children demonstrated a grade-dependent shift from use of audible private speech to use of silent private speech and from lower to higher levels of phonological awareness. In contrast, the development of MD was characterized not only by persistence of low phonological awareness but also by low levels of private speech internalization.

The most striking difference between children with and without MD, however, was related to the silence private speech-high phonological awareness combinations. Thus, this combination was used far more frequently by MN children than by MD children, and the difference between the two achievement groups became more marked as they progressed through primary school. More general analysis of the results revealed significant correlations between the children's developmental levels with respect to private speech and phonological awareness. (For further details from the study, see Ostad , 2013.)

Relationship Between Private Speech and Phonological Memory

The main line of research on phonological abilities focuses on phonological memory. Phonological memory is a form of memory that involves coding and storage of auditory information in short-term or working memory (Bishop, 2003; Swanson, Ashbaker, & Lee, 1996).

According to Wagner and Torgesen (1987), phonological memory deficits are one of the three main classes of phonological deficits. Arguably, phonological memory is important for solving arithmetic problems (Bull & Scerif, 2001; Fürst & Hitch, 2000). Consistent with this hypothesis, numerous studies have reported correlations between indicators of phonological skills and competence in mathematical computation (Adams & Hitch, 1997; Fuchs et al., 2006; Siegel & Ryan, 1989; Swanson & Sachse-Lee, 2001). In particular, phonological memory skills have been associated with the ability to do exact addition (Lemaire, Abdi, & Fayol, 1996), subtraction (Seyler, Kirk, & Ashcraft, 2003), and multiplication (Seitz & Schumann-Hengsteler, 2000). Furthermore, relationships have been reported between phonological memory and the ability to perform calculation procedures such as counting (Logie & Baddeley, 1987), retaining problem information (Fürst & Hitch, 2000) and holding interim results during counting (Berg, 2008). (For a review, see Raghubar, Barnes, and Hecht, 2010.)

In addition to correlational evidence from randomly selected samples, comparisons of children of differing mathematical abilities have shown that relative to children with grade-appropriate mathematical skills, children with MD tend to have poorer phonological memory whereas mathematically gifted children tend to have superior phonological memory (e.g., De Chambrier & Zeiger, 2018; Geary et al., 2000; Jordan et al., 2003; Mabbott & Bisanz, 2008; Ostad, 2013; Passolunghi & Cornoldi, 2008; Siegel & Ryan, 1989; Swanson & Jerman, 2006; Swanson & Sachse-Lee, 2001). These studies typically demonstrate that children with MD have lower digit span than children without MD. It is therefore surprising that no published study on cognitive correlates of primary school mathematics performance has investigated the contribution of phonological memory to private speech internalization. With this in mind, the main aim of our study (Ostad, 2015) was to determine how two indicators of phonological memory (forward digit span and backward digit span) are related to internalization of private speech and how this relationship varies with age and mathematical achievement levels.

Development of phonological memory and private speech was explored in two separate laboratory investigations*; age- and mathematical ability-related differences in phonological memory, and private speech* and *the relationship between phonological memory and private speech internalization.*

Assessment of Phonological Memory

Phonological memory is typically examined by serial recall tasks in which an individual is presented with a sequence of verbal items, such as spoken words or digits, and then asked to repeat them in either the same order (digit forward) or the reverse order (digit backward); for more details, see Swanson and Jerman (2006).

Several studies have hypothesised that factors of phonological memory may be loaded along a continuum (analogue and theoretical axis) from a passive storage system at one extreme, to an active storage system at the other (Cornoldi & Vecchi, 2000; Engle, Cantor, & Carullo, 1992; Passolunghi & Cornoldi, 2008; Swanson, 1994). In accordance with this hypothesis, digit forward tasks rely on a passive storage system and involve the recall of information without manipulating it in any way, which is closer to the passive pole. Conversely, digit backward tasks require more active processes and are those in which information is temporarily held while being manipulated or transformed, which is closer to the active pole (for more details, see Passolunghi and Siegel, 2004). The assessment of phonological memory skills in the present study was theoretically anchored to these hypotheses.

To standardise the measurements of digit span, we used two subtests of Math-Diagnostics (Ostad, 1987). The items in these subtests, a repetition of a dictated series of digits forward and digits backward, are comparable to the digit span from WISC-R (Undheim, 1977). Each series begins with two digits and keeps increasing in length, with two tasks of equal length.

Forward digit recall. The task requires the child to repeat a list of single-digit numbers in the same order as dictated the by the researcher.

Backward digit recall. The experimenter states a list of single-digit numbers and the child repeats them in reverse order. The procedure is similar to that of forward digit recall.

Table 4 presents descriptive information (means and standard deviations) about the performance of phonological memory measured as digit span forward and backward test results. The data are organised by grade group and by achievement group.

Tests of between-subject effects (ANOVAs) were performed to determine if the phonological memory test results of MD and MN children differed in all three grade groups. The analysis indicated significant differences in digit forward span test results across all three grade groups: F_{Gr2-3} (1, 42) = 74.606, $p <$ .01 (R^2 = .640); F_{Gr4-5} (1, 42) = 52.250, $p <$ .01 (R^2 = .554); F_{Gr6-7}(1, 44) = 40.631, $p <$ .01 (R^2 = .468). The corresponding analysis indicated significant digit backward span differences: F_{Gr2-3} (1, 42) = 63.117, p < .01 (R^2 = .600); F_{Gr4-5} (1, 42) = 81.301, $p <$ .01 (R^2 = .659); F_{Gr6-7} (1, 44) = 79.148, $p <$ .01 (R^2 = .643).

We performed two univariate ANOVAs (separate analyses for MD and MN children) with PSI scores as the dependent variable to determine whether PSI scores related differently to phonological memory test results between the two groups.

For the MD children, the analysis indicated a significant effect for grade group, F_{MD} (2, 67) = 10.494, $p <$ .01, a significant effect for DF (forward digital recall) test results, F_{MD} (1, 67) = 28.501, $p <$ .01, and a significant effect for DB (backward digital recall) test results, F_{MD} (1, 67) = 7.119, p <.01. The corresponding analysis of the MN children indicated significant effect for grade group, F_{MN} (2, 67) = 29.977, $p <$ .01, a non-significant effect for DF, F_{MN} (1, 67) = .010, p > .05, and non-significant effect for DB, F_{MN} (1, 67) = 0.737, p > .05.

Correlation analysis (Pearson) was performed at three levels: all participants, grade group, and achievement group. The results indicated significant correlations between PSI scores and DF test results when all children in the study were included, r(132) = .645, $p <$ .01. Furthermore, the analysis indicated significant correlations across all three grade groups: r_{Ggr1} (42) = .595, $p <$ .01; r_{Ggr2} (42) = .664, $p <$ .01; r_{Ggr3} (44) = .597, $p <$ .01. At the achievement group level, the analysis indicated significant correlations for MN children, r (65) = .405, $p <$ .01, and significant correlations for MD children, r(65) = .362, $p <$ .662.

Table 4
Means and Standard Deviations for Phonological Memory Skills by Grade and Mathematical Achievement Group

Phonological Memory Category	Ggr 1				Ggr 2				Ggr 3			
	MD		MN		MD		MN		MD		MN	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
DFS	4.41	1.09	7.23	1.07	5.77	1.07	8.00	0.98	5.78	1.57	8.22	0.95
DBS	2.18	0.85	4.77	1.27	2.68	0.84	5.59	1.26	2.78	0.90	5.87	1.49

Note. Ggr1= Grade group 1 (Grades 2-3); Ggr2 = Grade group 2 (Grades 4-5); Ggr3 = Grade group 3 (Grades 6-7); MD = Children with mathematical difficulties; MN = Children without mathematical difficulties; DFS = Forward digital recall score; DBS = Backward digital recall score.

The corresponding analysis for the correlations between PSI scores and DB test results were as follows: When all children in the study were included, $r(132) = .638$, $p < .01$. Furthermore, the analysis indicated significant correlations across all three grade groups: $r_{Ggr1}\ (42) = .511$, $p < .01$; $r_{Ggr2}\ (42) = .764$, $p < .01$; $r_{Ggr3}\ (44) = .639$, $p < .01$. However, at the achievement group level, the analysis indicated significant correlations for MN children, $r\ (65) = .529$, $p < .01$, and non-significant correlations for MD children $r(65) = .678$, $p > .05$.

There were significant differences between children with and without MD on both the forward and backward digit span test in all the grade groups included in the study. These findings are consistent with the conclusion of earlier studies; namely, that phonological memory predicts the development of mathematics skills and understanding and that phonological memory is a domain-general precursor of mathematical achievement (Anderson-Day & Fernyhoug, 2015; Geary et al., 2000; Jordan et al., 2003; Mabbott & Bisanz, 2008; Passolunghi & Siegel, 2004; Swanson & Jerman, 2006; Swanson & Sachse-Lee, 2001).

One of the most important insights from this study concerns the relationship between phonological memory and private speech internalization. We wanted to determine whether the skills profiles of children with and without MD suggested that the relationships between phonological memory and private speech internalisation are a function of mathematical ability.

Comparisons of children at various developmental stages revealed differences between the two achievement groups. In contrast to the MD children, children without MD showed an age-dependent increase in the strength of the positive association between private speech internalization and phonological memory. We also observed relations between private speech and performance on digit forward and digit backward in children with and without MD. Regression analysis indicated that the contribution of forward digit span to variance in private speech internalization varied with mathematical achievement. Whereas digit backward span made a similar independent contribution to variance in private speech internalization in children with and without MD, the independent contribution of digit forward span was markedly higher in the group of MD children. Thus, the study provides evidence that the relationship between phonological memory and private speech internalization lies closer to a passive storages system (Cornoldi & Vecchi, 2000; Engle, 2002; Passolunghi & Siegel, 2004) in MD children than in MN children. (For further details from the study, see Ostad, 2015.)

Discussion

As a whole, the results from this project not only confirm that private speech internalization is related to mathematical achievement; they also highlight possible parallels between the contributions of strategies, phonological awareness, and phonological memory to subsequent mathematical achievement. However, there is no simple explanation for the patterns of the results reported here. It is possible that there are reciprocal relationships among mathematical performance, private speech internalization, phonological awareness, and phonological memory. If this is the case, a plausible explanation for the relationships observed in these studies is that they are all underpinned by variation in phonological skills, which are a function of the quality of an individual's phonological representations (Geary et al., 2012; Mabbott & Bisanz, 2008; Matuga, 2003; Ostad, 2015).

Overall, the results seem to provide evidence against the hypothesis that private speech internalization is a general cognitive process that is causally related to mathematical achievement. In contrast, they are consistent with the hypothesis that mathematical achievement is causally related to phonological abilities – which underpin private speech internalization – rather than being directly related to private speech internalization. Based on our findings, we believe that more private speech activities do not necessarily mean better problem solving. For private speech activities to have a positive impact on problem solving, they need to be anchored in developmentally appropriate cognitive skills. This hypothesis remains to be directly tested.

In terms of implications for educational practice, classroom teachers should be encouraged to allow their students to utilize private speech in their mathematics learning processes. Furthermore, consistent with suggestions developed from earlier private speech investigations (Anderson-Day & Fernyhough, 2015; Chard & Dickson, 1999; Matuga, 2003; Ostad &Askeland, 2008; Winsler et al., 2000), the various roles private speech play in cognition should be a compulsory component in teaching training programmes.

Limitations of the reported studies suggest that the findings should be accepted with caution. First, in interpreting the results, it is important to bear in mind that the measurements involved in the studies were concurrent in nature, and the causal direction of the relationships analyzed is not known. Second, the studies did not include data to determine the complexity of cognitive processes hidden behind the private speech category silence (e.g. Butterworth, 2005; De Chambrier & Zeiger, 2018; Geary et al., 2012; Winsler et al., 2006). Third, the presented studies used simple number fact problems in addition as their point of departure in the assessment of private speech internalization. Future investigations should address these gaps in the private speech literature, and include a broader selection of mathematical problems, word problems and number fact problems in subtraction and multiplication problems.

References

Adams, J. W., & Hitch, G. J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology*, *67*, 21-38.

Anderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive function, phenomenology, and neurobiology, *Psychological Bulletin*, *141*(5), 931-965.

Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science, 14*, 255-259.

Badian, N. A. (1983). Dyscalculia and nonverbal disorder of learning. In H. R. Myklebust (Ed.), *Progress in learning disabilities* (vol. 5, pp. 235-264). New York, NY: Stratton.

Berg, D. H. (2008). Working memory and arithmetic calculation in children: The contributory roles of processing speed, short-term memory, and reading. *Journal of Experimental Child Psychology, 99*(4), 288-308.

Berk, L. E. (1986). Relationship of elementary school children's private speech to behavioral accompaniment to task, attention, and task performance. *Developmental Psychology, 22*, 671-680.

Berk, L. E. (1992). Children's private speech: An overview of theory and the status of research. In R. M. Diaz & L. E. Berk (Eds.), *Private speech: From social interaction to self-regulation* (pp. 17-53). Hillsdale, NJ: Erlbaum.

Berk, L. E., & Landau, S. (1993). Private speech of learning disabled and normally achieving children in classroom academic laboratory contexts. *Child Development*, *64*, 556-571.

Berk, L. E., & Winsler, A. (1995). *Scaffolding children's learning: Vygotsky and early childhood education*. Washington, DC: National Association for the Education of Young Children.

Bishop, A. G. (2003). Prediction on first grade reading achievement. A comparison of fall and winter kindergarten screenings. *Learning Disability Quarterly*, *26*, 189-200.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology*, *19*, 273-293.

Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry*, *46*, 3-14.

Carpenter, T. P., & Moser, J. M. (1982). The development of addition and subtraction problem solving skills. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 9-24). Hillsdale, NJ: Erlbaum.

Chard, D. J., & Dickson, S. V. (1999). Phonological awareness: Instructional and assessment guidelines. *Intervention in School and Clinic, 34*, 261-270.

Cohen, L., & Dehaene, S. (2000). Calculating without reading: Unsuspected residual abilities in pure alexia. *Cognitive Neuropsychology*, *17*, 563-583.

Cornoldi, C., & Vecchi, T. (2000). Mental imagery in blind people: The role of passive and active visuospatial processes. In M. Heller (Ed.), *Touch, representation, and blindness* (pp. 143-181). Oxford, UK: Oxford University Press.

Diaz, R. M., & Berk, L. E. (Eds.). (1992). *Private speech: From social interaction to self-regulation*. Hillsdale, NJ: Erlbaum.

De Chambrier, A. F., & Zeiger, P. (2018). Is a fact retrieval deficit the main characteristic of children with mathematical learning disabilities? *Acta Psycologica,* 190 (95 -102).

Dolcos, S., & Albarracín, D. (2014). The inner speech of behavioral regulation: Intentions and task performance strengthen when you talk to yourself as a you. *European Journal of Social Psychology, 44*, 636-642.

Dowker, A. (2005). *Individual differences in arithmetic. Implications for psychology, neuroscience and education*. London, UK: Psychological Press.

Durand, M., Hulme, C., Larkin, R., & Snowling, M. (2005). The cognitive foundations of reading and arithmetic skills in 7- to 10-year-olds. *Journal of Experimental Child Psychology, 91*, 113-136.

Engle, R. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19-23.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 972-992.

Flavell, J. H., Green, F. L., Flavell, E. R., & Grossman, J. B. (1997). The development of children's knowledge of inner speech. *Child Development*, *68*, 39-47

Fuchs, L. S., Fuchs, D., Campton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., & Schatschneider, C. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology, 98*, 29-43.

Fürst, A. J., & Hitch, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory and Cognition, 28*, 774-782.

Fuson, K. C. (1979). The development of self-regulating aspects of speech: A review. In G. Zivin (Ed.), *The development of* *self-regulation through private speech* (pp. 135-217). New York, NY: Wiley.

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114*, 345-362.

Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: A cross-sectional study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology*, *77*, 236-263.

Geary, D. C., Hoard, M., & Bailey, D. H. (2012). Fact retrieval deficits in low achieving children and children with mathematical learning disability, *Journal of Learning Disabilities*, *45*(4), 291-307.

Girbau, D. (2002). A sequential analysis of private and social speech in children's dyadic communication. *Spanish Journal of Psychology*, *5*(2), 110-118.

Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review*, 79, 329-343.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory.* Newbury Park, CA: SAGE

Hammervoll, T., & Ostad, S. A. (1999). *Basiskunnskaper i matematikk* [Standard Mathematics Performance Test]. Oslo, Norway: Universitetsforlaget.

Harris, K. R. (1986). Effects of cognitive-behavior modifications on private speech and task performance during problem solving among learning-disabled and normally achieving children. *Journal of Abnormal Child Psychology*, 14, 63-76.

Hecht, S. A., Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2001). The relation between phonological processing abilities and emerging individual differences in mathematical computation skills: A cross-sectional study from second to fifth grades. *Journal of Experimental Child Psychology*, *79*, 192-227.

Jordan, N., Hanich, L. B., & Kaplan, D. (2003). Arithmetic fact mastery in young children: A cross-sectional investigation. *Journal of Experimental Child Psychology*, *85*, 103-119.

Jordan, N. C., & Montani, T. O. (1997). Cognitive arithmetic and problem solving: A comparison of children with specific and general mathematical difficulties. *Journal of Learning Disabilities, 30*, 624-634.

Kohlberg, L., Yaeger, J., & Hjertholm, E. (1968). Private speech: Four studies and a review of theories. *Child Development, 39*, 691-736.

Krajewski, K., & Schneider, W. (2009). Exploring the impact of phonological awareness, visual-spatial working memory, and preschool quantity-number competence on mathematics achievement in elementary school: Findings from a 3-year longitudinal study. *Journal of Experimental Child Psychology*, *103*, 516-531.

Kulak, A. G. (1993). Parallels between math and reading disability: Common issues and approaches. *Journal of Learning Disabilities*, *26*, 666-673.

Lemaire, L., Abdi, H., & Fayol, M. (1996). Working memory and cognitive arithmetic: Evidence from the disruption of associative confusion effect. *European Journal of Cognitive Psychology*, *8*, 73-103.

Logie, R. H., & Baddeley, A. D. (1987). Cognitive processes in counting. *Journal of Experimental Psychology*, *13*, 310-326.

Mabbott, D. J., & Bisanz, J. (2008). Computational skills, working memory, and conceptual knowledge in older children with mathematics learning disabilities. *Journal of Learning disabilities*, *41*, 15-28.

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition*, *22*, 107-157.

Matuga, J. M. (2003). Children's private speech during algorithmic and heuristic drawing task. *Contemporary Educational Psychology*, *58*, 552-572.

Ostad, S. A. (1987). *Matematikk-diagnostikk* [Math-Diagnostics]. Oslo, Norway: Universitetsforlaget.

Ostad, S. A. (1997). Developmental differences in addition strategies: A comparison of mathematically disabled and mathematically normal children. *British Journal of Educational Psychology*, *67*, 345-357.

Ostad, S. A. (1998). Developmental differences in solving simple arithmetic word problems and simple number-fact problems: A comparison of mathematically normal and mathematically disabled children. *Mathematical Cognition*, *4*(1), 1-19.

Ostad, S. A. (1999). Developmental progression of subtraction strategies: a comparison of mathematically normal and mathematically disabled children. *European Journal of Special Needs Education*, *14*(1), 21-36.

Ostad, S. A. (2000). Cognitive subtraction in developmental perspective. Accuracy, speed-of-processing and strategy-use differences in normal and mathematically disabled children. *Focus on Learning Problems in Mathematics*, *22*(2), 18-31.

Ostad, S. A. (2013). Private speech in arithmetical calculation: Contributory role of phonological awareness in children with and without mathematical difficulties. *Journal of Learning Disabilities*, *46*(4), 291-303.

Ostad, S. A. (2015). Private speech in arithmetic calculation: Interaction with phonological memory skills in children with and without mathematical difficulties. *Annals of Dyslexia*, 65(2), 103-119.

Ostad, S. A., & Askeland, M. (2008). Sound-based number facts training in a private speech internalization perspective: Evidence for effectiveness of an intervention in grade 3. *Journal of Research in Childhood Education*, *23*(1), 109-124.

Ostad, S. A., & Sorensen, P. M. (2007). Private speech and strategy-use patterns. Bidirectional comparisons of children with and without difficulties in developmental perspective. *Journal of Learning Disabilities, 40*(1), 2-14.

Passolunghi, M. C., & Cornoldi, C. (2008). Working memory failures in children with arithmetical difficulties. *Child Neuropsychology*, *14*(5), 387-400.

Passolunghi, M. C., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology, 88*, 348-364.

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual differences, and cognitive approaches. *Learning and Individual Differences, 20*(12), 110-122.

Ramus, F., & Szenkovits, G. (2008). What phonological deficits? *The Quarterly Journal of Experimental Psychology, 61*(1), 129-141.

Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic, *Journal of Experimental Child Psychology*, *91*, 137-154.

Robinson, C. S., Menchetti, B. M., & Torgesen, J. (2002). Toward a two-factor theory of one type of mathematics disabilities. *Learning Disabilities Research & Practice*, *17*(2), 81-89.

Seitz, K., & Schumann-Hengsteler, R. (2000). Mental multiplication and working memory. *European Journal of Cognitive Psychology*, *12*, 552-570.

Seyler, D. J., Kirk, E. P., & Ashcraft, M. H. (2003). Elementary subtraction. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*, 1339-1352.

Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development*, *60*, 973-980.

Siegler, R. S. (1988). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development, 59*, 833-51.

Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.

Siegler, R. S., & Shrager, J. (1984). Strategy use in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origin of cognitive skills* (pp. 229-293). Hillsdale, NJ: Erlbaum.

Snowling, M. J. (2001). *Dyslexia* (2nd ed.). Oxford, UK: Blackwell.

Stahl, S. A., & Murry, B. A. (1994). Defining phonological awareness and its relatiuonship to early reading. *Journal of Educational Psychology*, *86*, 221-234.

Swanson, H. L., (1994). Short-term memory and working memory: Do both contribute to our understanding of academic achievement in children and adults with learning disabilities? *Journal of Learning Disabilities*, *27*, 34-50.

Swanson, H. L., Ashbaker, M. H., & Lee, C. (1996). Learning disabled readers' working memory as a function of processing demands. *Journal of Experimental Child Psychology, 68*, 377-400.

Swanson, H. L. & Jerman, O. (2006). Math disabilities: A selective meta-analysis of literature. *Review of Educational Research*, *76*(2), 249-274.

Swanson, H. L. & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities. Both executive and phonological processes are important. *Journal of Experimental Child Psychology*, *79*, 299-321.

Thevenot, C., Barrouillet, P., Castel, C., & Uittenhove, K. (2016). Ten-year-old children strategies in mental addition: A counting model account. *Cognition*, *146*, 48-57.

Torgesen, J. K. (1999). Phonological based reading disabilities: Toward a coherent theory of one kind of learning disability. In L. Spear- Swerling & R. J. Stenberg (Eds.), *Perspectives on learning disabilities* (pp. 231-262). New Haven, CT: Westview Press.

Undheim, J. O. (1977). *Håndbok. WISC-R* [Wechsler Intelligence Scale for Children: Examiner's handbook]. Trondheim, Norway: Norsk Psykologforening.

Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin, Trans.). Cambridge, MA: MIT Press. (Original work published 1934)

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*, 192-212.

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

Winsler, A., & Naglieri, J. A. (2003). Overt and covert verbal problem-solving strategies: Developmental trends in use, awareness, and relations with task performance in children aged 5–7. *Child Development*, *74*, 659-678.

Winsler, A., Carlton, M. P., & Barry, M. J. (2000). Age-related changes in preschool children's systematic use of private speech in a natural setting. *Journal of Child Language*, *27*, 665-687.

Winsler, A., Diaz, R. M., & Montero, I. (1997). The role of private speech in transition from collaborative to independent task performance in young children. *Early Childhood Research Quarterly*, *12*, 59-79.

Winsler, A., Feder, M., Way, E. L., & Manfra, L. (2006). Maternal beliefs concerning young children's private speech. *Infant and Child Development*, *15*, 403-420.

The International Academy for Research in Learning Disabilities

44th Annual IARLD Conference

Preconference Activities – June 18, 2020
Conference – June 19 and 20, 2020
The Education University of Hong Kong, Hong Kong

Conference Chair:

Professor Kevin Chung
Director, Centre for Child and Family Science,
The Education University of Hong Kong

Cruickshank Memorial Lecture:

by Professor Catherine McBride
Professor, Department of Psychology,
The Chinese University of Hong Kong

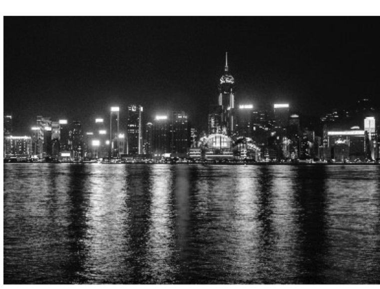
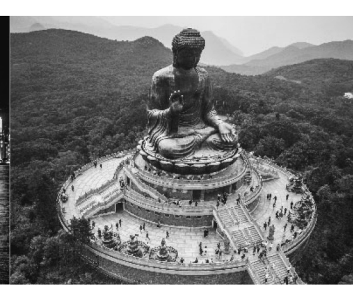
"The Write Stuff: What Do We Know About Dysgraphia?"

Proposal Submission Deadline:
February 29, 2020

Early Registration Ends:
April 30, 2020

Visit www.iarld.com for details

Conference Secretariat:
Centre for Child and Family Science
The Education University of Hong Kong
Email: ccfs@eduhk.hk
Phone: (852) 2948 8655


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 (i.e., 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 conveyance of 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/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: **ijrld@bc.edu** (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/2011/08/IJRLD-Call-for-Manuscripts.pdf

International Journal for Research in Learning Disabilities
Lynch School, Boston College
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)