Multidimensional etiology and individual differences in developmental dyslexia

Verwimp, C.T.

Publication date
2023

Document Version
Final published version

Link to publication

Citation for published version (APA):
Verwimp, C. T. (2023). *Multidimensional etiology and individual differences in developmental dyslexia*. [Thesis, externally prepared, Universiteit van Amsterdam].

MULTIDIMENSIONAL ETIOLOGY AND INDIVIDUAL DIFFERENCES

IN DEVELOPMENTAL DYSLEXIA

CARA VERWIMP

Multidimensional etiology and individual differences in developmental dyslexia

Cara Verwimp

This work was supported by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie (grant number 813546).

Printed by ProefschriftMaken
Cover design created with Canva

Multidimensional etiology and individual differences in developmental dyslexia

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op maandag 30 oktober 2023, te 14.00 uur

door Cara Tamara Verwimp
geboren te Herentals

Promotiecommissie

Promotor:	prof. dr. R.W.H.J. Wiers	Universiteit van Amsterdam
Copromotores:	dr. J. Tijms	Universiteit van Amsterdam
	dr. P.J.F. Snellings	Universiteit van Amsterdam
Overige leden:	prof. dr. K.R. Ridderinkhof	Universiteit van Amsterdam
	dr. B.R.J. Jansen	Universiteit van Amsterdam
	prof. dr. J.E. Rispens	Universiteit van Amsterdam
	prof. dr. M.P. Torppa	University of Jyväskylä
	prof. dr. S. Brem	University of Zürich

Faculteit der Maatschappij- en Gedragswetenschappen

Multidimensional etiology and individual differences in developmental dyslexia

ACADEMISCH PROEFSCHRIFT

1

General Introduction

Prologue

When someone asks about my research, I always start with a fascinating fact: “Did you know our brain is not made for reading? In fact, reading was quite recently invented by people, about 6000 years ago. Isn’t it weird that most children can learn to read so easily, and relatively few experience difficulties?”. Since most people assume reading is an innate ability, this often catches people off guard. In a world where literacy is more important than ever, being unable to read comes with many challenges, such as not understanding what is written on medical leaflets. Besides academic consequences, individuals with reading difficulties tend to have a higher risk of experiencing feelings of anxiety, failure, and depression (Mugnaini et al., 2009). It is, therefore, unsurprising that researchers have been investigating how children learn to read and how they should be taught to read for decades. The field of reading disabilities research has shifted from a core-deficit approach, in which one searched for the holy grail that causes dyslexia, to a multiple-deficit approach, in which various factors interact (Pennington, 2006). More recently, a shift from a categorical DSM-like perspective to a transdiagnostic approach that spans across traditional diagnostic boundaries happened as well (Astle & Fletcher-Watson, 2020; Cuthbert & Insel, 2010). Despite these shifts, most studies still examine these factors in isolation even excluding individuals with co-occurring difficulties. As a result, the homogeneity within individuals who experience reading difficulties is overestimated, and it remains unclear how factors within and across diagnostic boundaries interact with each other. The first section of this doctoral thesis aims at contributing to existing research by mapping the complexity of factors related to reading development by means of network analysis (Chapter 2) and shedding light on attentional processes that are involved in the development and integration of cross-modal letter-speech sound correspondences by integrating behavioral and neural measures (Chapter 3 and 4). The second section of this thesis focuses on processes that are involved in reading intervention, aiming to provide more insight into the significant reported inter-individual differences in intervention response and ameliorating interventions for struggling readers.

In this introduction, we present a brief overview of how we learn to read and the traditional definition of dyslexia, which has shifted toward a complex definition, followed by an overview of the most relevant deficits associated with dyslexia, including the role of executive functions. Finally, we will discuss how dyslexia can be remediated and which role digitalization can play in this remediation.

Learning to read

Around 5000 years ago, humankind invented the script. What was invented as tokens to keep track of goods slowly developed into contemporary scripts that provide the key to knowledge, learning, and technology (Landsbergen, 2021). Writing systems vary substantially, but the thing they have in common is that written symbols somehow represent spoken language. In order to read, one needs to crack the symbolic (e.g., alphabetic) code of their language. That is, one needs to understand that written symbols map onto language representations (Castles et al., 2018). In alphabetic languages, such as English and Dutch, these symbols typically represent individual sounds (i.e., phonemes), in contrast to syllabic and logographic languages, in which symbols refer to syllables (e.g., Japanese Hiragana), or elements that contain both meaning and sound (e.g., Chinese). Although oral language acquisition is thought to develop automatically from daily interactions between children and their caregivers, reading acquisition is an effortful and complex skill that requires formal instruction (Ehri et al., 2001).

According to Ehri (1995), learning to read comprises four phases. In the first phase, referred to as the pre-alphabetic phase, children start to recognize words visually. For example, by remembering the word *look* by the two eyeballs in the middle (Gough et al., 1992). At this point, children have no knowledge of the underlying alphabetic principle and are, therefore, considered non-readers. From the moment they start to recognize connections between some letters and speech sounds, they move toward the partial alphabetic phase. Children typically recognize the most salient letters of the word, or the letters a word begins and ends with. As knowledge of all letter-speech sound correspondences is incomplete, similarly spelled words are often confused (e.g., skin and spoon). In the full alphabetic phase, children know most of the letter-speech sound mappings, can sound out most familiar words, and accurately decode unfamiliar words. As a result, reading becomes more accurate, and similarly spelled words are less confused. Last, when letter-speech sounds mappings become more integrated and automatized, children enter the consolidated phase. Correspondences become consolidated in larger units, and children can read and recognize words quickly and accurately without sounding out every speech sound. It is important to note that, although reading happens automatically, children still rely on phonological information when retrieving a word from the mental lexicon (Frost, 1998; Zeguers et al., 2014). Studies have shown that children who struggle with reading often get "stuck" in the partial alphabetic phase and may need additional support to move on to the full alphabetic phase (e.g., Ehri, 1995).

How fast we move toward the consolidated phase depends on how transparent our language is (Schmalz et al., 2015), that is, how consistently speech sounds map onto visual symbols. In languages with consistent mappings, like Finnish, children typically learn to read accurately more quickly, as there are fewer exceptions on how a symbol should be pronounced. In deep languages such as English, learning to read can be more challenging as the pronunciation often depends on the position of a letter in the word (e.g., pronunciation of the letter *e* in the word *recipe*) and its morphological context (e.g., *nation* vs. *national*). Additionally, deep languages contain more irregular words that do not follow the standard pronunciation rules for that language. Essential to note is that phase models like Ehri's (1995) have been subject to criticism. Not all children pass through each phase and the proposed order of phases seems to depend on the teaching strategy used for reading instruction (Johnston & Thompson, 1989), and the language the child is learning to read in (Wimmer & Hummer, 1990). In more transparent languages, these results point toward individual variation, in which children can follow different paths toward proficient reading (Baron & Treiman, 1980).

There is considerable variability in how proficient reading should be defined. Some proposed that one is proficient when one can read accurately at a conversational rate with appropriate prosody (e.g., Hudson et al., 2005). Samuels (2006) argued that fluent reading is "decoding and comprehending at the same time" (p. 39) and suggested that decoding and comprehension are likely to occur together when processes operate automatically (LaBerge & Samuels, 1974). To be considered automatic, a process should happen fast and autonomously, without much conscious control or attention (Schneider & Chein, 2003). In proficient readers, familiar words are easily identified at first glance, better known as sight word reading (Ehri, 2005). Even if the reader does not intend to read the word, semantic information is extracted, as evidenced by the Stroop Effect (MacLeod, 1991). When word recognition happens automatically, and thus requires few cognitive resources, attention can be directed toward understanding what is read (LaBerge & Samuels, 1974).

Although most can learn to read relatively effortlessly, our brain is not hardwired for this. As children learn to read, existing brain areas are gradually reshaped into an integrated audiovisual network (Dehaene et al., 2015; LaBerge & Samuels, 1974; Romanovska & Bonte, 2021). Initially responsible for processing visual stimuli such as faces, visual brain areas become gradually connected with auditory brain areas. As a result, auditory areas start responding to written text in addition to spoken language (Bonte et al., 2017; Froyen et al., 2009; Kronschnabel et al., 2014), and visual areas become more specialized in recognizing our script (Amora et al., 2022; Brem et al., 2005; Maurer et al., 2005). Stronger connections are associated with fluent, effortless reading, whereas audiovisual integration is found to be reduced or less automatic in poor readers (Breznitz, 2002; Dehaene et al., 2015; Yap & van der Leij, 1993).

Children seem to easily acquire the knowledge of which speech sounds map onto which symbols. However, evidence showed that more than mere knowledge of these correspondences is required for fluent reading (Froyen et al., 2009). In line with more general accounts of skill acquisition (Schneider & Chein, 2003), complex skills such as reading take much exposure and practice to happen automatically. As Schneider & Chein (2003) noted, "*extended consistent training is required in order to develop automatic processing, while controlled processes can be established in a few trials and under varied mapping situations*." This aligns with brain studies that have provided insights into developing fluent reading skills. Although children knew all letters, studies by Froyen and colleagues (2008, 2009) found that beginning readers showed no neural signs of letter-speech sound integration after one year of formal reading instruction. Moreover, even after four years of reading, brain responses were not comparable to those of adults. Evoking reshaping in the brain such that perception of a letter automatically activates paired auditory information thus requires much exposure and practice, and can only be obtained when associations are highly overlearned (van Atteveldt et al., 2004, 2006). Given the complexity of processes involved in this reshaping of the brain that goes along with the transition from effortful, conscious decoding toward automatic, autonomous reading, it is not surprising that some individuals have more difficulties obtaining automatized reading than others, with the lowest end of the normal distribution typically being diagnosed with dyslexia (Shaywitz et al., 1992).

Developmental dyslexia and its definition

Since 2013, developmental dyslexia has been described as a specific learning disorder characterized by difficulties with accurate or fluent word recognition, poor spelling, and poor decoding abilities (American Psychiatric Association, 2013). This persistent reading disorder has a neurobiological basis and cannot be explained by a general learning deficit, sensory deficits, or inadequate instruction. Prevalence estimates vary between 5-15% of the population, depending on the criteria and exact definitions used (Peterson & Pennington, 2015; Snowling, 2013). In order to examine the persistence of reading and spelling difficulties, the Dutch education system employs a Response to Intervention framework. Children that get behind with reading receive gradually increasing reading support. When children lag behind in reading and spelling performance after receiving adequate formal reading and spelling instruction (Tier 1), they obtain more intensive, tailored support in the classroom (Tier 2), with even more rigorous, tailored support in smaller groups or individually for at least six months if necessary (Tier 3). When children perform below the 10th percentile compared to the relevant norm group on a standardized reading task on three consecutive assessment moments despite additional support, reading problems are considered to be severe and persistent (Tijms et al., 2021). As a result, children are referred to the clinical center for diagnostic assessment. A dyslexia diagnosis in the Netherlands depends on meeting multiple, specific criteria, including word reading performance 1.5 SD below average, or word reading performance 1.28 SD below average combined with spelling skills 1.5 SD below average (Tijms et al., 2021).

A dyslexia diagnosis is typically given after several years of reading instruction, often leading to negative consequences for children and their families and making it more challenging to make up for a significant gap in reading ability and experience at a later age (Vaughn et al., 2009). Consequences include poor educational attainment (Smart et al., 2017) and occupational difficulties (de Beer et al., 2014). However, dyslexia is also associated with various internalizing and externalizing problems (Wilmot et al., 2023). The most reported are depression and anxiety (Francis et al., 2019), which might be unsurprising as longitudinal research has provided evidence that depressive and anxious symptoms are often preceded by academic difficulties (Maughan et al., 2003). Other reported consequences include low self-esteem, high stress levels, and inefficient coping skills, sometimes even leading to somatic complaints (Wilmot et al., 2023). Besides the reported negative consequences of dyslexia in children, research has also shown a significant impact on the psychosocial well-being of their parents, including increased stress levels, and worrying about the well-being and future prospects of their child (Bonifacci et al., 2016; Craig et al., 2016; Matteucci et al., 2019).

Terminology in the field of dyslexia has been found to be inconsistent. Terms that are typically used are ‘reading impaired’, ‘reading disabled’ or ‘reading disordered’ (Snowling, 2000). Some of these terms have been criticized for implying that children with dyslexia are generally impaired, while other terms only suggest difficulties in acquiring reading skills. In fact, terms are interchangeably used without thinking about their implicit meaning (Snowling, 2000). Along with a plethora of terms, dyslexia as a label has come into question as well (Elliott & Grigorenko, 2014). Comparable to other neurodevelopmental disorders, the diagnosis is based on a somewhat arbitrary cut-off value on a continuous variable. As a result, criticism has been raised about whether children at the lowest tail of the distribution should be labeled, as this leads to artificial subgroups for which children who perform on the lowest end of the normal distribution can be classified as either poor readers or controls depending on the exact cut-off criteria. In addition, the traditional definition of dyslexia lacks a sufficient or necessary underlying deficit that can explain the heterogeneity in behavioral outcomes (Pennington et al., 2012). Although specific criteria are needed to facilitate healthcare policies, a shift toward a more multidimensional, transdiagnostic approach might be needed to account for individual differences and overlap between disorders (Astle & Fletcher-Watson, 2020).

Etiology of dyslexia: From a single to multifactorial view

Since the German ophthalmologist Rudolf Berlin introduced the term 'dyslexia' as *Wortblindkeit* in his book 'Eine besondere Art der Wortblindheit (Dyslexie)' (1887), numerous theories have been proposed to explain the development of dyslexia. One prevailing theory is that dyslexia is caused by a phonological deficit (Pennington, 2006). This theory implies that difficulties in representing, storing, and/or retrieving speech sounds underlie poor reading (Snowling, 1998). Research has consistently shown that individuals with dyslexia have difficulty discriminating and manipulating units of sounds in spoken language, which may consequently impact their ability to establish proper letter-speech sound mappings (Melby-Lervåg et al., 2012). Although this so-called phonological awareness is one of the strongest predictors of reading development, not all individuals with poor phonological abilities develop dyslexia, and not all individuals with dyslexia suffer from a phonological deficit (Boets et al., 2008). Moreover, phonological awareness seems to develop reciprocally with increased reading experience and, therefore, cannot be the only cause of reading difficulties (Bishop, 2006; Boets et al., 2010).

Two other common predictors of reading acquisition and development are rapid automatized naming (RAN) and verbal short-term memory (Caravolas et al., 2012; Swanson et al., 2009). Both are commonly included under the umbrella of phonological processing, as both require access to phonological codes (Melby-Lervåg et al., 2012). RAN tasks require rapidly naming an array of well-known visual stimuli aloud, commonly digits, letters, objects, or colors, as accurately as possible (Araújo et al., 2015). The rapid naming of visual items has been found to significantly predict later reading outcomes (Norton & Wolf, 2012). Verbal short-term memory, the capacity to store verbal information, is typically assessed using a word or digit span task (Swanson et al., 2009). Evidence indicated lower performance in poor readers on these memory span tasks; fewer elements could be recalled in poor readers compared to controls (Majerus & Cowan, 2016).

A more recent account proposed that dyslexia is not the result of a phonological deficit that prevents the establishment of letter-speech sound mappings per se but a deficit in storing automatized audiovisual objects in the brain (Blomert, 2011). Several studies have suggested that a deficit in the automatic integration of letters and speech sounds underlies slow and effortful reading and inadequate phonemic abilities during reading acquisition (Aravena et al., 2018; Blau et al., 2010; Kronschnabel et al., 2014; Žarić et al., 2014). These findings have been corroborated by neuroimaging studies that showed reduced differential activation for congruent versus incongruent letter-speech sound pairs in pre-readers at risk for dyslexia, as well as in children and adolescents with dyslexia compared to control groups (Karipidis et al., 2021; Wang et al., 2020; Žarić et al., 2015).

Other researchers argued that poor reading results from a more fundamental deficit in the visual or auditory domain. Examples include difficulties in processing auditory information (Nagarajan et al., 1999), a magnocellular deficit that results in low-level visual deficits in sensory temporal processing (Stein, 2001), impaired foveal and parafoveal processing of visual stimuli (Jones et al., 2012), or impaired control of eye movements (Premeti et al., 2022). However, all of them still imply that having a deficit in either component is sufficient for developing reading difficulties, although none can sufficiently explain the large variability observed among dyslexic readers. This either suggests that the actual underlying deficit is yet to be defined or that looking for a single deficit is inappropriate in the context of neurodevelopmental disorders. In addition, simply focusing on cause and effect seems to be inappropriate as well, given the reciprocal associations between many of these proposed causes and reading experience (Huettig et al., 2018).

As a single deficit model does not account for the heterogeneity of symptoms reported in dyslexic readers, the validity has come into question (Pennington, 2006; Pennington et al., 2012). Moreover, a single deficit explanation does not account for comorbidity or symptom overlap among disorders. As a result, the focus has shifted from deterministic models that were looking for the single underlying cause toward probabilistic models, in which a disorder results from complex interactions between symptoms (Pennington, 2006; van Bergen, van der Leij, et al., 2014). Multiple factors are thought to interact with other risk and protective factors, either in- or decreasing the risk of developing dyslexia. These factors and their interactions thus alter behavioral outcomes in dyslexia, in which no single factor is sufficient to develop dyslexia, resulting in a continuously distributed liability for reading failure. Moreover, in line with the Research Domain Criteria (RDoC) approach for investigating mental disorders, factors can cut across domains or constructs (e.g., cognitive, neurobiological, affective, genetic, and environmental) (Cuthbert & Insel, 2010; Insel et al., 2010). For example, specific genetic risk loci have been found to be associated with difficulties in working memory and phonological processing, which may modulate reading performance (Landi et al., 2013). Indeed, studies showed clear evidence that dyslexia is highly heritable, with a 3-4 times higher risk of developing dyslexia when a child has a dyslexic first-degree relative (Snowling & Melby-Lervåg, 2016). Typical environmental factors that are associated with literacy outcomes include socio-economic status, education level, home literacy environment, and parental reading skills (Becker et al., 2017; Khanolainen, 2020; Torppa et al., 2007; van Bergen, de Jong, et al., 2014). Studies reported less shared reading between fathers with dyslexia and their children than controls (van Bergen, de Jong, et al., 2014) and less frequent book reading by parents in at-risk families (Torppa et al., 2007). Given the potential variety of phenotypic manifestations and possibly distinct pathways toward reading difficulties, in-depth diagnostics are essential for identifying problems and providing concrete starting points for treatment.

Beyond the language domain: Executive functions in reading

Upon the shift toward a multifactorial view of dyslexia, one started to look into processes that cut across traditional diagnostic boundaries (Hawkins et al., 2016; Holmes et al., 2021). One of these processes are executive functions. Executive function difficulties are thought to be one of the main characteristics of ADHD (Hulme & Snowling, 2013), but have also been reported in individuals with dyslexia, such as difficulties in updating information and inhibiting distractors (Brosnan et al., 2002; Smith-Spark et al., 2016). This might be unsurprising as reading difficulties and attention deficit hyperactivity disorder (ADHD) co-occur more frequently in clinical samples than what would be expected by chance (Willcutt & Pennington, 2000). As a result, there has been considerable interest in understanding how these cognitive processes contribute to the development of complex skills, such as reading, as it not only influences reading development but might also impact responsiveness to intervention (Sesma et al., 2009).

Although there is a lack of agreement on the exact definition of executive functions, it is generally seen to encompass a wide variety of processes that coordinate cognition and facilitate the completion of goals (Jurado & Rosselli, 2007). Anderson (2002) defined executive functions as “*a collection of inter-related processes responsible for purposeful, goal-directed behavior,*” such as “*anticipation, goal selection, planning, initiation of activity, self-regulation, mental flexibility, deployment of attention, and utilization of feedback*” (Anderson, 2002). It is worth noting that the term “executive function” is used interchangeably with “cognitive control”, as both terms are meant to capture the same processes (Menon & D’Esposito, 2022). Where Baron (2004) listed more than 20 subdomains of executive functions, studies employing factor analysis suggested that these subdomains mainly load on three factors: inhibition, working memory, and shifting (Huizinga et al., 2006; Miyake et al., 2000). Inhibition refers to suppressing automatic responses in favor of a subdominant response, typically examined by the Stroop and stop-signal tasks. Working memory refers to the ability to update information stored in memory for a short time while doing something with it. Last, shifting is defined as the ability to switch between tasks and goals, typically assessed with the Wisconsin Card Sorting Task (Nyhus & Barceló, 2009) or the Trail-Making Test (Arbuthnott & Frank, 2000). Brain studies with frontal lobe injury patients observed that patients performed well in some executive functioning tasks while failing in others (Godefroy et al., 1999), suggesting that executive functions exist as modular, separable processes. However, Miyake et al. (2000) showed that they share some underlying commonality, possibly due to underlying inhibitory processes involved in all three executive functions.

The development of executive functions is tightly connected to the development of reading. Children at risk for dyslexia (i.e., with a first-degree relative with dyslexia) have been found to demonstrate challenges in selective attention and visuospatial short-term memory. Executive skills at the age of 4.5 predicted the development of dyslexia over and above the typical predictors of reading (Thompson et al., 2015). In school-age children, studies showed deficits in planning, whereas teenagers show more deficits in processing speed. These results indicate the lifelong development of executive functions, with some developing only later in life (Thompson et al., 2015). This is thought to be the result of the maturation of the frontal lobe, with increasing connections within and between the frontal lobe and other brain regions in older individuals (Fair et al., 2009). More general accounts of skill learning propose that automatic associative processes gradually replace controlled metacognitive processing as expertise grows (Schneider & Chein, 2003). When learning a complex skill, such as reading, directing attention toward relevant information and goals facilitates acquiring and consolidating new skills, such that the task can ultimately be executed automatically (LaBerge & Samuels, 1974; Yoncheva et al., 2010).

Besides difficulties that were reported in individuals with dyslexia in tasks specifically assessing executive functions, impaired executive functions might also directly interfere with the development of abilities that are needed for the acquisition and development of reading, such as efficiently discriminating between phonemes and establishing automatized letter-speech sound mappings (ten Braak et al., 2018; ten Oever et al., 2016). For example, to efficiently discriminate between phonemes, (auditory) attention should be directed toward the most informative acoustic dimensions for speech categorization (Francis & Nusbaum, 2002). According to Reynolds (2000), good readers are the ones who successfully allocate attentional resources to more resource-intensive processes, such as meta-comprehension, as essential reading skills are automatized. In contrast, poor readers allocate attentional resources to decoding and basic comprehension skills, leaving fewer resources available for higher-level processing.

On a neural level, brain lesions and neuropsychological studies have shown that executive processes can be attributed to two specific regions within the prefrontal cortex: the dorsolateral and orbitofrontal cortices (Iversen & Dunnett, 1990; cited in Welsh et al., 2006). Banfield et al. (2004) suggested that different brain regions mediate different executive processes. The dorsolateral prefrontal cortex and its connections with subcortical structures (e.g., basal ganglia and thalamic structures) are thought to be more involved in planning, strategic behavior, flexibility, and working memory, whereas the orbitofrontal prefrontal cortex and its subcortical structures are thought to be more involved in emotion regulation and self-monitoring. A specific process that was examined in this thesis is the monitoring of performance (Chapter 3). To monitor and evaluate the consequences of our behavior, we rely on internal and external feedback (Ferdinand & Kray, 2014). Performance monitoring is typically assessed using event-related potentials, although it also has been examined by heart rate changes (Crone et al., 2004; Fraga González et al., 2019). The error-related negativity (ERN) occurs around 100 ms after committing an error, and the feedback-related negativity (FRN) between 250 and 350 ms after obtaining external feedback, both having a frontocentral distribution. Both components are thought to be generated in the anterior cingulate cortex (ACC) (Ferdinand & Kray, 2014; Tamnes et al., 2013). Different hypotheses exist regarding the significance of these components, but according to the reinforcement theory, both the ERN and FRN are thought to be elicited by consequences that are worse than expected. That is, if the consequence of one’s action was not in line with what one was expecting, ERN and FRN amplitudes are higher, consequently inducing learning (Holroyd & Coles, 2002).

Studies reported a less efficient error detection system in individuals with dyslexia, in the sense that diminished ERN amplitudes were elicited after error commission compared to typical readers (Horowitz-Kraus, 2011; Horowitz-Kraus & Breznitz, 2008). In addition, a study that examined the feedback-locked FRN response showed lower FRN amplitudes in adolescents with dyslexia compared to peers (Kraus & Horowitz-Kraus, 2014). However, in the later trials of the task, FRN amplitudes in individuals with dyslexia increased to similar levels as controls, which the authors interpreted as a compensation mechanism that becomes apparent with the repetition of a task. Interestingly, studies that examine performance monitoring in individuals with dyslexia using linguistic tasks are scarce, so it remains unclear how performance monitoring is related to the acquisition and development of reading.

Evidence-based instruction and intervention

For decades, one has been discussing how we should teach children to read (Castles et al., 2018). This discussion was most apparent in Anglocentric countries. The field was switching between two opposing views: a phonics approach, in which letters and their speech sounds are taught explicitly, and a whole language approach, which favors the discovery of meaning through experiences in a literacy-rich environment (e.g., Goodman, 1967). A phonics method uses a stepwise approach, starting with explicit and systematic teaching of letter-speech sound mappings. Teaching children the alphabetic principle is thought to consequently give them access to the word's meaning (Castles et al., 2018). Whole language learning emphasizes the semantic properties of the word, therefore embedding it in its context. They believe that children will, as a result of reading and writing experience, naturally discover how letters are related to their sound (Bowers, 2020; Huang, 2014). As extensive national surveys have provided empirical evidence in favor of the phonics method (National Reading Panel, 2000; Rose, 2006), most countries nowadays rely on phonics to teach children how to read. However, it is essential to note that, according to Bowers (2020), meta-analyses typically compile studies using a whole-language approach and a sight-word reading approach. The difference in these methods lies in how they sort with context; where sight-word reading focuses on identifying words out of context, the whole-language approach emphasizes meaning using surrounding words.

In Dutch education, most children are already exposed to text before entering elementary school. Around the age of 6, children obtain formal reading instruction, although some of them are already introduced to phonemes and letters in kindergarten through playful activities (Reitsma & Verhoeven, 2016). Around 80% of Dutch elementary schools uses the *Veilig Leren Lezen* method (Learning to read safely; Mommers et al., 1994). This method focuses on mappings between spoken and written language while using sight words as concrete examples of how to use the mappings within a context. Dutch children typically obtain knowledge of all mappings within one year of formal reading instruction (Froyen et al., 2009).

Concerning interventions, a growing number of studies has investigated the effectiveness of interventions in remediating reading and spelling difficulties (Galuschka et al., 2014; Scammacca et al., 2015). Effective interventions typically comprise a combination of phonics instruction, decoding strategies, and reading and spelling activities (Partanen & Siegel, 2014; Tijms, 2007; Tilanus et al., 2016; Vaughn et al., 2009). Effective training involved in specialized treatment should be intensive, systematic, and explicit (Shaywitz et al., 2008), with an average duration between 45 and 65 hours according to the Dutch dyslexia protocol (Tijms et al., 2021). Interventions that target isolated components of reading are often ineffective (Galuschka et al., 2014). For example, a study of Pfost et al. (2019) failed to improve reading and comprehension skills after specifically targeting phonological awareness and letter knowledge.

Although most children with dyslexia obtain appropriate levels of reading accuracy compared to peers after intervention (Tijms, 2007; Tilanus et al., 2016), fluent reading is less susceptible to intervention, with some children even not responding at all (Torgesen et al., 2001). Torgesen and colleagues (2001) showed that even after 65.5 hours of treatment, children with dyslexia hardly increased in reading fluency, despite obtaining average accuracy levels. These results raised interest in whether reading fluency can be treated and which elements should be included in such an intervention. One line of research has examined the effect of repeated reading, in which children read the same words while obtaining feedback from teachers or peers (O'Connor et al., 2007). However, the improved reading rate for trained material did not successfully transfer to undertrained words (Berends & Reitsma, 2006). Another approach employed a reading acceleration program, in which sentences faded out letter by letter (Breznitz, 2006; Snellings et al., 2009). The fading rate was based on the child's reading rate, using comprehension questions to check whether the rate had to be decreased. Studies revealed mixed evidence, with some showing that text-fading training did not lead to a more significant reading fluency improvement compared to other existing training methods (Paige, 2011), whereas others found differential transfer effects on sentence reading fluency (Nagler et al., 2015). Last, Fraga González et al. (2016) examined the efficacy of a training program specifically focusing on developing automatized letter-speech sound mappings. The authors found more significant reading fluency gains in the trained dyslexic group compared to the waitlist and control conditions.

Another way to train letter-speech sound mappings by repeated practice and massive exposure is with gamified interventions. These have increasingly gained interest in recent years (Jaramillo-Alcázar et al., 2021; Lassault et al., 2022; Skiada et al., 2014; Yildirim & Surer, 2021), as they allow for targeted, individualized practice. Massive and repetitive training of letter-speech sound correspondences is thought to facilitate the neural shaping of our brain for fast integration of unique audiovisual orthographic–phonological objects (Blomert, 2011). In addition, game-based interventions are thought to maintain children's motivation and engagement (Prensky, 2003), which might be especially necessary for children who have to practice a compromised skill for a prolonged period, such as reading (Froyen et al., 2009; Stafford & Vaci, 2022). Various games have been developed in recent years in the context of reading. Some examples are GraphoGame (Lyytinen et al., 2009), ReadingTurbo (van Uittert et al., 2021), and JerenAli (Kartal & Terziyan, 2016). Most interventions were reported to enhance specific measures essential for reading but did not find a transfer to word reading fluency skills. The reported mixed evidence indicates the need for randomized experimental designs that provide insights into the effectiveness of the games and factors that might moderate this effectiveness (Kim et al., 2021). By addressing these issues, employing game-based interventions might hold significant promise to help individuals with reading difficulties in improving their reading skills.

Outline

In this thesis, we aimed to provide more insight into the multidimensional etiology of developmental dyslexia, more insight into neurocognitive mechanisms that underlie the proposed letter-speech sound automatization deficit that is thought to hinder fluent reading and how these can be translated into clinical practice, and more insight into how co-occurring ADHD symptoms alter the reading development. In *Chapter 2*, we employed a network analysis in an extensive clinical database to examine the complexity of the reading network, examining associations between cognitive, environmental, and demographic variables related to reading, aiming to identify intervention targets. In *Chapter 3*, we examined the role of cognitive control processes concerning the multisensory integration deficit frequently reported in individuals with dyslexia. We mimicked the initial formation of the neurocognitive reading network by an artificial letter-speech sound mapping learning task in a general school-aged sample. Children received either instructions that directed them toward the goal of the task or instructions that told them that the goal of the task should be discovered by themselves. Accuracy and reaction time during the learning task were used to map learning curves, allowing to examine the learning trajectory in detail instead of focusing on learning outcomes only. In *Chapter 4*, we built further on the previous chapter by examining the effects of instruction manipulation concerning neural feedback processes in children with dyslexia and typical readers. We aimed to shed light on the interaction between behavioral and neural cognitive control processes, i.e., goal-directed behavior, post-error slowing, and neural performance monitoring, as well as how behavioral ADHD measures impact letter-speech sound learnability. We aimed to identify factors involved in the typical and atypical integration of letters and speech sounds, which can possibly be targeted in intervention. The second section of this thesis sheds light on processes involved in reading intervention. In *Chapter 5*, we created a detailed window concerning the reciprocal associations between the development of letter-speech sound associations and word reading fluency within the treatment of dyslexia using a structural equation modeling approach. In *Chapter 6*, we examined whether online specialized reading treatment is as effective as face-to-face treatment in improving children's reading and spelling skills, as interventions massively moved to online platforms due to the global COVID-19 pandemic. Last, *Chapter 7* discusses the results of a randomized controlled trial. We examined the efficacy of a game-based intervention tool in improving letter-speech sound knowledge in pre-readers and the transfer to word reading. The main results of all chapters are summarized in *Chapter 8* and put into a broader perspective. Implications in a practical and clinical context are discussed, as well as future directions of research.

The present work is the result of an international collaborative effort between the University of Amsterdam (UvA), Rudolf Berlin Center (RBC), and the RID (clinical center for children with learning disabilities), conducted as part of a European Innovative Training Network (MSCA-ITN), named Neo-PRISM-C. The RID was essential in the recruitment of children with dyslexia. This work was funded by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Actions (MSCA, grant number 813546).

2

A network approach to dyslexia: Mapping the reading network

A version of this chapter is published as: Verwimp, C., Tijms, J., Snellings, P., Haslbeck, J. M. B., & Wiers, R. W. (2021). A network approach to dyslexia: Mapping the reading network. *Development and Psychopathology*, 1–15. https://doi.org/10.1017/S0954579421000365

Abstract

Research on the etiology of dyslexia typically uses a single core-deficit approach, failing to understand how variations in combinations of factors contribute to reading development and how this combination relates to intervention outcome. To fill this gap, this study explored links between 28 cognitive, environmental and demographic variables that are related to dyslexia by employing a network analysis using a large clinical database of 1257 elementary school children. We found two highly connected subparts in the network: one comprising reading fluency and accuracy measures and one comprising intelligence-related measures. Interestingly, phoneme awareness was functionally related to the controlled and accurate processing of letter-speech sound mappings, whereas rapid automatized naming was more functionally related to the automated convergence of visual and speech information. We found evidence for the contribution of a variety of factors to the (a)typical reading development, though associated with different aspects of the reading process. As such, our results contradict prevailing claims that dyslexia is caused by a single-core deficit. This study shows how the network approach to psychopathology can be used to study complex interactions within the reading network and discusses future directions for more personalized interventions.

A network approach to dyslexia: Mapping the reading network

Developmental dyslexia (henceforth dyslexia) is a neurodevelopmental disorder characterized by a failure to develop accurate and fluent reading (Ozernov-Palchik et al., 2016). It is estimated to affect 3-10% of children, depending on the exact definition and criteria used for diagnosis (Snowling, 2013). Children who lag behind in their early reading abilities often remain poor readers throughout their school years and beyond, resulting in a large gap with typically developing peers (Ferrer et al., 2015; Torgesen, 2002). Due to persistent reading problems and discouraging experiences, these children are at serious risk for negative psychosocial (e.g., anxiety, depressive thoughts and feelings of shame) and academic consequences (Hendren et al., 2018; Livingston et al., 2018; Mugnaini et al., 2009). Therefore, effective reading intervention is needed to prevent such long-term consequences.

Current interventions mainly focus on phonics instruction, comprising repeated reading practice and phoneme awareness training, along with decoding strategies (Fraga González et al., 2015; Galuschka et al., 2014; Scammacca et al., 2015). However, a meta-analysis of Galuschka and colleagues (2014) showed that dyslexia remediation only moderately improved foundational reading skills (mean $ES = .49$). Moreover, although interventions for dyslexia have been found to produce significant long-lasting improvements (Tijms, 2007), there is wide inter-individual variability in intervention response, with a substantial number of children with dyslexia that remains hampered in fluent reading (Torgesen et al., 2001). These results indicate the need for more insight into the etiology of dyslexia and how this relates to successful reading intervention.

Traditionally, the dominant view in psychopathology research was a single deficit model, which espoused that each neurodevelopmental disorder was associated with one specific, single underlying cognitive deficit, such as a phonological deficit in dyslexia (Pennington & Lefly, 2001). However, as the single deficit model failed to account for co-occurrence of and interindividual variability within developmental disorders, recent advances in psychopathology started to acknowledge developmental disorders as multidimensional (Astle & Fletcher-Watson, 2020; Thomas, 2020). That is, rather than being conceptualized as a latent condition that unidirectionally causes a set of symptoms, the modern view of psychopathology shifted toward a rather dimensional approach in which mental disorders involve complex interactions between symptoms. This is in line with the Research Domain Criteria approach of the National Institute of Mental Health (NIMH; Insel et al., 2010; Sanislow et al., 2019), which advocates that levels of information (from genomics to behavior) should be integrated, cutting across traditional diagnostic boundaries, in order to understand the full range of mental health and illness. It follows from this idea that multiple interacting risk factors within and across different levels (e.g., cognitive, neurobiological, affective, genetic and environmental), operate probabilistically by increasing (or decreasing) the risk that one develops a reading disorder (Pennington, 2006; van Bergen, van der Leij, et al., 2014).

In the present study, we aim to provide a more detailed window on the multidimensional character of dyslexia. Therefore, we drew on the pertinent literature to identify a wide range of factors that might be involved in the atypical reading development that characterizes dyslexia. In addition to poor reading and spelling skills, deficits in phoneme awareness and rapid automatized naming (RAN) are the two most frequently reported problems in dyslexia (Caravolas et al., 2012; Georgiou et al., 2012). Phoneme awareness refers to the ability to segment and manipulate the phonemes within a word (Gellert & Elbro, 2017), and is assumed to play an important role in the early reading phases of accurate word decoding. RAN reflects the ability to rapidly activate phonological codes from highly familiar visual stimuli, such as letters and digits, and is found to be especially associated with reading fluency (Araújo et al., 2015). As learning to read involves matching visual symbols (i.e., letters) to corresponding speech sounds, the formation of letter-speech sound mappings received growing interest in recent years (e.g., Blomert, 2011; Žarić et al., 2014). It has been shown that a deficit in developing automatized letter-speech sound associations results in dysfluent reading (Aravena et al., 2013; Blomert, 2011; Žarić et al., 2014). In addition, deficits in verbal memory and working memory have been reported in dyslexia (e.g., Scarborough, 1998; Swanson et al., 2009). Children with dyslexia were found to be impaired in recalling phoneme and digit sequences (i.e., verbal memory) and in storing and manipulating information (i.e., working memory; Swanson et al., 2009). Other predictors that have been put forth are broader language skills, such as expressive and productive vocabulary (e.g., Torppa et al., 2010). Slow vocabulary development and poor productive language skills appeared to be predictive for later reading disabilities (Snowling et al., 2003). Others suggest that the origin of dyslexia might be related to a deficit in visual processing. More specifically, dyslexics reported problems with spatial organization, visual perception and attention (Stein & Walsh, 1997).

All these factors are found to be related to (a)typical reading development. However, behavioral predictors only account for about half of the variance in later reading abilities. Family risk studies have shown that children with a first-degree relative who has dyslexia are more likely to develop dyslexia themselves (Harlaar et al., 2005; Torppa et al., 2010), indicating the importance of genetic risk factors. Several environmental factors, such as socio-economic status (SES), have also been reported to contribute to unique variance in the development of reading (e.g., Becker et al., 2017; Kovachy et al., 2015; Mascheretti et al., 2013). Parents with lower SES or a lower educational background tend to spend less shared reading time with their children or have fewer books at home, affecting early reading development (Dilnot et al., 2017; Hamilton et al., 2016). In addition, prematurity has been associated with developing deficits in reading abilities at a later age (Kovachy et al., 2015).

Recent literature thus suggests a variety of factors that contribute to the risk that one develops dyslexia. Although the literature on dyslexia is extensive, it remains unclear how variations in combinations of risk and protective factors contribute to the reading development and how this combination influences intervention outcome, since evidence mostly originates from parsimonious single deficit studies that examine one component (Astle & Fletcher-Watson, 2020). In those studies, children with diverse symptoms are routinely classified as either dyslexic or typically developing readers by employing an arbitrary threshold value. This method would be valid if dyslexia constituted a single condition, ignoring the heterogeneous nature of this developmental disorder. Given the abundance of mechanisms potentially involved in reading disorders and the heterogeneous nature of the behavioral outcome associated with dyslexia, we may gain better insight when dyslexia is conceptualized as a system of connected symptoms. This way, a variety of factors can be examined concurrently rather than conducting unimodal studies that fail to provide a clear understanding of how different factors are related (Borsboom, 2017; Borsboom & Cramer, 2013).

A promising modeling option toward this goal is a graphical model, previously applied in a variety of domains, such as genetics (Ghazalpour et al., 2006), sociology (Farasat et al., 2015), and psychology (Fried et al., 2019). From a statistical point of view, network analysis uses multiple regressions that are visually plotted. This approach improves on more conventional multivariate methods (e.g., stepwise multiple regression) in that it does not depend on the order in which variables are entered into the analysis. Moreover, this network technique allows rapid visualization of all these regression parameters. In such a graphical model, any conceivable variable (e.g., friends, symptoms, neurons) can be represented as circles (i.e., nodes) connected through lines (i.e., edges) that represent relations such as Facebook friendships, odd ratios or partial correlations, resulting in a network. Many psychological phenomena are considered to be caused by a large number of factors and the interactions between them. This approach allows one to tap into those interactions, which gives more insight into complex relations between factors at different levels (e.g., cognitive, environmental) (Borsboom & Cramer, 2013). Psychological networks are typically pairwise Markov Random Fields, which are characterized by undirected edges between pairs of nodes which represent conditional dependencies between pairs of nodes after controlling for all other nodes in the network (Epskamp et al., 2017). Edges between two variables therefore imply that there is a relationship between two nodes that cannot be explained by any other node in the network, akin to partial correlations. After estimating the network, core features of the network can be revealed by graph theory measures (Epskamp & Fried, 2018). Employing a network approach allows analyzing individual factors concurrently, which may reveal patterns that are currently neglected. Most important symptoms in the network and their interrelations can be identified, as well as differential associations between symptoms and risk factors. Therefore, more insight into these complex interactions and finding subgroups of individuals may inform the development of more efficient clinical assessments and personalized interventions (Ozernov-Palchik et al., 2016; Ziegler et al., 2020).

The present study thus aims to elucidate the complex relations between risk factors and symptoms in dyslexia. In addition, we aim to examine how intervention responsiveness can be seen in relation to symptoms and their interrelations. To this end, we employed a network analysis in a large sample of children with reading disabilities, aiming to examine I) how numerous variables including reading, spelling, several language and memory processes, SES and general abilities related to each other in the framework of reading disabilities and II) whether the amount of progress one made during reading intervention influenced other relations in the network. Twenty-eight reading-related variables, including cognitive, demographic and environmental measures, were used to estimate a general network structure in 1257 children with reading difficulties. With respect to the second research question, only children who received reading intervention ($n = 806$) were included in the analysis, with intervention progress as moderator. As this is the first network analysis study within the field of reading disabilities, this research was exploratory in nature and therefore no specific hypotheses regarding model structure were tested.

Method

Participants

Data was collected at a nationwide, clinical center for learning disabilities in the Netherlands (RID) during the years 2009-2019. This archival data set contained diagnostic data of 10.001 individuals on 176 measures, including reading- and spelling related measures, several language and memory processes, SES and general abilities. As the original database included participants aged from 6 to 33, many diverging tests were used based on one's age. To control for differences in test materials and because older individuals usually do not receive specialized reading treatment, we only selected participants within the elementary school age (i.e., between 8 and 14 years old). All children were native Dutch speakers that were referred to the clinical center for diagnostic screening for dyslexia because of severe and persistent reading disabilities at school (i.e., below the 10th percentile on standard reading measures or below the 10th percentile on spelling in combination with a score below the 16th percentile on reading) and resisted extra remedial support at school prior to referral. The dataset was completely anonymized before further analysis. This study obtained ethical approval from the ethics committee of the University of Amsterdam and is preregistered at the Open Science Framework (OSF; https://osf.io/mvzag).

Measures

Data cleaning

First, for each variable, a well-considered decision was made whether to include it in the final database or not, aiming to cover a wide range of theories in the field of reading (dis)abilities, including demographic, cognitive and environmental variables. For sake of brevity, this is reported in Supplementary Materials (Appendix A Table S1). Second, for the remainder of the variables, unreliable scores (i.e., scores that fell out of their score range) were removed and all participants with missing values on the selected variables were excluded through listwise deletion. To make it easier to interpret the network model, negatively scored variables, i.e., spelling fluency, rapid automatized naming and letter-speech sound speed, were reversed, such that for all variables higher scores represented better performance. Finally, as not all variables were norm-referenced, we controlled for the influence of age by regressing each non-norm-referenced variable on age. For these variables, residuals of this linear regression were used as scores in further analyses. An overview of all included variables can be found in Table 1.

Cognitive assessments

Reading and spelling measures.

Word reading. Reading was measured with the computerized reading task from the 3DM (Blomert & Vaessen, 2009). This word reading task comprised three different levels, i.e., high-frequency words, low-frequency words and pseudowords. Each level contained 75 words that were displayed on 5 sheets with 15 items on each. The difficulty level of each sheet increased systematically. Children had to read as many words as possible within a time-limit of 30 seconds per level. Reading fluency was computed as the number of correctly read words, with a maximum score of 75. Accuracy was computed as the percentage of correctly read words within this time-limit. As the task comprised three levels, three variables of reading fluency and three variables of reading accuracy were included in analyses, i.e., for high-frequent, low-frequent and pseudowords.

Spelling. Spelling was assessed with the computerized spelling task from the 3DM (Blomert & Vaessen, 2009). A word was both aurally and visually presented with a part of the word missing in the visual presentation. Children had to choose the missing part out of four visually presented alternatives by pressing the corresponding button. The task consisted of 54 items of which 18 items were phonetically transparent and 36 items required application of a spelling rule. Both spelling accuracy (% correct) and fluency (sec/item) were included in analyses.

Other cognitive measures.

Letter-speech sound mapping. Letter-speech sound mapping was measured with the letter-speech sound identification task from the 3DM (Blomert & Vaessen, 2009). Children had to match a speech sound to one out of four visually presented letter combinations (e.g., /oe/ and ‘ou’, ‘uu’, ‘o’, ‘oe’). Both accuracy (% correct) and response time (sec/item) were included in analyses as letter-speech sound accuracy and letter-speech sound speed respectively.

Phoneme awareness. Phoneme awareness was measured with the phoneme deletion task from the 3DM (Blomert & Vaessen, 2009). In this task, a pseudoword was aurally presented. The child was instructed to delete a speech sound (the beginning consonant, the end consonant, or a consonant within a consonant cluster) and to provide the resulting word. The score was computed as the percentage of correct responses.

Rapid automatized naming. Rapid automatized naming was measured with the digits and letters subtests of the rapid naming task from the 3DM (Blomert & Vaessen, 2009). Children had to name the items presented on the screen (either digits or letters, 15 items per set) as fast and accurately as possible. The test included two sets of digits and two sets of letters. The score was defined as the sum of the amount of time that was needed to complete all subtests.

Vocabulary. Both receptive and productive vocabulary were assessed with the Dutch version of the Peabody Picture Vocabulary Task (PPVT; Schlichting, 2005), and the Vocabulary subtest from the Wechsler Intelligence Scale for Children – Third Edition respectively (WISC-III; Kort et al., 2002). In the Peabody, a word was aurally presented, and children had to choose the correct image out of four alternatives. In the WISC, children needed to describe the meanings of words of increasing complexity. Both scores were defined as the number of correct items, with a maximum score of 240 for the PPVT and a maximum score of 70 for the Vocabulary subtest from the WISC.

Working memory. Working memory was measured with the Digit Span Backward subtest from the WISC-III (Kort et al., 2002). Children were required to memorize a sequence of digits and to repeat them in a reversed order. The score was defined as the number of correct repeated sequences with a maximum score of 14.

Verbal memory. Both short-term and long-term verbal memory were assessed with the 15 Words Test (Saan & Deelman, 1986), a Dutch version of Rey's Auditory Verbal Learning Test (AVLT; Rey, 1964). Children were required to recall as many words as possible from an aurally presented list containing 15 high-frequent Dutch nouns, comprising 5 learning blocks. After a 20-minute delay, children had to recall as many words as possible from the list they learned before. Short-term memory was calculated as the number of words one could recall during the fifth learning block, long-term memory was calculated as the number of words one could recall during 'delayed recall', with both maximum scores of 15.

Visuo-constructional abilities. Visuo-constructional abilities were indexed by the Block Design subtest from the WISC-III (Kort et al., 2002). Children had to replicate a two-dimensional pattern with one- or two-color (i.e., red and white) blocks within a specified time-limit. Each item was scored on accuracy and one could receive time bonus points with faster performance. The score was computed as the sum of all items, with a maximum score of 64.

Visual perception. Visual perception was measured with a subtest from *Groninger School Onderzoek* (GSO; (Kema & Kema-van Leggelo, 1987). For each item, four abstract figures were presented including two identical ones and two distractors. Children had to match the two identical figures. Children had to solve as many items as possible within a time-limit of 5 minutes. The score was defined as the number of correct items, with a maximum score of 50.

Intelligence. Intelligence was assessed with the WISC-III (Kort et al., 2002). As two standard subtests of the WISC-III were also used as proxies for other variables (i.e., block design for visuo-constructional abilities and vocabulary for productive vocabulary), these were removed from the total intelligence score as this would otherwise lead to spurious relations in the network. As a result, IQ was defined as the sum of the norm scores of the following subtests: similarities, information, comprehension, arithmetic, picture completion, picture arrangement, object assembly and substitution.

Attention. Attention was included as a binary variable that reflected the judgement of the parents whether their child had attention problems or not.

Demographic and environmental indicators

Birth weight (weight at birth in grams) and gestational age (weeks of gestation) were included as continuous measures. Furthermore, gender, family risk for dyslexia (i.e., having a first-degree relative with dyslexia), multilingualism (i.e., the use of more than one language at home) and whether children had to retake a grade after the third grade (i.e., grade retention), were included as binary variables. All these variables were obtained through questionnaires that were filled out by the parents. Socio-economic status (SES) was calculated as the mean disposable income per household, based on the postal code, and included as a continuous variable in the analysis.

Intervention training

Children with dyslexia followed an intensive computer-assisted training program provided by a well-trained psychologist. The training comprised a weekly 45-min one-to-one training session and three 15-min training-at-home sessions. Dutch letter-speech sound mappings were taught explicitly and consequently repeated intensively in order to obtain automatized L-SS associations. First, regular letter-speech sound correspondences were trained, and subsequently irregular letter-speech sound mappings were taught with increasing difficulty (i.e., first short vowels, then long vowels and then diphthongs). Children needed to pronounce the displayed vowels and were subsequently asked to identify the item by pressing the corresponding button on the screen. Errors were corrected by the tutor and by the computer screen following erroneous button presses. The total duration of the training was 40 sessions. For a more detailed description of the intervention program, we refer to Tijms (2011) or Fraga González et al. (2015).

Intervention data

Intervention data was used to examine whether intervention progress moderated the network. Here we consider moderation effects from a network perspective. However, the same effects could also be viewed from the perspective of running a regression on intervention outcome and testing the product interaction between all pairs of variables as predictors. As many children did not complete all 40 intervention sessions, intervention progress that was assessed halfway through the intervention period was used in this study for sake of power. More specifically, a difference score between word reading fluency prior to and after approximately 20 intervention sessions was included in the network analysis as a moderator variable. Word reading fluency was measured with the Dutch version of the One-Minute Test (Een-Minuut-Test; Brus & Voeten, 1973), in which children were required to read as many words as possible within one minute, of a list of 116 words of increasing difficulty. The score was defined as the number of words read correctly, with a maximum score of 116.

Table 1

Overview of Included Variables

Abbreviation	Meaning in the network	Task	Variable type
ReaFlH	Reading fluency in high frequent words	3DM (Reading Task)	Continuous
ReaFlL	Reading fluency in low frequent words	3DM (Reading Task)	Continuous
ReaFlP	Reading fluency in pseudowords	3DM (Reading Task)	Continuous
ReaAccH	Reading accuracy in high frequent words	3DM (Reading Task)	Continuous
ReaAccL	Reading accuracy in low frequent words	3DM (Reading Task)	Continuous
ReaAccP	Reading accuracy in pseudowords	3DM (Reading Task)	Continuous
SpFl	Spelling fluency	3DM (Spelling Task)	Continuous
SpAcc	Spelling accuracy	3DM (Spelling Task)	Continuous
LSSSpeed	Letter-speech sound speed	3DM (Letter-speech sound identification Task)	Continuous
LSSAcc	Letter-speech sound accuracy	3DM (Letter-speech sound identification Task)	Continuous
PhonAw	Phoneme awareness	3DM (Phoneme Deletion Task)	Continuous
RAN	Rapid automatized naming	3DM (Rapid automatized naming Task)	Continuous
ProdVoc	Productive vocabulary	WISC-III (Vocabulary)	Continuous
RecVoc	Receptive vocabulary	PPVT	Continuous
WorMem	Working memory	WISC-III (Digit Span Reversed)	Continuous
VerSTMem	Verbal short-term memory	15-woorden test	Continuous
VerLTMem	Verbal long-term memory	15-woorden test (Recall)	Continuous
VisConAb	Visuo-constructional abilities	WISC-III (Block Design)	Continuous
VisPer	Visual perception	GSO	

IQ	**Intelligence**	**WISC-III (similarities, information, comprehension, arithmetic, picture completion, picture arrangement, object assembly and substitution)**	**Continuous**
Att	Attention	-	Binary
BirWei	Birth weight	-	Continuous
GestAge	Gestational age	-	Continuous
Sex	Gender (0 = girl, 1 = boy)	-	Binary
FamRisk	Family Risk	-	Binary
MulLin	Multilingualism	-	Binary
GradeRet	Grade retention (after third grade)	-	Binary
SES	Income as proxy of socio-economic status	-	Continuous
IntProgr	Intervention Progress– difference in word reading fluency scores prior to and after intervention	One-minute test	Continuous

Statistical analysis

This dataset was analyzed in R (version 3.6.2; R Core Team, 2019). Continuous variables were checked for normality and normalized using the non-paranormal transformation prior to network estimation, implemented in the *huge*-package (version 1.3.4.1; Jiang et al., 2020). The network analysis involved the following steps: I) estimate the edges, II) calculate node predictability, III) visualize the network, IV) assess node centrality and V) assess edge weight accuracy. These steps are further described below for the two research questions separately.

The first research question aimed to examine the interrelations of various variables within the reading network. To this end, as our dataset contained different types of variables (i.e., binary and continuous), the *mgm*-package (version 1.2-11) was used to estimate a Mixed Graphical Model (MGM) with only pairwise associations, as described in Haslbeck and Waldorp (2020). *mgm* estimates MGMs using a nodewise estimation approach with a penalty based on the Least Absolute Shrinkage and Selection Operator regularization (LASSO) (Tibshirani, 1996). The strength of this penalty is controlled by a parameter λ, selected with the Extended Bayesian Information Criteria (EBIC) criterion (Haslbeck & Waldorp, 2020). The EBIC itself has a tuning parameter γ, which in practice controls the trade-off between sensitivity and precision, with a default of 0.25 (Haslbeck et al., 2020). Typically, the LASSO yields a parsimonious network in which small edges are shrunk to zero, hence only the most robust edges are presented. As spurious relationships are excluded, the LASSO has a low likelihood of false positives (Krämer et al., 2009). On the other hand, actual relationships that are present in the population may be omitted in the estimated network due to the applied penalization. Given the multivariate structure of dyslexia, some relations are expected to be small by nature and therefore are more likely to be shrunk to zero. Therefore, to allow for recovering small edges in the model, and particularly those between different domains, the gamma-value was set to 0 for the purpose of this paper (note that a gamma of 0 makes the EBIC equal to the BIC) (Schwarz, 1978). The nodewise regression approach returns two estimates for each pairwise interaction, which are combined into a single estimate using the AND-rule, which sets an edge to be present only if both estimates are nonzero (Haslbeck & Waldorp, 2020). All continuous variables were scaled to mean zero and standard deviation one, to avoid that penalization of a given parameter depended on the standard deviation of the associated variable (Haslbeck et al., 2020).

Second, node predictability, how well a node can be predicted by nodes it shares an edge with, was calculated for all nodes in the network (Haslbeck & Waldorp, 2018). Specifically, we reported the proportion of explained variance for continuous variables and the accuracy (or proportion of correct classification) for binary variables. The third step involved visualizing the estimated edges and the node predictability using the R-package *qgraph* (version 1.6.5; Epskamp et al., 2012), based on the *Fruchterman-Reingold algorithm* (Fruchterman & Reingold, 1991). As most of the measures were only collected during diagnostic assessment prior to intervention, this resulted in an undirected, weighted network that was based on cross-sectional data at a single time point. The network structure was analyzed with graph theory measures (i.e., centrality indices) using the *centralityPlot* function in the R-package *qgraph* (Epskamp et al., 2012). Three such measures are node strength, closeness and betweenness, quantifying the number and strength of the connections of the node of interest, how many indirect connections a node has with other nodes and how important a node is in the average path between two other nodes, respectively (Costantini et al., 2015). Finally, bootstrapping procedures were used to evaluate the edge weight accuracy of the estimated network. To ensure that the bootstrapped models were based on the same algorithm as the originally reported network model, the resample function implemented in the R-package *mgm* was used (Haslbeck & Waldorp, 2020). We ran hundred bootstrapped samples and plotted the resulting sampling distribution of all edges, indicating the 5% and 95% quantiles of the sampling distribution and the proportion of estimates whose absolute values were larger than zero (see Appendix A Figure S2). To investigate the robustness of our network obtained using a gamma-value of 0, the same estimation procedure was repeated with the default gamma of .25 and estimations of the two networks were compared (see Appendix A Figure S1).

For the second research question, we aimed to examine whether the network of children that made progress during reading intervention differed from those that did not. One way to do this is to split the dataset into two subgroups: responders to intervention vs non-responders. However, this required an arbitrary cut-off value at which the data would be split, and this implied that both networks had to be estimated on half the data, leading to reduced sensitivity to detect pairwise interactions and moderation effects (Haslbeck et al., 2020). Therefore, rather than dichotomizing our sample in responders versus non-responders, intervention progress was included as a continuous moderator to examine the moderating role of intervention progress on all individual edges of the reading network (i.e., three-way interactions; Haslbeck et al., 2020). Only participants who received reading intervention were selected from the total sample.

Intervention progress was defined as the difference between word reading fluency at the beginning of the intervention and after approximately 20 intervention sessions. The Moderated Network Model (MNM) was estimated with the R-package *mgm*, with intervention progress as moderator (Haslbeck et al., 2020). Node predictability was calculated for all nodes in the network (Haslbeck & Waldorp, 2018) and the reliability of estimates was examined. For a full tutorial see Haslbeck et al. (2020).

Results

Sample characteristics

Missing data percentages across variables ranged from 1.77 to 44.60 %. Excluding all participants with missing values on the selected variables through listwise deletion led to a final sample of 1257 children (696 boys), aged between 85 and 158 months (*M* = 114.00, *SD* = 14.42) with scores on 28 variables for the general network. Only selecting the children who received reading intervention led to a sample of 806 children (434 boys), aged between 85 and 158 months (*M* = 112.80, *SD* = 13.85) with scores on 29 variables. Raw scores of all included continuous variables for both networks are represented in Table 2.

Table 2

Descriptive Statistics for the Total and the Intervention (Sub)Sample

	Total sample (N = 1257)		**Intervention sample (n = 806)**	
	Mean (SD)	Range	Mean (SD)	Range
Age	114.00 (14.42)	85.00 - 158.00	112.80 (13.85)	85.00 - 158.00
ReaFlH	31.85 (10.53)	2.00 - 60.00	29.95 (10.35)	2.00 - 53.00
ReaFlL	24.04 (9.94)	1.00 - 48.00	22.09 (9.41)	2.00 - 48.00
ReaFlP	14.03 (5.61)	1.00 - 50.00	12.92 (5.22)	1.00 - 50.00
ReaAccH	94.54 (7.59)	13.00 - 100.00	93.86 (7.91)	40.00 - 100.00
ReaAccL	88.08 (13.49)	13.00 - 100.00	86.56 (14.17)	18.18 - 100.00
ReaAccP	75.00 (18.36)	6.67 - 100.00	73.31 (18.90)	6.67 - 100.00
SpFl	4.26 (1.14)	2.00 - 10.00	4.40 (1.18)	2.00 - 10.00
SpAcc	63.67 (14.85)	14.81 - 98.15	62.03 (14.83)	14.81 - 98.15
LSSSpeed	2.42 (0.68)	1.00 - 9.56	2.46 (0.64)	1.00 - 9.00
LSSAcc	88.02 (10.08)	3.00 - 100.00	87.06 (10.68)	4.00 - 100.00
PhonAw	45.75 (23.96)	0.00 - 100.00	41.27 (23.19)	0.00 - 100.00
RAN	20.11 (5.33)	11.00 - 70.00	20.88 (5.28)	12.00 - 64.00
ProdVoc	30.93 (6.46)	8.0 - 54.00	30.75 (6.34)	8.00 - 52.00

RecVoc	113.60 (11.51)	57.00 - 196.00	113.20 (10.56)	58.00 - 157.00
WorMem	4.12 (1.24)	2.00 - 10.00	4.08 (1.21)	2.00 - 10.00
VerSTMem	11.22 (2.04)	4.00 - 15.00	11.13 (2.06)	4.00 - 15.00
VerLTMem	9.90 (2.28)	2.00 - 15.00	9.80 (2.32)	3.00 - 15.00
VisConAb	39.93 (10.31)	6.00 - 66.00	39.45 (10.14)	6.00 - 65.00
VisPer	34.16 (8.08)	3.00 - 50.00	33.61 (8.02)	13.00 - 50.00
IQ	80.97 (12.76)	31.00 - 114.00	81.28 (12.34)	36.00 - 114.00
BirWei	3474.00 (672.59)	1050 - 8450	3472.00 (680.18)	1060 - 8450
GestAge	39.57 (2.10)	25.00 - 50.00	39.55 (2.16)	25.00 - 50.00
SES	2857.00 (1017.37)	1100 - 10000	2834.00 (951.35)	1300 - 10000
IntProgr	-	-	9.84 (6.62)	-20.00 - 33.00
Intervention period (months)	-	-	7.41 (1.78)	3.50 - 16.23

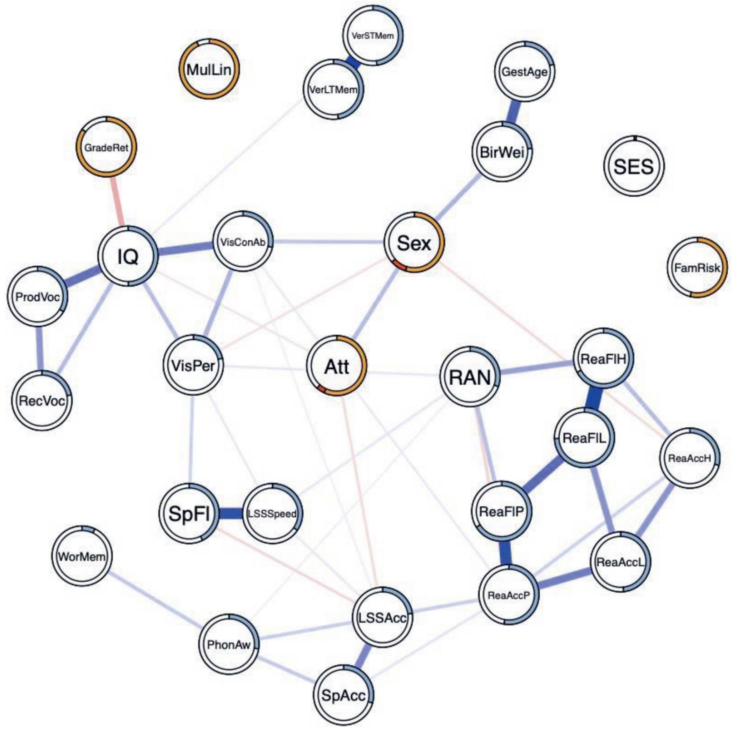
Note. Children in the intervention sample were selected from the total sample. Intervention period depended on the number of sessions. It is therefore possible that children differed in intervention period but received an equal number of intervention sessions.

General network

Network estimation led to the network presented in Figure 1. Each edge represents an undirected, statistical relationship, controlling for all other variables in the network. This means that edges do not imply causal relations but represent statistical relationships between variables, with stronger relationships being represented by thicker edges. In the case of relationships between two continuous variables, these can be interpreted as partial correlations. Absent edges do not imply that a variable is not marginally related to reading, but that a variable is not related given all the other variables in the network. Blue edges represent positive relations and red edges represent negative relations.

Figure 1

The General Network for Children With Reading Difficulties



ReaFlH = Reading fluency in high frequent words
ReaFlL = Reading fluency in low frequent words
ReaFlP = Reading fluency in pseudowords
ReaAccH = Reading accuracy in high frequent words
ReaAccL = Reading accuracy in low frequent words
ReaAccP = Reading accuracy in pseudowords
SpFl = Spelling fluency
SpAcc = Spelling accuracy
LSSSpeed = Letter-speech sound speed
LSSAcc = Letter-speech sound accuracy
PhonAw = Phoneme awareness
RAN = Rapid automatized naming
ProdVoc = Productive vocabulary
RecVoc = Receptive vocabulary
WorMem = Working memory
VerSTMem = Verbal short-term memory
VerLTMem = Verbal Long-term memory
VisConAb = Visuo-constructional abilities
VisPer = Visual perception
IQ = Intelligence
Att = Attention
BirWei = Birth weight
GestAge = Gestational age
Sex = Gender
FamRisk = Family risk
MulLin = Multilingualism
GradeRet = Grade retention
SES = Socio-economic status

Note. Positive associations are represented as blue edges and negative associations are represented as red edges in the network. The width of the edges is proportional to the absolute value of the edge weight. For continuous variables, the blue part of the ring indicates the percentage of explained variance. For binary variables, the orange part of the ring indicates the accuracy of the intercept model and the red part the additional accuracy achieved by all remaining variables. Hence, the sum of orange and red is the total accuracy of the full model.

Based on visual inspection, the network featured a highly connected subpart comprising reading accuracy and fluency variables at the right side of the network, with the strongest association between reading fluency for high and low frequent words (ReaFlH and ReaFlL respectively). This indicates that children with high fluency for high frequent words on average had higher scores on fluency for low frequent words and vice versa. Although reading fluency in low frequent words was connected to reading fluency in pseudowords, no direct edge appeared between fluency in high frequent and pseudowords. This relation was mediated by the fluency in low frequent words and by rapid automatized naming. In contrast, there did appear a direct relation between reading accuracy in high frequent words and reading accuracy in pseudowords. Although all reading accuracy variables were related to their reading fluency counterpart, edges between fluency and accuracy in low and high frequent words appeared to be much smaller than the edge between reading fluency and accuracy in pseudowords.

Rapid automatized naming (RAN) was positively connected to reading fluency in high frequent and pseudowords, but not to reading fluency in low frequent words. Regarding the reading accuracy variables, only a negative association between rapid naming and reading accuracy in pseudowords emerged, which means that children with higher automaticity in naming digits and letters, were less accurate in reading nonexistent words and vice versa. Furthermore, RAN was positively associated with phoneme awareness, visual perception and letter-speech sound speed.

Letter-speech sound speed had a large association with spelling fluency, which means that children who were more fluent in identifying correct letter-speech sound associations also were more fluent in spelling words. Children who were more accurate in identifying letter-speech sound associations were more accurate in spelling words but on average needed more time to spell words. These results are indicated by the positive edge between letter-speech sound accuracy and spelling accuracy and the negative edge between letter-speech sound accuracy and spelling fluency respectively. Simultaneously, letter-speech sound accuracy was positively associated with reading accuracy in pseudowords, suggesting that children that were better in accurately identifying corresponding letters and speech sounds were more accurate in decoding pseudowords and vice versa.

On the left side of the network visualization, intelligence was strongly associated with productive vocabulary and visuo-constructional abilities. Only this latter variable was directly connected to either reading fluency or accuracy measures, more specifically to reading accuracy in pseudowords. The remainder of this intelligence-related subpart was only connected to reading accuracy and fluency through visual perception, spelling fluency, letter-speech sound speed and RAN. In addition, the network showed that children with lower intelligence were on average more likely to retake a grade. Whether a child had to retake a grade or not was not directedly related to reading. Short-term verbal memory and long-term verbal memory on the one hand and gestational age and birth weight on the other hand, were also not directly related to reading accuracy or fluency. In contrast, a negative association between gender and reading accuracy in high frequent words appeared, which means that girls were on average more accurate in reading high frequent words. Gender was also related to attention, which in turn directly related to letter-speech sound accuracy. This latter association suggested that children with attentional problems were on average less accurate in identifying correct letter-speech sound correspondences. Finally, some variables were disconnected from the network: Socio-economic status, Family Risk and Multilingualism.

Predictability estimates, how well a node could be predicted by the variables it shared an edge with, are shown in the general network (see Figure 1). The proportion of explained variance and accuracy are reported for the continuous and binary variables, respectively. Highest predictability was observed for reading fluency measures (74.50%, 67.30% and 66.10% for low frequent, high frequent and pseudowords respectively). Lowest estimates were observed for family risk, multilingualism, and grade retention (all zero). Additionally, low predictability estimates were found for gender (0.88%), attention (0.46%), SES (0.12%) and working memory (0.71%). Low predictability implies that all other nodes together share nearly no variance with these variables.

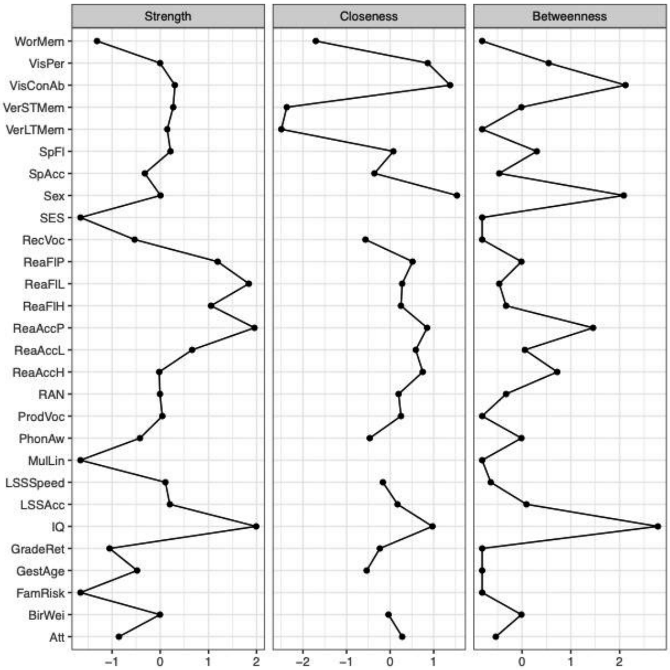
Re-estimating the models with a more conservative γ of 0.25 led to nearly identical models (correlations of adjacency matrices $r > 0.98$). However, some associations disappeared (note that these were all small ones). In the more conservative model, attention was no longer associated with gender, intelligence and letter-speech sound accuracy. In addition, the association between visuo-constructional abilities and letter-speech sound accuracy disappeared, and the latter was no longer associated with spelling fluency (see Supplementary Materials Appendix A Figure S1). Finally, the positive relation between letter-speech sound accuracy and speed disappeared. As these associations did not appear when the network was estimated with a higher tuning parameter, these might be spurious and therefore should be interpreted with caution.

Centrality. For sake of completeness, node strength, closeness and betweenness are reported as centrality indices. However, note that the meaning of closeness and betweenness is questionable in a psychological context given the more complex assumptions and a rather complicated interpretation of these two centrality measures (Bringmann et al., 2019). We, therefore, mainly focused our interpretations on node strength. The centrality plot for the general network is depicted in Figure 2, suggesting that nodes differed substantially in their centrality estimates. Reading fluency in low frequent words, reading accuracy in pseudowords and intelligence had the highest strength centrality score, whereas visual perception, visuo-constructional abilities and gender had the highest closeness. Visuo-constructional abilities and intelligence were also the variables with the highest betweenness centrality score in the network, with reading accuracy in pseudowords as runner-up. On the other hand, the variables that were disconnected from the network, i.e., SES, family risk and multilingualism had the lowest strength and betweenness, implying that these did not play an important role in the constitution of the network structure.

To determine whether these estimates were interpretable, the accuracy of edge weights was estimated in which we bootstrapped the model 100 times. The resulting sampling distribution is shown in Figure S2 (Appendix A). Most edges were estimated reliably as they were included in all or nearly all 100 bootstrapped samples. However, we observed considerable variability in edge parameters across the bootstrapped models, hence individual edges and their order should be interpreted with care.

Figure 2

Centrality Indices for The General Network (Shown as Standardized Z-Scores)



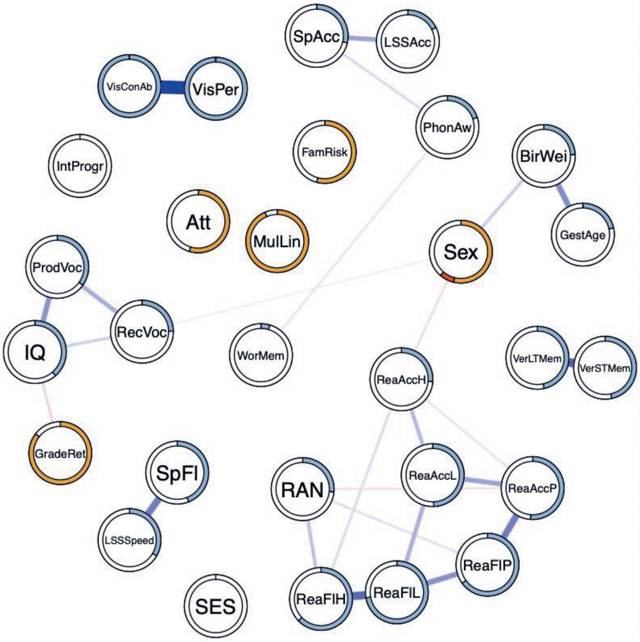
Intervention network

Children within the intervention sample were on average more fluent in reading after they attended the intervention program (Intervention progress: *M* = 9.84, *SD* = 6.62). Note that some children also read less fluent after reading intervention. Especially those who were severely hampered in fluent reading prior to intervention made the biggest progress. Including intervention progress as moderator led to the network depicted in Figure 3. Compared to the general network, a large number of associations attenuated or disappeared. At the right side of the network, the subpart comprising reading accuracy and fluency variables remained. In contrast to the general network (see Figure 1), RAN was no longer associated with phoneme awareness and letter-speech sound speed. Moreover, letter-speech sound and spelling fluency were no longer connected to the other variables in the network. Simultaneously, although better spelling accuracy was still associated with better phoneme awareness and better letter-speech sound accuracy, none of these three variables appeared to be associated with better reading performance. Last, no single edge between the subpart comprising intelligence-related measures and reading measures emerged, as the associations with the variables that funneled this relationship in the general network disappeared (visuo-constructional abilities and visual perception). No pairwise association between intervention progress and any other variable in the network appeared. In addition, no three-way interactions were found with intervention progress as moderator.

The bootstrapped models evaluating the edge weight accuracy of the intervention network are shown in Figure S3 (Appendix A). Again, most estimated edges were included in all or nearly all 100 bootstrapped samples and therefore estimated reliably. However, considerable variability in edge parameters was observed across the bootstrapped models, hence individual edges and their order should be interpreted with care.

Figure 3

The Intervention Network



ReaFlH = Reading fluency in high frequent words
ReaFlL = Reading fluency in low frequent words
ReaFlP = Reading fluency in pseudowords
ReaAccH = Reading accuracy in high frequent words
ReaAccL = Reading accuracy in low frequent words
ReaAccP = Reading accuracy in pseudowords
SpFl = Spelling fluency
SpAcc = Spelling accuracy
LSSSpeed = Letter-speech sound speed
LSSAcc = Letter-speech sound accuracy
PhonAw = Phoneme awareness
RAN = Rapid automatized naming
ProdVoc = Productive vocabulary
RecVoc = Receptive vocabulary
WorMem = Working memory
VerSTMem = Verbal short-term memory
VerLTMem = Verbal Long-term memory
VisConAb = Visuo-constructional abilities
VisPer = Visual perception
IQ = Intelligence
Att = Attention
BirWei = Birth weight
GestAge = Gestational age
Sex = Gender
FamRisk = Family risk
MulLin = Multilingualism
GradeRet = Grade retention
SES = Socio-economic status
IntProgr = Intervention progress

Note. Network displaying the relationship between variables in the framework of reading disabilities. Only children that received reading intervention were included in this sample. Intervention progress was included as moderator. Positive associations are represented as blue edges and negative associations are represented as red edges in the network.

Discussion

Typically, research on the etiology of dyslexia uses a single core-deficit approach (Astle & Fletcher-Watson, 2020). However, given the high number of children with dyslexia that remains hampered in fluent reading after intervention and the large inter-individual variability in intervention response, one needs to take the multidimensional character of dyslexia into account. To contribute to the lack of understanding of how multiple factors interact in the context of reading disorders, this study explored links between cognitive, environmental and demographic variables that are related to dyslexia by employing a network analysis using a large clinical database of 1257 elementary school children (55.37% boys). To our knowledge, this study was the first to elucidate the interrelations of various variables that are related to dyslexia using network techniques.

In accordance with approaches emphasizing the multidimensional character of developmental disorders (e.g., Astle & Fletcher-Watson, 2020; Cuthbert & Insel, 2010), our results revealed a complex network of multiple variables associated with different reading components. Based on visual inspection of variable centrality and interconnectedness, the general network suggested two subparts of the network: one comprising reading fluency and accuracy measures at the right side of the network, and one comprising intelligence-related measures at the left side of the network. Interestingly, our network showed that phoneme awareness (PA) appeared to be especially important in accurate reading, whereas RAN was mostly involved in fluent reading, corroborating the notion that PA and RAN play a role in different phases in the developmental pathway toward skilled reading (Vaessen & Blomert, 2010). More specifically, PA was functionally related to the controlled and accurate processing of speech sounds and letter-speech sound mappings (i.e., decoding), as well as to working memory, which are considered especially important in the early stages of reading (Vaessen & Blomert, 2010; Verhagen et al., 2008). RAN on the other hand was functionally related to the automated convergence of visual and speech information as well as to visual processing, important in more advanced phases toward reading fluency (Vaessen & Blomert, 2010). In the next section, some findings are discussed on a more detailed level.

In the general network, we found a subpart comprising reading fluency and accuracy measures, in which reading fluency in low frequent words and reading accuracy in pseudowords were found to be the most central. Strong relations between letter-speech sound measures and spelling performance emerged, suggesting that fluent and accurate letter-speech identification underlies orthographic representations that are needed in fluent and accurate spelling. However, it can be argued that the letter-speech sound identification task and spelling task of the 3DM assess similar processes, as both tasks tap letter-speech sound mapping (Fraga González et al., 2015). As we did not find substantial associations between accuracy of letter-speech sound mapping and reading fluency measures, our findings corroborate the notion that mere knowledge of letter-speech sound associations does not necessarily lead to fluent reading (Blomert, 2011).

The phonological theory of dyslexia claims that phoneme awareness, the ability to focus on and manipulate individual sounds in spoken words, is crucial for the establishment and automatization of letter-speech sound correspondences, which in turn underlie accurate and fluent word recognition (Pennington et al., 2012). In our network, a moderate direct association between PA and letter-speech sound accuracy was found, but no direct relation between PA and reading fluency. As mentioned earlier, our results suggest that phoneme awareness supports decoding, but especially in the early stages of reading development, as only a direct edge between PA and reading accuracy in pseudowords appeared. These results align with previous studies that showed that PA is mainly related to reading accuracy and nonword reading (e.g., Allor, 2002; Verhagen et al., 2008).

In contrast, RAN appeared to be especially important in fluency measures whereas associations with reading and spelling accuracy seemed to be limited after controlling for all other variables in the network. This is in line with earlier findings of Vaessen and Blomert (2010), which showed that the shift from slow phonological decoding to fast automatic word recognition is accompanied by a cognitive shift in PA and RAN contributions. More specifically, they found that PA was especially important in beginning readers, whereas RAN was more important in experienced readers. Moreover, high frequent words tend to become familiar more quickly than low frequent and pseudowords, leading to an earlier cognitive shift for high frequent than for low frequent and pseudowords. Therefore, as our sample contained children of different ages, ranging from beginning to more experienced readers, associations between PA and reading measures in high and low frequent words might be attenuated whereas these relations might be found when only beginning readers were included in the sample.

Previous research suggested that verbal memory is subsumed under phoneme awareness as both involve phonological processing. However, no association between verbal memory and phoneme awareness emerged in the current network. In contrast, there appeared a relation between phoneme awareness and working memory. We found that the relation between working memory and reading was mediated by phoneme awareness, which aligns with a recent study by Knoop-van Campen et al. (2018) that showed that working memory predicted reading efficiency via its relationship with phoneme awareness. It is postulated that this is particularly true for more difficult phoneme awareness tasks, such as the phoneme deletion task that was used in this study. This is because phonological representations of children with dyslexia may be intact, but not their access to these representations (e.g., Ramus & Szenkovits, 2008). According to these findings, children with a better working memory will be better in tasks in which they have to consciously manipulate phonological units of words.

Previous evidence remains inconclusive on whether reading and spelling skills are aspects of the same phenomenon, or that these should be treated as distinct constructs (e.g., Ehri, 1997). As we found moderate associations between spelling accuracy and reading measures but no direct associations between spelling fluency and reading measures, our results suggest that spelling and reading rely on partly different processes. Although high correlations have been reported in previous research, these have been mainly found in opaque orthographies (e.g., English), in which letter-sound mappings are highly inconsistent. In contrast, in more transparent orthographies (e.g., Dutch), reading accuracy easily reaches close-to-ceiling levels, even in poor readers, making reading fluency the most sensitive measure to assess reading performance. The lack of a direct association between spelling measures and reading fluency aligns with clinical practice, where dissociations between spelling and reading fluency impairments have been reported (Moll, Kunze, et al., 2014; Moll & Landerl, 2009; Wimmer & Mayringer, 2002).

No direct association between intelligence and reading fluency or accuracy emerged, suggesting that children with higher intelligence do not automatically read better, or that children with low intelligence do not automatically read worse. These findings corroborate Stanovich's (1996) earlier claim that there are no significant cognitive differences in fundamental cognitive processes that are the source of reading difficulties in children with high and low intelligence. This reflects the fierce criticism against the discrepancy criterion, in which diagnosis hinges upon normal intelligence. More specifically, according to the discrepancy criterion, differences between IQ and achievement are used to identify reading disabilities. In our network we observed that the influence of intelligence is mediated by visual perception, suggesting that children with higher visual perceptual abilities tended to be faster in identifying the correct letter-speech sound associations and were more fluent in selecting the correct letter during the spelling task. However, this relation may be mediated by an unmeasured variable, such as processing speed.

Children with better vocabulary knowledge were not characterized by better word reading skills, as no direct connection between vocabulary measures and reading accuracy or fluency emerged. This is in line with previous studies that only reported indirect effects of vocabulary on word reading skills (e.g., Torppa et al., 2010). In contrast, several studies reported a correlation between preschool vocabulary and phonological awareness as well as later reading abilities (Cooper et al., 2002; Kim et al., 2014), as an increase in vocabulary size may lead to better phoneme awareness. These findings suggest that vocabulary is especially important in early reading development. Moreover, vocabulary knowledge seems to be especially related to reading comprehension (Quinn et al., 2015), which was not assessed in this study. Therefore, not finding an association between our reading measures and vocabulary might be explained by using an age range in which the role of vocabulary size is presumably limited and not including reading comprehension as a variable in the network.

Current interventions mainly comprise repeated reading practice and phoneme awareness training, along with decoding strategies (Fraga González et al., 2015; Galuschka et al., 2014; Scammacca et al., 2015). Two major findings in intervention research are that (a) phonics instruction is effective but mostly in improving reading accuracy (Galuschka et al., 2014; Torgesen et al., 2001) and (b) repeated reading trainings often fail to be effective in children with severe reading disabilities (Galuschka et al., 2014). Network theory predicts that targeting certain nodes within the network will lead to a cascading in- or decrease in nodes this targeted node shares an edge with (Borsboom, 2017). According to the relations in our network, targeting phoneme awareness and L-SS mappings mainly leads to increased reading accuracy. For fluent reading, visual symbols need to be integrated with their corresponding speech sounds (Blomert, 2011; Žarić et al., 2014), in which a greater print-speech convergence (measured as spatial co-activation in brain areas) is correlated with better reading skills (Preston et al., 2016). This may explain why interventions that do not intensively address the automatization of letter-speech sound integration (i.e., L-SS speed), result in a substantial number of children with reading difficulties that remain hampered in fluent reading (Torgesen et al., 2001). At the same time, our network results provide insights into why children fail to profit from repeated reading training when their reading deficit is mostly related to the PA subpart. Consequently, network analysis seems to be a powerful technique in developing more tailor-made intervention programs.

The relatively high comparable node predictability values for reading fluency measures in all three levels (i.e., high frequent, low frequent and pseudowords) might indicate that these capture similar constructs or that these variables measure different concepts that strongly influence each other (Haslbeck & Fried, 2017). Nevertheless, high predictability values suggest that variables a node shares an edge with explain a considerable amount of variance of the particular node that is not predicted by the intercept model (Haslbeck & Fried, 2017). Hence, intervening on neighboring variables might positively influence the particular node. Therefore, given that rapid naming is so closely related to reading fluency in the network, one could argue that reading intervention should include rapid naming training. However, although there is abundance of evidence that showed that RAN is one of the best predictors of reading fluency (Araújo et al., 2015), studies failed to find a reliable impact on either RAN or reading measures after training rapid automatized naming as an isolated component (Conrad & Levy, 2011;

Fugate, 1997). In the framework of our network, RAN, or more in general the under-time pressure mapping of print and speech information, should be trained in the context of the connected reading variables, in order to transfer to benefits in reading. This may explain why studies that trained an isolated component failed to find significant improvement in reading performance (Kirby et al., 2010; Pfost et al., 2019). Incorporating under time pressure mapping of print and speech information in reading intervention may therefore be a line of research worth exploring (e.g., Pecini et al., 2019; Vander Stappen et al., 2020). Another closely related variable was reading in pseudowords. Interestingly, it was previously proposed to include pseudowords in remedial instruction in order to acquire better fluent reading skills (Fälth et al., 2015). In contrast, low predictability values (e.g., gender, attention, SES and working memory) indicate that variables are strongly influenced by factors that are not included in the network. In this case, one would rather look for additional variables as intervening on low predictable variables is likely to be inefficient.

Heritability of dyslexia has been estimated at approximately 60% (Harlaar et al., 2005) making family risk of dyslexia (i.e., having a first degree relative with dyslexia) an important predictor in developing atypical reading skills. However, in the present study, family risk does not seem to contribute to the other factors in the network. A potential explanation could be that family risk is particularly important in differentiating at-risk children from typically developing children before formal reading instruction starts, but that this predictive value disappears when children experience reading difficulties. Hence, as the current sample only comprises children with reading difficulties, the predictive value of family risk may have disappeared.

Including intervention progress as a continuous measure, i.e., the amount of progress made during reading intervention, led to a much sparser network compared to the general network. We did not find any association between intervention progress and any other variable in the network. In addition, how much a child progressed in intervention did not influence other pairwise relations in the network. In contrast to the general network, PA and L-SS speed were not connected to reading fluency measures. As current reading interventions mostly focus on phoneme awareness and letter-speech sound mapping (learning which alphabetic letters correspond to which speech sounds), one could argue that targeting these variables does not lead to better reading skills in these children according to the network. It is however important to note that the moderation model has roughly twice as many parameters as the pairwise model, and the available sample size was leading to reduced sensitivity for picking up edges in the network (Haslbeck & Waldorp, 2020). This is evidenced by the fact that many pairwise interactions present in the pairwise-only model have been set to zero in this model. This allows for the possibility that moderation effects of a similar magnitude as those pairwise interactions may exist in the population but could not be recovered with the present sample size.

Core deficit vs multiple deficit?

The prevailing single-deficit theory has led to discussions in the field about the core deficit in dyslexia (e.g., phoneme awareness versus RAN or phonological deficit versus letter-speech sound integration deficit). Although evidence has shown that a single mechanism cannot be considered as the "core" of dyslexia (O'Brien & Yeatman, 2020), most evidence still originates from single deficit studies that examine one component (Astle & Fletcher-Watson, 2020), and therefore relations between factors, especially those between different domains (e.g., cognitive, environmental) remain largely ununderstood. To this end, the current study employed a network analysis, a relatively new statistical technique that allows rapid visualization of regression parameters. We found that a variety of factors contributed to the (a)typical reading development, though associated with different aspects of the reading process. Although RAN and PA seemed to play a dominant role in our network, these factors were influenced through different pathways and hence suggested that the same behavioral reading outcome may arise from several different loci (i.e., equifinality; Thomas, 2020). Our study points toward the complexity of the developing reading network, in which a single phonological deficit does not account for all the observed heterogeneity in individuals with dyslexia, thus supporting the modern multiple-deficit approach that moves beyond a core-deficit. Our results replicated previous findings, such as the dominant role of PA and RAN in reading (Allor, 2002; Araújo et al., 2015; Verhagen et al., 2008), thereby validating our new network approach. Importantly, this study also provided insights into the differential associations between factors in the network and reading accuracy and fluency components, providing insights into why reading intervention fails to improve reading performance depending on children's individual deficits. To this end, we believe that this study is an important first step toward studying the multivariate pattern of associations in dyslexia in order to inform the development of more efficient clinical assessments and personalized interventions (Ozernov-Palchik et al., 2016; Ziegler et al., 2020).

Limitations and future directions

Although this study provides important insights into the complex interactions within the reading network of children with dyslexia, some limitations need to be taken into account. First, this study is limited by the collected variables. Since all reading-related variables were based on one test measure, the network may not be optimally representative. For example, a measure of phoneme deletion was included as a proxy of phoneme awareness, whereas this variable should ideally be conceptualized much broader. Moreover, the strong association between variables may have been due to the measures being part of the same test battery, as all reading fluency- and accuracy-related variables were measured with the 3DM Reading Test and visuospatial abilities, productive vocabulary and working memory were all subtests of the WISC-III. Accordingly, most of the variables that were disconnected from the network were included as binary factors, such as family risk, attention and multilingualism. However, these might provide more information when they are not polarized but captured on a dimension, as polarization ignores significant variation within the variable. Likewise, SES was conceptualized as the mean disposable income per household whereas SES is commonly conceptualized much broader, including parents' education and/or occupation. This might explain why these variables did not reveal associations with the rest of the network. Moreover, there are no clear guidelines about whether to combine items in one node instead of keeping them separately in network science. In this study, we thought it reasonable to model the three levels of reading fluency and accuracy (i.e., high frequent, low frequent and pseudowords) as separate variables, as previous studies showed that the contribution of underlying cognitive processes was modulated by word frequency (Vaessen & Blomert, 2010).

Second, despite our large sample size, many participants had to be excluded since *mgm* is not able to deal with missing values, potentially decreasing the power of our analyses. Especially for the intervention network, this resulted in a relatively small sample size ($n = 806$), which makes it unlikely to detect moderation effects as these tend to be quite small in observational studies in general (Haslbeck et al., 2020). We therefore chose to include intervention progress as a linear moderator rather than using a step function where the value at which the step occurs would be arbitrarily chosen and not estimated from the data (i.e., testing group differences between non-responders and responders), as this would have more power to detect small effects (but only if the true moderation effect is linear; Haslbeck et al., 2020).

Third, as our study was conducted in a clinical sample, our sample only comprised children with reading problems. Therefore, no strong claims could be made whether relationships differ between dyslexics and controls, although we believe that this study provides a comprehensive characterization of these factors for the dyslexic population.

Finally, most of the measures were only collected during diagnostic assessment prior to intervention. Hence, undirected networks were estimated that did not reveal the direction of the identified relationships. Future studies should employ a longitudinal design, which allows examining which nodes and edges are targeted by current interventions. Moreover, collecting data throughout reading development sheds a light on which risk and compensatory factors are important in the different phases toward skilled reading and how relationships between variables change over time. This way, one could also examine where differences occur between responders and non-responders, or at which stage networks of dyslexic readers start to differ from typically developing readers, providing important insights for future interventions. One may consider including factors across different levels (e.g., neural and genetic factors), as complex interactions between genes, environment, brain and behavior are far from understood in developmental disorders.

To sum up, this study is a first step in providing insights into the complex interactions within the reading network of children with dyslexia. We have shown that the network approach to psychopathology can be profitably used to study the multivariate pattern of associations in dyslexia, which is also applicable to other developmental disorders. Future research is needed to extend the current findings by including genetic and neurobiological data to get a comprehensive overview. Employing a time-series approach may also offer important insights into the highly dynamic development of reading, offering a roadmap of paths that can be targeted in reading intervention.

3

Goal-directedness enhances letter-speech sound learning and consolidation in an unknown orthography

A version of this chapter is published as: Verwimp, C., Snellings, P., Wiers, R. W., & Tijms, J. (2023). Goal-directedness enhances letter-speech sound learning and consolidation in an unknown orthography. *Child Development*. https://doi.org/10.1111/cdev.13901

Abstract

This study examined how top-down control influenced letter-speech sound (L-SS) learning, the initial phase of learning to read. In 2020, 107 Dutch children (53 boys, M_{age} = 106.845 months) learned eight L-SS correspondences, either preceded by goal-directed or implicit instructions. Symbol knowledge and artificial word-reading ability were assessed immediately after learning and on the subsequent day to examine the effect of sleep. Goal-directed children were faster and more efficient in learning a new script and had better learning outcomes compared to children who were not instructed about the goal of the task. This study demonstrates that directing children toward the goal can promote L-SS learning and consolidation, giving insights into how top-down control influences the initial phase of reading acquisition.

Goal-directedness enhances letter-speech sound learning and consolidation in an unknown orthography

Learning to read is essential for participation in society. Although our brain is not hardwired for reading, the vast majority of the population learns how to read relatively effortlessly. Previous studies have suggested that learning to read is a complex behavior that develops from dynamic interactions between multiple processes (Pennington, 2006; van Bergen, van der Leij, et al., 2014; Verwimp et al., 2021), but these processes have commonly been studied in isolation. It, therefore, remains unclear how top-down processes and subsequent consolidation contribute to the initial phases of reading acquisition. To fill this gap, this study employed an artificial letter-speech sound (L-SS) association task in a sample of 107 elementary school children to examine the effect of revealing and directing a child toward the goal of the task on the initial learning and consolidation of new L-SS correspondences, aiming to provide a better understanding of the processes that influence the first stage of learning to read.

Learning to read

Learning to read is a complex multisensory process, which entails learning the underlying regularities of a writing system. In an alphabetic writing system, one of the regularities that has to be learned is how letters map onto units of speech (e.g., the letter A maps onto the sound /a/). How fast one can learn these mappings depends on the orthographic transparency of a given script, i.e., the degree of regularity in L-SS correspondences (Seymour et al., 2003). In more transparent languages, such as Dutch, it takes approximately one year of formal reading instruction to acquire the knowledge of these correspondences (Vaessen & Blomert, 2010).

Accurate reading can be obtained relatively fast, but mere knowledge of these correspondences is not sufficient for fluent reading. Brain potential and neuroimaging studies have shown that once L-SS correspondences become highly overlearned, the visual symbol automatically elicits an auditory referent (Blau et al., 2010; van Atteveldt et al., 2007). As a result, brain regions that are involved in multisensory processing and speech sound processing respond differently to congruent pairs (e.g., a - /a/) compared to incongruent pairs (e.g., a - /t/), referred to as the congruency effect. The emergence of this automatic neural sensitivity for congruent and incongruent pairs has been argued to be a fundamental building block for successful fluent reading and appears to be persistently less automatic in individuals with reading difficulties (Blomert, 2011; Dehaene et al., 2015; Žarić et al., 2014). Understanding the mechanisms underlying the formation of these associations is thus of great importance for two main reasons. First, to understand individual differences in reading development, and second, to ultimately assist all readers to become proficient, as instructional support can be better matched to individual needs.

Influence of instructions in a learning context

Although there has been substantial interest in how we acquire letter-speech sound mappings (Karipidis et al., 2021; Xia et al., 2022), it remains heavily debated whether individuals are able to learn these regularities without explicitly drawing attention to that what has to be learned. One approach advocates that learning to read should be accompanied by explicit, systematic instructions (Castles et al., 2018; Rastle et al., 2021). When expertise grows, this gradually shifts toward an implicit form of learning in which associating a letter with its speech sound becomes automatic without the need for explicit instructions. An alternative approach advocates that many complex behaviors, like reading, should be learned in a minimally guided environment, in which learners should discover essential information on their own based on repeated exposure and incidental experience (e.g., Bruner, 1961; Krahenbuhl, 2016).

Based on this latter approach, policymakers and educators are investing more into technology-based interventions that use playful, repetitive practice to obtain automatized skills (e.g., GraphoGame; Saine et al., 2011). Although these technology-based interventions are thought to be more motivational and engaging, questions are raised regarding their compatibility with learning (Graesser et al., 2009), as they often direct learners by performance prompts (e.g., perform better than others or obtain a higher score to reach the next level). However, prompting learners to develop specific skills or master new knowledge has been found to promote better learning, as our brain can better select the inputs that are most pertinent to our behavioral goal at that time (Talsma et al., 2010). Accordingly, instructions that reveal or direct one to the goal of the task (hereafter referred to as goal-directed instructions) may facilitate the development of new knowledge and consequently foster the transfer of learning, whereas instructions that lack the rationale behind that what needs to be learned (hereafter referred to as implicit instructions) may complicate filtering out irrelevant distractors, hindering the establishment and consolidation of newly learned information (Erhel & Jamet, 2016).

Although cross-modal integration (i.e., integration of information from two or more different sensory modalities) has often been characterized as an implicit, automatic process, recent findings point toward the role of higher-level cognitive functions, such as top-down control (i.e., the mechanism of attentional filtering to minimize distraction of irrelevant stimuli; Talsma et al., 2010; van Atteveldt et al., 2007). Selectively directing attention to relevant graphemes and phonemes is thought to facilitate the formation of integrated representations, which are consequently stored in multisensory brain regions through repeated practice and offline sleep consolidation for fast and automatic retrieval (Klinzing et al., 2019; Stein & Stanford, 2008). Remarkably, it remains largely unclear how top-down control contributes to successful integration of letters and speech sounds and subsequent consolidation in the general population, as these processes are commonly examined in isolation. More insight into these processes is highly relevant for effective reading instruction and therapeutic remediation strategies.

Artificial learning design

Revealing the goal of a task and explaining in what context the new knowledge can be used (i.e., goal-directed instructions) is expected to influence the learning rate of the child when learning to associate letters and speech sounds. However, as most studies do not directly address the actual process of learning L-SS mappings, little is known about how goal-directed instructions influence the learning of new letter-speech sound correspondences and how this contributes to the offline consolidation of this new knowledge. Although the studies previously discussed provided insights into letter-speech sound learning, results are commonly influenced by prior letter knowledge and reading experience, as native graphemes and phonemes were used. To control for this, recent studies have developed an artificial L-SS training paradigm, in which participants had to learn how unknown symbols correspond to known speech sounds (e.g., Aravena et al., 2013; Karipidis et al., 2017; Rastle et al., 2021). Such a design allows for a more controlled mapping of the initial phase of learning to read. Moreover, the fact that such a paradigm is devoted to learning rather than to the level of skill already obtained makes it a useful tool to predict individual differences in reading performance and future gains in reading intervention (Aravena et al., 2018; Horbach et al., 2018).

Previous artificial learning studies have shown benefits of explicit instructions in learning a new script (Aravena et al., 2013; Rastle et al., 2021). In the study of Rastle et al. (2021), adults learned to read words printed in two unknown, artificial alphabets. One group received explicit instructions on the underlying regularities of the writing system whereas the other group had to discover these regularities through text experience alone. The authors found that in contrast to the explicitly instructed participants, only 20% of participants in the discovery-learning group obtained high levels of task performance, even after 18 hours of training. In the study of Aravena and colleagues (2013), an artificial L-SS training paradigm was used to learn correspondences between visual Hebrew symbols and Dutch speech sounds. They found that explicit instructions were more efficient in initial L-SS binding, especially when a new orthographic rule had to be learned. In addition, they showed that children with dyslexia were as accurate as typically developed readers but were more prone to errors when applying this knowledge in an under-time pressure reading task. However, different game designs were used in the different conditions, allowing for the possibility that differences between the conditions are partially caused by the different game designs. Moreover, to address the actual process of these mappings, it is highly relevant to examine the learning curve during the training rather than only the behavioral outcome after the learning phase.

Current study

To address these issues, this study used a computerized artificial L-SS learning task similar to that of Aravena et al. (2013), aiming to mimic the initial formation of the neurocognitive reading network. Given that most experiments have been performed in adults (e.g., Rastle et al., 2021) or made group comparisons between individuals with and without dyslexia (e.g., Aravena et al., 2013), this study was performed in children with a wide range of reading levels to get more insight into individual differences in reading development.

The goal of this study was twofold; First, we wanted to examine whether goal-directed instructions influenced the learning and consolidation of L-SS correspondences. We provided the two groups with the same learning conditions, but either preceded by instructions that revealed the goal of the task and directed children toward that goal (i.e., goal-directed instructions) or instructions that prompted children to discover the goal of the task on their own (i.e., implicit instructions). Accuracy and reaction time were measured during the learning phase, allowing us to map the learning curve. Afterward, the instrumental use of the artificial correspondences was examined in a reading task within the artificial script. A congruency task in which children were required to rate an audiovisual presentation as congruent or incongruent served as a measure of L-SS integration. Contrary to most studies, the influence of the instructions on the retention of the L-SS knowledge was assessed 24 hours after the learning task.

For this part, the following research questions were assessed: I) Are children who are directed toward the goal of the task faster in learning the L-SS correspondences? II) Does this instruction manipulation result in better knowledge of the newly learned script? III) Does the instruction manipulation lead to differences in the consolidation of newly learned L-SS correspondences and IV) Are L-SS correspondences better integrated in children who were directed toward the goal of the task?

We hypothesized that both groups would improve in accuracy and reaction time during the L-SS learning task, but that children who received goal-directed instructions would be faster in learning a new script, resulting in a steeper learning curve. As a consequence, we expected that symbols and speech sounds would be better integrated in participants who received goal-directed instructions, reflected in the performance on the reading task, in which symbols and speech sounds needed to be integrated to use these correspondences under time pressure. In the congruency task, we expected that children who received implicit instructions would show no or a weak conflict-related reaction time. This is, we expected no difference in reaction time between congruent and incongruent trials, which may be interpreted as a weaker L-SS integration. Last, given that overnight sleep has been proposed to benefit integrating newly encoded information and memory consolidation (Klinzing et al., 2019), we expected that the new correspondences would be better integrated in both groups on the subsequent day and therefore would lead to a better performance in the reading task. However, working on the assumption that goal-directed instructions affect the quality of processing during the learning phase, we hypothesized that this increase in reading performance would be most apparent in children who received goal-directed instructions.

In the second part of this study, we wanted to establish how valid and reliable such an artificial learning design is. For this we examined whether indexes of our L-SS task were related to the reading performance in the Dutch script. In addition, children were tested with an alternative version of the task containing different L-SS correspondences, approximately three weeks after the first version to assess test-retest reliability.

Taken together, this study aims to provide a better understanding of how top-down control influences L-SS learning and subsequent consolidation, potentially of great importance for effective reading instruction and therapeutic remediation strategies.

Methods

Participants and procedure

Participants were recruited through two primary schools in Amsterdam (The Netherlands) to participate in an artificial letter-speech sound learning task to mimic the initial formation of the neurocognitive reading network. Amsterdam is highly diverse in cultures and ethnicities, but as current research questions did not address socio-demographics, this information was not collected or used to recruit participants. Parents were informed through a digital letter about the goal of the study. To capture the wide range of individual differences in reading levels, all children were eligible to participate when they were native Dutch speakers (both mono- and multilingual) in grade 3 or 4 (approximately aged between 8 and 10), without severe (uncorrected) visual and/or hearing problems. No further exclusion criteria were applied. A total of 107 elementary school children (53 boys) aged between 92-136 months (*M* = 106.845, *SD* = 8.876) took part in this study after obtaining active informed consent of the parents. Data was collected over a four-day period (see Figure 1). On the first day, children completed the computerized artificial learning task, which consisted of three blocks that were devoted to learning the L-SS correspondences, and one testing block (i.e., congruency task). The artificial learning task took approximately 30 minutes in total and was conducted individually. Afterwards, all children completed three tasks that were related to the new artificial script: a productive symbol knowledge task and two one-minute reading tests within the artificial script (real words and pseudowords). To assess word reading skills in the native language, a one-minute reading test within the Dutch script was conducted as well. Last, all children who were tested on the same day (with a maximum of 12) were positioned in a quiet room at the children's school and received 20 minutes to fill in all 20 items of the Raven's Colored Progressive Matrices (Raven, Court, & Raven, 1984). On the second day, to assess the retention of the L-SS knowledge, a shortened version of the congruency task, the symbol knowledge task, and the two word reading tasks within the artificial script were conducted. This took approximately 10 minutes. After approximately three weeks, an alternative version of the learning task (Version B; mapping new symbols onto new speech sounds) was conducted in the same sample to examine the test-retest reliability of this Dutch artificial L-SS learning paradigm, followed by a retention session on the subsequent day as well. All data was collected within a period of three months (September 2020 - November 2020). This study was approved by the ethics committee of the University of Amsterdam. Analyses were either pre-registered at OSF (https://osf.io/2cge6) or described as exploratory in the current study.

Artificial letter-speech sound learning task

This study used an adaptation of the training paradigm used in Aravena et al. (2013), Fraga González et al., (2015) and Guerra (2022). All children completed a computerized artificial letter-speech sound learning task, programmed in Psychopy2 (Peirce et al., 2019). The goal of the task was to learn eight new letter-speech sound correspondences that consisted of eight unknown characters from the BACS-1 alphabet (Vidal et al., 2017) that were linked to eight Dutch speech sounds (for an overview of all L-SS correspondences, see Table 1). Phonemes for Version A and B were selected such that enough words could be created for the subsequent artificial reading tests, with each version containing four consonants, two vowels, and two diphthongs. In addition, the mean phoneme duration was kept constant between the two versions. The learning experiment comprised four blocks, of which the first three aimed at learning the correspondences and the last block aimed at testing the correspondence knowledge. For the first three blocks, each trial started with a fixation cross presented with equiprobable durations of 500, 750 or 1000 ms, followed by the audiovisual presentation. Each speech sound was accompanied by two black artificial characters that appeared on a white background, one on the left-hand and one on the right-hand side of the screen. Children had to press the left or right button to indicate whether the left or right character corresponded to the speech sound (50% chance level). The position of the corresponding symbol was randomly determined at each trial. Each trial was followed by feedback such that children could learn the correct correspondences during the task. When no response was given after 4000 ms, children were encouraged to respond faster with an image of a snail. In the first block, four out of eight L-SS pairs were presented. The other four pairs were presented in the second block. The pairs that were presented in either the first or second block were kept constant across children, with each speech sound occurring 18 times. In the third block, all eight L-SS pairs were presented, with each speech sound occurring 7 times. In the fourth block (i.e., congruency task), either congruent (learned correspondence) or incongruent (new correspondence) L-SS pairs were presented on the screen. Children had to press the left or right button to indicate whether the pair was congruent or incongruent. As this block served as a testing block rather than a learning block, children received no feedback. However, children saw a blue dot in the middle of the screen every time they pressed a button, such that they knew that the computer had recorded their answer. The order of items was randomized for all blocks as well as which symbol was presented as the non-corresponding distractor, without presenting the same speech sound twice in a row. Figure 1 shows an overview of a trial in the learning and congruency task.

Task instructions were given verbally in a standardized manner prior to the learning task. In the goal-directed condition, the participants received information about the goal of the task, i.e., learning as many symbols as possible to crack a secret code in the end. Participants in the implicit condition were told to play a computer game of which the goal would become clear during the task itself. The message participants saw on the screen during breaks was also manipulated; children who received goal-directed instructions were prompted to learn more symbols to be able to crack the code, whereas children in the implicit condition just received the message to press the spacebar to continue. All participants were told which two keys to use during the task and received the same computerized feedback when pressing the right or wrong answer during the learning phase. Participants were randomly allocated to one out of two experimental conditions.

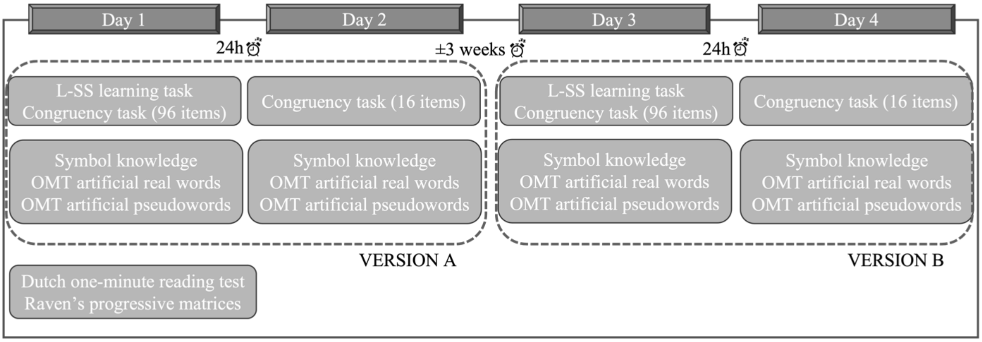
Table 1

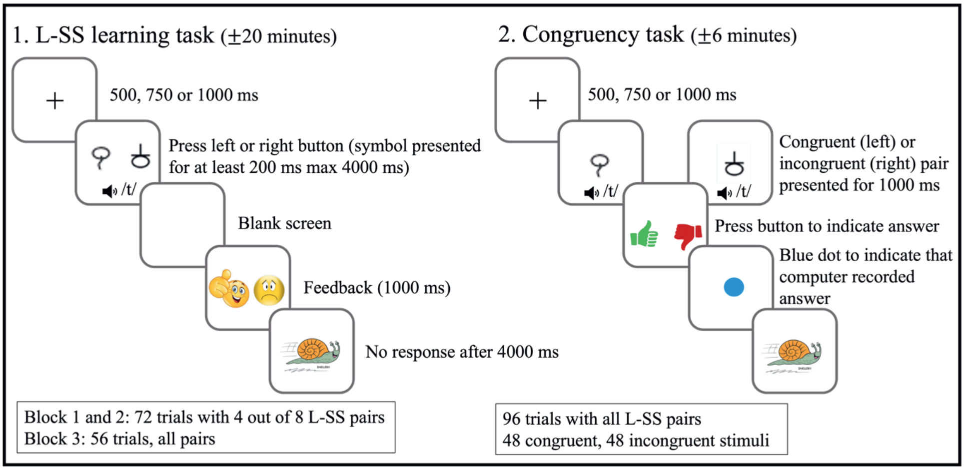
Letter-Speech Sound Combinations Adapted from the BACS-1 Alphabet (Vidal et al. 2017) for Version A and B Separately

Version A								
Letter	♁	♀	∋	⊥	∟	ɋ	⊐	∠
Speech sound (IPA)	[ɑʊ]	[t]	[z]	[ɛɪ]	[ɛ]	[n]	[f]	[ɔ]
Phoneme duration (ms)	505	194	516	527	387	734	303	383
Version B								
Letter	┤	♉	Γ	Ω	⊾	ϕ	δ	↓
Speech sound (IPA)	[ɑ]	[Ø:]	[ɪ]	[u]	[p]	[r]	[s]	[ʊ]
Phoneme duration (ms)	386	515	486	369	314	549	491	392

Figure 1

Overview of Experimental Design





Note. OMT = One-minute reading test. Children had to learn eight new letter-speech sound (L-SS) correspondences with eight artificial symbols from the BACS-1 alphabet (Vidal et al., 2017) that were matched to a Dutch phoneme. Consequently, they had to indicate whether the presented audiovisual stimulus was congruent or incongruent. The L-SS and congruency task were exactly the same for both conditions but were preceded by either goal-directed or implicit instructions to manipulate goal-directedness. The learning task was followed by three tasks that were related to the artificial script. A shortened version of the congruency task and the L-SS learning related tasks were repeated on the subsequent day to assess retention. An alternative version of the learning task (Version B) was conducted in the same sample approximately 3 weeks later to measure test-retest reliability.

Outcome measures

Letter-speech sound learning during the training

Both accuracy and reaction time of letter-speech sound learning were recorded during the training. The accuracy score was determined by the total number of correct responses divided by the total number of items (%). Reaction time was defined as the average reaction time of the correct responses.

Passive letter-speech sound integration

Both accuracy and reaction time were assessed as a proxy of passive letter-speech sound integration during the congruency task, in which children had to indicate whether a presented L-SS pair was correct (learned) or incorrect (not learned). Accuracy scores and reaction times were computed for congruent and incongruent items separately. Accuracy was determined by the total number of correct responses divided by the total number of the items (%) and reaction time by the average reaction time of correct responses. A shortened version of this task (16 trials) was repeated on the subsequent day to assess the influence of sleep on the passive L-SS integration.

Productive symbol knowledge

The experimenter presented children with a form containing the eight artificial symbols. While pointing at one of the symbols, children were asked to name the letter out loud. Children repeated the symbol knowledge task as a measure of retention on the subsequent day. The score was determined by the number of speech sounds that were named correctly (maximum score = 8).

Word reading rate in artificial orthography

Two lists of 14 monosyllabic words within the artificial orthography were constructed with increasing difficulty: one resulting in real Dutch words when all symbols were correctly decoded, one resulting in pseudowords (see Appendix B Table S4). The words were arranged in one column on two different papers. Children needed to read as many words as possible within one minute. The score was determined as the number of words read correctly within one minute (for both versions maximum = 14).

Word reading in Dutch

The One-minute test was used as a measure of word reading skills in Dutch (Brus & Voeten, 1973). Children needed to read as many words as possible within a time-limit of one minute. The score was determined by the number of words read correctly within one minute (maximum = 116).

Non-verbal IQ

Non-verbal IQ was assessed with a time-limited version of the Raven's Coloured Progressive Matrices (Hamel & Schmittmann, 2006; Raven et al., 1984). Raw data was used as a measure of non-verbal IQ (maximum = 36).

Data analysis

Before the statistical analyses, reaction times were outlier-corrected using a two-step procedure; First, values below 100ms were considered unconscious (i.e., guessing behavior) and therefore removed. Second, reaction times deviating more than 3SDs from the individual mean of each subject were excluded. Based on this procedure a maximum of three trials per block per participant were removed. Only trials with correct answers were included in the reaction time analysis. Furthermore, children were only included when they had all data available that was needed to answer a certain research question. As a result, for the first research question, three children had to be excluded as they had missing data due to technical issues. For the second research question, 16 additional children had to be excluded as consolidation data was missing due to technical problems of one computer. For the third and the last research question, no children had to be excluded.

For the first part of this study, we aimed to examine how goal-directed instructions influenced the learning of new L-SS correspondences. To examine differences in learning progress between the two conditions, accuracy and reaction time during the learning task were averaged across 4 bins of 18 trials for Block 1 and 2, and of 14 trials for Block 3. The percentage of correct answers and reaction time was computed per block for each bin and submitted to a repeated-measures MANOVA with Group (Goal-directed versus Implicit) as between-subjects factor and Time (Bin 1 versus Bin 2 versus Bin 3 versus Bin 4) as within-subjects factor. Participants were removed from the analysis if a certain bin did not have at least one value left (*n* = 1). To examine the influence of instruction on the passive L-SS integration immediately after learning and on the subsequent day, accuracy and reaction times following either congruent or incongruent trials on Day 1 and Day 2 were compared between children who received goal-directed or implicit instructions with a factorial MANOVA, with Group (Goal-directed versus Implicit) as between-subjects factor and Congruency (congruent trials versus incongruent) and Day (Day 1 versus Day 2) as within-subjects factors. To examine differences in L-SS knowledge immediately after the learning task and on the subsequent day, a factorial MANOVA was conducted with symbol knowledge and word reading rate within the artificial orthography (Real words and pseudowords) as dependent variables, Condition as between-subjects variable and Day (Day 1 versus Day 2) as within-subjects variable.

For the second part of this study, we aimed to examine the validity and reliability of this Dutch artificial L-SS learning paradigm. To examine the external validity, a one-tailed Pearson correlation test was conducted between reading rate in Dutch and reading rate in the artificial orthography (real words and pseudowords). As we wanted to compare reading fluency in both scripts and fluent reading requires the knowledge of all symbols, only children who obtained full mastery of the new symbols were included in this analysis (*n* = 32). Last, test-retest reliability was computed by correlating the outcome measures (i.e., symbol knowledge and the two reading tasks within the artificial script) of the original version of the task (Version A) to the outcome measures of the alternative version (Version B). Correlations between the two versions were computed for the outcomes immediately after the learning task (Day 1) as well as for the outcomes after one night of sleep (Day 2).

The significance level for all analyses was set at .05. Given the increased risk for type-1 error when conducting multiple tests, p-values were False Discovery Rate corrected (FDR; Benjamini & Hochberg, 1995) within each set of outcomes associated with a certain research question (for research question 1, 2 and 3 separately). Pairwise comparisons between levels of main effects were only performed when interactions involving our manipulation of interest were found to be significant after FDR correction. Partial Eta squared effect sizes were computed and reported as well.

Results

Participant characteristics

The goal-directed condition comprised 54 participants with a mean age of 106.201 months (*SD* = 8.690) and the implicit condition comprised 53 participants with a mean age of 107.50 months (*SD* = 9.215). Participants' characteristics are shown in Table 2. No significant differences in age, intelligence or reading fluency in Dutch were found between the two conditions. Although gender distribution was not significantly different between the two conditions, gender could still influence the results given possible attentional and motivational differences between girls and boys. Re-running the analyses with gender as an additional between-subject factor did not influence our main outcomes.

Table 2

Participant Characteristics

	Mean (SD)		**Group comparison**[b]
Characteristic	GD group[a]	Implicit group	
n	54	53	
Gender (M:F)	23:31	31:22	$X^2(1) = 2.106, p = .147$
Age	106.201 (8.690)	107.500 (9.215)	*F*(1,105) = 0.544, *p* = .462
IQ	31.500 (3.155)	31.170 (3.751)	*F*(1,105) = 0.243, *p* = .623
Reading fluency Dutch	58.574 (14.565)	59.925 (13.701)	*F*(1,105) = 0.244, *p* = .622

Note. [a]GD = Goal-directed; [b] Nominal data were investigated using Pearson's chi squared tests, continuous data were investigated using ANOVAs.

Letter-speech sound learning during the training

A 2 (Condition) x 4 (Time) repeated measures MANOVA with accuracy and reaction time as dependent variables was performed for Block 1, 2 and 3 separately. Mean accuracy scores and reaction times for each bin are shown in Table 3 for the two conditions separately and statistical values are reported in Table S1 (Appendix B). For Block 1, we found a main effect of Condition, a main effect of Time, and a significant interaction effect between Condition and Time. For Block 2, we found a main effect of Time, a significant interaction effect between Condition and Time, but no main effect of Condition. For Block 3, there appeared a main effect of Condition, but no main effect of Time or interaction effect between Condition and Time. After FDR correction, only the main effect of Time and the interaction effect between Condition and Time in the first and the second block remained significant ($p_s <$ .003). To interpret these effects, RM ANOVAs were conducted for accuracy and reaction time separately.

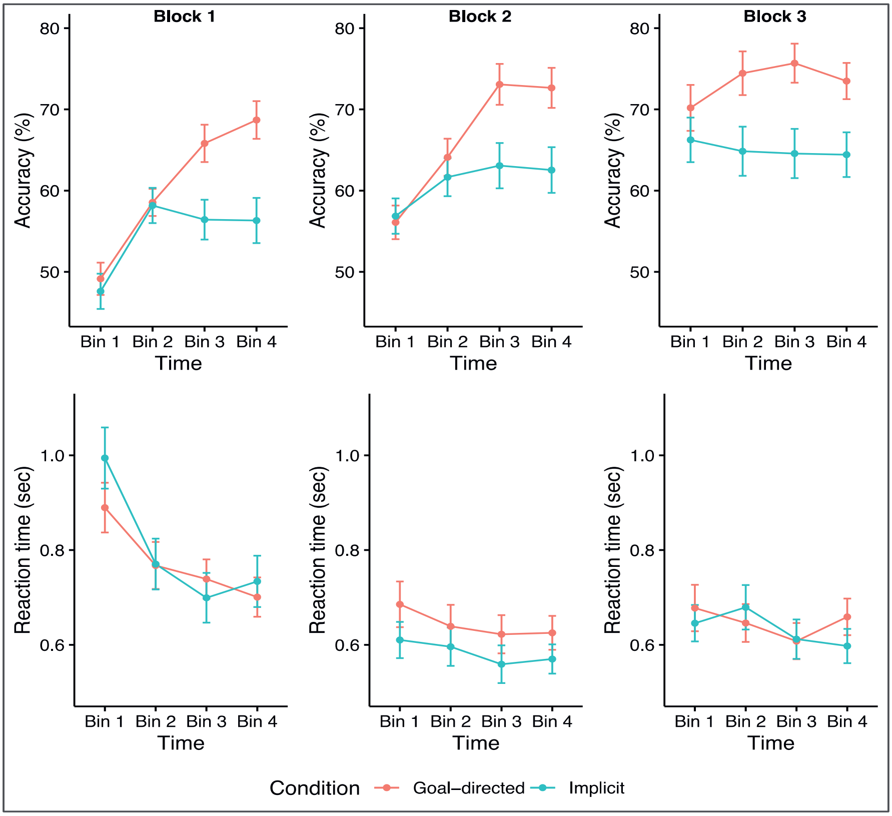
Accuracy. Children who received goal-directed instructions prior to the learning task were on average more accurate in responding compared to children who received implicit instructions. This effect was found for Block 1 and Block 3*,* but not for Block 2. For Time, we found a main effect for Block 1, as well as for Block 2, but not for Block 3, indicating that children learned the most during the first two blocks, in which all symbols were new. Last, significant interactions between Condition and Time were found for Block 1 and Block 2, but not for Block 3. Post-hoc tests showed that during Block 1, children who received goal-directed instructions started to differ from children who received implicit instructions from the third bin onwards (*p*s < .007). This is, children who received implicit instructions slightly increased in accuracy but tend to stagnate relatively close to the chance level, whereas children who were directed toward the goal increased further in accuracy until the last bin (see Figure 2). The same results appeared for Block 2.

Reaction time. For reaction time, no differences between the two conditions were found across all blocks. However, we found a main effect of Time in the first and the second block, but only the Time effect in Block 1 remained significant after FDR correction. No interaction effects were found ($p_s >$ .219) and therefore no follow-up tests were conducted.

To sum up, children who received goal-directed instructions prior to the training increased more in accuracy compared to the children who received implicit instructions in Block 1 and 2. Although children in the implicit group seemed to perform equally well in the first bins of the first two blocks, differences in learning performance became apparent in the later trials. In the third block, where all eight L-SS correspondences came together, children who received goal-directed instructions were on average more accurate compared to children who received implicit instructions. Regarding reaction time, children became faster in the first few trials, but then remained stable during the rest of the learning phase in both conditions. This suggests that this time effect is mainly due to becoming familiar with the task rather than a substantial influence of the instructions on the reaction time during the learning phase.

Figure 2

Learning Curves During the Letter-Speech Sound Mapping Task



Note. Accuracy (percentage correct) and reaction time (in seconds) during the learning task for the goal-directed and implicit condition are presented separately, averaged across 4 bins of 18 trials for Block 1 and 2, and of 14 trials for Block 3 to map the learning curves. Children only learned four out of eight L-SS pairs in the first block, and four new L-SS pairs in the second block, explaining why they start again around an accuracy level of 55% in the second block. In the third block, all eight L-SS pairs were presented together. Error bars represent standard errors.

Table 3

Mean (SD) Accuracy and Reaction Time (RT) of Correct Responses by Condition and Learning Block for Each Bin

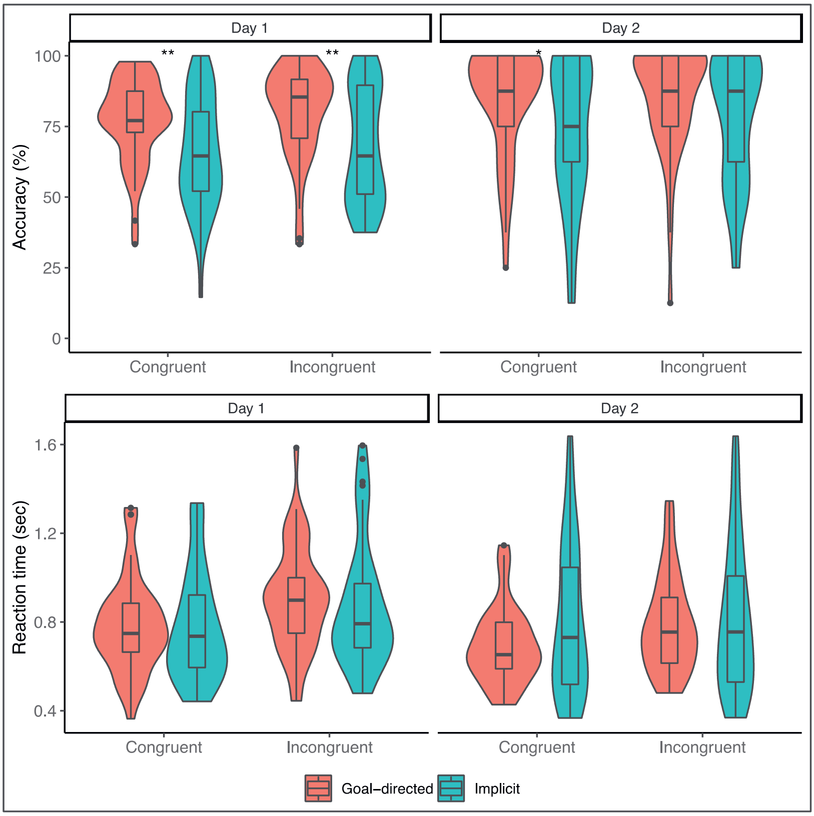
Condition		**Goal-directed group**		**Implicit group**	
		Accuracy (%)	RT	Accuracy (%)	RT
Block 1	Bin 1	49.14 (14.32)	0.890 (0.379)	47.60 (15.53)	0.994 (0.461)
	Bin 2	58.55 (12.02)	0.767 (0.358)	58.17 (15.49)	0.770 (0.384)
	Bin 3	65.81 (16.59)	0.739 (0.299)	56.43 (17.51)	0.699 (0.375)
	Bin 4	68.70 (16.72)	0.701 (0.299)	56.32 (19.95)	0.734 (0.387)
Block 2	Bin 1	56.09 (14.93)	0.685 (0.348)	56.86 (15.50)	0.610 (0.273)
	Bin 2	64.10 (16.45)	0.639 (0.327)	61.66 (16.79)	0.596 (0.289)
	Bin 3	73.08 18.16)	0.622 (0.291)	63.07 (19.96)	0.559 (0.285)
	Bin 4	72.65 (17.73)	0.625 (0.258)	62.53 (20.09)	0.570 (0.220)
Block 3	Bin 1	70.19 (20.39)	0.678 (0.353)	66.25 (19.59)	0.646 (0.274)
	Bin 2	74.45 (19.41)	0.646 (0.288)	64.85 (21.61)	0.679 (0.334)
	Bin 3	75.69 (17.34)	0.608 (0.276)	64.57 (21.66)	0.612 (0.297)
	Bin 4	73.49 (16.09)	0.659 (0.278)	64.43 (19.61)	0.597 (0.258)

Passive letter-speech sound integration immediately after and one day after training

A factorial MANOVA with accuracy and reaction time with Condition (Goal-directed vs Implicit) as between-subjects variable and Congruency (congruent vs incongruent trial) and Day (Day 1 vs Day 2) as within-subjects variables was conducted. Results are visualized in Figure 3 and statistical values are reported in Table S2 (Appendix B). This revealed a significant main effect of Condition, Congruency and Day, but no significant interaction effects (all p_s > .071). Separate ANOVAs for accuracy and reaction time showed that children who received goal-directed instructions were on average more accurate (M = 81.76%) in the congruency task compared to the children who received implicit instructions (M = 70.93%). We found a significant main effect of Day, meaning that children were more accurate in the congruency task on the second day. No interaction effects including our manipulation of interest, i.e., instruction, reached significance (all p_s > .345) and therefore no follow-up tests were conducted. For reaction time, we found a significant main effect of Congruency, with children being faster in responding to congruent trials. This was especially true for children who received goal-directed instructions. In addition, children became faster in responding the next day, but this seemed to be especially true for the incongruent trials. However, these interaction effects did not remain significant after FDR correction and no main effect of Condition was found. Therefore, no follow-up tests were conducted.

Figure 3

Performance During the Congruency Task



Note. Accuracy and reaction time for congruent and incongruent trials during the congruency task immediately after the learning phase (Day 1) and after one night of sleep (Day 2) for both conditions separately. $^{**}p < .01$, $^{*}p < .05$.

Symbol knowledge and active letter-speech sound integration immediately after and one day after training

A factorial MANOVA with symbol knowledge and reading within the artificial script (Real words and pseudowords) as dependent variables, Condition (Goal-directed vs Implicit) as between-subjects variable and Day (Day 1 vs Day 2) as within-subjects variable revealed a significant effect of Condition and Day (for statistical values see Appendix B Table S3). Moreover, a significant interaction effect between Condition and Day was found. To interpret these effects, ANOVAs were conducted for symbol knowledge, reading artificial Dutch words and reading artificial pseudowords separately.

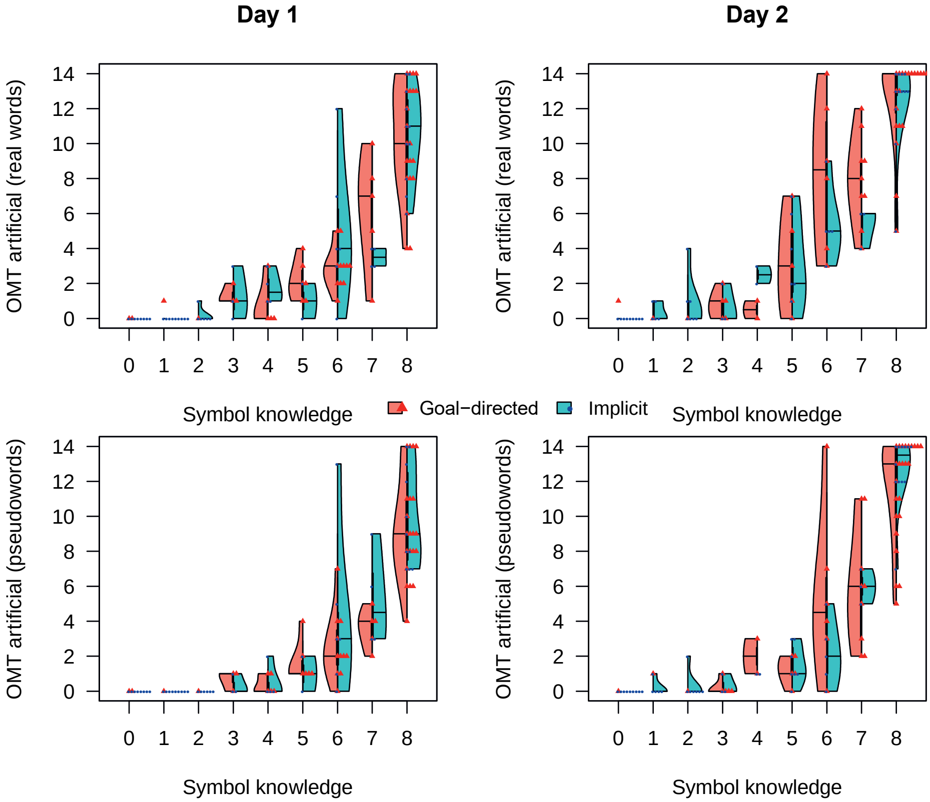
As shown in Table 4, children who received goal-directed instructions had a better knowledge of the newly learned symbols compared to those that received implicit instructions. Moreover, they seemed to be more fluent in using this knowledge in reading words within the artificial script (real words and pseudowords). A significant main effect of Day indicated that children on average knew more symbols on the second day, and were more fluent in applying this knowledge after one night of sleep in both artificial reading tests. A significant interaction effect between Condition and Day for both reading tasks indicated that this time effect was most apparent in children who received goal-directed instructions compared to their implicitly instructed peers.

However, knowledge of the symbols is required to be able to decode the artificial words. Based on visual inspection of Figure 4, especially children who knew more than 5 symbols were able to decode the words within the artificial script. To this end, we conducted an exploratory analysis in which we only selected children who had learned 5 or more symbols on Day 1. In the goal-directed group, 79.63% of the children met this criterium, whereas only 43.40% of the implicitly instructed children met this criterium. An ANOVA within this subgroup with reading within the artificial script (Real words and pseudowords) as dependent variable, Condition (Goal-directed vs Implicit) as between-subjects variable and Day (Day 1 vs Day 2) as within-subjects variable revealed again a significant main effect of Day. This is, children in both conditions profited from one night of sleep. By visual inspection of Figure 4, this seemed to be especially true for the goal-directed group, but this interaction effect between Day and Condition did not reach statistical significance (see Appendix B Table S3).

In sum, children who received goal-directed instructions on average learned more symbols than their implicitly instructed peers. Although there seemed to be a difference between the conditions in applying the new knowledge in a time-limited reading task as well, these results were mainly driven by the differences in symbol knowledge. As can be seen in Figure 4, especially children who learned more than 5 symbols obtained better scores on the reading tasks, for both the implicit and the goal-directed condition. In both groups, children became more fluent in applying this knowledge after one day of sleep. Although this time effect seemed to be more pronounced for children who were directed toward the goal during the learning phase, this effect was non-significant.

Figure 4

Reading Scores in the Artificial Script in Relation to Symbol Knowledge



Note. Scores on the one-minute reading tests (OMT) within the artificial script resulting in real, Dutch words (top) or pseudowords (bottom). Results are visualized for the two conditions separately and show that especially children who knew more than 5 symbols were able to decode the words within the artificial script. On average, children became better in applying the learned knowledge after one night of sleep (Day 2; right) compared to immediately after the learning phase (Day 1; left).

Table 4

Means (SD) of Outcome Measures for Both Conditions Separately

Condition	**Goal-directed group**		**Implicit group**	
	Day 1	Day 2	Day 1	Day 2
Symbol knowledge	6.093 (2.121)	6.333 (2.101)	3.925 (3.018)	4.113 (2.991)
OMT real words	5.574 (4.657)	8.278 (5.304)	3.283 (4.469)	4.302 (5.033)
OMT pseudowords	4.741 (4.327)	6.981 (5.368)	3.019 (4.461)	3.736 (5.211)

Note. OMT = One-minute reading test

External validity and reliability of the artificial learning paradigm in Dutch

External validity

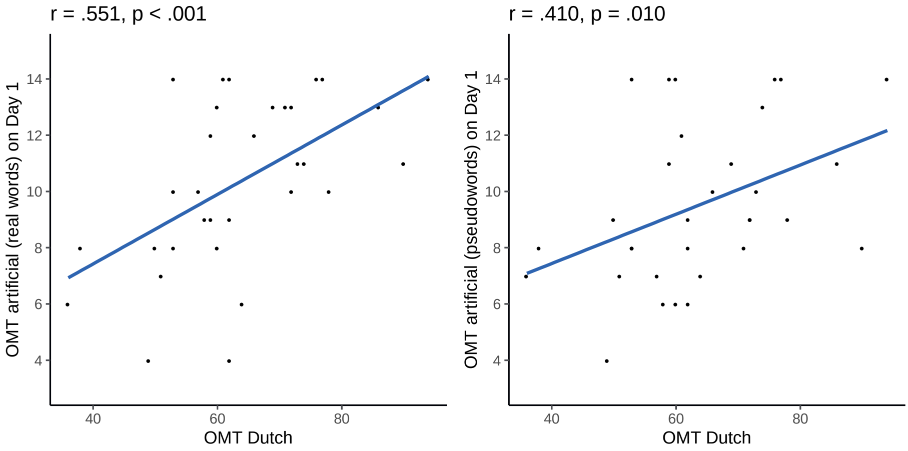
To examine the external validity of our results, we compared reading fluency within the artificial script immediately after the learning task (Day 1) with the typical reading skills assessed with the OMT. As fluent reading requires knowledge of the symbols, only children who obtained full mastery of the new symbols were included in this analysis (n = 32). A one-tailed Pearson correlation test revealed a moderate correlation between reading rate in Dutch and reading rate in the artificial orthography for both real words (r = .551, $t(30)$ = 3.612, p < .001) and pseudowords (r = .410, $t(30)$ = 2.459, p = .010), shown in Figure 5. These results indicated that children who were better readers in Dutch tended to read better in the new artificial script as well, validating generalizations of our findings to reading in Dutch.

Test-retest reliability

To examine the reliability of our artificial learning task, an alternative version of the task (Version B) was conducted in the same sample approximately 3 weeks later (M_{days} = 25.53, SD_{days} = 6.71). We conducted a one-tailed Pearson correlation test between the outcome measures from the original version and the alternative version of the task. For symbol knowledge assessed immediately after the learning phase, we found a significant correlation between the two versions (r = .560, $t(104)$ = 6.900, p < .001). That is, children with better scores on the original version of the task also obtained better scores in the alternative version. Correlations between word reading fluency within the artificial script after learning symbols were much lower (r = .289, $t(104)$ = 3.082, p = .003 for Dutch words, r = . 301, $t(104)$ = 3.224, p = .002 for pseudowords). Correlating the outcome scores between the two versions after one night of sleep (Day 2) revealed a significant correlation for symbol knowledge (r = .388, $t(104)$ = 4.296, p < .001). Moreover, correlating the word reading fluency tasks within the artificial script on Day 2 for the original version and the alternative version of the task revealed significant correlations for both Dutch words (r = .405, t(104) = 4.515, p < .001) and pseudowords (r = .421, t(104) = 4.727, p < .001). Correlations between the learning blocks of the original and the alternative version of the task are presented in Supplementary Materials (Appendix B Table S5).

Figure 5

Correlations Between Reading Rate in Dutch and the Artificial Orthography



Note. OMT = one-minute reading test. Correlations with the artificial orthography resulting in Dutch words (left) and pseudowords (right) immediately after the letter-speech sound learning task (Day 1).

Discussion

This study aimed to shed a light on the role of top-down control in associative cross-modal learning and subsequent consolidation by manipulating the manner of instruction. We found that children who were directed toward the goal of the task were faster and more efficient in learning a new script and had a better learning outcome compared to their peers that had to rely on implicit instructions, suggesting that letter-speech sound learning is more than merely mapping letters onto speech sounds using associative, statistical processes. Our results contribute to the long-standing debate on the role of top-down processes in acquiring new knowledge such as letter-speech sound correspondences.

The results of the current study demonstrate the benefit of directing children toward the goal of the task prior to the learning process. More specifically, we found that goal-directed instructions significantly influenced the rate at which children learned new letter-speech sound mappings. Differences in learning performance already became apparent in the second half of the first two blocks and persisted in the third block of the learning task in which all eight L-SS correspondences were presented. Although accuracy in implicitly instructed peers seemed to increase during the first trials of the learning phase, implicit learners stagnated around an accuracy level of 65%. These results suggest that implicitly instructed children were able to learn the new knowledge to some extent, but that learning was faster and more efficient when the learning process was preceded by clear, explicit instructions.

Regarding response latencies, we did not find any evidence for differences between the two conditions during the learning phase. Previous studies showed that the quality of audiovisual integration of L-SS correspondences in the brain is reflected in the time course of the neural activation of target units, and consequently manifests in the behavioral response latencies during identification (Blomert, 2011). However, in the first phase of learning to read, decoding is effortful and non-automatic, which is gradually replaced by fluent and effortless reading after much exposure and reading practice (Karipidis et al., 2021; Romanovska & Bonte, 2021). As automatic integration of letters and speech sounds might take up to 2 years of reading instruction, our 20-minutes learning task might not be sufficient to elicit differences in response times. In the congruency task, in which children had to indicate whether a congruent or incongruent trial was presented, children were faster in identifying congruent trials. Although most studies examined this congruency effect on a neural rather than on a behavioral level (e.g., Xu et al., 2018; Žarić et al., 2014), this is in line with some older studies (Dijkstra et al., 1993; Herdman et al., 2006) that interpreted this effect as multisensory facilitation during processing congruent grapheme-phoneme stimuli.

Instruction manipulation led to differences in symbol knowledge immediately after the learning phase, with children who received goal-directed instructions on average knowing more symbols than their implicitly instructed peers. The difference in applying this new knowledge in a time-limited reading task seemed to be mainly driven by the differences in symbol knowledge. Especially children who had learned more than five symbols obtained better scores on the reading tasks, suggesting a certain threshold that was needed to read the monosyllabic words. Hence, our results suggest that goal-directed instructions were beneficial for faster and more efficient learning and led to a better knowledge of the symbols, but we did not find statistical evidence for better application of the knowledge on top of the symbol knowledge effect.

To get more insight into how top-down control influenced the consolidation of newly learned L-SS correspondences, we examined the effect of one night of sleep on the outcome measures. In the passive letter-speech sound integration task, both children who received goal-directed instructions and their implicitly instructed peers were more accurate in deciding which trials were congruent or incongruent on the second day. In addition, they became faster in making this decision. Although especially children who received goal-directed instructions were the ones that became faster, there was no significant effect of the instruction manipulation on the accuracy. However, as children who received goal-directed instructions already reached high accuracy levels on the first day, immediately after the learning phase, they could not increase as much compared to their implicitly instructed peers on the second day, possibly explaining why we failed to find this effect. For the active letter-speech sound integration task, in which children needed to read words that were written within the artificial script, children read significantly more words on the second day. One explanation might be that children remembered words from the previous session and therefore were more fluent on the subsequent day. However, as we also found evidence for decreased response times during the congruency task after one night of sleep, these results can be explained by the consolidation of the new knowledge rather than a word recognition effect. This is, memory for newly learned information is enhanced following a period of sleep, making it easier to retrieve information (Klinzing et al., 2019; Mazza et al., 2016) and therefore resulting in improved performance in the next session. Recent findings of Wang et al. (2022) showed that even short naps facilitated the acquisition and application of letter-speech sound mappings in preschool children. In our study, both conditions benefitted from offline sleep consolidation, but this effect seemed to be pronounced for the goal-directed group. This finding suggests that goal-directedness might help with better integration of letters and speech sounds after already one night of sleep. Bitan and Booth (2012) reported similar findings, namely, that participants whose attention was directed toward the correspondence between individual artificial letters and their corresponding Latin phonemes benefited the most from offline improvement.

The second aim of our study was to examine the external validity and test-retest reliability of the used artificial learning paradigm. We found that children who were better readers in Dutch tended to be more fluent in decoding the new artificial script as well. This was especially true for decoding the artificial words that resulted in a real Dutch word. The same result was found in the study of Aravena et al. (2013), implying that our findings can be applied to reading in the Dutch language and highlighting the applicability of artificial script learning paradigms in studying individual differences in early reading development. To examine the reliability of our artificial learning task, we computed correlations between the outcome measures of the original version of the task and the alternative version that was conducted approximately three weeks later. We found significant correlations between the two versions for symbol knowledge, the reading outcomes on Day 1 and the reading outcomes on Day 2. However, the correlation between the two reading tasks on Day 1, immediately after learning, was rather low. Skimming the data revealed high variability in children's outcomes. Some children learned more symbols after the alternative version and obtained higher reading scores after learning the second script and likewise, children who learned fewer symbols obtained lower reading scores in the alternative version. Others learned as many symbols in the alternative version as in the original version and obtained similar reading scores in both versions. As argued before, knowledge of the symbols is needed to accurately decode the words, and especially children who learned more than 5 symbols were able to decode the words. As a result, although symbol knowledge in both versions was moderately correlated, reading scores might have increased exponentially as a result of an increase in symbol knowledge. For example, children who had learned two symbols in the original version and three symbols in the alternative version did most likely obtain similar reading scores in both versions, whereas children who had learned five symbols in the original version and six symbols in the alternative version most likely increased exponentially in their reading outcome, therefore returning low correlations between the two versions. Another explanation might be that some children remembered the task and therefore were faster in applying the new knowledge the second time and therefore obtained higher reading scores. Last, although symbols and speech sounds were carefully matched (i.e., diphthongs and monotones in both versions), the two versions might differ in difficulty. As we wanted to construct artificial combinations that resulted in real Dutch (or pseudo)words, the word lists slightly differed in positions of vowels, consonants, and diphthongs, possibly resulting in differences in how easy these could be decoded. These observations are merely anecdotal and therefore need further research to be confirmed. A future study might want to examine the test-retest reliability with the same version and a longer period in between to control for memory effects. Despite these considerations, these trends nicely highlight individual differences in learning to read and we therefore believe that such an artificial learning design is a promising platform to investigate early reading skills and a potential tool for the prevention of reading difficulties.

Our findings relate to the major debate concerning the role of instructions in skill learning. Constructivists suggested for decades that people learn best in a minimally guided environment, based on the principles of discovery learning (e.g., Bruner, 1961). There are however empirical and theoretical grounds for questioning this. The current study demonstrated that directing children's attention toward the goal of the task shortens the course of acquisition and eventually leads to a better learning outcome. This means that in a rather transparent language, i.e., Dutch, children do better when they receive direct instructions rather than when they need to discover underlying regularities on their own. In contrast, technology-based interventions which are recently gaining field often rely on implicit, statistical association mechanisms without a clear goal that is related to the new knowledge. Our results suggest that, even in serious games, implementing instructions that direct the learner toward the goal and stress mastery of knowledge might be a key design factor to obtain the most efficient learning. In learning to read, explicit instruction is required to direct attention toward visual and auditory information, after which visual and auditory information is combined into audiovisual objects in multisensory brain regions (Romanovska & Bonte, 2021; Stein & Stanford, 2008). Repeated practice consequently feeds implicit learning mechanisms and ensures that audiovisual objects are stored in the neocortex for fast and automatic retrieval (Klinzing et al., 2019). This was also found in the work of Aravena et al. (2013), in which the authors argued that at least some explicit preparation is required to strengthen the benefits of implicit, associative training. Likewise, Rastle et al. (2021) showed that very few adults who had to rely on discovery learning performed on the same level as the adults who received explicit instructions in a task where they had to learn how to read novel words printed in two artificial alphabets, even after 18 hours of training. From a neural perspective, this learning process is accompanied by an inverted developmental U-curve in cortical responses to text and audiovisual stimuli, with maximal activations in beginning readers when reading is effortful that slowly decrease when reading becomes automatized and fluent (Fraga González et al., 2021). From an educational perspective, Wouters and van Oostendorp (2013), argued that clear instructions help learners to use their cognitive capacity efficiently and therefore improve learning. A recent review by McTigue et al. (2020) corroborated this notion. The authors synthesized 28 empirical studies that examined the effect of playing GraphoGame, an adaptive serious game that promotes sound-symbol connections to prevent reading difficulties, and factors that moderated the outcome measures. Results suggested that adult involvement was a critical parameter when individuals had to learn from a serious game. More specifically, they argued that adult involvement helped learners to efficiently select and organize new information, which seems to have a similar effect as the goal-directed instructions in the current study.

Although this study provides some important insights into how top-down control contributes to letter-speech sound learning and subsequent consolidation, some limitations need to be taken into account. First, this study did not assess common predictor measures such as phonemic awareness (PA) and rapid automatized naming (RAN). However, previous studies suggested that the outcome of a comparable learning task still uniquely contributed to the variance of reading performance (Gellert & Elbro, 2017; Horbach et al., 2015). The same is true for short-term memory. Although verbal short-term memory is needed to memorize speech sounds and to merge sounds into whole words (Gathercole et al., 2006), studies showed that a sound-symbol paradigm was positively associated with reading performance over and above short-term memory. Second, our results suggest that sleep consolidation increases letter-speech sound knowledge. However, we did not include any measure to quantify sleep quality, which might have influenced the consolidation process. Likewise, results might be influenced by motivation. Previous studies have shown that prompting children to pursue a specific goal can considerably impact intrinsic motivation (Barron & Harackiewicz, 2001), and consequently the quality of learning. In line with this, our goal-directed instructions, i.e., learn the symbols to crack a secret code at the end, might have been more motivating for the child compared to the implicit instructions. Including a subjective measure to assess children's motivation during the task in future studies would allow us to get a more detailed insight into the dynamic interplay of attentional and motivational influences and strengthen the conclusions about the benefit of goal-directed instructions. Moreover, although gender distribution was not significantly different between our goal-directed and implicit instruction condition, gender could still influence the results given possible attentional and motivational differences between girls and boys. A future study should use stratified randomization, which might prevent gender imbalance between the two conditions. Third, all our participants received several years of formal reading instruction, meaning that they were already aware that letters correspond to sounds. This may have helped them during the letter-speech sound learning task, although the learning curves showed that learning the new correspondences was nontrivial. Last, due to our design, it was not possible to examine the influence of goal-directed instructions on the application of the new knowledge without accounting for differences in symbol knowledge. Although we were interested in the full range of individual differences in reading levels, we had to select a subgroup of children who obtained full mastery of the new symbols to compare reading fluency within the artificial and Dutch script, as fluent reading requires the knowledge of all symbols. Moreover, for children who could not learn a substantial number of the symbols, it was rather difficult to perform the reading tasks. As our reading tasks comprised short, monosyllabic words, children could guess the word when they for example knew two out of three symbols, whereas some would immediately say that they did not know the answer. It is not known to what extent our reading outcomes were influenced by such personality traits, as for example introversion and performance anxiety. Although this again nicely highlights individual differences between participants, future studies might want to employ either a longer learning task to ensure learning or an adaptive design in which participants need to obtain a predefined level of performance before moving to follow-up tasks (e.g., Karipidis et al., 2017).

In sum, our findings contribute to understanding the mechanisms that are associated with typical and atypical audiovisual integration at the early reading stage as well as to the long-standing debate concerning the role of top-down control in cross-modal learning. We showed that goal-directed instructions significantly influence how children learn new letter-speech sound correspondences after only 20 minutes of learning. In addition, we showed that children in both conditions profited from offline sleep consolidation, but this effect was most apparent in the goal-directed group. The influence of top-down control on letter-speech sound binding highlights this mechanism as a potential contributor to the atypical audiovisual integration in individuals with dyslexia, and thus, appears pertinent for intervention research. Our results also suggest that learning to read should be considered a multidimensional process influenced by other mechanisms such as top-down control and motivation, although most dyslexia research still focuses on one single underlying deficit and dyslexia diagnoses typically exclude co-occurring cognitive or neurological deficits. Future work should test this paradigm in a clinical sample comprising children with dyslexia and ADHD to explore the role of goal-directedness in learning to read on a behavioral and neural level and include test moments after a longer period to examine retention. Understanding such mechanisms is particularly relevant to educational and clinical practice, which are recently benefiting from new tools based on implicit associative learning (e.g., serious games). Better identification of individual differences in these mechanisms is of great importance for literacy policy, and for timely, effective therapeutic remediation strategies, as these can be better matched to the needs of the child.

4

Behavioral and neural correlates of cognitive control during artificial letter-speech sound learning in typical and atypical readers

Abstract

Despite the high co-occurrence of dyslexia and ADHD, it remains ununderstood how attentional processes influence the initial formation of the neurocognitive reading network. To fill this gap, we mimicked the initial phase of reading acquisition using an artificial letter-speech sound (L-SS) learning paradigm in which 71 school-aged children with dyslexia and 59 controls were instructed to learn eight new L-SS correspondences. Importantly, children with co-occurring ADHD were included in both groups. We assessed three relevant correlates of cognitive control associated with learning; (1) Goal-directed behavior was manipulated by giving either instructions that directed children toward the goal of the task (i.e., goal-directed instructions) or instructions that prompted children to discover the goal of the task on their own (i.e., implicit instructions). (2) Post-error slowing was examined in addition to behavioral learning outcomes, and (3) feedback-related negativity (FRN) ERPs were assessed as a measure of performance monitoring. Most children were able to learn the new script, but with large inter-individual variability and without main effects of instruction or reading group (children with dyslexia vs. controls). Interestingly, attentional measures were associated with learning outcomes and reading performance in the native language, suggesting that attentional difficulties are not only related to reduced reading performance but may also limit the acquisition and development of alphabetic knowledge. Further research should consider individual learning indices to clarify how cognitive control mechanisms contribute to the reading development, as this is highly relevant to understand and remediate reading difficulties.

Behavioral and neural correlates of cognitive control during artificial letter-speech sound learning in typical and dyslexic readers

Learning to read is essential in society. Although reading might seem straightforward, 3 to 10% of the population struggles persistently to develop fluent reading skills (Snowling, 2013), known as developmental dyslexia. Some studies have proposed a failure to develop fluent, automatized letter-speech sound (L-SS) mappings as the most proximal cause of dyslexia (Aravena et al., 2013; Blomert, 2011; Breznitz, 2002). In line with complex skill acquisition in general, cognitive control mechanisms contribute to the formation and integration of these mappings and the development of fluent reading skills (Segers et al., 2016; ten Braak et al., 2018). However, despite the high co-occurrence of ADHD and dyslexia, cognitive control processes and reading are commonly examined in isolation (but see e.g., Arrington et al., 2014; Cartwright, 2012; ten Braak et al., 2018). As such, how these processes influence the initial formation of the neurocognitive reading network remains poorly understood. To fill this gap, the current study employed an artificial letter-speech sound learning task to mimic the initial phase of learning to read and examined three correlates of cognitive control associated with learning. More insight into these correlates and how they relate to letter-speech sound learning can guide tailored support for struggling readers instead of employing a one-fits-all approach.

Learning the alphabetic principle

The first step of learning to read is discovering the alphabetic principle, that is, children learn that written symbols represent speech sounds (Castles et al., 2018). Knowledge of these letter-speech sound (L-SS) correspondences allows children to decode words aloud and subsequently blend them together. Along with learning which letters correspond to which speech sounds, a neural, visual system for processing written language integrates with an innate system for processing spoken language (Froyen et al., 2009; Romanovska & Bonte, 2021; van Atteveldt et al., 2004). Studies have shown that full integration and automatization of these letters and speech sounds is required to become a fluent reader (Blau et al., 2010; Blomert, 2011; Žarić et al., 2014). Furthermore, diminished L-SS automatization has been reported in individuals with dyslexia, including reduced neural activity while processing congruent letter-speech sound pairs (Blomert, 2011). However, there is no consensus on the different processes contributing to learning letter-speech sound couplings, and their implication in dyslexia remains understudied.

Learning to read involves directing attention toward the written stimuli in order to visually perceive words and link them to the corresponding speech sounds, requiring higher-level cognitive processes better known as executive functions (Booth et al., 2010; Farah et al., 2021; Horowitz-Kraus, 2014; Varvara et al., 2014). It is, therefore, not surprising that recent theories highlight the crucial role of cognitive control in reading (see Farah et al., 2021). However, studies on developmental disorders typically assess these factors in isolation, using restrictive ranges of measures that do not go beyond diagnostic boundaries (Astle & Fletcher-Watson, 2020). In fact, individuals with co-occurring attentional difficulties are commonly excluded in studies on reading disabilities, despite the high co-occurrence between reading and attentional difficulties (McGrath & Stoodley, 2019; Willcutt et al., 2010), underestimating the large inter-individual variability within groups. In addition, given the substantial variability in responsiveness to reading intervention (Torgesen et al., 2001), it is highly relevant to shed light on how cognitive control mechanisms contribute to reading acquisition and development.

Cognitive control and its mechanisms

Cognitive control mechanisms can be defined as mechanisms that are used to adjust information to adapt and optimize performance (Ridderinkhof et al., 2004). In the context of learning, cognitive control mechanisms are needed to select relevant information for the context and, subsequently, select actions in relation to their outcomes (Ridderinkhof et al., 2004). A widely used test to assess these mechanisms is the Wisconsin Card Sorting Test (WCST; Booth et al., 2010; Horowitz-Kraus, 2014; Horowitz-Kraus et al., 2018). Participants are instructed to sort cards according to various sorting principles (e.g., color, shape, number). Participants must deduce the sorting rule based on feedback they obtain regarding their accuracy and thus must either continue or change their sorting behavior. Studies have reported more errors, longer reaction times, and fewer completed categories in children and adults with dyslexia compared to their typically reading peers using the WSCT (Booth et al., 2010; Horowitz-Kraus, 2014; Horowitz-Kraus et al., 2018). These findings were attributed to larger switch costs between tasks in individuals with dyslexia compared to controls (Poljac et al., 2010), working memory impairments (Horowitz-Kraus, 2014), or lower baseline performance due to a top-down processing impairment (Kraus & Horowitz-Kraus, 2014).

In order to learn from previous mistakes, it is essential to be aware of erroneous responses, either by internal or external feedback (Dion & Restrepo, 2016; Narciss et al., 2014). Studies postulated that individuals with dyslexia are less efficient in processing external feedback cues, given that their reading errors tend to be inconsistent (Breznitz, 1987; Horowitz-Kraus & Breznitz, 2008). This consequently interferes with their potential to learn from previous mistakes. On a neural level, feedback processing is typically examined with an event-related potential (ERP) marker known as the feedback-related negativity (FRN; e.g., Dion & Restrepo, 2016; Groen et al., 2007; Hauser et al., 2014). The FRN is a negative going ERP, time-locked to the presentation of feedback, mostly recorded at fronto-central recording sites. The FRN typically peaks between 200 and 350 ms after feedback onset and has been found to be larger (i.e., more negative) following negative feedback compared to positive feedback in various learning and gambling paradigms (Arbel & Fox, 2021; Ferdinand et al., 2016). Moreover, studies have shown that a greater difference between the amplitude of the FRN following negative versus positive feedback is associated with higher accuracy and better learning (Hämmerer et al., 2011; van der Helden et al., 2010). In the context of dyslexia, Kraus & Horowitz-Kraus (2014) found lower FRN amplitudes (i.e., less negative) for children with dyslexia than their typically developing peers, but only for the early trials of the task. The authors interpreted this as a lower baseline performance in individuals with dyslexia that could be rapidly adjusted by repetitive training. Literature examining feedback processing in individuals with dyslexia is, however, scarce. In addition, most studies that examined feedback processing used gambling tasks or tasks that assess executive functions, so it remains unclear how feedback processing relates to reading development. It is, therefore, essential to better understand how processing feedback relates to the initial phase of reading, i.e., mapping symbols onto speech sounds.

The present study

The present study mimicked the initial formation of the neurocognitive reading network by employing an artificial symbol-sound learning task in which children learned to associate unknown letter-like symbols with Dutch speech sounds (Aravena et al., 2013; Fraga González et al., 2015; Guerra, 2022; Verwimp et al., 2023). Given the high co-occurrence of attentional difficulties and dyslexia (Willcutt & Pennington, 2000), children with ADHD were included in both groups to investigate reading within a broader developmental context. Goal-directed behavior was manipulated by giving either instructions that directed children toward the goal of the task (i.e., goal-directed instructions) or instructions that prompted children to discover the goal of the task on their own (i.e., implicit instructions). We examined post-error slowing during the learning task as an additional behavioral cognitive control index. To investigate feedback processing, we examined the FRN to positive and negative feedback in the symbol-sound learning task. In addition, we examined associations between ERP measures, behavioral performance, and attentional measures to determine the role of individual learning differences.

Based on previous research that examined the influence of instruction manipulation on letter-speech sound learning (Aravena et al., 2013; Verwimp et al., 2023), we expected to find a main effect of instruction. That is, both children with dyslexia and typical readers were expected to benefit from goal-directed instructions by showing greater accuracy gains during the L-SS learning task compared to their implicitly instructed peers. However, we expected that dyslexic readers would perform worse than typical readers when applying this knowledge in a time-restricted reading task within the artificial script that was administered after the learning task. Given the reported performance monitoring deficit in individuals with dyslexia (Kraus & Horowitz-Kraus, 2014), we expected to find decreased FRN amplitudes in children with dyslexia, especially in the implicit condition. Last, we hypothesized that high ADHD measures would negatively impact the letter-speech sound learning outcomes.

Method

Participants and procedure

Seventy-one children with dyslexia (38 girls) and fifty-nine (29 girls) typical readers participated in this study. The average age of the dyslexic group was 9.82 years (*SD* = 0.96), and of the typical readers 9.22 years (*SD* = 1.29). Children with dyslexia (DR) were recruited via a nationwide clinical center for children with learning disabilities in the Netherlands. They had been referred to the center after showing severe and persistent reading difficulties at school, identified as performance below the tenth percentile on standard reading measures. Typically developing readers (TR) were recruited via local primary schools or were acquaintances of the children with dyslexia. Children were excluded when they reported uncorrected sight problems or hearing loss. All children scored above 75 on a non-verbal intelligence test (Raven's 2 Progressive Matrices; Raven & Raven, 2020). Children with co-occurring attentional difficulties were included in both groups to investigate reading within a broader developmental context. All children were native Dutch speakers. Participants' characteristics of the DR group and the TR group are reported in Table 1. Typical readers were slightly younger than children with dyslexia, but there was no age difference between children in the goal-directed or implicit condition (*p* = .68). Ethical approval was obtained from the Ethics board of the University of Amsterdam (2020-DP-12243), and all parents or caretakers actively signed informed consent before children participated.

The data analyzed in the current study was part of a longer testing session which comprised an electrophysiological (EEG) and behavioral part and took approximately 3 hours.

EEG acquisition took place in a temperature-regulated room at the University of Amsterdam. The complete procedure and how the EEG system works was first explained to the children to make them feel at ease. Children were seated in a comfortable chair where the capping and gelling occurred. The EEG session consisted of four different tasks, of which one will be described and analyzed in the current study. Participants were seated at approximately 80 cm distance from the computer screen, with the lab assistant sitting behind them to monitor compliance with the experimental procedure. At the same time, the experimenter controlled the EEG recording in an adjacent room. Two conventional reading(-related) measures were assessed to check whether children were correctly assigned to the dyslexic or control group. Word reading skills were assessed with a Dutch version of the One-minute test (*Een-minuut-test*, EMT; Brus & Voeten, 1973). Phoneme awareness was assessed with the *Fonemische Analyse Test* (FAT-R; De Groot et al., 2015). Finally, to assess behavioral symptoms of ADHD in children, parents filled out a short Dutch questionnaire (*ADHD-Vragenlijst*; Scholte & Van der Ploeg, 2005). Between the tasks, children had breaks that varied according to their needs, and all received a small present at the end of the session.

Table 1

Participants' Characteristics and Cognitive Assessments

	Mean (SD)				**DR vs TR**
Characteristic	**Goal-directed group (n = 66)**		**Implicit group (n = 64)**		
	Dyslexic readers	Typical readers	Dyslexic readers	Typical readers	
N	37	29	34	30	
Gender (M:F)	21/16	16/13	14/20	13/17	
Age (in years)	9.76 (0.99)	9.15 (1.27)	9.90 (0.95)	9.25 (1.36)	$F(1,126) = 9.14, p < .01$
Non-verbal IQ^A	100.86 (10.98)	100.14 (12.92)	96.06 (10.22)	98.59 (13.06)	$F(1,125) = 0.001, p = .98$
Word reading Dutch (raw)	32.32 (12.15)	58.07 (17.04)	34.06 (9.94)	67.24 (16.71)	$F(1,126) = 140.86, p < .01$
Phoneme awareness (T^B)	31.06 (8.92)	43.82 (11.81)	32.79 (7.84)	46.12 (12.31)	$F(1,124) = 52.21, p < .01$
AVL inattention (raw)	6.72 (4.72)	4.32 (3.41)	7.00 (4.58)	4.52 (3.51)	$F(1,122) = 11.11, p < .01$
AVL hyperactivity (raw)	5.14 (4.16)	3.86 (4.58)	6.53 (5.35)	4.66 (3.43)	$F(1,122) = 4.44, p < .05$
AVL total (raw)	18.94 (11.95)	13.36 (12.07)	20.06 (13.01)	14.31 (8.35)	$F(1,122) = 8.08, p < .05$

Note. ^{A}IQ-scores (M = 100, SD = 15). ^{B}T-scores (M = 50, SD = 10). AVL = ADHD questionnaire. Subclinical and clinical cut-off scores for the separate scales, i.e., AVL inattention and hyperactivity, are 12 and 16, respectively. For the total score, the subclinical cut-off score is 36 and the clinical cut-off score is 48. Only main effects of group (dyslexic readers vs. typical readers) are reported in this table for sake of brevity

Artificial learning paradigm

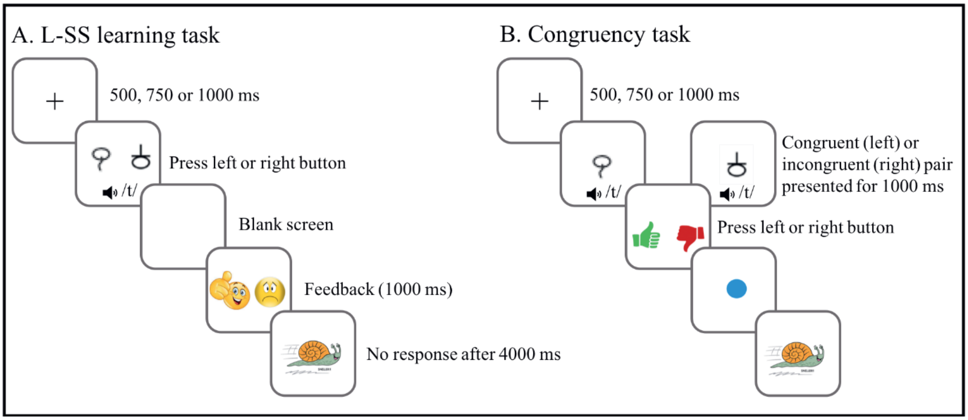
To mimic the initial formation of the neurocognitive reading network, children performed a learning task involving a set of eight L-SS correspondences (see Table 2). The stimuli comprised eight artificial symbols from the BACS-1 alphabet (Vidal et al., 2017) matched to a Dutch phoneme (adapted from Aravena et al., 2013; Fraga González et al., 2015; Guerra et al., In revision). The task was programmed in and presented with Psychopy3 (Peirce et al., 2019). In each trial, a fixation cross was presented with equiprobable durations of 500, 750 or 1000 ms, followed by an audiovisual presentation (Figure 1A). Two black artificial symbols were presented on a white screen, one on the left side and one on the right side of the screen, together with an aurally presented speech sound. Children were asked to indicate which symbol corresponded to the speech sound by pressing the left or right button (50% chance level). This button press was followed by visual feedback (a happy face for correct responses and a sad face for incorrect responses), presented for 1000 ms to indicate whether the answer was correct. If children did not respond within the required response time (4000 ms), a picture of a snail appeared on the screen to encourage them to answer faster.

The learning task took approximately 25 minutes and consisted of three learning blocks and one testing block. In the first block, four out of eight L-SS correspondences were presented, with each speech sound occurring 18 times (72 trials in total). In the second block, the four other pairs were presented. In the third block, all eight L-SS pairs were presented, with each speech sound occurring 8 times (56 trials in total). The fourth block served as a testing block, in which either congruent (learned correspondence) and incongruent (random correspondence) L-SS pairs were presented (Figure 1B). Participants pressed the left or right button to indicate whether the pair was congruent or incongruent. Each speech sound occurred four times, two times with the corresponding symbol and two times with a randomly paired symbol (32 trials in total). In the fourth block, children received no feedback whether the answer was correct, but a blue dot appeared in the middle of the screen to indicate that the computer had recorded their answer. No EEG was recorded during this block, so the outcome only served as a behavioral measure for passive L-SS integration.

Immediately after the learning task, children were presented with a sheet of paper containing the eight artificial symbols and were asked to name them aloud (max score = 8). Subsequently, they were presented with two lists of 14 monosyllabic words written within the artificial script. One list resulted in high-frequency Dutch words when correctly decoded and the other resulted in pseudowords. Children were instructed to read as many words as possible within one minute (max score = 14 for each list).

Figure 1

Overview of the Experimental Design



Note. L-SS = letter-speech sound.

To manipulate cognitive control, children received standardized, oral instructions that either revealed the goal of the task (hereafter referred to as goal-directed instructions), or implicit instructions that prompted them to discover the goal of the task by themselves. Goal-directed instructions prompted children to pay close attention to the presented symbols and speech sounds and to learn as many symbols as possible to crack the secret code. Children in the implicit instruction condition were told to play a computer game of which the goal of the task would become clear during the game. The instructions were repeated on the computer screen during the breaks. Children were randomly allocated to one out of two experimental conditions. All participants were told which two keys to use during the game and received the same computerized feedback during the learning phase. For more details on the paradigm, we refer to Verwimp et al. (2023).

Table 2

Letter-Speech Sound Combinations Adapted From the Brussels Artificial Character Set-1 Alphabet (Vidal et al., 2017)

Symbol								
Speech sound (IPA)	[ɑʊ]	[t]	[z]	[ɛɪ]	[ɛ]	[n]	[f]	[ɔ]
Phoneme duration (ms)	505	194	516	527	387	734	303	383

EEG acquisition and preprocessing

Continuous EEG was recorded using a 64-channel Biosemi ActiveTwo system (Biosemi, Amsterdam, Netherlands) at a 1024 Hz sample rate. The 64 Ag/AgCl scalp electrodes were applied using an elastic electrode cap and positioned across the scalp according to the 10-20 international system. Four external flat-type active electrodes were used to record vertical and horizontal electro-oculogram, and two were placed on the left and right mastoids for offline reference. In line with the manufacturer's recommendations, electrode offsets were kept below 50 kΩ (see https://www.biosemi.com/faq.htm for details).

EEG data was preprocessed with EEGLAB (Brunner et al., 2013). The continuous EEG data was resampled to 256 Hz and digitally filtered using a basic FIR filter (high pass 1 Hz and low pass 70 Hz). Trials containing excessive non-stereotyped artifacts were removed. Channels that showed non-stereotyped artifacts across the EEG recording were removed and interpolated ($M = 0.60$, range = 0-3) after Independent Component Analysis (ICA). ICA using the "runica" algorithm (Makeig et al., 1997) was used to remove ocular, muscle, and line noise artifacts. After removing artifactual components, all data were filtered with a 30 Hz low-pass filter, re-referenced to the average of the 64 scalp electrodes, and segmented into 1200 ms epochs starting 200 ms before the onset of the feedback stimuli and ending 1000 ms following the feedback. Baseline correction was performed on the first 200 ms of the epoch.

Epochs were averaged separately for positive and negative feedback for each condition (goal-directed vs. implicit) and group (DR vs. TR). To examine whether feedback processing changed during the task, this was done separately for the early and late trials of the task. More specifically, the first half of the trials of blocks 1 and 2 were considered early trials and were averaged together for positive and negative feedback separately, and the second half of the trials were considered late trials and were averaged together as well. The average number of total trials per participant was 134.30 ($SD = 9.25$, range = 95-144), with an average of 82.03 positive feedback trials ($SD = 17.16$) and 52.26 negative feedback trials ($SD = 15.42$). Children with dyslexia did not differ significantly from typical readers in trial counts (*t*(113.50) = -0.02, *p* = .98), and neither did the two instruction conditions (*t*(117.35) = 0.95, *p* = .34).

We used a semi-automatic peak detection method to quantify the amplitudes and latencies of our components of interest. The FRNp and FRNn were initially quantified as the difference between the maximum positive peak between 150-250 ms and the maximum negative peak between 250-350 ms (hereafter referred to as the peak-to-peak FRN), measured at electrode site Fz in line with previous literature (Hämmerer et al., 2011; Roos et al., 2015). However, to examine whether the effect was not driven by a P2 effect, we also extracted mean amplitudes for these two peaks separately. Mean amplitudes were defined as the mean voltage over a 50-ms window centered around the maximum positive peak between 150-250 ms (hereafter referred to as the P2), and around the maximum negative peak between 250-350 ms (hereafter referred to as the FRN). Last, mean amplitudes for an additional late negativity (hereafter referred to as the LN) were defined as +- 50 ms around the most negative peak between 400-600 ms after feedback onset. For 6 children with dyslexia (3 goal-directed, 3 implicit) and 10 typical readers (3 goal-directed and 7 implicit), at least one of the peaks was manually selected outside the predefined time windows (all within a range of 28 ms).

Data analysis

Behavioral measures

From the initial sample of 130 children, data of 3 participants was excluded from the behavioral analysis because they did not complete the learning task. For the behavioral responses during the L-SS learning task, reaction times were first outlier-corrected. Trials with reaction times of less than 100 ms were removed as these were considered to result from guessing behavior. Second, reaction times deviating more than 3 SDs from the individual mean of each participant were excluded (Geurts et al., 2008). A maximum of three trials per participant was removed based on this procedure. Only trials with correct answers were included in the reaction time analysis.

First, we examined how goal-directed instructions influenced the learning of new L-SS correspondences in children with dyslexia and their typically developing peers. To map the learning curves, trials were averaged across 4 bins of 18 trials for Block 1 and 2, and of 14 trials for Block 3. Accuracy (% correct) and reaction time were computed for each bin per block and submitted to two separate mixed ANOVAs with Accuracy and Reaction time as dependent variables, Condition (Goal-directed versus implicit) and Group (DR versus TR) as between-subjects variables, and Time (Bin 1 versus Bin 2 versus Bin 3 versus Bin 4) as within-subjects variable.

As a behavioral index of cognitive control, we examined whether post-error slowing was influenced by goal-directed instructions in children with dyslexia and typical readers. We conducted a factorial ANOVA with post-error slowing as dependent variable, and Condition and Group as between-subjects variables. Only trials with correct responses, either following a correct or an incorrect trial, were taken into account in line with previous literature (Roos et al., 2015).

Second, we examined the influence of instruction on the learning outcomes of the L-SS task. Two mixed ANOVAs were conducted with mean accuracy and reaction time of block 4 (testing block) as dependent variables, Condition (goal-directed vs implicit) and Group (DR vs TR) as between-subjects variables, and Congruency (congruent vs incongruent trials) as within-subjects variable. Furthermore, a mixed ANOVA with symbol knowledge as dependent variable, and Condition (goal-directed vs implicit) and Group (DR vs TR) as between-subjects variables was conducted. As reading requires knowledge of the symbols, a subgroup analysis was conducted with only children who obtained full mastery of the new symbols ($n = 31$), to examine the effect of instructions on applying the new knowledge in a reading task.

Last, to examine the external validity of our artificial task, we conducted a one-tailed Pearson correlation test between reading rate in Dutch and reading rate in the artificial orthography (real words and pseudowords) within the subgroup of children who obtained full mastery of the new symbols ($n = 31$).

Electrophysiological measures

Data of 8 participants was excluded from the EEG data analysis; three children did not complete the learning task, data of three other children was not saved properly due to technical issues, and data of two children was removed due to excessive artifacts. ERP amplitudes and latencies were inspected for extreme outliers based on the interquartile range (IQR) method (Schwertman et al., 2004). Values higher than 3 times the IQR above the upper quartile, or lower than 3 times the IQR below the lower quartile were considered extreme outliers. Based on this criterion, data points of two participants were excluded. Thus, final ERP analyses were based on data from 63 children with dyslexia (34 goal-directed, 29 implicit), and 56 typical readers (27 goal-directed, 29 implicit). To assess the ERP waves elicited by feedback stimuli, peak-to-peak FRN amplitudes and latencies were submitted to a mixed ANOVA with Group and Condition as between-subjects variables and Valence (positive vs. negative feedback) and

Phase (early vs. late) as within-subjects variables. To check whether results were not driven by P2 differences, P2 mean amplitudes and FRN mean amplitudes were also analyzed separately with mixed ANOVAs. Last, mean late negativity (LN) amplitudes were submitted to a mixed ANOVA with Group and Condition as between-subjects variables and Valence and Phase as within-subjects variables.

We conducted a correlational analysis to further understand how behavioral and ERP markers of learning were related and how these were consequently related to reading ability. To reduce the dimensionality of the data, we first conducted a principal component analysis with an oblique promax rotation. The resulting factor scores were used in a correlational analysis. We controlled for age and corrected p-values using FDR correction for multiple tests. All analyses were conducted in RStudio version 2022.07.2 with a significance threshold of 0.05 (R Studio Team, 2022).

Results

Letter-speech sound learning

Behavioral learning curves. Children varied considerably in their learning performance. On average, children obtained a total accuracy of 62.06% (*SD* = 12.98%) across the three learning blocks, which was significantly better than chance level ($t(126) = 10.47$, $p < .001$). These findings indicated that children were able to learn the artificial letter-speech sound pairs. Based on descriptive values, children who obtained goal-directed instructions attained slightly higher accuracy scores (DR: *M* = 63.00%, *SD* = 14.70%; TR: *M* = 62.50%, *SD* = 12.80%) compared to implicitly instructed children (DR: *M* = 60.90%, *SD* = 12.00%; TR: *M* = 61.70%, *SD* = 12.40%).

To examine differences in their learning curves, two 2 (Condition) x 2 (Group) x 4 (Time) mixed ANOVAs with accuracy and reaction time as dependent variables were performed for Block 1, 2 and 3 separately (see Figure 2). To control for age-related effects, we used a median split of age as an additional between-subjects factor in the model (younger children: *M* = 8.58 years, *SD* = 0.61; older children: *M* = 10.58 years, *SD* = 0.66). Mean accuracy and reaction time scores for each bin are shown in Table 3 and statistical values are reported in Table S1 (Appendix C). For accuracy, this revealed a main effect of Time in all three blocks (*p*s < .001), but no significant main effect of condition, group or any interaction effect was found (*p*s > .05). In addition, a main effect of Age was found in all three blocks ($p < .01$), with older children responding on average more accurate than younger children. Post-hoc tests showed that in block 1, TR who were directed toward the goal of the task significantly increased in accuracy from bin 1 to 2 (*p* < .001 after FDR correction (*)), but not from bin 2 to 3 (*p* = .91*) or bin 3 to 4 (*p* = .37*). The same pattern appeared for block 2, with an increase in accuracy from bin 1 to 2 (p < .01*), but not from bin 2 to 3 (*p* = .26*) and bin 3 to 4 (*p* = .41*). In block 1, DR who were directed toward the goal of the task significantly increased in accuracy from bin 1 to 2 (*p* = .02*) and bin 2 to 3 (*p* < .01*), but not from bin 3 to 4 (*p* = .79*). In block 2, they only increased significantly in accuracy from bin 1 to 2 (*p* < .01*), but not from bin 2 to 3 (*p* = .06*) or bin 3 to 4 (*p* = .09*). In the implicit condition, both TR and DR only increased significantly from bin 1 to bin 2 (*p*s = .03* and < .01* for TR and DR respectively) in the first block. For block 2, both groups increased in accuracy from bin 1 to bin 2 (*p* = .04* and .05* for TR and DR respectively), and implicitly instructed TR significantly increased in accuracy from bin 2 to 3 (*p* < .01*). For block 3, only the difference in accuracy between bin 2 and 4 in the implicitly instructed DR remained significant after FDR correction (*p* = .03*).

Across the three learning blocks, children on average responded after 0.84 seconds (*SD* = 0.27). Based on descriptive values, DR and TR who obtained goal-directed instructions responded equally fast (DR: *M* = 0.85 sec, *SD* = 0.26; TR: *M* = 0.85 sec, *SD* = 0.30). In the implicit condition, DR responded slower than TR (DR: *M* = 0.87 sec, *SD* = 0.27; TR: *M* = 0.79 sec, *SD* = 0.27). Results of a mixed ANOVA revealed a main effect of Time for Block 1 for reaction time. Post-hoc tests showed that TR who were directed toward the goal of the task significantly decreased in reaction time from bin 1 to 2 (*p* = .01*). Implicitly instructed DR decreased in reaction time from bin 1 to 2 (*p* < .01*) and from bin 3 to 4 (*p* = .04*). Besides the main effect of time, a Condition x Group x Time effect was found for all blocks (*p*s < .01). To interpret these effects, mixed ANOVAs were conducted for children with dyslexia and typical readers separately. For both groups, only the time effect for block 1 remained significant after FDR correction.

In sum, although we did not find significant differences between the groups and conditions, post-hoc tests suggested that children with dyslexia and typical readers slightly differed in their learning trajectories. More specifically, goal-directed typical readers already became more accurate in the first bin of the first block, but then stagnated around an accuracy level of 60%, whereas goal-directed children with dyslexia needed slightly more time to obtain an similar accuracy level as typical readers. In addition, for the DR group it appeared that in the implicit condition, children only increased in accuracy in the early trials, but then stagnated around an accuracy level of 58%, whereas goal-directed children further increased until 62%.

In TR, the two conditions showed similar learning patterns. Regarding reaction time, the response pattern of goal-directed DR was more stable than implicitly instructed DR, whereas the opposite pattern was seen in typical readers. That is, implicitly instructed DR responded slightly slower in the first trials of the first block, but then became as fast as their goal-directed peers. For typical readers, goal-directed children responded slightly slower in the first trials of the first block, but then became as fast as their implicitly instructed peers.

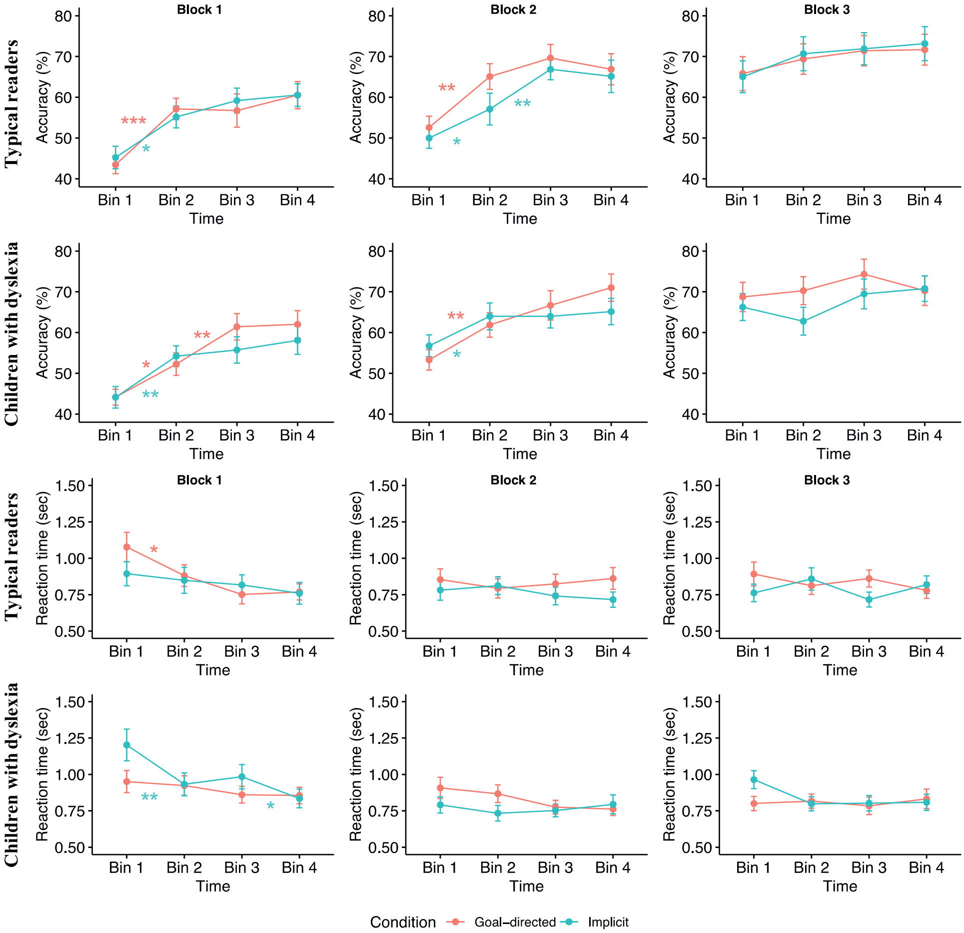
Table 3

Mean (SD) Accuracy and Reaction Time of Correct Responses by Condition and Group for Each Block

Condition		Goal-directed group (n = 65)				Implicit group (n = 62)			
		DR (n = 37)		TR (n = 28)		DR (n = 33)		TR (n = 29)	
Accuracy		*Mean (SD)*	*Range*	*Mean (SD)*	*Range*	*Mean (SD)*	*Range*	*Mean (SD)*	*Range*
Block 1	Bin 1	0.44 (0.12)	0.11-0.67	0.43 (0.12)	0.22-0.67	0.44 (0.15)	0.17-0.83	0.45 (0.15)	0.22-0.72
	Bin 2	0.52 (0.17)	0.17-0.89	0.57 (0.14)	0.39-0.83	0.54 (0.15)	0.33-0.89	0.55 (0.14)	0.28-0.94
	Bin 3	0.61 (0.20)	0.22-1.00	0.57 (0.22)	0.21-0.94	0.56 (0.19)	0.22-0.89	0.59 (0.16)	0.33-0.94
	Bin 4	0.62 (0.20)	0.17-1.00	0.61 (0.18)	0.22-0.94	0.58 (0.20)	0.28-0.94	0.61 (0.15)	0.33-0.94
Block 2	Bin 1	0.53 (0.15)	0.17-0.83	0.53 (0.15)	0.28-0.78	0.57 (0.15)	0.17-0.83	0.50 (0.14)	0.33-0.83
	Bin 2	0.62 (0.18)	0.28-1.00	0.65 (0.17)	0.39-0.94	0.64 (0.19)	0.28-1.00	0.57 (0.21)	0.22-0.94
	Bin 3	0.67 (0.22)	0.28-1.00	0.70 (0.18)	0.22-1.00	0.64 (0.16)	0.33-1.00	0.67 (0.14)	0.44-0.89
	Bin 4	0.71 (0.20)	0.28-1.00	0.67 (0.20)	0.28-1.00	0.65 (0.19)	0.28-0.94	0.65 (0.21)	0.17-0.94
Block 3	Bin 1	0.69 (0.22)	0.14-1.00	0.66 (0.22)	0.21-1.00	0.66 (0.19)	0.29-1.00	0.65 (0.21)	0.36-1.00
	Bin 2	0.70 (0.21)	0.14-1.00	0.69 (0.20)	0.29-1.00	0.63 (0.20)	0.21-1.00	0.71 (0.22)	0.29-1.00
	Bin 3	0.74 (0.22)	0.29-1.00	0.71 (0.20)	0.29-1.00	0.70 (0.21)	0.21-1.00	0.72 (0.21)	0.21-1.00
	Bin 4	0.70 (0.22)	0.29-1.00	0.72 (0.20)	0.36-1.00	0.71 (0.18)	0.36-1.00	0.73 (0.22)	0.21-1.00
Reaction time		*Mean (SD)*	*Range*	*Mean (SD)*	*Range*	*Mean (SD)*	*Range*	*Mean (SD)*	*Range*
Block 1	Bin 1	0.95 (0.46)	0.20-1.92	1.08 (0.53)	0.36-2.32	1.20 (0.62)	0.35-2.65	0.89 (0.44)	0.27-2.05
	Bin 2	0.92 (0.41)	0.28-1.72	0.88 (0.40)	0.36-2.19	0.93 (0.45)	0.25-1.95	0.85 (0.48)	0.31-2.34
	Bin 3	0.86 (0.35)	0.22-1.59	0.75 (0.34)	0.32-1.87	0.98 (0.48)	0.32-2.29	0.82 (0.37)	0.29-1.57
	Bin 4	0.86 (0.34)	0.32-1.74	0.77 (0.29)	0.38-1.63	0.83 (0.37)	0.34-1.82	0.76 (0.40)	0.17-1.61
Block 2	Bin 1	0.91 (0.44)	0.39-2.53	0.85 (0.39	0.39-2.53	0.79 (0.32)	0.33-1.39	0.78 (0.37)	0.27-1.69
	Bin 2	0.87 (0.37)	0.36-2.28	0.79 (0.35)	0.29-1.62	0.73 (0.31)	0.17-1.59	0.81 (0.33)	0.19-1.38
	Bin 3	0.78 (0.29)	0.29-1.43	0.82 (0.35)	0.19-1.42	0.75 (0.25)	0.27-1.40	0.74 (0.33)	0.27-1.33
	Bin 4	0.76 (0.27)	0.34-1.28	0.86 (0.40)	0.22-1.66	0.79 (0.37)	0.27-1.69	0.72 (0.28)	0.26-1.26
Block 3	Bin 1	0.80 (0.30)	0.16-1.38	0.89 (0.43)	0.19-1.88	0.96 (0.35)	0.49-1.81	0.76 (0.32)	0.34-1.43
	Bin 2	0.82 (0.29)	0.32-1.39	0.81 (0.32)	0.31-1.44	0.80 (0.28)	0.39-1.36	0.86 (0.42)	0.27-2.01
	Bin 3	0.78 (0.37)	0.26-1.64	0.86 (0.31)	0.35-1.64	0.80 (0.30)	0.34-1.53	0.72 (0.27)	0.12-1.34
	Bin 4	0.83 (0.41)	0.22-2.16	0.78 (0.29)	0.29-1.38	0.81 (0.32)	0.23-1.49	0.82 (0.33)	0.26-1.42

Figure 2

Learning Trajectories During The Letter-Speech Sound Learning Task



Note. ***p ≤ .001, **p ≤ .01, *p ≤ .05.

Post-error slowing. There was a significant main effect of Stimulus (*F*(1,123) = 16.01, *p* < .001), indicating that participants on average responded slower after incorrect trials than after correct trials. However, we did not find a significant main effect of Condition nor Group (*p*s > .53) nor any interaction effect (*p*s > .56).

Letter-speech sound integration and symbol knowledge. On average, children obtained a total accuracy of 69.40% (*SD* = 20.10) during the testing block, which was significantly better than chance level (*t*(126) = 12.38, *p* < .001). To examine whether the instruction manipulation influenced the L-SS learning outcomes, a mixed ANOVA with Condition (Goal-directed vs implicit) and Group (DR vs TR) as between-subjects variables, and Congruency (congruent vs incongruent) as within-subjects variable was conducted for accuracy and reaction time separately. Age (younger vs. older children) was added as a between-subjects variable. Statistical values are reported in Table S2 (Appendix C). For accuracy, we did not find any main effect, two-way or three-way interaction effects including Condition or Group (*p*s > .13). However, we found a main effect of Age ($p < .001$), indicating that older children ($M = 75.70\%$, $SD = 18.90$) were on average more accurate than younger children ($M = 63.80\%$, $SD = 19.60$). For reaction time, a significant main effect of Congruency revealed that children were on average faster in responding to congruent trials ($M = 0.99$ sec, $SD = 0.42$) than to incongruent trials ($M = 1.10$ sec, $SD = 0.45$) ($p < .001$). No other main effects or interaction effects reached statistical significance (*p*s > .06), so no follow-up tests were conducted.

On average, children learned 4.87 symbols ($SD = 2.72$) and could accurately decode 4.23 artificial words resulting in real Dutch words ($SD = 4.61$) and 4.05 artificial words resulting in pseudowords ($SD = 4.65$) (see Table 4). We found no significant main effects of either condition or group on how many symbols participants could actively name aloud after the learning task (*p*s > .72). In addition, no significant interaction effect between condition and group was found ($p = .93$). However, a main effect of Age revealed that older children on average learned more symbols ($M = 5.78$, $SD = 2.62$) compared to younger ones ($M = 3.98$, $SD = 2.53$) ($F(1,119) = 17.41$, $p < .001$). A subgroup analysis with only children who reached full mastery of the symbols ($n = 31$, DR: 10 goal-directed, 8 implicit; TR: 6 goal-directed, 7 implicit) did also not yield any differences between either conditions or groups in applying the new knowledge in a time-limited reading task (*p*s > .48). In addition, we did not find a main effect of Age within this subgroup analysis ($p = .87$), suggesting that once children were able to learn all symbols, older children were not better in decoding the artificial words than younger children.

Last, a one-tailed Pearson correlation test in this subgroup showed a correlation between reading rate in Dutch and reading rate in the artificial orthography for both real words ($r = .35$, $p = .03$) and pseudowords ($r = .42$, $p < .01$), indicating that proficient readers in Dutch also tended to perform better when decoding the new, artificial script.

Table 4

Means (SD) of Outcome Measures For Both Conditions Separately

Condition	Goal-directed group (*n* = 65)		Implicit group (*n* = 62)	
	DR (*n* = 37)	TR (*n* = 28)	DR (*n* = 33)	TR (*n* = 29)
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Symbol knowledge	4.73 (3.01)	4.86 (2.52)	4.88 (2.74)	5.03 (2.66)
OMT real words	3.59 (4.50)	4.68 (4.73)	4.36 (4.64)	4.45 (4.76)
OMT pseudowords	3.30 (4.35)	4.21 (4.55)	4.33 (4.81)	4.52 (5.07)

Note. OMT = One-minute reading test within the artificial script.

Electrophysiological measures

Peak-to-peak FRN. The feedback-locked ERPs for each condition and group separately at recording site Fz are displayed in Figure 3. The averaged data showed a negative deflection in the ERP waveform following feedback which appeared to be more negative following positive feedback than following positive feedback. In addition, the amplitude difference between positive and negative feedback appeared to be greater in children with dyslexia.

A mixed ANOVA revealed no main effect of either Group (*F*(1,111 = 0.06, *p* = .81) or Condition (*F*(1,117) = 0.86, *p* = .36). A main effect of Valence indicated that peak-to-peak amplitudes were on average higher following negative feedback stimuli compared to positive stimuli (*F*(1,111) = 62.65, *p* < .001) and higher in the first half of the trials of blocks 1 and 2 (i.e., early phase of the learning task) compared to the second half of the trials of blocks 1 and 2 (i.e., late phase) (main effect Phase: *F*(1,111) = 15.95, *p* < .001). In addition, we found an interaction effect between Condition and Phase (*F*(1,111) = 6.06, *p* = .02) and Condition and Group (*F*(1,111) = 6.11, *p* = .02). Post-hoc tests revealed that for children who received goal-directed instructions, peak-to-peak amplitudes were on average higher in early trials (*M* = 15.90, *SD* = 6.61) compared to later trials (*M* = 14.1, *SD* = 6.44) (*p* < .001), but these did not differ significantly in implicitly instructed children (early: *M* = 15.70, *SD* = 5.68; late: *M* = 15.30, *SD* = 5.66) (*p* = .27). In addition, implicitly instructed typical readers on average showed higher peak-to-peak amplitudes (*M* = 16.80, *SD* = 6.04) compared to their goal-directed instructed peers (*M* = 14.00, *SD* = 7.01) (*p* < .001), whereas the opposite pattern appeared in children with dyslexia (implicit: *M* = 14.20, *SD* = 4.94; goal-directed: *M* = 15.80, *SD* = 6.12) (*p* = .03). Last, there was no main effect of Age (*F*(1,111) = 1.12, *p* = .29), but we found a three-way interaction effect between Group, Valence, and Age (*F*(1,111) = 4.65, *p* = .03). To further interpret this effect, mixed ANOVAs were conducted for children with dyslexia and typical readers separately, with Age as between-subjects variable and Valence as within-subjects variable. For

TR, peak-to-peak amplitudes were higher for negative feedback than for positive feedback (main effect Valence: *F*(1,54) = 25.11, *p* < .001). The same was found in DR (main effect Valence: *F*(1,61) = 32.79, *p* < .001), but this was especially true in younger children (interaction effect Valence x Age: *F*(1,61) = 4.47, *p* = .04).

Regarding latency, a mixed ANOVA with peak-to-peak latencies revealed an interaction effect between Group and Valence (*F*(1,111) = 5.25, *p* = .02). More specifically, in typical readers, the time between the positive P2 peak and the negative FRN peak did not differ significantly following positive or negative feedback (*p* =.12). In children with dyslexia, the time between the two peaks was significantly longer after obtaining positive feedback compared to negative feedback (*p* < .01). In addition, we found a significant interaction effect between Age and Valence (*F*(1,111) = 4.21, *p* = .04). In younger children, the time between two successive peaks was longer after negative feedback (*M* = 118.00) than after positive feedback (*M* = 113.00), although this difference was not significant (*p* = .11). In older children, this time was significantly longer after positive feedback (*M* = 115.00) than after negative feedback (*M* = 106.00) (*p* < .01).

P2. A mixed ANOVA with mean P2 amplitudes revealed a main effect of Valence, with positive feedback on average eliciting a smaller P2 than negative feedback (*F*(1,111) = 171.13, *p* < .001). In addition, an interaction effect between Valence and Phase indicated that the P2 amplitude became smaller in later phases of the task after obtaining negative feedback (*p* = .02), but this effect did not appear for positive feedback (*p* = .55) (*F*(1,111) = 4.91, *p* = .03). Last, an interaction effect between Age and Valence suggested that although negative feedback elicited a more pronounced P2 amplitude than positive feedback for both age groups, this effect was larger in younger children (*F*(1,111) = 4.87, *p* = .03).

Regarding latency, the P2 peak appeared earlier for positive feedback compared to negative feedback (main effect Valence: *F*(1,111) = 41.23, *p* < .001). A Group x Valence interaction effect indicated that this effect was especially apparent in dyslexic readers (*F*(1,111) = 12.03, *p* < .001). In addition, a Group x Condition interaction effect indicated that when children received goal-directed instructions, P2 latencies of children with dyslexia and typical readers did not significantly differ from each other (*p* = .39). In implicitly instructed children, the P2 peak appeared earlier in typical readers compared to children with dyslexia (*p* < .001).

FRN. A mixed ANOVA with mean FRN amplitudes revealed a main effect of Valence (*F*(1,111) = 26.75, *p* < .001) and Phase (*F*(1,111) = 7.88, *p* < .001), with more negative amplitudes for positive (*M* = -9.28, *SD* = 5.00) compared to negative feedback (*M* = -7.45, *SD*

= 5.46), and more negative amplitudes in the early trials of the task (*M* = -8.79, *SD* = 5.30) compared to the late trials (*M* = -7.94, *SD* = 5.29). No interaction effects were found (*p*s > .33). In addition, we found a three-way interaction effect between Group, Age, and Valence (*F*(1,111) = 5.29, *p* = .02). To further interpret this effect, mixed ANOVAs were conducted for children with dyslexia and typical readers separately, with Age as between-subjects variable and Valence as within-subjects variable. For DR, FRN amplitudes were more negative for positive feedback than for negative feedback (main effect Valence: *F*(1,61) = 0.42, *p* < .001). The same was found in TR (main effect Valence: *F*(1,654) = 9.17, *p* < .01), but this was especially true in younger children (interaction effect Valence x Age: *F*(1,54) = 5.06, *p* = .03).

Regarding latency, FRN peaks appeared earlier following positive feedback compared to negative feedback (main effect Valence: *F*(1,111) = 6.47, *p* = .01), and earlier in older children compared to younger ones (main effect of Age: *F*(1,111) = 5.66, *p* = .01). In addition, we found a main effect of Phase, indicating that peaks appeared earlier in early trials of the task compared to late trials *F*(1,115) = 6.12, *p* = .02). No two-way or three-way interaction effects were found (*p*s > .11).

Late negativity measures. A mixed ANOVA revealed a main effect of Group (*F*(1,111) = 4.88, *p* = .03), Valence (*F*(1,111) = 19.12, *p* < .001) and Phase (*F*(1,111) = 20.34, *p* < .001). That is, mean amplitudes were on average more negative for typical readers compared to children with dyslexia, more negative following positive feedback compared to negative feedback, and more negative in early trials compared to late trials.

Regarding latency, the late negativity peak appeared earlier after positive feedback than negative feedback (*F*(1,111) = 46.49, *p* < .001). In addition, we found a three-way interaction effect between Group, Valence, and Phase (*F*(1,111) = 4.57, *p* = .04). To further interpret this effect, factorial ANOVAs were conducted for children with dyslexia and typical readers separately, with Valence and Phase as within-subjects variables. For both groups, this revealed a significant main effect of Valence, with the LN peak occurring earlier after positive feedback than after negative feedback (*p*s < .01), but no interaction effect between Phase and Valence (*p*s > .30).

Figure 3

Feedback-Locked ERPs For Each Condition and Group Separately

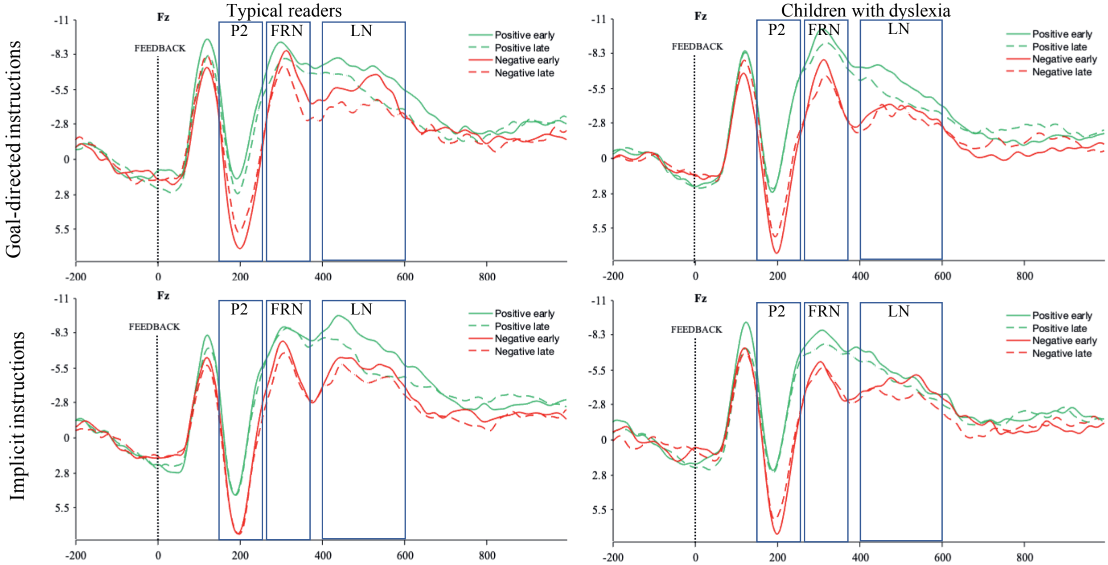
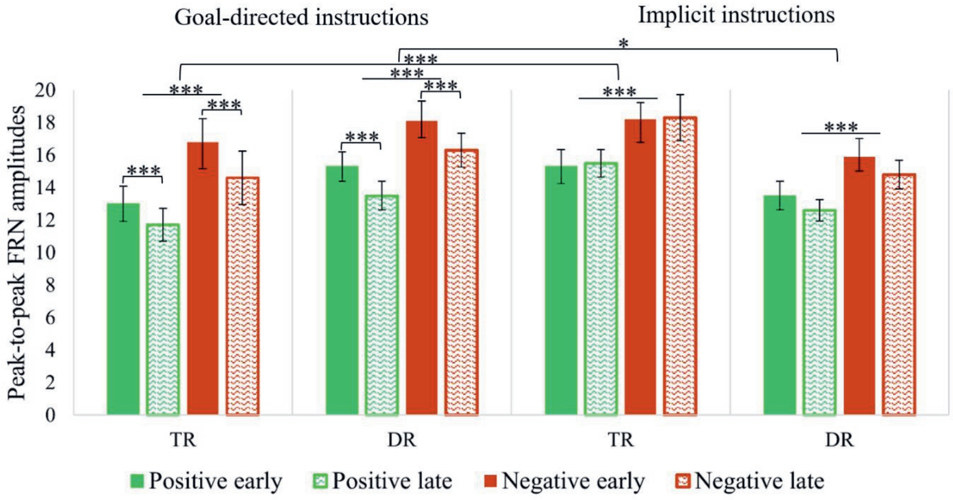


Figure 4

Mean Peak-to-Peak FRN Amplitudes For Each Condition and Group Separately



Note. FRN = feedback-related negativity. TR = typical readers. DR = children with dyslexia. ***p ≤ .001, **p ≤ .01, *p ≤ .05. Error bars represent standard errors.

Correlation analyses

To further investigate how individual differences in behavioral and ERP markers of learning were related, we performed correlational analyses between learning outcomes, reading-related measures, attentional and neural measures. First, to reduce the dimensionality of the data, we ran a principal component analysis with oblique Promax rotation. This yielded eight factors with eigenvalues above 1 (see Appendix C Table S3). As only one variable loaded on component 8, we chose for a seven-factor solution. Together, these components explained 74% of the variance. Factor loadings are reported in Table S4 (Appendix C).

Correlations between the principal components were corrected for age and are presented in Table 5. On a behavioral level, attention (component 3), was negatively associated with the learning outcomes of the L-SS task (component 4) (*r* = -.30) and with reading-related measures (component 5) (*r* = -.68). Regarding associations between our neural and behavioral measures, attention measures were positively associated with late negativity amplitudes (component 7) (*r* = .23). The associations between late negativity amplitudes and reading-related measures (*r* = -.22) and FRN amplitudes (component 2) and L-SS learning outcomes and reading measures (*r* = .21) did not survive FDR correction.

Table 5

Correlation Matrix of Principal Components Comprising Learning Outcomes, Reading-related Measures, Attention Measures and ERP Measures

	RC1	RC2	RC3	RC4	RC5	RC6	RC7
RC1: P2-FRN		.18*	.10	-.11	-.07	.07	**.92*** **
RC2: FRN	.18*		-.05	.21*	.21*	**.64*** **	.19*
RC3: Attention	.10	-.05		**-.30*** **	**-.68*** **	-.01	**.23* **
RC4: L-SS	-.11	.21*	**-.30*** **		**.49*** **	.21*	.01
RC5: Reading	-.07	.21*	**-.68*** **	**.49*** **		.12	-.22*
RC6: Latency	-.07	**.64** **	-.01	-.21*	.12		-.06
RC7: LN	**.92*** **	.19*	**.23* **	.01	-.22*	-.06	

Note. Correlations are corrected for age. ****p* < .001, ***p* < .01, **p* < .05. P-values that remained significant after FDR correction are shown in bold. FRN = Feedback-related negativity. L-SS = letters-speech sound learning outcomes. LN = late negativity.

Discussion

In this study, we sought to examine how cognitive control processes relate to the mapping of written onto auditory language representations. Children with and without dyslexia performed a learning task in which they learned eight new letter-speech sound mappings to mimic the initial phase of reading acquisition. Children obtained either instructions that directed them toward the goal of the task (i.e., goal-directed instructions) or instructions that directed them to discover the goal of the task on their own. During this task, we assessed post-error slowing as a behavioral index of cognitive control and feedback-related negativity ERPs as a measure of performance monitoring. Our results revealed large inter-individual differences in learning outcomes, but we did not find any consistent significant group differences between instruction conditions or reading groups (children with dyslexia vs. typical readers) on a behavioral or neural level. However, our results suggest that, on average, children with dyslexia followed a slightly different learning trajectory compared to their typical reading peers, which was also influenced by instructions. Results are discussed in light of the reported audiovisual integration and performance monitoring deficits in dyslexia. More insight into these neurocognitive mechanisms is highly relevant for understanding the mechanisms that underlie dysfluent reading and how this can be translated into clinical practice and education.

Learning a new artificial script

Previous studies suggested that the ability to learn associations between written and spoken language representations is related to reading ability (Aravena et al., 2013, 2018; Horbach et al., 2018). Aravena et al. (2013) used an artificial learning paradigm in which children were instructed to learn correspondences between Hebrew symbols and Dutch speech sounds. Interestingly, although children with dyslexia and typically reading peers did not differ in symbol knowledge, children with dyslexia performed worse than their peers when applying the new knowledge in a time-limited reading task. In a later study, the same authors showed that such a learning task could differentiate individuals with dyslexia from typical readers (Aravena et al., 2018). In a study by Horbach et al. (2015), outcomes of a symbol-sound learning paradigm in preschool children predicted reading fluency and comprehension three years later. Their task predicted later reading performance over and above the established predictors phonological awareness, rapid automatized naming, short-term memory, and environmental measures (Horbach et al., 2015).

In the current study, we did not find significant differences between children with dyslexia and typical readers in their ability to learn new correspondences on a group level. This is in contrast to a recent study by Guerra et al. (In revision), who reported more shallow learning trajectories in children with dyslexia compared to typical readers. However, post-hoc tests in the current study suggested that children with dyslexia and typical readers slightly differed in their learning trajectories, as typical readers already became more accurate in the early trials of the task (i.e., bin 1 vs. bin 2) whereas children with dyslexia needed more time to reach a similar accuracy level (i.e., most improvement from bin 2 to bin 3). In addition, typical readers were slightly younger than children with dyslexia. As we found that older children were, on average, more accurate during the learning task than younger children, it is not unlikely that age-matched TR would have outperformed DR. This age difference warrants consideration in future research.

In an adaptive learning task, Karipidis et al. (2018) reported that children with dyslexia needed more time to learn six L-SS correspondences compared to typical readers. Interestingly, in the study of Guerra et al. (In revision), children with dyslexia responded less accurately compared to typical reading peers in the last two-thirds of each block of the learning task. Their task consisted of only 48 trials per block compared to the 72 trials per block in the current study. It might, therefore, be possible that due to the more extended design in the current study, children with dyslexia caught up with their peers. These results corroborate the notion that children with dyslexia are not necessarily impaired in initially setting up the associations but rather in storing them as automated, audiovisual objects (Blomert, 2011; Fraga González et al., 2015).

Concerning response latencies, the most improvement appeared during the first block. More specifically, children became faster from bin 1 to bin 2 but then remained relatively stable. Although this was true for children with dyslexia and typical readers, instructions had a differential effect on the two groups. Implicitly instructed DR responded slightly slower in the first trials of the first block, but then became as fast as their goal-directed peers. For typical readers, goal-directed children responded slightly slower in the first trials of the first block but then became as fast as their implicitly instructed peers. As proposed by Blomert and Vaessen (2009), faster responding indicates better integration of multisensory objects. However, response latencies tend to decrease steadily until Grade 6, despite accuracy performance reaching ceiling levels (Blomert & Vaessen, 2009). Differences in response times depending on instructions in the current study may be attributed to various reasons. First, directing attention toward the goal of the task might increase the workload by paying attention to the low-level features of the symbols that need to be learned, leading to slower responses (Paas & Van Merriënboer, 1994). In contrast, uncertainty about what to focus on might pressure one's self-beliefs, potentially leading to slower responses (Edgar et al., 2019). Further investigations should illuminate the underlying processes at work.

In contrast to Aravena et al. (2013), children with and without dyslexia did not differ in applying this new knowledge in time-limited reading tasks within the artificial orthography. It is essential to note that the learning task Aravena et al. (2013) used comprised two blocks of 30 minutes each. After 60 minutes of learning, 71.4% of their sample mastered the unknown Hebrew symbols. In our study, 60% of the sample learned more than half of the eight symbols, but only 24% of learners obtained full mastery. In addition, their explicit condition was preceded by an introduction of all correspondences before starting the learning task, and symbols were linked to words that were represented by symbols, in line with how letters are introduced at schools. Given that instructions in the current study did not follow this systematic phonics approach and that the learning blocks were shorter, it is highly likely that letter-speech sound correspondences were not yet integrated as audiovisual objects, which is a prerequisite for fluent decoding. As we wanted to keep the number of trials consistent across participants for the EEG part of this study, we used a fixed design instead of an adaptive task. An adaptive design offers the advantage of identifying and repeating individual errors such that more difficult mappings can be rehearsed until a certain accuracy level is obtained before continuing with the subsequent reading tasks. Future studies should examine how employing such an adaptive design can be used in studies dependent on a specific number of trials to obtain sufficient data quality, such as in EEG research.

Children, on average, responded faster to congruent trials than incongruent trials, but we did not find significant differences between children with dyslexia and typical readers in either accuracy or reaction time during the congruency task. This is in contrast to Guerra et al. (In revision), who reported differences in accuracy, which the authors interpreted as difficulties in learners with dyslexia with applying the learned information in a new context. Similar to the current study, children first had to identify the corresponding symbol to the presented speech sound between two answer alternatives and then had to apply this knowledge to indicate whether a presented letter-speech sound pair was congruent or incongruent. It should be noted that Guerra et al. (In revision) used an asynchronous task, in which the auditory stimulus was only presented after the visual stimulus disappeared, whereas in the current study these were presented simultaneously. The asynchronous presentation might have placed higher demands on their working memory, which has been found to be impaired in dyslexic readers (Arrington et al., 2014; Breznitz, 2002; Swanson et al., 2009).

Interestingly, attention measures were significantly associated with learning outcomes of the learning task. More specifically, children with higher attention factor scores (including inattention, impulsivity, and hyperactivity) obtained lower learning performance. Likewise, attention measures were negatively related to reading performance in the native language. These results suggest that attentional difficulties are not only related to reduced reading performance but may also limit the acquisition and development of alphabetic knowledge. This gives support to the multiple deficit model in which dyslexia is not caused by one, single deficit but by an interaction of multiple risk and protective factors (Pennington, 2006; van Bergen, van der Leij, et al., 2014; Willcutt et al., 2010). Providing targeted interventions that take into account this multidimensionality, such as including executive function training along with phonics instruction, might increase intervention outcomes.

Processing feedback in the brain

The lack of a consistent pattern of reading errors in individuals with dyslexia might interfere with their ability to learn from previous mistakes. Considerable research has proposed a performance monitoring deficit in individuals with dyslexia, typically assessed with error-related negativity potentials (Horowitz-Kraus, 2016a; Horowitz-Kraus & Breznitz, 2008, 2013; Van De Voorde et al., 2010). On a behavioral level, one would expect learners to respond slower after committing an error, referred to as post-error slowing (Smulders et al., 2016). We found that children, on average, responded slower after an error than after a correct response, but we did not find a main effect of either condition or reading group, suggesting that children with dyslexia, on average, were not impaired in detecting this .error.

Kraus & Horowitz-Kraus (2014) reported diminished brain activation following negative feedback in adolescents with dyslexia compared to their typically reading peers, especially in the early trials of a Wisconsin Card Matching Task (WCMT). The present study found no overall group differences between children with dyslexia and typical readers concerning peak-to-peak FRN amplitudes. It is essential to note that the WCMT tackles different cognitive mechanisms than a two-choice learning task, which might explain differences in results. The task specifically targets set shifting, as participants must apply a particular sorting rule and then switch flexibly to a new rule contingent upon the experimenter's feedback. The WCMT does not include specific stimulus-task associations, contrary to the task used in the current study. Negative feedback in the WCMT has to be followed by a series of processes; inhibiting the old response, identifying a new sorting rule, comparing the experimenter's feedback to the information stored in working memory, and so on (Kohli & Kaur, 2006). A deterministic two-choice learning task might in that sense be more straightforward as correct answer alternatives are provided.

When considering instruction conditions, DR and TR showed a different pattern. In children with dyslexia, peak-to-peak amplitudes were higher in children who obtained goal-directed instructions compared to those who obtained implicit instructions. In typical readers, we found the opposite effect: higher peak-to-peak amplitudes in implicitly instructed learners compared to those who obtained goal-directed instructions. The variety of methods used to define the FRN complicates the interpretation of results. While most studies used the most negative peak in a specified time window or the peak-to-peak amplitude with the preceding P2 peak (e.g., Kraus & Horowitz-Kraus, 2014; Roos et al., 2015), others used factors scores of principal component analyses or calculated area measurements (Arbel & Fox, 2021; Bellebaum & Daum, 2008). If FRN amplitudes reflect the number of cognitive resources allocated to the feedback, as proposed by Ferdinand & Kray (2014), we would expect that higher amplitudes represent better learning, suggesting improved feedback processing in children with dyslexia when obtaining goal-directed instructions. Lower amplitudes might represent either less efficient processing of feedback, or an earlier shift toward internal mechanisms (Holroyd & Coles, 2002). In good learners, we would, therefore, expect that FRN amplitudes rapidly decrease toward later trials of the task. In the context of the learning task used in the current study, the commission of an error is likely to be internally detected when associations between letters and speech sounds become integrated. As a consequence, external feedback does not carry much information, yielding lower amplitudes.

We found that, on average, peak-to-peak FRN amplitudes were higher in early trials of our learning task than in late trials. This aligns with the reinforcement learning theory that suggests that amplitudes are higher following unexpected feedback and decrease with learning as there is less need to attend to expected outcomes (Holroyd & Coles, 2002). However, this trend was the most apparent when children obtained goal-directed instructions. That is, amplitudes decreased in later trials when children obtained goal-directed instructions, but did not decrease significantly when they obtained implicit instructions. This was also in line with their behavioral learning trajectories that showed that implicit learners especially became more accurate during the early trials, whereas goal-directed learners still increased in later trials as well.

Peak-to-peak amplitudes were, on average, higher following negative feedback stimuli compared to positive feedback stimuli. This was true for both DR and TR, but in children with dyslexia this was especially true in younger children. This aligns with neuroimaging studies that typically reported more negative amplitudes after negative feedback (Arbel & Fox, 2021; Ferdinand et al., 2016), indicating that our monitoring systems may be more sensitive to feedback when it hinders the current task goal (Hämmerer et al., 2011). However, developmental studies argued that stronger responses do not necessarily lead to better learning

(Ferdinand & Kray, 2014). As prefrontal brain areas responsible for feedback processing are still developing along with the brain's ability to extract information from negative feedback (Dion & Restrepo, 2016), children are less efficient in using feedback in adapting behavior accordingly (Ferdinand & Kray, 2014). This is why children, on average, need more trials to learn from feedback compared to adolescents and adults (Eppinger et al., 2009). Very few studies examined the association between the FRN and learning in children (but see e.g., Arbel & Fox, 2021; Groen et al., 2007; van der Helden et al., 2010). We found a small positive association between FRN and learning outcomes, suggesting that more positive FRN factor scores (i.e., less negative amplitudes), were associated with higher learning outcomes, but this effect did not survive FDR correction. A study by Arbel & Fox (2021) found that especially the FRN elicited by positive feedback was predictive of learning retention in children. More specifically, in younger children (*M* = 8.4 years), smaller FRN to positive feedback was related to better learning retention and in older children (*M* = 10.5), larger FRN was related to better retention. As we used a principal component analysis to reduce the dimensionality of our data, our results are based on factor scores combining neural responses following negative and positive feedback. As a result, whether higher amplitudes to either positive or negative feedback contribute to more efficient learning in an artificial learning context is yet to be examined.

Influence of instructions

Previous findings using a similar behavioral learning paradigm in a school-aged sample indicated a significant impact of goal-directed instructions on the learning trajectory and the subsequent consolidation of the new knowledge (Verwimp et al., 2023). In the present study, we did not find a main effect of instructions on the learning trajectories, symbol knowledge or application of the new symbols in a time-limited reading task. However, it appeared that implicitly instructed children with dyslexia only increased in accuracy in the early trials, but then stagnated around an accuracy level of 55%, whereas in the goal-directed condition their performance further increased in accuracy. This pattern is in contrast with typical readers, as goal-directed and implicitly instructed typical readers showed similar learning patterns, although goal-directed instructed children seemed to learn faster in the second block. In our previous study (Verwimp et al., 2023), the goal-directed group obtained an accuracy level of 68.70% at the end of block 1 (*SD* = 16.72), compared to 61.50% (*SD* = 19.00) in the current study (averaged for DR and TR). The implicit group obtained an average accuracy level of 56.32 at the end of block 1 (*SD* = 19.95), compared to 59.50% (*SD* = 15.50) in the current study. Goal-directed children thus performed worse whereas the implicit group performed slightly better compared to our previous study, yielding no significant main effect of condition. These contradictory findings might be explained by the lab environment, as children were found to perform better in more comfortable and familiar settings. Goal-directed instructions are thought to activate selective, goal-directed behavior. Uncomfortable situations, like wearing an EEG cap in a university lab, might have interfered with the hypothesized effect of goal-directed instructions. This was previously reported by Romanovska et al. (2021), who found reduced behavioral effects in the MRI scanner compared to results outside the scanner. The authors argued that contextual factors might especially reduce attentional task focus in children who are still developing their attentional skills. Another reason for contradictory findings in the current study compared to previous studies might be that some children might have been less motivated to obtain the goal, as the learning task was part of a larger test session that lasted approximately 3 hours. As negative feedback has a more aversive effect on children than on adolescents or adults, and especially slow learners obtain much negative feedback, it can demotivate them and affect their attention on the task (Lukie et al., 2014; Peters et al., 2014).

Clinical implications and future directions

Children are exposed to different forms of feedback daily, at school and during most clinical interventions. Our findings suggest that, although we failed to find any main effects of reading group, children with dyslexia appear to follow a different trajectory toward audiovisual learning. Therefore, school and clinical treatment programs should be aware of individual differences in learning ability and internal and external feedback processes, instead of employing a one-fits-all approach. In addition, previous studies have shown that the P300, a stimulus-locked electrophysiological response, is associated with adjustments in learning behavior (Arbel & Fox, 2021). Therefore, it might be interesting to examine the P300 in a trial-by-trial analysis, that is, whether higher P300 amplitudes lead to corrective behavior in the next trial.

Although we specifically aimed to have a heterogenous sample, better representing the heterogeneity reported in individuals with dyslexia, this might obscure relevant differences in the current study. For example, Roos et al. (2015) reported differential associations between FRN amplitudes and task performance in children with low impulsivity compared to children with high impulsivity. The current study found substantial variability within learning trajectories and outcomes. Therefore, examining individual learning indices on top of performing group analyses is highly relevant. Future studies should further investigate individual differences in audiovisual learning in relation to differences in cognitive control processes, as this is highly relevant to understand and remediate the potential L-SS integration deficit reported in individuals with dyslexia.

Conclusion

Our study did not show an overall letter-speech sound learning deficit in children with dyslexia. However, we found substantial inter-individual learning differences in both children with dyslexia and typical readers. This nicely aligns with the behavioral heterogeneity reported in reading acquisition and development. Accordingly, while some might suffer from a binding deficit, this is not necessarily true for all dyslexic readers. In addition, our findings suggest a differential influence of instructions on behavioral and neural outcomes in individual with dyslexia and typical readers. As learning outcomes were related to reading ability within the native language, such an artificial learning paradigm allows us to further examine factors contributing to the transition from letter-speech sound integration to consolidated audiovisual objects. Further research should consider individual learning indices to clarify how cognitive control mechanisms contribute to the reading development, as this is highly relevant to understand and remediate reading difficulties.

5

Reciprocal associations between letter-speech sound integration and word reading fluency during dyslexia intervention

Abstract

Specialized reading interventions are effective in improving reading and spelling skills. However, despite intervention efforts, a significant proportion of children remains dysfluent in reading, indicating significant inter-individual differences in intervention response. To better understand how current reading treatment interacts with enhanced reading performance, we used a path analysis to investigate the reciprocal associations between letter-speech sound (L-SS) integration, i.e., the first step of learning to read and the foundation of current reading treatments, and word reading fluency during dyslexia intervention in 3676 Dutch school-aged children. Results showed that changes in L-SS integration were associated with changes in word reading fluency. Accurate L-SS mapping was the most predictive at the start of the intervention. Fluent, automatized L-SS mapping was the most predictive toward the end, which might be associated with the developmental shift from effortful to fluent decoding. Especially individuals with poor L-SS integration at the beginning of the intervention benefitted the most, suggesting that children whose reading difficulties were not associated with poor L-SS integration might need a slightly different approach to intervene on their reading problems. Processing speed played a limited role in this intervention, as only a small negative association with word reading fluency appeared. Current findings provide valuable insights into underlying mechanisms of reading development and have implications for developing effective interventions for individuals with dyslexia.

Reciprocal associations between letter-speech sound integration and word reading fluency during dyslexia intervention

In a world where literacy is more important than ever, being unable to read fluently comes with significant challenges. Around 5-10% of the population is estimated to remain impaired in reading throughout life, referred to as developmental dyslexia (Snowling, 2013). Specialized reading interventions have been found to be effective in improving reading and spelling skills. However, significant inter-individual differences in intervention response have been reported, with a significant proportion of children remaining dysfluent in reading, even after intervention (Galuschka et al., 2014). To better understand these differences in intervention response and adapt the intervention to individual needs, we need to gain better insights into the mechanisms of change that drive intervention-induced reading gains. This study used a structural equation modeling approach to examine the reciprocal associations between letter-speech sound integration, i.e., the first step of learning to read and the foundation of current reading treatments, and word reading fluency during dyslexia intervention in an extensive Dutch sample. Letter-speech sound (L-SS) integration and word reading fluency were assessed at multiple time points throughout the intervention, making it possible to examine how changes in letter-speech sound integration are associated with changes in word reading fluency.

To develop efficient word-recognition skills, one must learn which written symbols correspond to which spoken language representations (Castles et al., 2018). The acquisition of these associations is the initial step of learning to read and takes about one year of formal instruction in more transparent languages such as Dutch (Froyen et al., 2009). However, knowing how letters map onto speech sounds is insufficient to achieve fluent reading (Blomert, 2011; Ehri, 2005). For fluent reading to occur, letters and speech sounds must be integrated into fully automatized, audiovisual objects (Blomert, 2011). This gradual increase toward automatization is accompanied by re-shaping of innate auditory and visual brain areas into an integrated audiovisual reading network (Breznitz, 2002; Romanovska & Bonte, 2021), requiring much exposure and practice. The activity in this network is found to be reduced, or developmentally delayed, in individuals with dyslexia (Blau et al., 2010; Blomert, 2011; Žarić et al., 2014), suggesting that a lack of automatic neural sensitivity for L-SS pairs may contribute to persistent dysfluent reading.

Based on the L-SS automatization deficit commonly reported in individuals with dyslexia, targeted reading intervention programs have been developed. The most effective intervention methods comprise systematic instruction of letter-speech sound correspondences along with decoding strategies (Galuschka et al., 2014). After obtaining explicit instructions, the associations are systematically repeated to enhance and integrate the L-SS associations through exposure (Froyen et al., 2009; Xue et al., 2006). Because this training gradually increases the complexity of the L-SS associations over a year, the associations can become automatized, eventually leading to more fluent reading (Fraga González et al., 2015).

Despite the importance of L-SS automatization in fluent reading, the extent to which letter-speech sound integration predicts intervention progress appears less clear-cut. This is expected because intervention progress is commonly measured by progress in specific reading or spelling measures in which progress is the most noticeable, such as how accurately one reads. Letter-speech sound strength is usually only measured during diagnostic assessment, and when measured multiple times during the intervention it is commonly examined in relatively small sample sizes leading to low power to detect subtle changes during the intervention period (e.g., Žarić et al., 2015). As a result, it remains to be examined how changes in letter-speech sound integration relate to improved reading fluency.

To fill this gap, the present study aimed to create a detailed window concerning the development of L-SS associations within the treatment of dyslexia. We employed a structural equation modeling approach using an archival dataset acquired from a reading intervention program in the Netherlands for children diagnosed with dyslexia. Letter-speech sound integration strength and word reading fluency were measured at three time points during the intervention, allowing us to examine the reciprocal relation between these two variables. We had two main research questions: I) How does L-SS integration growth relate to word fluency growth? And II) How does processing speed measured at T0 contribute to word fluency growth? Based on the importance of L-SS integration in fluent reading, we hypothesized that an increase in L-SS integration would be positively associated with increased word reading fluency. In addition, we expected that children with high L-SS scores at T0 would benefit less from the intervention. Last, since processing speed is thought to facilitate the acquisition of new information (Kyllonen et al., 1991), we expected that processing speed would be positively associated with L-SS integration growth and word fluency growth.

Method

Participants and procedure

An extensive clinical database (*N* = 13,793) was collected from 2014-2020 at a clinical center for learning disabilities in the Netherlands. All children experienced severe reading difficulties at school and resisted additional remedial support and were, therefore, referred to the clinical center for diagnostic screening for dyslexia. Participants were included based on the following criteria: I) obtained formal reading instruction at school, and therefore only children older than 72 months were included; II) intelligence score higher than or equal to 80 as assessed with the WISC-III (Kort et al., 2002); III) severe and persistent reading problems, defined as at least 1.5 standard deviations below the average on a standardized reading test; IV) children had received at least 15 intervention sessions halfway through the treatment program and 30 sessions at the end of the program. Last, children with unreliable scores, i.e., scores that fell out of their score range, were removed from the sample. The final sample comprised 3676 children aged between 76 and 150 months (*M* = 103.64, *SD* = 13.54; 1713 girls (46.60%)). Children had received an average of 18.19 intervention sessions halfway through the program (*SD* = 2.37) and an average of 36.60 sessions at the end of the program (*SD* = 4.66). All participating children were native Dutch speakers who had normal or corrected to normal vision and hearing and went to regular elementary schools in the Netherlands. This study obtained ethical approval from the ethics committee of the university (no. 2016-DP-7127) and all parents actively consented to use the clinical data for research purposes.

Intervention

Children followed an intensive computer-based intervention program provided by a psychologist who was specialized in dyslexia treatment. The intervention consisted of a weekly 45-minute individual training session and four to five 15-minute homework sessions, which both lasted for approximately 40 weeks.

The intervention was constructed based on the principles of general skill acquisition (Chein & Schneider, 2012; Davydov, 1995). Each (letter-speech sound) element was taught explicitly at first, followed by intensive repetition in order to obtain a transition from accurate, controlled to associative, automatic processing. Sessions consisted of two main components: an instructional and a practice part. During the instructional part, explicit teaching of the letter-speech sound correspondences took place with the goal to achieve accurate mastery of the learned associations. During the practice part, computer-based training was used to provide extensive exposure to the specific letter speech sound associations that were taught during the instructional part. This high exposure aimed at stimulating the automatic integration of letters and speech sounds (Fraga González et al., 2015; Tijms, 2011).

The program started with consistent mappings between phonemes and their corresponding graphemes, followed by progressively more complex and inconsistent mappings. For this purpose, a reconfigured touchscreen was used consisting of icons for each

Dutch speech sound. Each icon showed the standard letter or letter-cluster of the corresponding speech sound. In addition, the touchscreen included several icons to indicate the type of phoneme (e.g., ‘long vowel’), syllable icons (e.g., ‘stressed syllable’), and rule icons to perform an operation (e.g., delete a selected grapheme) in the case of inconsistent mappings (Fraga González et al., 2015; Tijms, 2011).

The practice of these consistent and inconsistent letter-speech sound mappings was also done in the context of reading and spelling exercises. For example, individual words were presented on the computer screen and the child was instructed to pronounce the word sound by sound until they could pronounce the complete word. Practice was adjusted to the individual acquisition rate by adapting the presentation rate to the child’s performance level. Children continued with the next phase of the training when a minimum of 80% of the items was executed correctly. For a more detailed description of this intervention program, we refer to Tijms (2007) or Fraga González et al. (2015).

Outcome measures

Five subtests of the 3DM Differential Diagnostics for Dyslexia were individually administered at the clinical center to assess children’s reading and spelling skills (Blomert & Vaessen, 2009) at three time points: before the intervention (T0), halfway through the intervention (T1), and at the end of the intervention period (T2). These tests have been used to construct three outcome measures: L-SS integration, word reading fluency, and processing speed.

L-SS integration

Three 3DM subtests were used as a proxy of L-SS integration: L-SS identification, L-SS discrimination, and computerized spelling (Blomert & Vaessen, 2009). For the L-SS identification task, a speech sound was auditorily presented, after which the child was required to match the sound to one of the four presented letters or letter combinations by pressing the corresponding button. Internal consistency of the accuracy scores varied between .59 and .75, depending on grade (see Blomert & Vaessen, 2009). Split-half coefficients for the speed scores varied between .77 and .86. For the L-SS discrimination task, children were required to decide whether a presented speech sound was congruent or incongruent with a simultaneously presented letter (e.g.., /a/ and ‘a’ versus /a/ and ‘e’). Internal consistency of the accuracy scores varied between .79 and .89, and split-half coefficients for the speed scores varied between .76 and .89, depending on the grade. During the computerized spelling task, a word was aurally presented while a part of the word was visually presented on the computer screen. Children were instructed to fill in the missing part by choosing one out of four presented letters or letter combinations by pressing the corresponding button. Internal consistency of accuracy scores varied between .75 and .82, split-half reliability of the speed scores varied between .83 and .90. For the three tasks, accuracy (% correct) and speed (sec/item) were calculated. Based on these three subtests, two L-SS integration factor scores were computed, one for accuracy and one for speed. Factor loadings are reported in Table S1 (Appendix D).

Word reading fluency

This subtest consisted of three subtasks: high-frequency words, low-frequency words, and pseudo-words. For each subtask, 75 words were visually presented on five successive screens with 15 items each. Each subtask started with monosyllabic words and gradually built up to more complex four-syllable words. The experimenter pressed a button when children finished reading a screen such that they could continue with the next one, with the goal of reading as many words as possible within a time limit of 30 seconds. The sum of the number of correctly read words in the three subtasks was used as a measure for word reading fluency, with a maximum of 225. Test–retest reliability coefficients varied between .87 and .94 (Blomert & Vaessen, 2009).

Processing speed

The basic reaction time task of the 3DM was used as a measure of processing speed. The screen was divided into four squares, and in one, an animated figure appeared. The goal was to press the corresponding button of the square with the figure as fast as possible. The score was calculated as the mean reaction times of all correct trials. Split-half coefficients of the task varied between .84 and .91, depending on grade (Blomert & Vaessen, 2009).

Statistical analysis

To create a detailed overview of the reciprocal associations between letter-speech sound integration and the development of word reading, we conducted a path analysis with Amos in IBM SPSS Statistics (Version 26). To find the best-fitting model, we followed a stepwise approach. First, four nested models ranging from most elementary to most complex were estimated using our outcome variables of interest (letter-speech sound integration (accuracy and speed), word reading fluency and processing speed), with age and IQ as covariates. Model 1 only included word reading fluency (at T0, T1, and T2) and the associations with age and IQ. For model 2, the variables L-SS speed and L-SS accuracy were added, including only direct associations with word reading fluency. In model 3, associations between L-SS speed and L-SS accuracy measures were also added. Last, in model 4, processing speed was added with connections to word reading fluency and both L-SS measures. A combination of fit indices was used to select the superior model; root mean square error of approximation (RMSEA), Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), and Normed Fit Index (NFI). For the first three, lower values indicated a better fit (Schermelleh-Engel et al., 2003). RMSEA values lower than 0.05 were qualified as 'very good fit', 0.05-0.08 as 'good fit', 0.08-0.10 as 'mediocre fit', and > 0.10 as 'poor fit'. For NFI, higher values indicated a better fit, with values > 0.95 considered 'very good fit', 0.90-0.95 'good fit', 0.80-0.90 'mediocre fit' and < 0.80 'poor fit'. In addition, a comparative analysis was used to compare the model's χ^2 and degrees of freedom to the most complex model (Model 4), with significance indicating evidence in favor of the more complex model. To compute a χ^2 difference test, the different models must be nested, meaning that models can transform into each other by adding or eliminating paths (Schermelleh-Engel et al., 2003). Therefore, Model 4 was used as a parent model, and the different paths were constrained to obtain Models 1, 2, and 3. Differences in degrees of freedom, χ^2, and their significance are reported in Table 2. Cramer's V was reported to interpret the size of the difference between models, with a higher value indicating a more significant difference.

Second, to investigate whether our covariates age and IQ had any additional explanatory value in the model, we estimated four models using the superior model of the previous step. More specifically, the first model included both covariates, the second included only age as a covariate, the third included only IQ as a covariate, and the fourth included no covariates. RMSEA-CFI, NFI, AIC, and BIC were used to select the best-fitting model given the data.

Once we identified the best model, we conducted a path analysis to examine the predictive values of processing speed and L-SS speed and accuracy on word reading fluency before the L-SS training (T0), during the training (T1), and after the training (T2). A significance level of 0.001 was used for all tests due to the large number of participants (Shreffler & Huecker, 2022).

Results

Descriptive characteristics

The descriptive characteristics of the sample are presented in Table 1. Outcome measures were assessed before treatment (T0), halfway through the treatment (T1), and after treatment (T2). All variables, except age and IQ, were age-normed T-scores (*M* = 50, *SD* = 10). As expected, the average word reading fluency score before the intervention was poor, on average two standard deviations below the mean. In addition, all L-SS measures were below average as well. A paired sample t-test revealed a significant improvement in children's word reading fluency scores during the treatment, providing evidence for the effectiveness of the intervention (*t*(3638) = -73.29, *p* < .001, Cohen's d = 1.22).

Table 1

Means and Standard Deviations of Variables of Interest

	T0		T1		T2	
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
Age	3676	103.64 (13.54)	3676	109.25 (13.54)	3676	115.76 (13.63)
IQ	3676	101.91 (11.58)				
PS	3665	59.64 (10.62)				
N sessions			3676	18.19 (2.37)	3676	36.60 (4.66)
WRF	3667	30.90 (5.66)	3666	35.57 (6.74)	3647	37.98 (7.72)
L-SS-D A	3667	42.39 (9.61)	3669	51.27 (9.17)	3465	52.92 (9.06)
L-SS-I A	3667	40.19 (11.27)	3667	49.85 (9.75)	3468	51.08 (9.33)
SP A	3665	37.11 (8.15)	3669	40.89 (8.72)	3651	46.61 (9.10)
L-SS-D S	3667	43.08 (10.11)	3665	49.68 (10.20)	3463	50.95 (10.39)
L-SS-I S	3667	40.92 (9.72)	3663	49.90 (10.31)	3466	50.81 (10.62)
SP S	3665	36.44 (9.05)	3665	37.99 (9.02)	3649	38.88 (10.24)

Note. N = number of participants; PS = processing speed; N sessions = number of attended intervention sessions; WRF = word reading fluency; L-SS-D A = letter-speech sound discrimination – accuracy; L-SS-I A = letter-speech sound identification – accuracy; SP A = computerized spelling – accuracy; L-SS-D S = letter-speech sound discrimination – speed; L-SS-I S = letter-speech sound identification – speed; SP S = computerized spelling – speed. Tasks were assessed before intervention (T0), halfway through the intervention (T1) and at the end of the intervention (T2).

Model comparisons

First, four nested models were estimated and compared based on a combination of fit indices. Model 1 ($\chi2(48) = 7703.637$, $p < .001$) only contained associations between word reading fluency outcomes at the different time points and age and IQ as covariates. Fit indices (see Table 2) indicated poor fit given the data. In Model 2 ($\chi2(32) = 752.784$, $p < .001$), we added letter-speech sound speed and accuracy to the model, with only associations between letter-speech outcomes and word reading fluency. This model had a good fit based on the RMSEA and a very good fit based on the NFI. Model 3 ($\chi^2(24) = 441.369$, $p < .001$), which contained associations between letter-speech sound speed and accuracy as well, had a good fit based on the RMSEA and very good fit based on the NFI. Finally, fit indices of Model 4 ($\chi^2(22) = 398.782$, $p < .001$), which included processing speed, indicated a good and very good fit given the RMSEA and NFI, respectively. BIC and AIC values were lower for Model 4 than for the other three models. Second, each model was compared to the most complex one (Model 4) with a χ^2 difference test. Each comparison was significant, indicating that Model 4 was superior to all other models (see Table 2). However, it is important to note that Cramer`s V associated with the comparison between model 3 and model 4 is small, suggesting that the contribution of processing speed to the model is, although significant, relatively small.

New models were constructed to determine the added value of the covariates age and IQ in this model. Model 4A was the original model, including both covariates, Model 4B included only age as a covariate, Model 4C included only IQ as a covariate, and Model 4D included no covariates. As the models were non-nested, only the fit indices were used to determine the fit. Model 4A ($\chi^2(22) = 398.782$, $p < .001$) had a good fit based on the RMSEA and very good fit based on the NFI. Fit indices of Model 4B ($\chi^2(16) = 335.443$, $p < .001$), 4C ($\chi2$ $(16) = 174.793$, $p < .001$), and 4D ($\chi2$ $(10) = 133.985$, $p < .001$) also indicated a good fit based on the RMSEA and a very good fit based on the NFI. In addition, BIC and AIC values were the lowest for Model 4D, indicating that the model without age and IQ was the most appropriate and, therefore, used to estimate the path coefficients.

Table 2

Fit Indices of all Estimated Models

	Fit indices							Model comparison				
Model	**χ^2**	**df**	***p***	**RMSEA**	**NFI**	**AIC**	**BIC**	**Comparison**	**$\Delta\chi^2$**	**Δdf**	***p***	**Cramer's V**
1	7703.637	48	< .001	0.208	0.540	7787.637	7787.935	M4 vs. M1	7304.855	26	< .001	0.997
2	752.784	32	< .001	0.078	0.955	868.784	869.196	M4 vs. M2	354.002	20	< .001	0.219
3	441.369	24	< .001	0.069	0.974	573.369	573.838	M4 vs. M3	42.587	2	< .001	0.076
4	398.782	22	< .001	0.068	0.976	534.782	535.265	-	-	-	-	-
4A	398.782	22	< .001	0.068	0.976	534.782	535.265					
4B	335.443	16	< .001	0.074	0.979	457.443	457.842					
4C	174.793	16	< .001	0.052	0.989	296.793	297.192					
4D	**133.985**	**10**	**< .001**	**0.058**	**0.991**	**243.985**	**244.315**					

Note. χ^2 = chi-square; df = degrees of freedom; RMSEA = Root mean square error of approximation; NFI = Normed fit index; AIC = Akaike information criterion; BIC = Bayesian information criterion. First, models 1 to 4 were estimated with increasing complexity. Models were compared to the most complex one (Model 4) with a χ^2 difference test. Second, models 4A to 4D were estimated to examine whether age and IQ contributed significantly to the model fit. The best-fitting model is indicated in bold

Correlations and predictive values

First, we calculated the correlations between our outcomes of interest at the start of the intervention. The correlation between letter-speech sound accuracy and word reading fluency was medium ($r = 0.294$, $p < .001$), and the correlation between letter-speech sound speed and reading fluency was small ($r = 0.194$, $p < .001$). In addition, processing speed was moderately associated with letter-speech sound speed ($r = 0.396$, $p < .001$), but the correlations between processing speed and L-SS accuracy and word reading fluency were very small with the former being non-significant ($r = 0.014$, $p = .397$, and $r = 0.091$, $p < .001$, respectively).

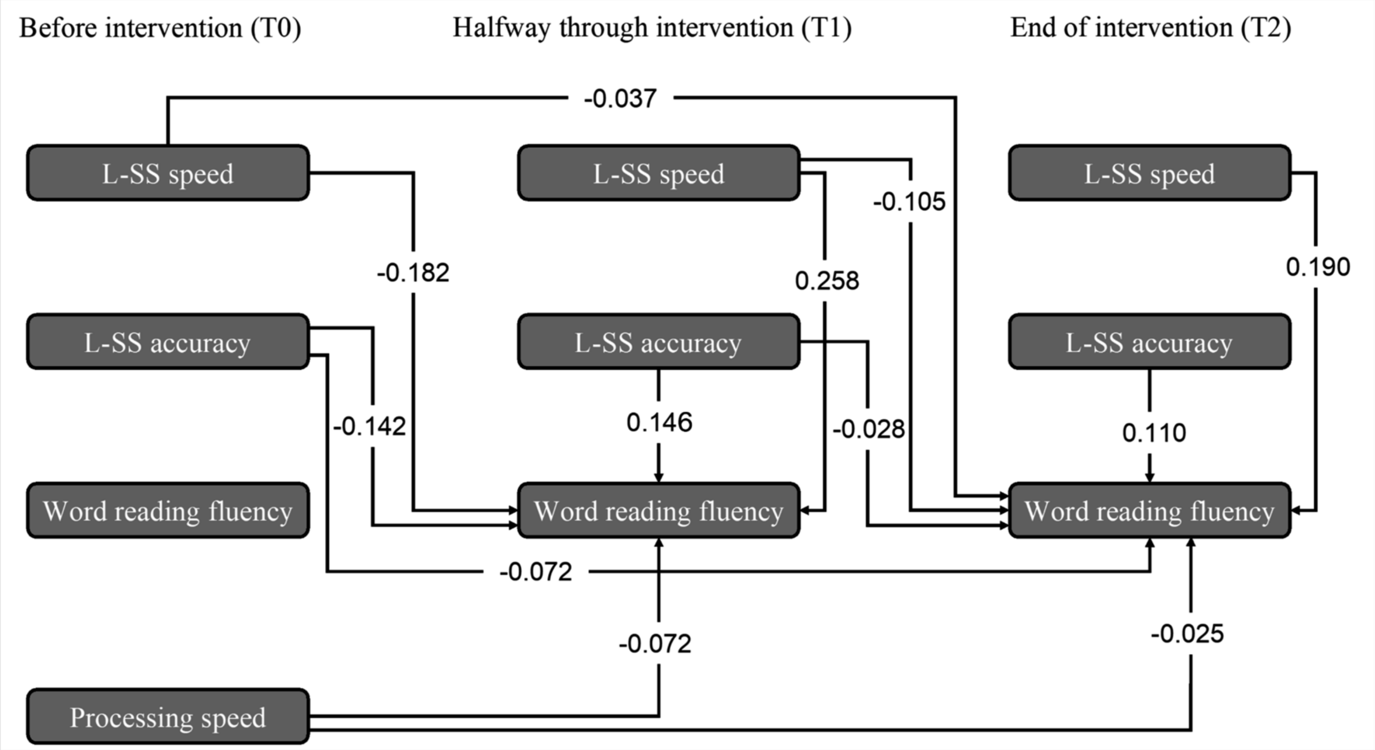
A path analysis was conducted to obtain estimates for all paths of the superior model (4D). The complete list of coefficients is presented in Table S2 (Appendix D), with the most important paths regarding our research questions discussed here (see Figure 1). Letter-speech sound accuracy and letter-speech sound speed at T0 were both negatively associated with word reading fluency at T1 (after controlling for associations with word reading fluency at T0), with standardized estimates of -0.142 and -0.182, respectively (*p*s < .001). Interestingly, our letter-speech sound measures at T0 became less predictive for word reading fluency at the end of the intervention (T2), although still significant (standardized estimates of -0.072 for accuracy and -0.037 for speed). That is, children with lower letter-speech sound integration scores before the intervention improved more on word reading during both the first half and the second half of the intervention sessions.

Letter-speech sound accuracy and speed at T1 were positively associated with word reading fluency at T1 ($\beta = 0.146$ and $\beta = 0.258$ for accuracy and speed, respectively). That is, children who increased in letter-speech sound accuracy and speed also increased in word reading fluency. A similar result was found for letter-speech sound measures and word reading fluency outcomes at T2. That is, letter-speech sound accuracy and speed at T2 were positively associated with word reading fluency at T2 ($\beta = 0.110$ and $\beta = 0.190$ for accuracy and speed, respectively). Letter-speech sound speed at T1 significantly predicted word reading fluency at T2 ($\beta = -0.105$, $p < .001$), but the coefficient of letter-speech sound accuracy at T1 on word reading fluency at T2 was non-significant ($\beta = -0.028$, $p = .018$).

Last, processing speed, only measured at T0, was negatively associated with word reading fluency at T1 ($\beta = -0.072$). However, the association between processing speed and word reading fluency at T2 was non-significant ($\beta = -0.025$, $p = .006$).

Figure 1

Overview of the Most Important Paths Regarding our Research Questions



Note. The displayed values are standardized estimates. A complete list of standardized and unstandardized path estimates is presented in Appendix D Table S2. L-SS = letter-speech sound

Discussion

Although specialized reading interventions effectively improve reading and spelling skills, children widely differ in their intervention response, with some remaining dysfluent after treatment (Galuschka et al., 2014; Torgesen et al., 2001). Instead of only focusing on the efficacy of a specific intervention program, it is of great importance to study why and for whom treatment works (Kazdin, 2004). To better understand and address these individual differences in susceptibility to intervention, more detailed insights in the intervention mechanisms of change are needed. The strength of the cross-modal integration of letters and speech sounds is repeatedly reported as an essential factor in reading fluency development and reading deficits (Blomert, 2011). Accordingly, an important part of reading intervention is focused on developing integrated letter-speech sound mappings. We, therefore, examined the dynamics of change in reading fluency vis-à-vis letter-speech sound processing during intervention in the current study. We employed a path analysis in 3676 Dutch elementary-school children. Letter-speech sound integration and word reading fluency were assessed before the intervention, halfway through the intervention, and at the end of the intervention, allowing us to examine how letter-speech sound integration and word reading fluency were associated throughout a dyslexia intervention program.

First, we found that children, on average, improved their positions within the distribution of reading performance of an age-related normative sample, providing evidence for the effectiveness of the intervention. At the start of the intervention, word reading fluency was stronger associated with letter-speech sound accuracy than with letter-speech sound speed. Letter-speech sound speed reflects how fast one can process the speech sound and associate it with the corresponding letter combination, thought to be related to automaticity. In the developmental pathway toward integrated processing of letters and speech sounds, developing accurate mapping between letters and speech sounds is the initial phase that slowly shifts toward prolonged development of increasingly automated processing of the mappings (Vaessen & Blomert, 2010). As dysfluent reading is the most persistent characteristic of dyslexia in more transparent languages such as Dutch (de Jong & van der Leij, 2003; Struiksma et al., 2009), children in our sample might still be in a more initial stage of establishing accurate mappings. As a result, the controlled and accurate processing of L-SS mappings at the start of the intervention is likely to be more predictive of how many words a child can read than how fast one can process these mappings.

Letter-speech sound speed and accuracy before intervention were negatively associated with word reading fluency halfway through the intervention. That is, children with lower letter-speech sound integration scores before the intervention improved more on word reading during the first half of the intervention sessions. Letter-speech sound speed and accuracy at T0 became less predictive for word reading fluency at the end of the intervention, which might be associated with the finding that children grew the most on word reading fluency during the first half of the intervention. In addition, our results indicated that stronger gains in letter-speech sound speed and accuracy were associated with stronger gains in word reading, with letter-speech speed becoming more predictive for word reading fluency growth than letter-speech sound accuracy halfway through and at the end of the intervention. Our results thus align with the developmental shift observed in reading; word reading fluency was stronger associated with L-SS accuracy at the start of the intervention but becomes more associated with L-SS speed toward the end of the intervention.

As previous studies reported that processing speed is related to learning new skills (Kyllonen et al., 1991), we hypothesized that children's processing speed would predict their intervention progress. Processing speed was positively associated with letter-speech sound speed at the start of the intervention, but the association with word reading fluency was relatively small. In addition, regarding the influence of processing speed during the intervention, we only found a minimal association with reading fluency at T1. Interestingly, children in the current sample performed above average on processing speed, although previous studies reported a processing speed deficit in individuals with dyslexia (de Oliveira et al., 2014; Moll, Ramus, et al., 2014). The nature of the task can possibly explain contradictory results. Previous studies argued that children with dyslexia showed especially difficulties in processing verbal information (Moll, Ramus, et al., 2014), whereas we used a visual processing speed task. In addition, as suggested by de Oliveira et al. (2014), processing speed might be especially important in more complex reading tasks. For example, when reading and comprehending longer texts, it is essential to process a greater amount of information efficiently. Future studies should address how the results of the current study generalize to text reading and, ultimately, reading comprehension.

Interestingly, our findings align with those of a previous study on the same intervention program (Žarić et al., 2015). The authors used an oddball paradigm to study neural ERP markers of letter-speech sound integration at the start of intervention and after 20 sessions. Results showed changes toward faster neural letter-speech sound processing after 20 sessions that were correlated with higher letter-speech sound performance on behavioral measures (same as in our study). These results suggest that, although it may take years to fully develop the neural integration of letters and speech sounds (Froyen et al., 2009; Žarić et al., 2014), changes in the integration of letters and speech sounds that can be measured at the behavioral level are highly associated with changes at the neural level.

Although the current study provides important insights into underlying mechanisms of change during dyslexia intervention, some limitations must be considered. First, current reading interventions typically comprise decoding strategies along with phoneme awareness training (Galuschka et al., 2014). Phonological ability is one of the strongest predictors of reading development but also seems to develop with increased reading experience (Bishop, 2006; Boets et al., 2010). Although we specifically aimed to focus on letter-speech sound integration as an underlying mechanism for fluent reading improvement, it is interesting to examine reciprocal associations between phoneme awareness growth and L-SS integration throughout the intervention as well. Second, previous studies reported the influence of homework compliance on treatment outcomes (Mausbach et al., 2010). Despite instructing children to complete a strict number of sessions, there may have been variability in the number of sessions completed, potentially explaining individual differences in intervention outcomes.

In sum, our findings demonstrated that integrating letters and speech sounds is an important mechanism of change within dyslexia treatment. At the start of the intervention, letter-speech sound accuracy was the most predictive of word reading fluency, gradually being replaced by letter-speech sound speed throughout the intervention. Individuals with poor letter-speech sound integration at the beginning of the intervention benefitted the most, suggesting that children whose reading difficulties were not associated with poor L-SS might need a slightly different approach to target reading problems. This study contributed to a better understanding of the mechanisms of change within an intervention, so we can more optimally adapt interventions to individual needs.

6

The COVID generation: Online dyslexia treatment equally effective as face-to-face treatment in a Dutch sample

Abstract

Due to pandemic-induced lockdown(s) in 2020, dyslexia treatment was forced to move to online platforms. The purpose of this study was to examine whether Dutch children who received online treatment progressed as much in their reading and spelling performance as children who received the usual face-to-face treatment. To this end, 254 children who received treatment-as-usual were compared to 162 children who received online treatment with Bayesian methods. The advantage of a Bayesian approach is that it can provide evidence for and against the null hypothesis whereas frequentist approaches only provide evidence against it. We found that children in the online treatment condition received slightly fewer treatment sessions but progressed equally after controlling for the number of sessions compared to the treatment-as-usual condition. These results have clinical and practical implications as they show that reading treatment can be successfully delivered online.

The COVID generation: Online dyslexia treatment equally effective as face-to-face treatment in a Dutch sample

Developmental dyslexia, a neurodevelopmental disorder characterized by inaccurate and dysfluent reading, has prevalence rates of around 7% worldwide (Snowling, 2013). Much progress has been made in remediating these severe reading difficulties, with most children attaining reasonable levels of reading accuracy after an intensive treatment. However, as a result of the pandemic-induced lockdown(s) in 2020, this treatment had to shift to online platforms. As it is of great clinical interest to know whether children were able to have their needs met under these circumstances, the current study investigated whether children who received online reading treatment progressed as much on conventional reading and spelling measures as children who received the common face-to-face treatment.

Dyslexia treatment most commonly comprises systematic instruction of letter-speech sound correspondences and decoding strategies, along with applying these skills in reading and writing activities (Galuschka et al., 2014). These sessions are intense, systematic, and explicit, with an average length of 50 to 80 hours (Torgesen, 2005). Previous studies have shown that Dutch reading treatment programs can result in significant long-lasting improvements in children with reading difficulties (e.g., Tijms, 2007, 2011; Tilanus et al., 2016). During this phonics-based program, children are provided with weekly one-to-one treatment sessions at a clinical center in which intensive tutoring is followed by extensive practice. However, the physical distancing regulations as a result of the COVID-19 pandemic in 2020 forced health centers and schools to close for extended periods, making the weekly treatment sessions at the clinical center impossible. In order to still provide individuals with the needed support, health services in the Netherlands moved to online platforms. The major unplanned shift to tele-practice was unforeseen and posed great challenges. Professionals provided services such as diagnostic assessment, intervention, and supervision using online telecommunication tools.

Some evidence is available that reading instruction provided via online videoconferencing tools can be equally effective as in default face-to-face contexts (Furlong et al., 2021). However, dyslexia treatment differs from standard classroom instruction, and children with dyslexia cannot be assumed to be equally susceptible to online support as typically developing readers. Dyslexia frequently co-occurs with other developmental problems, including behavioral and affective problems (Margari et al., 2013), requiring specialized therapeutic skills from their therapists for effective support. It is therefore of interest to examine to what extent therapists are able to provide effective treatment for dyslexia via online channels. Scientific evidence on online provided dyslexia treatment is very scarce, aside from a case study (Wright et al., 2011) and a pilot study (Kohnen et al., 2021). Although both studies suggest that dyslexia treatment can be effectively provided via online channels, small sample sizes and the lack of a control group of children presented with default face-to-face treatment limits the interpretation of these findings. It is therefore critical to report evidence-based findings that compare traditional face-to-face methods to online treatment, offering insights into whether children with dyslexia were able to have their needs met under these circumstances and to examine groups at special risk.

To this end, this study used a clinical database of the RID, a nationwide clinical center for learning disabilities in the Netherlands, to investigate whether reading treatment delivered via tele-practice is as effective as in-person therapy. Children's treatment progress after an online dyslexia treatment collected during the pandemic was compared with treatment-as-usual data that was collected before the pandemic. Three assessments were conducted: one prior to treatment (T0, baseline), one after approximately 20 sessions (T1, intermediate), and one after approximately 40 sessions (T2, end). Treatment progress was measured by gains in reading and spelling measures.

Method

Participants and procedure

Data was collected at RID, a nationwide clinical center for learning disabilities in The Netherlands, over the period 2018-2021. All children were native Dutch speakers (mono- or multilingual) and went to regular elementary school. Children had been referred to the center because of severe and persistent reading disabilities at school (i.e., below the 10th percentile on conventional reading measures or below the 10th percentile on spelling in combination with a score below the 16th percentile on reading) and resisted extra remedial support at school prior to referral. This study consisted of an observational, pretest-treatment-posttest design. All children had a differential diagnostic baseline assessment after which they followed the specialized reading treatment program. Children in the face-to-face treatment condition (n = 254) finished their specialized reading treatment before the first lockdown of the COVID-19 pandemic in The Netherlands (March 2020), whereas all children in the online condition (n = 162) started their first session around the first lockdown and followed the entire treatment program online. Treatment progress was assessed halfway through the treatment period and after finishing the treatment. This study obtained ethical approval from the ethics committee of the University of Amsterdam.

Outcome measures

Reading

Reading was measured with a Dutch word decoding test (DMT; Verhoeven, 1991). The task consisted of three types of words: monosyllabic CVC words, monosyllabic words with consonant clusters, and polysyllabic words. All words were judged to be familiar to 6-year-old children (Schaerlaekens et al., 1999). For each type of word, a sheet containing 150 words organized in rows was presented and children were instructed to read as many words as possible within 1 minute. The total score was the sum of all correctly read words on the three sheets.

Spelling

Spelling was measured with a Dutch spelling test (PI dictee; Geelhoed & Reitsma, 1999). This task consisted of 135 items. For each item, a sentence was read aloud to the child and the word that the child had to write down was repeated. Words increased in difficulty and syllabic complexity but were mainly consisting of only one morpheme. The total score was the number of correctly spelled words.

Treatment

Face-to-face treatment

Following the dyslexia protocol in The Netherlands, children received a phonics-based, computer-aided treatment. The treatment consists of approximately 40 sessions in which all Dutch letter-speech sound (L-SS) mappings are first explicitly taught by a therapist and consequently provided with high exposure by digital tools in order to obtain a transition from accurate, controlled to associative, automatic processing. Children received weekly, 50-minute one-on-one sessions at the clinical center and had to practice at home five times a week for 15-20 minutes.

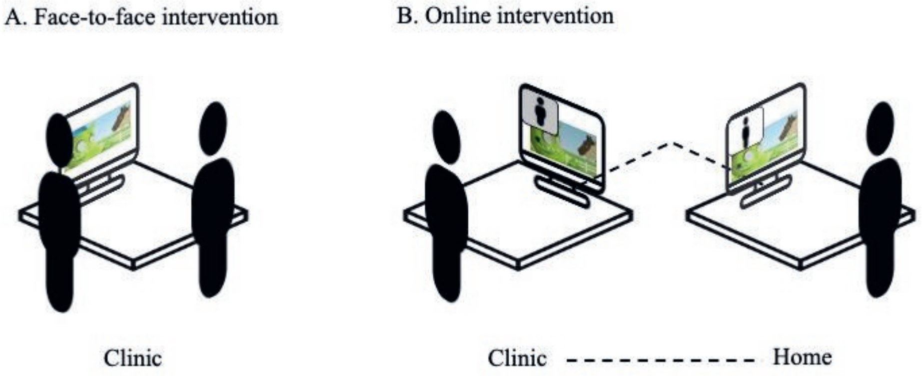
The L-SS mappings were taught step-by-step. That is, regular mappings were taught first followed by increasingly more complex, irregular L-SS mappings (i.e., first short vowels, then long vowels, then diphthongs). After the mapping was introduced, a touchscreen containing buttons corresponding to each Dutch speech sound was used. The touchscreen included several icons to indicate the type of phoneme (e.g., ‘long vowel’), syllable icons (e.g., ‘stressed syllable’), and rule icons to perform certain operations (e.g., delete a selected grapheme). A Dutch word was aurally presented, and children were instructed to pronounce the presented vowels and to identify the item by pushing the corresponding buttons on the touchscreen. Each button press produced the matching sound to direct attention to the matching of letters and speech sounds. Erroneous responses were corrected by the tutor and the computer screen. Reading and spelling were also practiced in an explicit, structured manner. The focus during the treatment gradually shifted from easy word structures into complex words and loan words. To achieve mastery of all elements in the treatment program, children only proceeded to the next phase when they obtained an accuracy rate of at least 80%. The individual acquisition rate was considered by adapting time-constraints to the level of children's performance.

Online treatment

The online treatment comprised the same components and used the same computer-aided software as the face-to-face treatment. The fundamental difference was that instead of weekly face-to-face sessions at the clinic, children logged in from home and the weekly sessions with the therapist were provided via an online videoconferencing platform (Webex). During the online sessions, the therapist and child could communicate via a video-audio connection, while being able to practice using the computer-aided software (see Figure 1). Session duration was identical to that of the face-to-face condition, as was the five-time/week practice sessions at home.

Figure 1

Intervention Set-Up of the Face-To-Face Treatment (Left) and the Online Treatment (Right)



Data analysis

First, children with missing reading or spelling scores were removed from the current sample (n = 13). Baseline differences in age, intelligence, number of sessions, and baseline scores of word reading and spelling were then examined with univariate Bayesian ANOVAs. In addition, we examined whether the two conditions differed in the compliance with homework assignments (% of homework sessions completed) as an index of treatment adherence. Second, to examine whether children who received online treatment improved equally compared to children who received face-to-face treatment, we performed a Bayesian Analysis of Covariance with word reading and spelling scores after treatment as dependent variables, treatment mode (face-to-face vs online) as a fixed factor and baseline word reading and spelling scores as covariates. Bayesian approaches have the advantage of estimating evidence for both the null and alternative hypothesis, whereas frequentist methods only provide evidence against the null hypothesis (Dienes, 2014). The Bayesian ANCOVA compared 4 models with varying predictors of the dependent variables: (1) a null model, (2) a model containing only treatment mode as a predictor, (3) a model containing only baseline scores as a predictor and (4) a model containing both treatment mode and baseline scores as predictors. We examined which model was the best among possible models by comparing Bayesian Factors (BFs). We used Jeffrey's benchmarks to interpret the strength of evidence; BFs between 1 and 3 were considered anecdotal, BFs between 3 and 10 as moderate, between 10 and 30 as strong, between 30 and 100 as very strong and > 100 as decisive evidence for a given model relative to another. Last, in order to examine whether children really improved their reading and spelling scores during treatment, standard reading and spelling scores of pre- and post-test were compared with a Bayesian paired sample T-test. Analyses were run in JASP (JASP team, 2022) and therefore the default priors of JASP were used in the analyses. Results were visualized in R Studio version 1.2.5033 (R Studio Team, 2022). Analyses were not pre-registered.

Results

Baseline measures

In total, 254 children who received face-to-face treatment and 162 children who received online treatment were included in the analyses. The descriptive characteristics of the two groups are shown in Table 1. The Bayesian ANOVAs indicated that for age, the data were 4.74 times more likely to occur under the model including treatment compared to the null model. This means that children differed slightly in their age at the moment they started the treatment in the two conditions. Children who received the face-to-face treatment were on average 8.62 (0.80) years old and children who received the online treatment were on average 8.85 (0.85) years old. In addition, children in the face-to-face treatment group received slightly more treatment sessions (*M* = 48.21, *SD* = 5.68) compared to the online group (*M* = 44.15, *SD* = 5.41), with the data being more than 100 times more likely to occur under the model including treatment. This difference arose at random as in the Netherlands, dyslexia care is reimbursed at the municipality level. Municipality's resources affect the number of reimbursed sessions, and the number of children receiving either face-to-face or online treatment differed across municipalities. As a result, this yielded an at random difference in number of sessions attended in the face-to-face condition compared to the online condition. Therefore, the number of sessions was included as a covariate in the analyses. For intelligence and for reading scores, the data were approximately 8 times more likely to occur under the null model compared to the model including treatment and approximately 4 times more likely for spelling and phoneme awareness to occur under the null model. Last, for number of homework sessions, the data were 5 times more likely to occur under the null model. In other words, it was most probable that there were no differences between the two conditions in these baseline measures.

Table 1

Baseline Measures for Both Conditions Separately

Variable	**Face-to-face (*n* = 254)**		**Online (*n* = 162)**	
	Mean	SD	Mean	SD
Age	8.62	0.80	8.85	0.85
N sessions	48.21	5.68	44.15	5.41
N reimbursed sessions	48.38	5.65	44.60	5.23
Intelligence	100.11	11.73	99.47	12.14
Homework sessions (%)	93.06	7.67	92.30	7.58
Reading (raw scores)	80.04	39.90	81.38	38.99
Spelling (raw scores)	38.43	18.54	40.83	20.72

Note. N sessions = Number of sessions. % of homework sessions was used as an index of treatment adherence. For reading and spelling assessments, raw scores were used.

Treatment

To examine differences in reading and spelling outcomes for children who received face-to-face and online treatment, we performed a Bayesian ANCOVA. Intervention outcomes on reading and spelling outcomes are depicted in Figure 2. As we wanted to control for number of treatment sessions as well, eight different models were compared with varying predictors of the outcome variables: (1) a null model, (2) a model only containing treatment mode, (3) a model only containing baseline scores, (4) a model only containing number of sessions, (5) a model containing baseline scores and number of sessions, (6) a model containing treatment mode and number of sessions, (7) a model containing baseline scores and treatment mode and (8) a model containing baseline scores, treatment mode and number of sessions as predictors.

For reading scores, only model 5 had its model odds increased after observing the data ($BF_M = 50.57$). This means that the model including the baseline score and number of treatment sessions as predictors was the most probable (*P*(M|data = 0.88), with the observed data being 7.4 times more likely under this model than the model containing the baseline score, number of sessions and treatment mode. To account for model uncertainty, we performed Bayesian model averaging of the effects of all three predictors. The data was > 100 times more likely under models containing baseline scores and number of sessions as predictors of reading progress, but only 0.136 times as likely when including treatment mode. In sum, only baseline scores (mean effect = 0.95, 95% credible interval = [0.88,1.01]) and number of sessions (mean effect = 0.92, 95% credible interval = [0.46,1.38]) impacted reading scores after treatment and importantly, whether a child received face-to-face or online treatment had no effect.

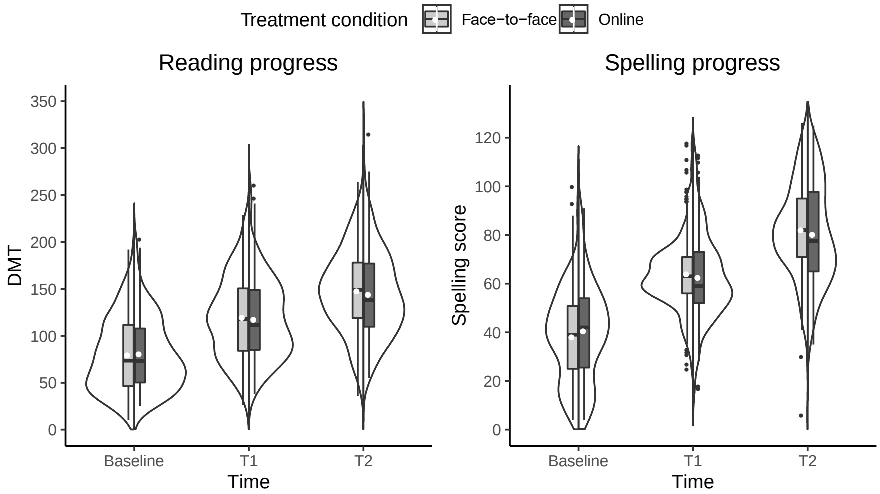
For spelling scores, model 5 (including baseline score and number of sessions as predictors), and model 7 (including baseline score, number of sessions and treatment mode as predictors) increased in model odds after observing the data ($BF_M = 40.52$ and $BF_M = 1.21$ for model 5 and 7 respectively). The model excluding treatment mode as a predictor was the most probable (*P*(M|data) = 0.85) with the observed data being 5.8 more likely under this model than the model containing all three predictors. To account for model uncertainty, we again performed Bayesian model averaging to the effects of all three predictors. The data was > 100 times more likely under models containing baseline scores and number of sessions as predictor of reading progress, but only 0.173 times as likely when including treatment mode. In sum, only baseline scores (mean effect = 0.78, 95% credible interval = [0.69,0.82]) and number of sessions (mean effect = 0.58, 95% credible interval = [0.34,0.81]) impacted spelling scores after treatment and importantly, whether a child received face-to-face or online treatment had no effect.

Finally, we performed Bayesian paired sample t-tests with age-normed T-scores (*M* = 50, *SD* = 10) to examine whether children improved their positions within the distribution of reading and spelling performance of an age-related normative sample. As the two conditions did not differ significantly, we pooled the data of the face-to-face treated and online group for this analysis. We found decisive evidence (BFs > 100) that standard scores at posttest were greater compared to pretest scores. This was true for both reading (pre: *M* = 31.19, *SD* = 4.24; post: *M* = 33.84, *SD* = 6.78) and spelling scores (pre: *M* = 27.38, *SD* = 6.05; post: *M* = 36.85, *SD* = 10.17).

Note that the improvement in reading scores after the intervention is less than reported in previous studies (e.g., Tijms, 2007; 2011). This is likely to result from the three DMT sheets with increasing complexity being averaged together. While most children improve in reading easy words, improvement might be less apparent in more complex polysyllabic words. To test this hypothesis, we did a follow-up analysis of the three subtests separately. However, T-scores for the separate sheets were not available. Instead, we used the reported categories for each subtest. Categories ranged from A to E, with A being the highest level and E being the lowest. Generally, levels A to C all comprise 25% of the data. The last quartile is divided into two categories, with level D containing 15% of the data, and level E containing the lowest 10%. Children who were referred to the clinical center for dyslexia intervention were thus thought to perform in the two lowest categories. Based on descriptive data, 355 children performed in the lowest category for the first DMT sheet, 377 children performed in the lowest category for the second sheet, and 367 performed in the lowest category for the third sheet. 46%, 31% and 29% of children performed in at least one category higher after the intervention for the three sheets respectively. Using a Bayesian Wilcoxon Signed-Rank Test, we found decisive evidence (BFs > 100) that children increased their reading performance on all three DMT sheets after the intervention.

Figure 2

Intervention Outcomes for the Face-To-Face Condition and Online Condition Separately



Note. Reading (left) and spelling (right) progress for children who received the face-to-face treatment and the online treatment separately. The white dots represent the group means. DMT = *Drie-minuten test* (three-minutes test).

Discussion

As a result of the COVID-19 pandemic, many health services had to move to online platforms. Although previous studies have reported that similar results can be obtained through online literacy assessments compared to face-to-face assessments, results indicating the same for online provided reading treatment are lacking. The current study used a clinical Dutch sample to evaluate whether children who received online dyslexia treatment during the ongoing COVID-19 pandemic improved as much as their peers who received the usual face-to-face treatment.

We compared 254 children who received treatment as usual to 162 children who received online treatment with Bayesian methods. All children in the face-to-face treatment condition finished their treatment before the pandemic-induced lockdown in March 2020, whereas all children in the online treatment condition started their treatment after this date and received all treatment sessions through Webex. Reading and spelling scores at baseline and after approximately 40 treatment sessions were used to examine treatment progress. Children in the online condition received slightly fewer treatment sessions compared to the face-to-face group as a result of at random differences in reimbursed sessions per municipality but correcting for number of sessions revealed that children in the online treatment condition progressed equally compared to the treatment-as usual-condition. Importantly, the two conditions did not differ in treatment compliance assessed by percentage of completed homework assignments, which has been found to be an important predictor of intervention outcomes (Mausbach et al., 2010). These results have clinical and practical implications as they show that reading treatment can successfully be delivered online.

Besides the successful applicability of tele-practice in speech- and language assessments (Taylor et al., 2014), two studies previously showed promising results for specialized reading treatment as well (Kohnen et al., 2021; Wright et al., 2011). However, small sample sizes and a lack of a treatment-as-usual condition limited the interpretation of these findings. Results of the current study corroborates the notion that providing reading intervention via online platforms does not significantly impact intervention outcomes. Although it is of great importance to consider that some children do not have access to digital resources required for virtual treatment, tele-practice has promise for improved access to health services in rural areas or in special conditions such as pandemics. In relation to this, access to services that are not available in certain geographical areas is within one’s reach. Even when living abroad, individuals can obtain support in their native language. In addition, providing online treatment increases flexibility and reduces travel time. Furlong and Serry (2022) employed a cross-sectional survey in speech-language pathologists and found that even though many therapists did not provide tele-practice before the pandemic, most of them would continue delivering this service when all physical distancing regulations are abolished. Whether motivation in children receiving online treatment is equally high still needs to be examined in future studies, as well as whether some children benefit more from online treatment than others. Since dyslexia frequently co-occurs with other developmental problems, such as ADHD (Margari et al., 2013), it is probable that some children benefit more from a face-to-face approach instead of online treatment.

There are some limitations in the current study that need to be considered. First, due to the physical distancing requirement from March 2020 onwards, all children who received reading treatment automatically received online treatment, and therefore participants were not randomly allocated to a treatment condition. As a result, confounding factors could have introduced a bias. As a next step in developing the evidence base, a randomized controlled trial should be conducted to validate our results. Second, the standard face-to-face treatment provided in the Netherlands already includes the use of digital tools. This possibly makes the difference between the online provided treatment and face-to-face treatment smaller compared to face-to-face treatments that do not use digital tools. Therefore, the generalizability of our findings to other treatment methods and languages should still be assessed in future studies.

In sum, this study provides support to the small body of literature that supports the equivalence of online provided dyslexia treatments compared to traditional face-to-face methods in a Dutch clinical sample. Children who received the online treatment progressed equally compared to their face-to-face treated peers. Future studies should examine whether some children benefit more from online treatment than others and whether our results can be generalized to other treatment methods and languages.

7

A randomized proof-of-concept trial on the effectiveness of a game-based training of phoneme-grapheme correspondences in pre-readers

A version of this chapter is published as: Verwimp, C., Snellings, P., Wiers, R. W., & Tijms, J. (2023). A randomised proof-of-concept trial on the effectiveness of a game-based training of phoneme-grapheme correspondences in pre-readers. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12821

Abstract

Learning which letters correspond to which speech sounds is fundamental for learning to read. Based on previous experimental studies, we developed a serious game aiming to boost letter-speech sound (L-SS) correspondences in a motivational game environment. The goal of this study was to determine the efficacy of this game in training L-SS correspondences in pre-readers. Additionally, an extended version of the game was developed given the importance of handwriting in audiovisual integration. We established whether including a motoric component in the game boosted the letter-speech sound training on top of the effect of the game without the motoric component. One-hundred forty-five kindergartners were randomly allocated to play either the standard audiovisual version of the game, the motoric version, or a control math game. All children were pre- and post-tested on L-SS knowledge and reading accuracy. We found that playing the game enhanced pre-readers' L-SS knowledge, but not reading accuracy, after a short, intensive intervention period of three weeks. However, children who played the motoric version of the game did not differ significantly from either the standard or the control condition. This game was efficient in training L-SS correspondences in pre-readers. These results suggest that this game might be useful as a preventive evidence-based intervention for at-risk children in kindergarten who might benefit from a head start before learning how to read. Future studies are needed to examine whether a longer intervention period results in L-SS knowledge being translated into reading skills.

A randomized proof-of-concept trial on the effectiveness of a game-based training of phoneme-grapheme correspondences in pre-readers

The importance of reading is uncontested in today's society. Learning to read involves successfully linking visual to auditory information. Learning which letters correspond to which speech sounds takes up to one year of formal reading instruction in Dutch (Vaessen & Blomert, 2010). However, automatizing this process, a fundamental building block for fluent reading, takes much longer and requires much exposure. Based on previous experimental studies, we developed a serious game aiming to boost letter-speech sound correspondences in a motivational game environment. The purpose of the present study was to investigate the efficacy of this game and the effect of implementing a motoric component in the game in children who have not learned how to read yet, which is of great importance to examine whether this game can serve as a preventive intervention for developing severe reading difficulties in the future.

Although the vast majority of the population learns how to read relatively effortlessly, three to ten percent of children worldwide experiences severe difficulties with reading acquisition (Snowling, 2013). These difficulties are typically identified after several years of formal reading instruction, that is, after the most effective time for reading intervention has passed, referred to as the dyslexia paradox (Ozernov-Palchik & Gaab, 2016). As a result of this late diagnosis, the gap in reading performance between children with reading difficulties and their typically developing peers tends to widen over time (Vaughn et al., 2009). In addition, children suffer from adverse long-lasting effects, such as feelings of shame, anxiety, loneliness, and depressive thoughts (Hendren et al., 2018; Livingston et al., 2018; Mugnaini et al., 2009; Xiao et al., 2022). Therefore, it is of great importance to identify and prevent reading difficulties at the earliest age possible.

The hallmark of learning to read in an alphabetic language, such as Dutch, is the acquisition of the alphabetic principle. That is, one needs to understand how written language (graphemes) relates to spoken language (phonemes; Castles et al., 2018). When learning how to read, explicit instruction is required to direct attention toward visual and auditory information, after which this information is combined into audiovisual objects in multisensory brain regions (Stein & Stanford, 2008). After explicitly teaching these correspondences, children slowly shift toward automatized word reading (Karipidis et al., 2021; Romanovska & Bonte, 2021). Subsequently, children start to decode unknown words autonomously and create orthographic representations for successfully decoded words, known as the 'self-teaching theory' (Share, 1995). Although these correspondences can commonly be learned within one year of formal reading instruction in the Netherlands, developing fully automatized associations involves a substantially more protracted developmental pathway that takes up to about 4 years and requires much exposure (Froyen et al., 2009). Behavioral and neuroimaging studies have reported that strong integration of letters and speech sounds is associated with proficient reading whereas poor integration is associated with dysfluent reading (Blau et al., 2009; Froyen et al., 2009, 2011; Žarić et al., 2014). The difficult pathway toward this L-SS automatization in children with reading difficulties indicates the need for high exposure to and repeated practice of these correspondences.

Providing intensive teacher support to provide high exposure for all at-risk children is economically and practically unfeasible. One way to provide this support in different settings such as schools and at home is with game-based interventions (Lassault et al., 2022; Neumann, 2018; Patel et al., 2021; Richardson & Lyytinen, 2014; Vanden Bempt et al., 2021). Game-based interventions have increasingly gained interest in recent years to train academic skills (Jaramillo-Alcázar et al., 2021; Lassault et al., 2022; Skiada et al., 2014; Yildirim & Surer, 2021). These games commonly comprise multimodal features (e.g., sounds, animations, text) that stimulate young children's attention and enhance motivation (van de Ven et al., 2017). This might be especially necessary for children who are struggling to master basic academic skills, such as reading, and need to remain motivated to practice the compromised skill for a prolonged period (Froyen et al., 2009; Stafford & Vaci, 2022). In this context, games are the ideal combination of massive, targeted exposure while maintaining the child's motivation and engagement (Prensky, 2003). In addition, immediate and continual feedback prompts children to update their knowledge and improve their learning outcomes (Muis et al., 2015), without the need for an external resource such as a teacher or peer. Their adaptive algorithms allow for individualized, targeted practice, as they provide enough cognitive stimulation while minimizing failure experiences, making it appropriate for each learner's developmental stage. Altogether, these features make game-based interventions a low-cost and practical tool to support individual learning in an educational context (Jaramillo-Alcázar et al., 2021; Lämsä et al., 2018).

Although many games targeting math and literacy skills have been developed in recent years, studies reporting on their effectiveness and features that moderate this effectiveness are scarce and yield mixed results (J. Kim et al., 2021). A recent review by McTigue et al. (2020) has synthesized findings from 28 studies that examine the effectiveness of GraphoGame, a game that was initially developed for dyslexia prevention but further developed as an intervention focusing on the connections between written and spoken language representations

(Lyytinen et al., 2009). Given that behavioral and neuroimaging studies have shown that strong integration of letters and speech sounds is associated with proficient reading, one would expect that intensive training of these connections would lead to enhanced word reading performance (Blau et al., 2009; Froyen et al., 2009, 2011; Žarić et al., 2014). Multiple studies reported a significant increase in sub-lexical skills, such as letter knowledge or phonological processing (McTigue et al., 2020). In addition, studies that involved high adult interaction (i.e., adults providing technological and motivational support) reported better word reading skills (g = 0.48). However, estimation of the overall mean effect size (g = -0.02) indicated that a transfer to word reading was not consistently found, especially in studies in which the game was played in solitude. To get more insight into which game features exactly contribute to better learning, Wouters et al. (2013) proposed to compare learning outcomes of learners who played different versions of the same game, referred to as a value-added approach (Mayer, 2011). In the current study, we employed such a value-added approach to examine a feature that is assumed to support prereaders in shifting from letter-speech sound knowledge to accurate word decoding: a sensori-motoric component.

Successful letter-speech sound learning is accompanied by functional specialization of the left ventral occipitotemporal cortex, a core region for fast recognition and processing of print (Brem et al., 2010; I. Karipidis et al., 2017; Pleisch et al., 2019). Studies have shown greater brain activation when letter learning was combined with handwriting practice than with other types of practice including typing and visual recognition training (James & Atwood, 2009; James & Engelhardt, 2012). These results suggest that motor experience may facilitate the integration of auditory and visual word form information, and consequently change visual processing (Guan et al., 2021; James, 2017). Moreover, studies reported that handwriting training did not only shorten the acquisition course (Bosse, 2015; Guan et al., 2011; James, 2017; Wiley & Rapp, 2021), but this knowledge was also extended to untrained tasks such as word reading (Wiley & Rapp, 2021). Although this sensori-motor approach has already been proposed to be implemented in education when teaching children to recognize shapes such as letters (Montessori, 1912), studies examining the effect of this motoric component on top of extensive exposure in the context of a game-based reading intervention remain scarce.

The serious game KlankKr8 was developed as a collaboration between a national clinical center for children with learning difficulties and technical partners based on previous experimental studies (e.g., Aravena et al., 2013, 2018; Guerra, 2022). As most game-based interventions concerning reading considered an English-speaking or Finnish-speaking sample (e.g., McTigue et al., 2020), the game used in the current study was developed in Dutch.

Regarding orthographic consistency, Dutch is a semi-transparent language because it has few rules and a small proportion of irregular words (Schmalz et al., 2015). This makes it less complex than English which has a great number of irregular words that cannot be sounded out phonetically, but more complex than Finnish, which is highly consistent in how speech sounds are mapped onto visual symbols (Schmalz et al., 2015). This game systematically and explicitly introduces Dutch letter-speech sound correspondences in a highly engaging game environment. Children first become familiarized with Dutch letters and speech sounds after which they are massively exposed to the correspondences. Besides the standard audiovisual version of the game, which specifically aims to train letter-speech sound mappings, we developed a version in which children practiced the motoric movement associated with the letters as well. The aim of the current study was to examine the efficacy of these two game versions in improving letter-speech sound knowledge, potentially giving more insights into how we can intervene on the letter-speech sound automatization deficit associated with reading difficulties. As proposed by Wouters et al. (2013) and Mayer (2011), we followed a value-added approach in which one compares a standard version of a game to an enhanced version. That is, we aimed to examine the effect of adding a motoric component to the game on the learning outcomes, on top of the effect of the standard audiovisual version.

To this end, we used a randomized controlled trial with three conditions: 1) the standard audiovisual version of the game, 2) the motoric version of the game and 3) a control game (mathematical training). In line with the dyslexia paradox (Ozernov-Palchik & Gaab, 2016), we aimed to examine the effect of the game before children received formal reading instruction, and therefore this study was conducted in kindergartners. All children were tested prior to the intervention period during the first half of the second year of kindergarten. Measures included letter-speech sound knowledge, phoneme awareness, and reading accuracy. Children were randomly allocated to one of three conditions and played the game for an intensive period of 3 weeks. After the intervention period, letter-speech sound knowledge and reading accuracy were assessed again to quantify the effect of the intervention. We expected that children who played either the standard version or the motoric version of the game would improve more in letter-speech sound knowledge compared to the control condition and that playing the motoric version (i.e., the enhanced version of the game) would even lead to stronger gains given the importance of handwriting in facilitating the integration of auditory and visual word form information. In addition, we aimed to examine whether this game could enhance reading accuracy.

Method

Participants

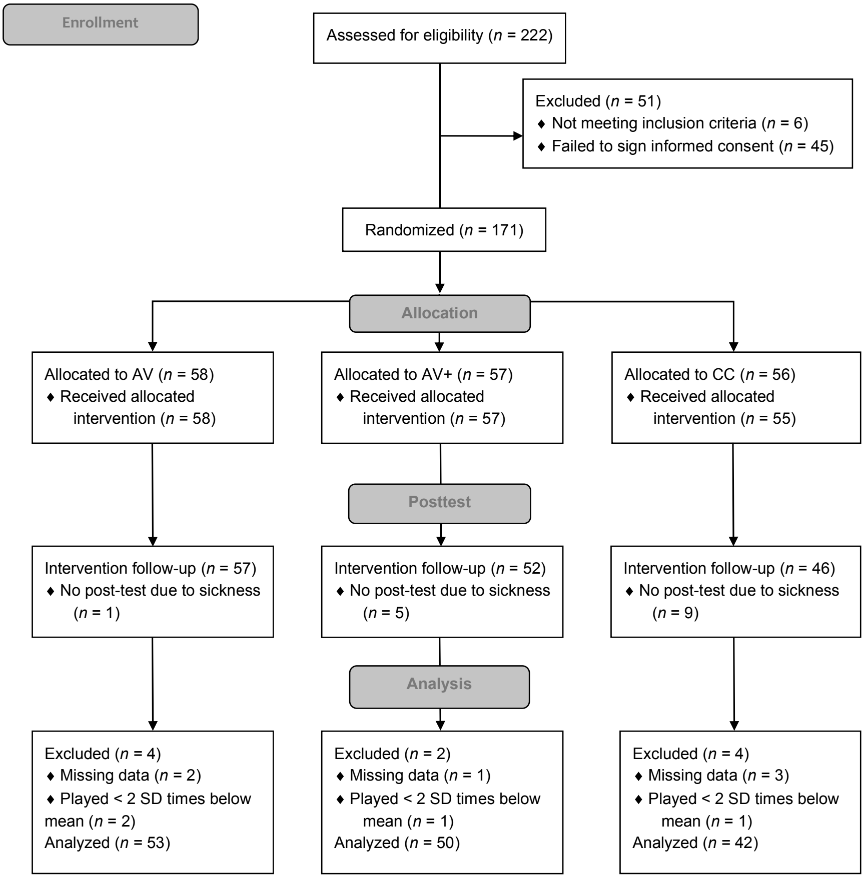
Based on an a priori power analysis for a one-way fixed effect analysis of covariance with 3 levels, 42 cases per cell were required to obtain a power of 0.90. In anticipation of a possible attrition rate of 15%, we aimed to include 50 participants in each condition (*N* = 150). Six primary schools in the Amsterdam metropolitan area participated. The only inclusion criteria were that children were required to be in the second year of kindergarten (i.e., just before receiving formal reading instruction) and had Dutch as their mother tongue (mono- or multilingual). Schools sent information letters to all parents of children in the 2nd year of kindergarten (*N* = 222). After receiving active consent from both parents and their children, children were randomly assigned to one of three conditions using a computerized, within-classroom randomization to control for the influence of teacher and school. In total, 177 parents and children gave consent, of which 6 children had to be excluded because they were still in the first year of kindergarten (*n* = 5) or already in Grade 1 (*n* = 1). This resulted in 58 children in the standard audiovisual condition (AV), 57 children in the motoric condition (AV+), and 56 children in the control condition (CC). As can be seen in Figure 1, some children had to be excluded from the final sample due to missing data or because they did not have enough playtime. The final sample comprised 145 children in the last year of kindergarten (66 boys) with a mean age of 63.56 months (*SD* = 4.47).

Design and procedure

The intervention was a single-blind randomized controlled trial in which children were randomly allocated to one out of three conditions. This study was approved by the ethical committee of the local university and pre-registered at Netherlands Trial Register (Trial NL9604). All children were tested individually prior to the intervention period (i.e., baseline) during the first half of the second year of kindergarten, including measures of letter-speech sound knowledge, phoneme awareness and reading accuracy. The individual sessions lasted approximately 20 minutes and were conducted by trained research assistants who also guided the intervention to reduce the burden on schools. After the baseline measures, children played one of the serious games for a period of 3 weeks in total (15min/5 days/week). Intervention sessions took place in a spare classroom during regular school hours. Tablets with headphones (IMG stageline MD-5000DR) were positioned in a spare room and children were brought in class-by-class. Each child had a pre-created profile for either the experimental game or the mathematical control game to ensure that children played under the same profile for the entire intervention period. The tablet games closed themselves after 15 minutes to ensure that all children had the same amount of playtime. Post-tests were done by research assistants who did not guide the intervention for that specific classroom, and these research assistants were therefore blinded to the intervention condition of the child. Parents and children were also blinded to intervention assignment but were debriefed at the end of the trial period. Children allocated to the control condition received the opportunity to play the serious game after the trial period ended.

Figure 1

Flowchart Showing Enrolment and Allocation of Participants (Adapted From Moher et al., 2010).



Assessment battery

Outcome measures

Letter-speech sound knowledge was measured with a letter-speech sound identification task that was programmed in PsychoPy3 (Peirce et al., 2019). Children heard a Dutch speech sound through headphones which was accompanied by two Dutch letters (or letter combinations) on a computer screen; one on the left side and one on the right side of the screen (see Figure 2). They had to indicate which letters corresponded to the sound by pressing the left, yellow ('A' key) or right, blue button ('L' key). Children received no feedback on whether the answer was correct, but a spaceship appeared on the side of the selected answer to indicate that the computer had recorded their answer (see Figure 2). If no response was given after 4000 ms a picture of a snail appeared prompting the child to respond faster. Each Dutch speech sound, 44 in total, was presented 4 times, resulting in a total of 160 items with a break in between. As the task was self-designed, no norms were available and, therefore, raw scores were used in the analyses. The score was calculated as the number of correct items with a maximum score of 160.

Reading ability. Children were asked to read as many words as possible within one minute. The task consisted of 20 carefully selected words with increasing difficulty that were all known to Dutch toddlers according to a list of basic vocabulary in kindergarten in the Netherlands (*Basiswoordenlijst Amsterdamse Kleuters*; Mulder et al., 2009). As the task was self-designed, no norms were available and, therefore, raw scores were used in the analyses. The score was calculated as the number of correctly read words within the time limit of one minute, with a maximum score of 20.

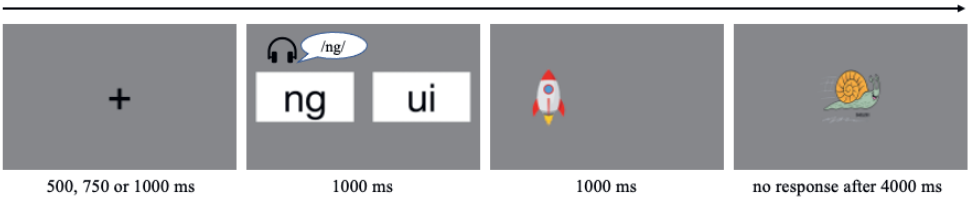
Baseline measures

Phoneme awareness was measured with a beginning phoneme identification task (Boets et al., 2010). Each item was visually presented, followed by four pictures. Children had to indicate which items had the same beginning phoneme as the target item. All items represented high frequent one-syllable Dutch words and were named for the child before they were prompted to give the answer. To prevent guessing, the distractor alternatives contained a correct, semantically related, phonologically related, and non-related answer. The task consisted of ten items preceded by two practice items to which the researcher gave feedback ensuring that the child understood the task. The score was defined as the number of correctly answered items with a maximum score of 10. Cronbach's alpha was 0.59 at age 5.4 (Boets et al., 2010).

Non-verbal intelligence was measured with Raven's 2 Progressive Matrices (CPM-2; Raven et al., 1998). Children had to indicate the missing element in a pattern out of five answer alternatives. The task comprised three sets of 12 items each with increasing complexity. Small groups of children were seated in a silent room and received 20 minutes to fill out all 36 items individually after they received feedback from the researcher on the practice items to ensure that they understood the task. The number of correctly answered items out of 36 was norm-referenced afterwards. The reliability coefficient based on a Fisher's z-transformation was 0.82.

Figure 2

Task Design of the Letter-Speech Sound Knowledge Task



Interventions

Audiovisual only condition (AV)

In the first condition, children played the standard version of the game. The purpose of this game was to learn letter-speech sound correspondences with high exposure in a motivational environment. The game took place in a space environment in which an astronaut systematically introduced all Dutch speech sounds (44) and their corresponding letter-representations to the child with increasing difficulty. The game comprised 8 levels, built up as different star constellations. Each constellation consisted of multiple speech sounds children had to learn, with the first star system comprising short vowels, and the more complex speech sounds like /eeuw/ appearing in the later levels. The astronaut first introduced a letter with its corresponding speech sound and explained how it should be pronounced. A variety of intonation patterns was introduced for each speech sound, resembling the different pronunciations in real life. Afterwards, children had to tap the bullets with the corresponding letters (see Figure 3A). Immediate feedback was provided; when the correct bullets were tapped, a green smiley appeared, whereas a red smiley appeared when a wrong bullet was tapped (Figure 3C). By tapping the correct letters children collected stardust which made their start grow. The game was adaptive in nature meaning that speech sounds that elicited errors or slower response rates were repeated and subsequently presented more often. In addition, misidentified letters were temporarily presented more frequently as distracter stimuli (see e.g., Lyytinen et al., 2009). This implies that children were provided with exercises at their own level, that is, neither too easy which could result in less engagement, nor too difficult which may cause feelings of failure. In addition, the multimodal features of the game in combination with external motivational components, such as time bonuses and rewards for obtained levels, contributed to the enjoyableness of the game.

Audiovisual condition with motoric component (AV+)

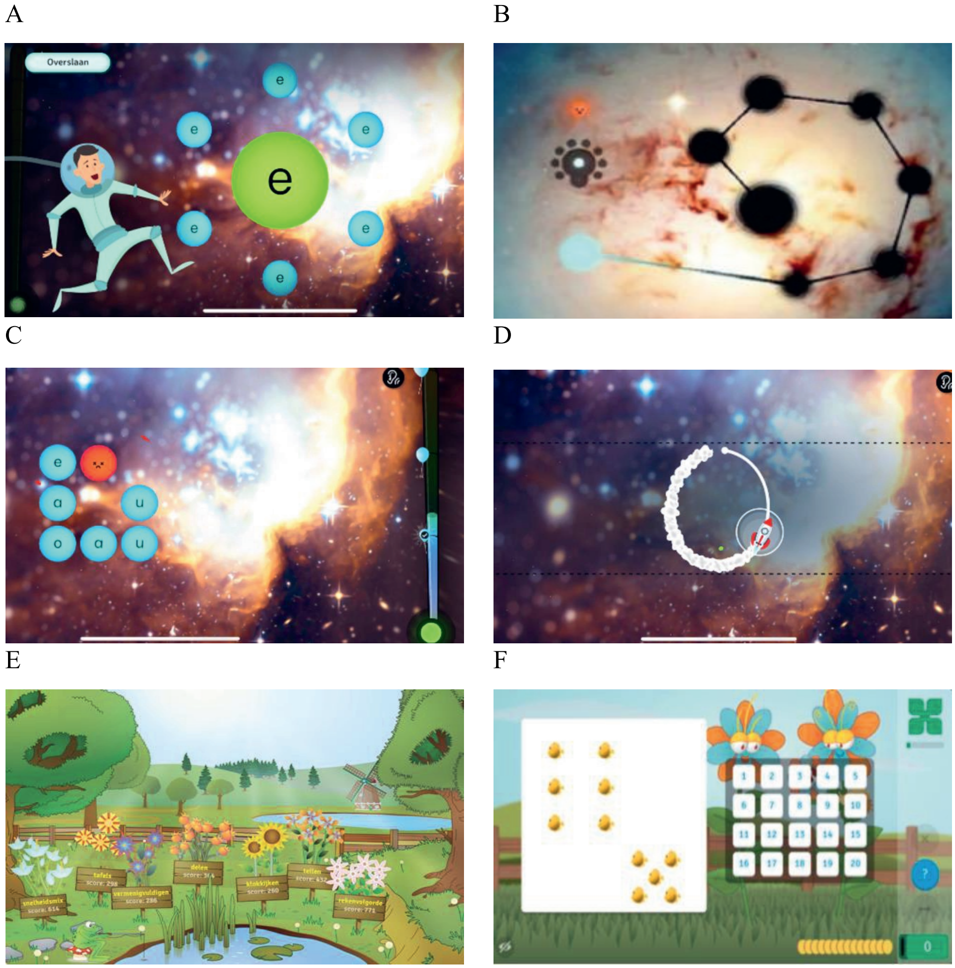
In the second condition, children played an extended version of the game (hereafter referred to as motoric condition (AV+)). Getting familiar with the Dutch speech sounds and tapping the corresponding bullets was accompanied by instructions on how a certain letter had to be written. This writing motion was first introduced by the astronaut after which children had to write the letters on the tablets themselves. Considering the premature fine motoric skills of kindergartners, children used their index finger to draw the letters on the tablet instead of using a pen or pencil. This writing motion was guided by a spaceship that followed the movement of the child's finger while writing the letter. When children did not follow the correct shape of the letter, the rocket stopped until they continued following the correct shape of the letter (see Figure 3D).

Mathematical control condition (CC)

Math Garden [*Rekentuin*] was used as a control condition that did not involve any learning of letters or reading while being exposed to a tablet-based game with an equivalent amount of playtime. In this game, children were learning basic mathematics by watching educational videos and making assignments with numbers and shapes (Straatemeier, 2014). These mathematical skills included counting, comparison skills, and basic arithmetic operations. The game took place in a garden in which each plant represented a game (see Figure 3E). This game was also adaptive in nature, and children saw their plants grow when their ability increased.

Figure 3

Examples of Screens in the Experimental Games (Klankkr8) and the Control Game



Note. (A) The astronaut introduces the child to the letter ‘e’. (B) Example of a star constellation that comprised different speech sounds. (C) Children saw a sad smiley when they tapped the wrong (non-corresponding) bullets. (D) Example of the writing exercise that was included in the motoric version of the game. Children had to follow the pattern to write the letters. (E) Home screen of Math Garden. Different plants represent different games. (F) Example of the counting game in Math Garden. Children had to count the little fish and fill in the correct number.

Data analysis

Children were excluded when post-test data was missing due to illness (*n* = 14) or when they played the intervention 2 SD sessions below the mean (< 9 sessions, *n* = 4). In total, the analysis included 53 children who played the standard audiovisual version of the game, 50 who played the motoric version, and 42 children who played the mathematical control game. To examine differences in baseline measures between the three conditions, univariate ANOVAs were conducted. Differences in gender distribution between the three conditions were examined with a Chi-Square test. As a primary question, we examined the effect of the game on letter-speech sound knowledge. According to the pre-registered analysis and in terms of power (Van Breukelen, 2006), we conducted a univariate ANCOVA with L-SS outcome at post-test (T1) as the dependent variable and the intervention as factor while controlling for L-SS knowledge at pre-test (T0). As a secondary question, we wanted to examine whether children could read short, easy words after this short intervention period. A similar ANCOVA was therefore repeated with reading accuracy at post-test (T1) as the dependent variable and the reading score at pre-test (T0) as covariate. Assumptions of linearity, homogeneity of regression slopes, homogeneity of variances, and normality were checked before conducting the statistical analyses. All assumptions were met except the assumption of normality for the reading scores. However, as each condition contained more than 30 participants and deviations from normal in reading score distribution were minor, normality can be assumed according to the Central Limit Theorem (Kwak & Kim, 2017). Significant results were followed up with Tukey post-hoc comparisons. Having a minimal number of intervention sessions was not pre-registered and therefore we examined whether including the four children who played the game less than 9 times in the analysis yielded similar results.

As an exploratory analysis, we explored associations between in-game measures and children's pre- and post-test measures. We first computed correlations between the cognitive tasks (i.e., L-SS knowledge and reading accuracy at pre- and post-test), accuracy in tapping the correct bullets (%) and the highest obtained level in the game (max = 50). In addition, we wanted to examine whether the in-game accuracy score and highest obtained level mediated the relation between L-SS score at pre- and post-test, and between the reading accuracy score at pre- and post-test. To this end, two separate mediation analyses were performed using the *lavaan* package (Rosseel, 2012) in R Studio (RStudio Team, 2022) with the two pre-test scores (L-SS knowledge and reading accuracy) as predictors, post-test scores as outcomes and game accuracy and highest obtained level as mediators. A bootstrap analysis with 1000 replications was applied to estimate the total, direct, and indirect (i.e., mediated) effects and their 95% CIs.

For these analyses, only children who played either the standard audiovisual version of the game or the motoric version were included (*n* = 103). The two conditions were analysed in one mediation model considering inadequate statistical power to detect mediation effects due to low sample sizes when these two conditions would be analysed separately (Fritz & MacKinnon, 2007). For all analyses an alpha of .05 was used to examine statistical significance.

Results

Baseline measures

Univariate ANOVAs did not reveal differences in age or non-verbal intelligence between the three conditions (*ps* > .63). There were no significant differences in gender distribution across the three groups ($X^2(2)$ = 4.29, *p* = .12). Comparing the pre-test scores did not reveal any differences in L-SS knowledge, phoneme awareness, and reading ability between the conditions (*ps* > .38). Participants' characteristics are shown in Table 1 and descriptive statistics of the outcome measures at baseline and post-test in Table 2.

Table 1

Participant Characteristics for the Standard Audiovisual Condition (AV), the Motoric Condition (AV+) and the Control Condition (CC) Separately

	Mean (SD)			Group comparison
Characteristic	**AV**	**AV+**	**CC**	
n	53	50	42	$X_{(2)}$ = 1.34, *p* = .51
Gender (M:F)	19:34	23:27	24:18	$X_{(2)}$ = 4.29, *p* = .12
Age (months)	63.37 (4.86)	63.40 (3.71)	64.00 (4.87)	*F*(2,134) = 0.469, *p* = .63
Intelligence	97.34 (13.30)	96.79 (12.86)	96.28 (12.21)	*F*(2,132) = 0.096, *p* = .91
N play[a]	12.72 (1.41)	12.72 (1.42)	12.45 (1.55)	*F*(2,141) = 0.355, *p* = .70
PA[b]	4.98 (3.01)	4.58 (2.98)	4.12 (2.94)	*F*(2,141) = 0.978, *p* = .38

Note. [a]Number of intervention sessions [b] Phoneme awareness

Intervention effects

Intervention effects for L-SS knowledge and reading accuracy are depicted in Figure 4. The L-SS score at pre-test was significantly related to the L-SS score at post-test (*F*(1,141) = 242.80, *p* < .001, η_p^2 = .63). In addition, we found a significant effect of condition on the L-SS score at T1 after controlling for the L-SS score at T0 (*F*(2,141) = 4.82, *p* = .009, η_p^2 = .06). Tukey post hoc tests showed that the covariate-adjusted mean of the AV condition (*M* = 99.88) was significantly higher than the mean of the CC condition (M = 90.59) (mean difference = 9.29, t = 3.03, p = .008, d = 0.63). However, the covariate-adjusted mean of the AV+ condition (M = 94.06) did not differ significantly from the CC (mean difference = 3.47, t = 1.12, p = .51, d = 0.23) or the AV condition (mean difference = -5.82, t = -1.99, p = .12, d = -0.39). Corresponding confidence intervals are reported in Table S1 (see Appendix E). Including children who played the game less than 9 times in the analysis yielded similar effects ($F(1,145) = 251.23$, $p < .001$, $\eta_p^2 = .63$ and $F(2,145) = 4.47$, $p = .013$, $\eta_p^2 = .06$ for covariate and condition respectively). Although we controlled for the effect of teacher and school by using within-classroom randomization, children in the same school might be more similar than children in other schools. To examine whether it was needed to account for this school clustering with a multilevel model, we checked whether there was significant variation across schools. A baseline model with only the intercept included was fitted and compared to a model that allowed intercepts to vary across schools. The model with varying intercepts was not significantly better than the model with the fixed intercept ($X^2(1) = 3.671$, p = .06), justifying the use of a general linear model (Field et al., 2012).

For reading accuracy, we conducted an ANCOVA with the reading score at post-test (T1) as the dependent variable and the intervention as factor while controlling for reading accuracy at pre-test (T0). The reading score at pre-test was significantly related to the reading score at post-test ($F(1,141) = 695.83$, $p < .001$, $\eta_p^2 = .83$), but we did not find any significant effect of condition on the reading score at T1 after controlling for reading accuracy at T0 ($F(2,141) = 0.54$, $p = .59$, $\eta_p^2 < .01$). Including children who played the game less than 9 times in the analysis yielded similar effects ($F(1,145) = 721.76$, $p < .001$, $\eta_p^2 = .83$ and $F(2,145) = 0.61$, $p = .54$, $\eta_p^2 < .01$ for covariate and condition respectively). The model with varying intercepts was not significantly better than the model with the fixed intercept ($X^2(1) = 0.141$, p = .71), justifying the use of a general linear model (Field et al., 2012).

Table 2

Descriptive Statistics at Baseline and Post-Test for the Standard Audiovisual Condition (AV), the Motoric Condition (AV+) and the Control Condition (CC) Separately: Mean (Standard Deviation, Range)

		Condition		
Task		AV	AV+	CC
L-SS[a]	Pre	87.77 (18.33; 62-136)	90.44 (21.96; 52-138)	88.98 (18.03; 69-135)
	Post	98.62 (24.55; 49-150)	95.44 (25.83; 55-149)	90.52 (22.25; 56-136)
Reading[b]	Pre	2.42 (4.57; 0-20)	2.48 (4.16; 0-17)	2.07 (3.20; 0-13)
	Post	3.11 (4.75; 0-19)	3.46 (5.22; 0-20)	2.62 (4.02; 0-13)

Note: L-SS = letter-speech sound knowledge. Maximum test score: [a] = 160; [b] = 20

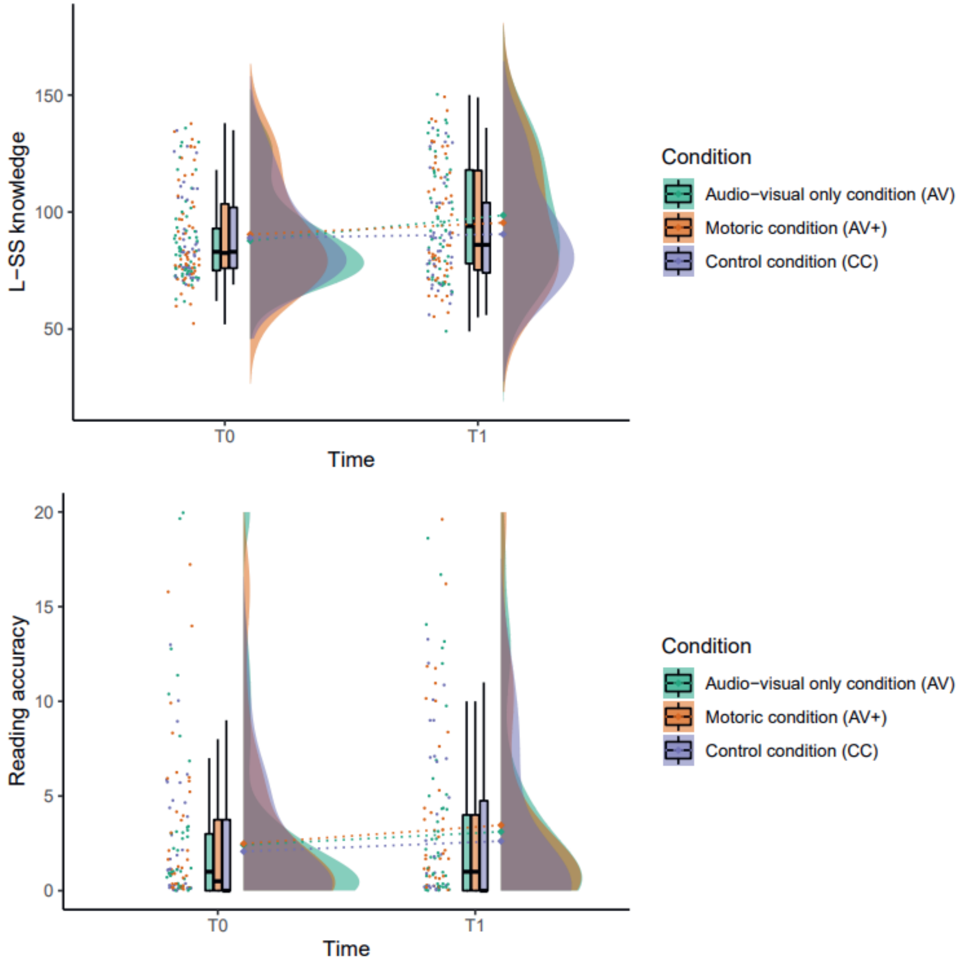
In-game measures

To better understand how intervention progress was related to game performance, we examined the correlations between L-SS and reading accuracy at pre- and post-test on the one hand, and in-game accuracy (%) and the highest obtained level out of 50 on the other hand. For these analyses, only children who played either the standard audiovisual version of the game or the motoric version were included (n = 103). All correlations are reported in Table S2 (Appendix E). Pre- and post-test scores of L-SS knowledge and reading accuracy were significantly correlated to in-game accuracy. That is, not only children who had higher post-test scores, but also higher pre-test scores were more accurate in the game. Likewise, the highest obtained level in the game was positively associated with L-SS and reading scores, indicating that children who completed more levels on average had higher L-SS and reading scores and vice versa, children who had higher baseline scores completed more levels in the game.

A mediation analysis showed that the L-SS score at pre-test positively predicted the L-SS score at post-test (b = 0.97, z = 13.83, p < .001). The association between L-SS knowledge at pre- and post-test became less pronounced, but still highly significant, when including the two mediators (see Figure 5). Analysing the indirect effects indicated that the highest obtained level mediated the relationship between L-SS at pre-test and post-test (b = 0.10, z = 2.44, p = .02), with the 95% bias-corrected confidence interval based on 1000 bootstrap samples entirely above zero (0.03 to 0.20). Although, the 95% bias-corrected confidence interval of the indirect effect through in-game accuracy was also entirely above zero (0.01 to 0.12), this effect was not significant (b = 0.06, z = 1.83, p = .07). For the reading scores, the reading score at pre-test positively predicted the reading score at post-test (b = 1.04, z = 15.44, p < .001). However, none of the mediators significantly mediated the relationship between reading accuracy at pre- and post-test (game level: *b* = -0.02, *z* = -1.42, *p* = 0.16; in-game accuracy: *b* = 0.02, *z* = 1.08, *p* = .28).

Figure 4

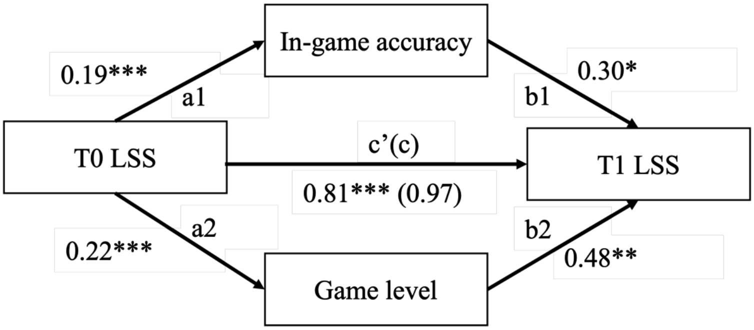
Raincloud Plots Depicting Intervention Progress in L-SS Knowledge (Up) and Reading Accuracy (Down) Across the Three Conditions



Note. In the boxplots, lines indicate medians and the areas above and below the medians indicate first and third quartiles. L-SS = letter-speech sound knowledge.

Figure 5

Mediation Model Including L-SS Knowledge at Pre- and Post-Test With in-Game Accuracy and Highest Obtained Game Level in the Game



Note. Coefficients are unstandardized. The direct effect without the mediators is presented in parentheses. ***p ≤ .001; **p ≤ .01; *p ≤ .05.

Discussion

The goal of this study was twofold. First, we established the efficacy of the game KlankKr8 in training letter-speech sound correspondences in pre-readers. Second, we examined whether including a motoric component in the game boosted the letter-speech sound training on top of the effect of the game without the motoric component. Regarding the first research question, we found evidence for the hypothesized effect of the game in pre-readers; playing the game enhanced pre-readers' letter-speech sound knowledge after a short intervention period of 3 weeks. However, children who played the motoric version of the game did not differ significantly from either the standard or the control condition.

We tested the effect of playing the game on two main outcomes, namely, letter-speech sound knowledge (near transfer) and reading accuracy (far transfer). We found a significant increase in L-SS knowledge in children who played the standard audiovisual version of the game compared to children who played the control game. This is in line with our hypothesis as the game was developed in order to train L-SS correspondences. This finding suggests that KlankKr8 is able to boost L-SS correspondences in younger children who did not receive formal reading instruction yet, and might therefore be useful as a preventive evidence-based intervention for at-risk children in kindergarten who might benefit from a head start before learning how to read.

Second, although the game boosted L-SS knowledge, this knowledge did not transfer to reading short, easy words, as none of the conditions had higher scores for reading accuracy after the intervention period. Although this is in contrast with earlier studies that found a positive effect on word reading skills after training letter-speech sounds correspondences (Fraga González et al, 2015), this is corroborating the findings of a recent review of McTigue and colleagues (2019). McTigue et al. (2019) found that GraphoGame leads to an improvement in a variety of reading subskills depending on the study, but rarely led to an increase in word reading. It should be noted that fully integrating letters and speech sounds takes up to one year of formal reading instruction and takes even longer to become fully automatized (Froyen et al., 2009). As the participants in our study were in the last year of kindergarten, they did not receive any reading instruction yet. In accordance with the self-teaching hypothesis (Share, 1995), a minimal number of mappings is needed before children are able to autonomously decipher words (Perry et al., 2019), which might not be obtained after three weeks of intervention. In addition, in Dutch schools, children are commonly getting familiar with how letters relate to words from the second half of the last year of kindergarten onwards or even later. However, our intervention took place in the first half of the last year of kindergarten, implying that most children had no understanding of how L-SS knowledge should be used to decode words. Although the intervention led to an increase in L-SS knowledge, it is therefore not completely unexpected that three weeks of intervention is not sufficient for transferring this knowledge to word reading. Future studies might want to consider implementing a longer intervention period, in a period in which letters are getting introduced at school, to examine the transfer to word reading. Implementing the game at the time when informal letter and reading instruction starts, that is, at the second half of the 2nd year of kindergarten in Dutch schools, children might be more goal-directed toward applying the L-SS knowledge in decoding words (see also Verwimp et al., 2023), possibly leading to better word reading skills. In addition, it is worth considering using an alternative task such as the word-specific orthographic knowledge task as used in Lassault et al. (2022). In such task, children have to choose the correct word amongst incorrect answer alternatives after the word was presented auditorily, which more closely resembles the skills practiced in the game used in the current study.

The second aim of this study was to examine the added value of including a motoric component in the game compared to the standard version. We hypothesized that children who played the motoric version would have better L-SS knowledge compared to children who played the standard version, as handwriting helps young children to understand and recognize letters (James, 2017). Handwriting movements have been found to facilitate the integration of auditory and visual word form information (Guan et al., 2011, 2021), and therefore can act as a scaffold for coupling auditory and visual word forms. Our results showed that children who played the motoric version of the game did not perform significantly better than children in the control condition. However, it is important to note that they also did not perform significantly worse than children in the standard condition. This finding can be explained as follows. First, we wanted to have the same amount of screen time across the three conditions, therefore all games automatically closed after playing for 15 minutes. For children who played the standard version, this time was mainly occupied with tapping the corresponding bullets. For children who played the motoric version of the game, this time was divided into tapping bullets and writing the letters themselves. As the process of writing the letters is significantly slower than tapping the corresponding bullets, children who played the motoric version of the game in the end thus had less L-SS tapping which might explain why they did not improve as much in L-SS knowledge compared to the standard condition. Second, writing requires fine motor skills that might not have been fully developed in children at this age. The development of handwriting starts around the age of 5 but continues to develop between the age of 6 and 9 (Feder & Majnemer, 2007). Although we aimed to control for the possibly immature fine motoric skills in kindergartners by using their fingers to write instead of a pencil or pen, directing cognitive capacity toward controlling the motoric movement might have hindered the development of specific memory traces contributing to the integration of letters and speech sounds.

In an exploratory analysis, we examined whether in-game measures, i.e., accuracy during the game and highest obtained level, were related to letter-speech sound knowledge and reading accuracy at pre- and post-test. We found that children with higher post-test scores, but also higher pre-test scores, were the ones who were more accurate in the game. Likewise, children who completed more levels on average had higher L-SS and reading scores and vice versa, children with higher baseline scores completed more levels in the game. A mediation analysis revealed that only the highest obtained level mediated the relation between the L-SS score at pre- and post-test. It is however important to note that children could independently decide whether to continue with the next level or replay the previous level, leading to huge variability in game progress. This could explain why in-game accuracy does not mediate the relationship between the L-SS score at pre-test and the L-SS score at post-test, as children who kept playing the early, easier levels in the game probably obtained higher accuracy scores but were not exposed to the more complex speech sounds and therefore did not improve in L-SS knowledge. Another explanation can be that children who obtained higher levels, and thus progressed further in the game, were the ones who were more motivated and engaged more, leading to greater benefits of the intervention.

The present results add to the growing evidence of game-based interventions as promising tools to support reading development, and the potential to prevent reading difficulties in more transparent languages such as Dutch. However, some limitations to this study need to be addressed. First, as this study served as a proof-of-concept study, we have chosen to implement a short, intensive intervention period. A longer intervention might, however, be needed to examine the transfer of L-SS integration to word reading and to see the effect of the motoric component. Moreover, previous research has suggested that the intervention, especially in children who are at risk of reading difficulties, is the most effective when it happens in short sessions, multiple times a day, to accumulate the memory trace of the newly learned knowledge (e.g., Juhani Lyytinen et al., 2021). In addition, as this study was carried out during the COVID-19 pandemic, some children were absent multiple times during the intervention or were in quarantine when post-tests took place. This resulted in more missing data than anticipated. Providing a longer intervention with multiple sessions a day might lead to better consolidated L-SS correspondences and missing a day of intervention might therefore have less effect on the results. Second, the post-test only took place immediately after the intervention. A future study might want to include a longitudinal follow-up moment to examine whether this intervention still has an effect after a particular amount of time or ultimately, whether playing this game in kindergarten can prevent reading difficulties in elementary school. It is however important to note that the current game design is specifically intended for building automatized letter-speech sound mappings, which is the hallmark of fluent reading (Castles et al., 2018). Only when your ability to decode words is effortless and automatic, you can direct your cognitive capacity to comprehend what is written. Reading comprehension requires the integration of meaning across sentences, making use of contextual cues and inferences based on an individual's general knowledge (Muijselaar et al., 2017). Training reading comprehension thus requires support on a different level, implemented in other game designs (e.g., ComprehensionGame, https://comprehensiongame.com/info/). Third, as we wanted to keep the amount of screen time constant between the conditions, children in the condition with the motoric component had less time for tapping the corresponding bullets compared to the standard condition, possibly explaining why they did not improve on L-SS knowledge. To control for this, it is recommended to implement a second part of the game in the standard audiovisual condition as well, in which children are passively exposed to letters without writing the letters, ensuring that the two conditions have a similar amount of L-SS tapping. Last, the reliability of the phoneme awareness task in the current study was somewhat lower than the reliability of other measures. Phoneme awareness tasks have been found to be very difficult for kindergarten children, often exhibiting floor effects (Catts et al., 2009). Nevertheless, the phoneme awareness task in the current study has been found to load highly on the construct phoneme awareness (Boets et al., 2010), and was only used as a baseline measure in the current study and not as the outcome of interest.

To summarize, the present study showed that KlankKr8 is efficient in training letter-speech sound correspondences in pre-readers. These results suggest that this game might be a useful evidence-based intervention for at-risk children in kindergarten. It may prevent reading problems later in life and may especially be beneficial for at-risk children who need a head start before learning how to read. Future studies can shed light on the optimal duration of the intervention, to determine whether a longer intervention period results in L-SS knowledge being translated into reading skills, and whether using this game in kindergarten can prevent the development of severe reading difficulties.

8

General Discussion

Despite shifting from a core-deficit approach toward a multidimensional approach, most studies still examine factors as isolated components and exclude individuals with co-occurring difficulties. This leads to an overestimation of the homogeneity within individuals who experience reading difficulties, and it remains to be seen how factors within and across diagnostic boundaries interact with each other. The overarching aims of this dissertation were to provide:

1. a better understanding of the multidimensional etiology of developmental dyslexia
2. more insight into neurocognitive mechanisms that underlie the proposed letter-speech sound automatization deficit needed for fluent reading
3. how neurocognitive mechanisms that underlie the proposed letter-speech sound automatization deficit can be translated into clinical practice
4. more insight into how co-occurring attentional difficulties alter reading development.

This chapter summarizes experimental findings and puts them in a broader theoretical context. Finally, we discuss the implications as well as limitations of the present studies and directions for future research.

Summary of findings

In *Chapter 2*, we mapped the complexity of links between cognitive, environmental, and demographic factors related to reading development using network analysis. In this network analysis, we discovered two interconnected subparts: one that included reading fluency and accuracy measures and another that consisted of intelligence-related measures. Interestingly, phoneme awareness was linked to the precise and accurate processing of letter-speech sound (L-SS) mappings. Rapid automatized naming was more closely associated with the automatic integration of visual and speech information. Our findings suggest that a complex network of factors contributes to typical and atypical reading development, each influencing different aspects of the reading process. Consequently, our research challenges the commonly held belief that a single underlying deficit causes dyslexia. By utilizing the network approach to psychopathology, we have demonstrated how we can take into account interactions within the reading network, which holds great promise for more individualized interventions in the future.

In *Chapter 3*, we shed light on attentional processes involved in developing and integrating cross-modal letter-speech sound correspondences. One hundred and seven children learned eight new symbols corresponding to Dutch speech sounds. The learning task was either preceded by instructions that revealed the learning goal of the task and directed children toward that goal, or instructions that told them that the goal would become evident during the task itself. Assessments of artificial word-reading ability and symbol knowledge were conducted immediately following the learning session and on the subsequent day to examine the effect of sleep. Children who received goal-directed instructions learned the new script faster and more efficiently and had better learning outcomes than their implicitly instructed peers. Our results highlight that directing children toward the learning goal can promote L-SS learning and consolidation, providing insights into how top-down control impacts the initial reading skill acquisition phase.

The same artificial learning task was used to examine learning differences between children with dyslexia and typical readers in mapping visual symbols onto spoken language representations, as reported in *Chapter 4*. We shed light on the interaction between behavioral and neural cognitive control processes, i.e., goal-directed behavior, post-error slowing, and neural performance monitoring, as well as how behavioral ADHD measures impact letter-speech sound learnability. Our results showed that most children could learn the new script to some extent, with significant inter-individual differences, but independent of the instruction condition or reading group (children with dyslexia vs. typical readers). This suggests that while some individuals have deficits in letter-speech sound mapping and/or feedback processing, we did not find evidence to assume this was true in all dyslexic learners. This aligns with the multiple deficit model that proposed that no single deficit is sufficient to cause dyslexia, but multiple interacting factors might increase the risk. Learning outcomes were related to reading ability within the native language, validating the use of such an artificial learning paradigm. This paradigm allows us to further examine factors that contribute to the transition from letter-speech sound integration to consolidated audiovisual objects and how these relate to individual learning indices.

The second section of this thesis discusses mechanisms involved in reading intervention, and shed light on the previously reported significant inter-individual differences in intervention response. In *Chapter 5*, we created a detailed window on the reciprocal associations between the development of letter-speech sound integration and word reading fluency within a highly influential and standardized Dutch dyslexia treatment program. A path analysis revealed that the L-SS training program significantly improved L-SS associations and word reading fluency in children with dyslexia, with more automatized L-SS associations leading to more fluent word reading. During the intervention, processing speed only had a small negative relationship with word reading fluency, and age and IQ did not contribute significantly to the model. Interestingly, children with weaker L-SS integration at the beginning of the intervention were the ones that benefitted more. The intervention's second half was crucial in achieving the full potential for word reading fluency growth.

In *Chapter 6*, we examined whether online specialized reading treatments are as effective as face-to-face treatments in improving children's reading and spelling skills, as interventions massively moved to online platforms because of the global COVID-19 pandemic. Using Bayesian methods, 254 children who received treatment-as-usual were compared to 162 children who received online treatment. The results showed that children in the online treatment group had slightly fewer treatment sessions but made similar progress in their reading and spelling abilities, after controlling for the number of sessions, when compared to those who received treatment as usual. Interestingly, the conditions did not differ in the percentage of homework sessions they completed, suggesting that moving interventions to an online format does not affect treatment adherence in this context.

Last, in line with the digitalization of treatment programs, *Chapter 7* discusses the results of a short, intensive, randomized controlled trial in which we examined the effectiveness of a game-based intervention tool. The standard game version was developed to boost letter-speech sound (L-SS) correspondences in a motivational game environment. In this new extended version of the game, we added a motoric handwriting component, given the importance of handwriting in audiovisual integration. We found that pre-readers who played the game improved in L-SS knowledge after a three-week intervention period, but this did not transfer to word reading ability. Children who played the motoric version of the game did not outperform children in the control condition but did also not underperform when compared to children in the standard condition. These findings suggest that this newly developed and motivating game could be a useful preventive intervention for at-risk children in kindergarten to improve L-SS knowledge before learning to read. Further research is necessary to explore whether more extended intervention periods allow for transfer to reading skills.

Synthesis of results

Letter-speech binding in typical and atypical readers

Previous studies showed empirical support for a letter-speech sound binding deficit as the proximal cause of dyslexia (Aravena et al., 2013; Fraga González et al., 2015; Žarić et al., 2015). In accordance with approaches that emphasize the multidimensional character of developmental disorders (e.g., Astle & Fletcher-Watson, 2020; Cuthbert & Insel, 2010), our results revealed a complex network of multiple factors associated with different reading components. Phoneme awareness, the ability to manipulate individual sounds in spoken words, was associated with the accurate processing of letters and speech sounds, which was in turn associated with accurate reading in pseudowords. On the other hand, the automated convergence of letter and speech sounds was functionally related to rapid automatized naming, which was in turn associated with fluent reading. As only a direct association between reading accuracy in pseudowords and our letter-speech sound accuracy measure appeared, our results extend previous findings that accurate processing of letters and speech sounds is especially important in earlier phases of the reading development when words are sounded out letter by letter, but that mere knowledge of L-SS associations is not sufficient for fluent reading (Blomert, 2011). Instead, according to our network model and in line with previous studies, automatic integration of visual symbols with their corresponding speech sounds underlies fluent reading.

In addition to storing a unified percept of this audiovisual object (Blomert, 2011), evidence suggested that children with dyslexia are also impaired in learning letter-speech sound mappings (Aravena et al., 2013; Horbach et al., 2018). However, evidence for a letter-speech sound learning deficit mainly comes from neuroimaging studies (Blomert, 2011; Žarić et al., 2014; but see Breznitz, 2002; Yap & van der Leij, 1993) and most behavioral studies solely examined the learning outcomes (Aravena et al., 2013; Law et al., 2018). As a result, it remains unclear how individuals with and without dyslexia differ in their learning trajectories. Using an artificial paradigm allowed us to examine differences in learning new letter-speech sound mappings. In a general school-aged sample, we found that directing children toward the learning goal of the task significantly impacted the learning trajectory. That is, children were faster and more efficient in learning the new mappings compared to their peers who obtained implicit instructions. Our results showed the influence of top-down control on letter-speech sound learning, suggesting that this mechanism may act as potential contributor to the acquisition and development of reading.

To further our understanding of the influence of cognitive control processes in learning to read, we used the same paradigm in a clinical sample comprising children with dyslexia compared to typical readers. Children with co-occurring ADHD were included in both groups. Our study (Chapter 4) showed that, although children with dyslexia and typical readers followed a slightly different learning trajectory, there were no differences in either accuracy or response time during the learning task at a group level. Regarding the learning outcomes, we examined whether children with dyslexia and typical readers differed in identifying congruent versus incongruent stimuli pairs. On a group level, children responded faster to congruent stimuli pairs than incongruent ones. This aligns with some older studies that interpreted this as multisensory facilitation during processing congruent audiovisual stimulus pairs (Dijkstra et al., 1993; Herdman et al., 2006). This also corroborates findings of several masked priming studies that showed that recognition of a target word was faster when it was preceded by a prime word that shared characteristics with the target word compared to when it was preceded by an incongruent stimulus word (Brysbaert, 2001; Zeguers et al., 2014). Responding more accurately to congruent pairs might also be facilitated by the nature of the task. Each speech sound had only one corresponding symbol, whereas it could be paired with seven non-corresponding symbols. It might take fewer cognitive resources to identify the corresponding pair compared to the mismatched one, which could have various combinations. However, in contrast to Guerra et al. (In revision), we did not find significant differences between children with dyslexia and typical readers in either accuracy or reaction time. Contradictory results can be explained by two significant differences between the experimental designs used in the current study and that of Guerra and colleagues (In revision). First, Guerra et al. (In revision) used an asynchronous task. Symbols appeared on the computer screen for 1000 ms, and speech sounds were presented only 500 ms after the symbol disappeared. According to the Synchronization Theory proposed by Breznitz (2006), an asynchrony in processing speed of visual-orthographic and auditory-phonological information underlies word reading difficulties in individuals with dyslexia. According to this theory, synchronization of information coming from the visual-orthographic and auditory-phonological systems is crucial for accurate decoding (Breznitz, 2006). Asynchronous presentation of visual and auditory information might have widened the gap in processing the two modalities, leading to lower performance in children with dyslexia in the study of Guerra et al. (In revision). In addition, the ability to sustain attention is critical in tasks where symbols and speech sounds are presented asynchronously. In our study, symbols and speech sounds appeared simultaneously. The absence of differences between children with dyslexia and typical readers in our study compared to Guerra et al. (In revision) might, therefore, either point toward widened asynchrony in processing speed of visual and auditory information or difficulties in sustained attention instead of differences in multisensory integration per se. Second, in contrast to our study, Guerra et al. (In revision) still provided children with feedback on their performance during the congruency task. Therefore, children continued to show progress in learning, evidenced by differences in accuracy that became especially apparent in the last two-third of the congruency block. These results suggest that differences between children with dyslexia and typical readers might especially appear with continued practice.

In the subsequent reading tasks, in which children were instructed to decode artificial words using the newly learned symbol knowledge, we did not find significant differences between children with dyslexia and typical readers. This contrasts with a previous study by Aravena et al. (2013). It is important to note that in the study by Aravena and colleagues (2013), most children could name all letters after the training. However, their training consisted of two 30-minute learning blocks. After each block, symbol knowledge was assessed, and after the two learning blocks (i.e., 1 hour of learning in total), the reading task within the artificial script was assessed. In the current study's 20-minute design, typical readers did not obtain high scores in word reading, similar to children with dyslexia. It can, therefore, be argued that the developmental shift from letter-speech sound integration to automatization, which is assumed to underlie fluent reading, did not yet happen in our study given the relatively short learning phase (i.e., 20 minutes). Differences may become more apparent in longer learning tasks or when a dynamic assessment is used to ensure that children reach a certain accuracy level before continuing with subsequent reading tasks. In addition, it is worth examining whether differences in artificial decoding, as reported by Aravena et al. (2013), also appear in non-time-limited tasks. As studies reported the influence of perceived consequences of dyslexia on the actual reading ability (Bazen et al., 2022; Elbro, 2010), it might be that children with dyslexia experience more pressure under time-constrained conditions compared to typical readers given previous reading failure, and, therefore are more prone to errors. Furthermore, neuroimaging studies might be more sensitive to early differences in letter-speech sound integration, even in the absence of behavioral differences. Karipidis et al. (2017) showed that brain responses to learned audiovisual mappings became apparent after less than 30 minutes of learning. The authors found an association between individual learning rate and congruency effects in the brain. More specifically, faster learning led to more pronounced congruency effects in the left inferior temporal gyrus, that is, more pronounced activity for congruent versus incongruent audiovisual stimuli.

In sum, although children with dyslexia slightly differed in learning trajectories, we found no overall significant differences in learning new letter-speech sound mappings or learning outcomes in contrast to previous studies. However, the learning outcomes of our artificial task were related to word reading ability in children's native language. This indicated that such an artificial design holds promise to further investigate the developmental shift from letter-speech sound integration to consolidated audiovisual objects and, importantly, identify factors that impede or facilitate this process. Moreover, the fact that such a paradigm is devoted to learning rather than to the level of skill already obtained makes it a useful tool to predict individual differences in reading performance and future gains in reading intervention (Aravena et al., 2018; Horbach et al., 2018). Future work might consider to test this paradigm using multiple learning sessions or using an adaptive design in order to obtain a certain accuracy level before continuing with the subsequent reading tasks. To examine learning retention, it might be of great interest to include test moments after a longer period. In addition, one might consider to familiarize children with the lab environment or to examine performance with and without EEG cap to shed light on the influence of unfamiliar testing conditions.

Role of executive functions

In order to behave adaptively and update previously learned information, one needs to direct attention according to strategic goals. Previous studies have reported difficulties in executive functions in individuals with dyslexia, such as verbal working memory, task switching, and attentional control (Thompson et al., 2015). However, these studies typically assessed performance in non-linguistic tasks, limiting our understanding of how these relate to the proposed letter-speech sound automatization deficit reported in individuals with dyslexia (Blomert, 2011). In our network (Chapter 2), attention was negatively associated with letter-speech sound accuracy and positively associated with reading accuracy in pseudowords. However, as discussed in Chapter 2, these associations disappeared when estimating a more conservative network (i.e., which assumedly yields fewer spurious edges). We attributed this result to the binary attention measure, in which parents indicated whether their child suffered from attentional difficulties, ignoring significant variation within the variable. To obtain more insight into how attentional problems impact reading development, we employed an artificial learning task in which children were required to learn new L-SS correspondences. This task was thought to mimic the initial formation of the neurocognitive reading network. Several components that are typically classified under the term ‘executive function’ were assessed. First, we manipulated instructions in order to manipulate goal-directedness during learning. In a general school-aged sample, including a broad range of reading levels, we found clear evidence in favor of instructions that revealed the goal of the task and directed learners toward that goal (Chapter 3). Besides faster and more efficient learning, goal-directed children appeared to profit more from consolidation during sleep compared to their implicitly instructed peers. These results corroborated the findings of Bitan & Booth (2012), that reported differences in offline improvement depending on the instruction type. However, the authors used a different experimental design, in which one condition was presented with the complete target word written in the unknown orthography while hearing the pronunciation (i.e., whole-word instruction), and the other condition was presented with the whole-word instruction preceded by the individual letters with their corresponding Latin translation (i.e., letter instruction). The authors argued that it is easier to focus on small segments rather than memorizing large units in the early stages of reading development. As reading unfamiliar words is thought to be facilitated by letter-by-letter decoding, the letter instruction condition appeared to be more effective than incidental learning of letters. According to Cartwright (2012), this might be because clear instructions help children to manage complex features of spoken and printed language, facilitating the development of word reading and comprehension. This is also in line with the advantage of explicit classroom instruction that is needed for beginning readers (de Graaff et al., 2009; Shapiro & Solity, 2008).

To further deepen our understanding of cognitive control mechanisms in learning to read, we conducted the same task in children with dyslexia compared to typically developing peers (Chapter 4). We were specifically interested in how children with and without dyslexia processed feedback and how instruction manipulation would alter feedback processes. On a behavioral level, we did not find a main effect of instruction. That is, children who obtained goal-directed instructions were not overall faster and more efficient in learning compared to their implicitly instructed peers, and did not obtain higher scores on outcomes measures. Comparing the learning trajectories between the two studies revealed that goal-directed children performed worse than in the previous study. The implicit condition performed slightly better, although still stagnating around an accuracy level of around 65%, yielding no significant differences between the conditions. The following reasons can possibly explain these contradictory results. First, goal-directed instructions are thought to activate selective, goal-directed behavior. Uncomfortable situations, like wearing an EEG cap in a university lab, might have interfered with attentional mechanisms that typically maintain goal-directed behavior. This aligns with recent studies that suggested that emotional and motivational factors modulate attentional selection and, thus, related task performance (Vuilleumier, 2015). Likewise, as the learning task was part of a more extensive test session that lasted approximately 3 hours, some children might have needed more motivation to obtain the goal. A growing body of evidence showed that even when one is not aware of it, intrinsic motivational properties could influence cognitive processes (Aarts et al., 2008). The current design does not allow for disentangling emotional and motivational influences from the attentional influence on cross-modal learning. Future studies should consider including subjective measures to assess these aspects, giving us a more detailed insight into their dynamic interplay.

Although children with dyslexia, overall, did not significantly differ from typical readers in learning and using new letter-speech sound correspondences, there seemed to be a trend of a benefit of goal-directed instructions in children with dyslexia. In this light, it is important to raise the following points. First, learning trajectories suggested that children with dyslexia who obtained implicit instructions did not learn much after the second bin, stagnating around an accuracy level of below 70%. In contrast, implicitly instructed typical readers still increased in later bins, showing almost similar learning trajectories to their goal-directed peers. This indicates that typical readers were almost equally efficient in learning without knowing the goal of the task. In addition, it is important to note that typical readers were slightly younger than the dyslexic group. The age range from 7-12 years old is a critical period in developing neural mechanisms for explicit learning, with explicit mechanisms developing later in childhood (Parkin & Streete, 1988; Zelazo et al., 1996). Therefore, although we did not find significant group differences in the current sample, it is possible that an age-matched typical reading sample would have outperformed children with dyslexia in the goal-directed condition. Second, children with dyslexia will likely be more familiar with testing situations. The children included in the current study experienced one or more diagnostic assessment sessions besides additional reading support and testing at school. Therefore, it is not unlikely that typical readers who are less exposed to these test environments perform below their actual ability. However, it might have been easier for some to perform in a quiet one-on-one lab environment compared to noisy classrooms. Guerra et al. (2021) proposed that auditory attention is one of the underlying factors of difficulties in dyslexic readers. The authors showed that louder background noise led to slower reading and disrupted reading comprehension. Especially in children with dyslexia with co-occurring attentional difficulties, noisy classroom environments might highly impact their reading ability, potentially affecting their grades. Our results also indicated associations between attention measures (including inattention, impulsivity, and hyperactivity) and learning outcomes of our artificial learning task. These results suggest that attentional difficulties may also limit the acquisition and development of alphabetic knowledge, on top of the influence of reading performance. Thus, although children with dyslexia seemed to perform equally well in our artificial learning task, it must be borne in mind that a quiet lab environment does, unfortunately, not represent reality.

For efficient learning, we need to evaluate the consequences of our behavior and adjust it accordingly (Ridderinkhof et al., 2004). Individuals with dyslexia tend to make inconsistent reading errors and often fail to correct them, even after obtaining feedback (Breznitz, 1987). This led to the hypothesis that an impaired performance monitoring is associated with difficulties to obtain fluent reading skills (Horowitz-Kraus & Breznitz, 2008). On a neural level, several studies reported lower ERP amplitudes after error commission in individuals with dyslexia in children (Horowitz-Kraus, 2011; Kraus & Horowitz-Kraus, 2014; Van De Voorde et al., 2010) and adults (Horowitz-Kraus, 2016b; Horowitz-Kraus & Breznitz, 2008). Although both teaching strategies and treatment programs highly rely on processing external feedback, studies examining feedback-locked ERPs in the domain of dyslexia research are very scarce. To fill this gap, we examined feedback-locked ERPs in children with dyslexia and typical readers during an artificial learning task.

ERP results showed no overall main effects between the two groups. In a previous study, Horowitz-Kraus (2011) reported more similar error-related negativity amplitudes in adolescents with dyslexia and typically reading peers than in adults with dyslexia compared to their peers. Another study by Kraus & Horowitz-Kraus (2014) suggested that differences in feedback-locked ERP responses between children with dyslexia and typical readers might also be diminished compared to previously reported results in adolescents. This is likely the result of the still-developing anterior cingulate cortex, which is the source of performance monitoring mechanisms (Gerber et al., 2009). However, we found that peak-to-peak feedback-related negativity (FRN) amplitudes decreased in the later task trials but only when children received goal-directed instructions. One could interpret decreased amplitudes as fewer attentional resources that are directed toward feedback stimuli due to fatigue and reduced motivation, or alternatively that feedback becomes less informative as learning increases (Holroyd & Coles, 2002). Concerning the behavioral learning curves, implicitly instructed learners tend to stagnate after the second bin, whereas goal-directed learners still increase in accuracy toward the end of the task, so we argue that the latter explanation with feedback becoming less informative is the most likely. In addition, we found higher peak-to-peak FRN amplitudes in children with dyslexia who obtained goal-directed instructions compared to their implicitly instructed peers, whereas the opposite pattern appeared in typical readers; goal-directed children showed lower peak-to-peak amplitudes than implicitly instructed learners. In children with dyslexia, goal-directed instructions might have led to higher brain responses as participants paid closer attention to feedback when they knew how to use it. In contrast, lower peak-to-peak to amplitudes in goal-directed typical readers compared to their implicitly instructed peers might either indicate that typical readers do not optimally profit from goal-directed instructions, or that typical readers do not necessarily need goal-directed instructions to learn efficiently. The former might result from the slight age difference between children with dyslexia and typical readers. As explicit mechanisms are still developing, goal-directed instructions in younger children might have a different effect compared to older children. Another suggestion is that FRN amplitudes decrease over learning due to participants shifting toward an internal error detection mechanism system (i.e., the error-related negativity). When participants learn, external feedback becomes less informative, as participants can internally detect when they commit an error (Holroyd & Coles, 2002). The fact that the interaction between these components seems to depend on the specific task further complicates clear conclusions at this time (Luft, 2014). Future studies should therefore consider examining the ERN and FRN in the same group of individuals with dyslexia to fill this knowledge gap.

Digitalization of intervention

With the increase in digital use, questions have been raised about whether interventions can be provided through digital media or online tools. First, many serious games have been developed, claiming to be effective learning tools without providing empirical evidence. To contribute to the existing body of research, we examined the effectiveness of a game-based intervention in prereaders. The game “Klankkr8”, which is a further adaptation of the earlier game “Kosmos Klikker”, was developed as a collaboration between RID (clinical center for learning difficulties in the Netherlands), the Rudolf Berlin Center of the University of Amsterdam, TU Delft, and Game Tailors. Krankkr8 systematically introduces Dutch letter-speech sound mappings and then provides training of these mappings using massive exposure. Baseline measures were assessed at the beginning of the last year of kindergarten, and the short, intensive intervention was provided immediately after (see Chapter 7). We provided evidence for a behavioral increase in L-SS knowledge in prereaders who played the game’s standard audiovisual version of the game. However, improved L-SS knowledge did not transfer to word reading skills. This aligns with a recent review by McTigue et al. (2020). The authors reviewed the effect of “GraphoGame”, another game aimed at improving elementary reading skills, on word reading in 19 studies. Although most studies reported a significant increase in sub-lexical skills, only studies involving adults that provided technological and motivational support reported generalization to word reading measures. In the Netherlands, children are only getting familiar with what reading entails from the second half of the last year of kindergarten onwards, meaning that most of them did not know how to use the letters for reading. Children might be more goal-directed toward applying the L-SS knowledge in decoding words when such a game is offered at the time when reading instruction starts at school. In addition, knowledge might more easily translate to word reading if the game illustrates how specific speech sounds are embedded within Dutch words.

Another way to induce a shift from letter-speech sound knowledge to accurate reading was thought to be handwriting. Evidence showed that providing handwriting practice to letter learning leads to better integrated audiovisual objects (Guan et al., 2021; James, 2017). The benefit of handwriting is hypothesized to follow various routes. First, handwriting increases activation in brain areas essential for proficient reading (James & Engelhardt, 2012). Second, the visual output of the handwriting movement might facilitate neural specialization for letters and words (James & Atwood, 2009). Third, handwriting requires greater attention to the shape of the letter, potentially facilitating the integration of multisensory information (Talsma et al., 2010). Contrary to our expectations, our experimental group did not show significant improvement compared to the control condition. Interestingly, they did not perform significantly worse than the standard audiovisual condition. We attributed the lack of a significant improvement to two major aspects. First, as we kept the amount of play time between the three conditions constant, children in the motoric condition (i.e., including handwriting) had less time to practice tapping the bullets and thus less exposure to the L-SS correspondences. Second, learning to write is a gradual process that requires actual practice, and fine motor skills are still developing in prereaders (Feder & Majnemer, 2007). As a result, directing cognitive capacity toward the handwriting movement might have hindered the development of integrated audiovisual objects. Therefore, instead of concluding that handwriting does not contribute to efficient coupling of letter-speech sound mappings, we argue that one should be cautious with excluding classical pen and paper writing from literacy education. A longer intervention period might be needed to see the effect of handwriting practice on the integration of letters and speech sounds. In addition, future studies can shed light on the effect of handwriting in children who have already developed fine motoric skills, and might want to examine the effect of using a touchscreen pencil instead of using their fingers to better represent real writing practice at school.

As a result of the global COVID-19 pandemic in 2020, specialized reading treatments massively moved to online platforms, raising questions concerning the effectiveness of treatment provided via online platforms compared to conventional face-to-face sessions. Although tele-practice has been used in language assessments long before the pandemic (e.g., Grogan-Johnson et al., 2013), strong evidence indicating similar intervention outcomes in online versus face-to-face reading treatment is scarce. A pilot study by Kohnen et al. (2021) examined whether children with poor reading skills could improve their reading scores through online provided intervention. In addition, a case study by Wright et al. (2011) reported on intervention gains after a remotely provided treatment on reading ability and comprehension in a 10-year-old child. Both studies reported enhanced reading performance after the intervention, but small sample sizes and the lack of control conditions limit the generalizability of their findings. By using Bayesian methods, our results suggest that online intervention is equally effective as the traditional face-to-face method. Children were instructed to complete five homework sessions each week. We found that both conditions completed a similar number of homework sessions. As treatment adherence has been found to be an important predictor of intervention outcomes (Mausbach et al., 2010), our results suggest that online treatment can be equally effective in a reading treatment context, as long as treatment adherence is guaranteed.

In sum, the results reported in this thesis are promising in relation to digitized interventions. Digital resources might be significant in preventing and intervening on reading difficulties. A multimodal approach that combines digital learning with traditional teaching methods might be the most promising as it allows for personalized, flexible learning while providing systematic instruction and support, but further research is needed to consider individual differences and limitations of digital tools. Since dyslexia frequently co-occurs with other developmental problems, such as ADHD (Margari et al., 2013), future studies should examine whether some children benefit more from online treatment than others. In addition, a randomized controlled trial should be conducted to validate whether our results can be generalized to other treatment methods and languages.

Complexity of dyslexia: Core versus multiple deficit

The dominant framework in dyslexia research, the phonological deficit hypothesis, has been criticized for its inability to account for co-occurring disorders and for not explaining why some individuals with phonological difficulties do not develop dyslexia. The second chapter of this dissertation aimed to map the complexity of the reading network by exploring the associations between cognitive, environmental, and demographic variables related to dyslexia. Our results indicate the variety of processes involved in reading, with different factors being associated with different aspects of the reading process. Moreover, although not included in the network, there also seems to be a developmental shift in executive functions (Farah et al., 2021), potentially yielding differential associations between components of executive functions and different phases of reading development. Processes such as attention, inhibition, and visuospatial short-term memory seem to be most predictive in prereaders, whereas beginning readers struggle more with working memory, planning, and shifting (Farah et al., 2021). Overall, this highlights the need for more comprehensive and nuanced models that account for the variety of factors involved in reading development and dyslexia.

Although we used a cross-sectional approach, each individual might have a personal network with its own nodes and edges (Borsboom, 2017). This aligns with the idea of neurodiversity, in which all individuals (also the ones that are assumed to be 'typically developing'), differ to some extent. Differences in the network might give rise to individual differences in observed symptoms. In networks with strongly connected nodes, one deficit might quickly affect other nodes. In contrast, in a weakly connected network, the same deficit does not significantly influence other nodes (Borsboom, 2017). According to these individual networks, findings that dyslexic readers perform worse on a particular task compared to typical readers do not automatically mean that this mechanism causes dyslexia or that similar outcome measures between the groups do automatically mean that there is no impairment. Instead, it nicely aligns with the behavioral heterogeneity reported in dyslexia. For example, while some might suffer from a binding deficit, this is not necessarily true for all dyslexic readers, potentially explaining mixed results (e.g., Aravena et al., 2013; Law et al., 2018; Nash et al., 2017). In addition, it may relate to the wide variability in intervention response. Concerning our network model (Chapter 2), intervening on phonics and L-SS mappings mainly leads to enhanced accurate reading performance. A more detailed insight into how L-SS mappings and word reading fluency influence each other (Chapter 5) revealed that treatment focusing on letter-speech sound mapping indeed improved letter-speech sound accuracy and speed, which improved in concert with reading fluency. These results highlight letter-speech sound integration as an important mechanism of change during the intervention. Interestingly, this training was especially effective for children who start the intervention with weaker L-SS associations, as they showed greater improvement during the training. In contrast, children might show reduced responsiveness when their reading deficit is mainly related to phoneme awareness. This heterogeneity within a sample or subgroups for which the interrelationships between the factors differ can be captured by the network model approach. Measuring an individual at consecutive time points offers an understanding of person-specific dynamics, providing valuable insights into individual strengths and pitfalls which can potentially be targeted in individualized support (Epskamp et al., 2018).

Despite the observed complexity in reading acquisition and development, research tends to hold on to parsimonious models that fail to account for the various factors that are associated with reading development. As argued by Cramer et al. (2010), "*the disorder as a whole cannot be fully understood by analyzing its individual components*". Core-deficit accounts will likely remain appealing given their relative simplicity (Astle & Fletcher-Watson, 2020), but the field requires more complex analysis methods. Our network results aligned with earlier theories, suggesting the valuable application of such complex analyses. Network theory holds the premise to model the dynamic nature of disorders by employing a time-series approach, which allows examining the developmental shift from ‘typical’ toward ‘disordered’, as well as how a specific intervention modifies the network. Shedding light on the highly dynamic development of reading by employing such complex models potentially offers a roadmap of paths that can be targeted in early, tailored reading intervention.

Implications and future directions

The studies described in this thesis mostly rely on cross-sectional designs. However, using time-series network approaches may provide essential insights into the dynamic development of reading. Such an approach can offer a roadmap of paths that can be targeted in reading intervention, as well as indicate when networks of poor and good readers start to differ, and how brain maturation and developmental changes affect performance. For example, attentional processes might be fundamental at the beginning of reading acquisition or in specific stages of reading development (e.g., when children shift from letters to words or from words to sentences). By better understanding attentional processes and their role in learning to read, we can identify children who struggle with reading, and who require additional support to promote successful reading outcomes. Likewise, a recent study by Harris et al. (2022) argued that impaired performance monitoring might result from poor reading skills instead of a performance monitoring deficit leading to reading difficulties. More specifically, the authors suggested that in more complex linguistic tasks, the ERP response depends upon the knowledge of the task. For example, in a reading task, not noticing a reading error might be the result of reduced orthographic knowledge of individuals with dyslexia, therefore leading to reduced ERP amplitudes after error commission (Harris et al., 2022). In our artificial learning task, it is, therefore, highly relevant to compare neural measures of good learners versus poor learners to account for differences in task knowledge. Longitudinal designs can provide more insight into the reciprocal associations between feedback processing and reading development.

The effectiveness of reading interventions for children with dyslexia has been well-established. However, it is essential to consider that there are many inter-individual differences among children with dyslexia, and what works for one individual may not work for another. Likewise, many different underlying patterns can lead to dysfluent reading. A shift in diagnosis and treatment is needed to address these issues. For example, co-occurring disorders such as ADHD and Developmental Language Disorder are often excluded from specialized reading treatment. The primary goal of diagnostic assessment is still to provide insights into the severity and the persistence of the reading difficulties along with the specificity of the reading and spelling problems. However, shedding light on children's inhibiting and protective factors that play a role in the development and maintenance of the difficulties can provide clinicians with early identification signatures for improved diagnosis and intervention. Future studies should examine how phonics-based dyslexia treatment is generalizable to children with co-occurring difficulties and how it should be adapted to their needs. In addition, it is essential to consider compensatory mechanisms as well, as they might decrease the liability of developing dyslexia (Pennington, 2006; van Bergen, van der Leij, et al., 2014).

The digitalization of interventions is of great value in the 21st century and has significant advantages. First, most children in Western countries grow up with digital resources and are already familiar with digital tools. Using digital tools during intervention might, therefore, facilitate motivation and engagement (Neumann, 2018). Second, families living in rural areas have limited access to health providers, or visiting a clinic multiple times a month is unfeasible due to poor public transport or heavy traffic (e.g., Kohnen et al., 2021). Moreover, as there is less need to rent spaces to provide therapy, providing online intervention can be accompanied by reduced intervention costs, which is especially important for people of low socio-economic classes. Another advantage is reduced travel time for patients and therapists, potentially contributing to general mental health and quality of life (Roberts et al., 2011). However, online interventions also come with disadvantages, such as disparities in internet access, less control on what the child pays attention to during intervention sessions, and required caregiver supervision in young children (Arnold et al., 2022; Zemlak et al., 2021). This might also be the case in children with attentional problems, whereas children with social anxiety might experience online interventions as more comfortable than real-life experiences. Future studies should shed light on for whom online-provided interventions are more or less suited.

General conclusion

The studies included in this doctoral thesis expanded our understanding of the complexity of factors involved in the acquisition and development of reading. In a broad, school-aged sample, we found evidence for the benefit of goal-directed instructions in acquiring and consolidating new letter-speech sound mappings, shedding light on how top-down control influences the initial phase of reading development. Employing the same task in children with dyslexia compared to typical readers indicated wide inter-individual learning differences and outcomes. In addition, although previous research reported a performance monitoring impairment in individuals with dyslexia, we did not find significant differences at the group level. Interestingly, behavioral attentional measures were associated with learning outcomes and reading within the native language, suggesting that attention does not only influence reading performance but may also impact the acquisition of reading skills. Examining the underlying mechanisms of change during a specialized dyslexia intervention revealed that especially children with poor letter-speech sound integration at the start of the intervention improved the most from a phonics-based treatment. In line with the multiple deficit model, we argue that while some deficits are present in some, they are absent in others. Future research should step away from the search for the 'core' deficit underlying each neurodevelopmental disorder and aim to capture complex interrelationships that move beyond diagnostic boundaries. We have shown how new complex analysis techniques, such as network analysis, can profitably be used to study these complex multivariate interrelationships. Considering this multidimensionality in targeted interventions, such as including executive function training in children with attentional difficulties, may increase intervention outcomes. Understanding the complex nature of typical and atypical reading development is particularly relevant to educational and clinical practice, which are recently benefiting from new developments such as game-based or online interventions. Better identification of individual differences and understanding why and for whom these interventions work is of great importance for literacy policy, and for timely, effective therapeutic remediation strategies, as these can be better matched to children's needs.

Appendix A

Appendix to Chapter 2

Table S1

Overview of all Variables in the Original Clinical Database and Rationale why Variables Were Included in the Final Dataset

#	**Variable**	**More information**	**Included?**	**Reason**
1.	**Client ID**	Unique Client Identification Code	No	Not relevant and privacy sensitive information.
2.	**Cycle**	Number of times a client received a cycle of care at the clinic	No	Administrative variable. Not relevant for the network analysis.
3.	**DBC start date**	DBC = diagnosis-treatment combination. Start date of insured health care trajectory.	No	Administrative variable. Not relevant for the network analysis.
4.	**DBC end date**	End date of insured health care trajectory.	No	Administrative variable. Not relevant for the network analysis.
5.	**Year start DBC**	Year in which insured health care started.	No	Administrative variable. Not relevant for the network analysis.
6.	**Start date remediation**	Start treatment	No	Used to calculate the duration of the intervention period but not included in the network analysis.
7.	**Outtake decision**	Diagnostic code indicating whether or not treatment would be covered by the Dutch basic Health Care Act (and under which type of arrangement).	No	Not relevant for the network analysis.

8.	**Outtake date**	Date on which the outtake of the diagnostic assessment took place.	No	Not relevant for the network analysis.
9.	**Date diagnostic assessment**	Date of the first test session of the diagnostic assessment.	No	Not relevant for the network analysis.
10.	**End date diagnostic assessment**	Date of the last test session of the diagnostic assessment.	No	Not relevant for the network analysis.
11.	**Remediation status**	Categorical variable that indicated whether client already received remediation, was on a waiting list to receive remediation, stopped with remediation etc.	No	Administrative variable. Not relevant for the network analysis.
12.	**Acceptance for diagnostic assessment status**	Categorical variable that indicated whether client was accepted for diagnostic assessment	No	Administrative variable. Not relevant for the network analysis.
13.	**Date acceptance for diagnostic assessment**	Date on which client was accepted for diagnostic assessment.	No	Administrative variable. Not relevant for the network analysis.
14.	**Reason acceptance for diagnostic assessment**	Reasons for accepting (or not accepting) client for diagnostic assessment.	No	Administrative variable. Not relevant for the network analysis.
15.	**Location diagnostic assessment**	At which location diagnostic assessment was conducted, as data was collected at a nationwide, clinical center.	No	Clients’ postal code was used as a sociodemographic variable in the network analysis. As provision of care was strictly protocolized over locations, we did not expect this information to provide extra information.

16.	**Year registration client**	Year at which the client was registered at the clinic.	No	Administrative variable. Not relevant for the network analysis.
17.	**Insurance mother company**	Name of insurance mother company	No	Administrative variable. Not relevant for the network analysis.
18.	**Insurance name**	Name of insurance	No	Not relevant for the network analysis.
19.	**UZOVI-code insurance**	Unique healthcare insurer identification	No	Administrative variable related to insurance company. Not relevant for the network analysis.
20.	**Age**	Age in months	Yes	To control for differences in scores as a result of age.
21.	**School name**	Name of school	No	Privacy sensitive information and not relevant for the network analysis.
22.	**ZIP school**	Postal code of school	No	Clients’ postal code was used as a sociodemographic variable in the network analysis. As children in the Netherlands commonly go to a school in their neighborhood, we did not expect this variable to provide extra information above that of clients’ postal code.

23.	**School grade**		No	We did not expect this variable to provide extra information beyond that of clients' age.
24.	**School type**	Indicated the type of school: elementary, secondary, higher education	Yes	To select only children of elementary school to control for differences in test materials.
25.	**Hometown**	Town where the child lived	No	Privacy sensitive information.
26.	**Home ZIP**	Postal code home address	No	This variable was used to calculate the mean disposable income per household as a proxy of socio-economic status but home ZIP itself was immediately removed after calculation as this is privacy sensitive information.
27.	**Gender**		Yes	To examine gender differences.
28.	**Name child**		No	Privacy sensitive information and not relevant for the network analysis.
29.	**Birth date**		No	Used to calculate a variable age based on the birth date and the date of diagnostic assessment. Birth date itself was removed from our dataset as this is privacy sensitive information.
30.	**BSN number**	Citizen service number	No	Administrative variable and privacy sensitive. Not relevant for the network analysis.

31.	**Type of care**	Categorical variable that indicated whether care was reimbursed or private	Yes	To include only children that received reimbursed care in the sample.
32.	**Reading problems elementary school (yes/no)**	Qualitative information on school history of the child, provided by the parents.	No	We did not expect this variable to provide extra information above the reading test scores we included in the network analysis (see variable 146).
33.	**Spelling problems elementary school (yes/no)**	Qualitative information on school history of the child, provided by the parents.	No	We did not expect this variable to provide extra information above the spelling test scores we included in the network analysis.
34.	**Multiplication tables problems elementary school (yes/no)**	Qualitative information on school history of the child, provided by the parents.	No	As the focus was on dyslexia (i.e. reading and spelling problems), we did not include qualitative information on math skills.
35.	**Math problems elementary school (yes/no)**	Qualitative information on school history of the child, provided by the parents.	No	As the focus was on dyslexia (i.e. reading and spelling problems), we did not include qualitative information on math skills.
36.	**Extra attention reading elementary school (yes/no)**	Qualitative information on school history of the child, provided by the parents.	No	This variable was not informative as all children in the study received extra support at school prior to referral.
37.	**Remedial teaching reading (yes/no)**	Qualitative information on school history of the child, provided by the parents.	No	This variable was not informative as all children in the study received extra support at school prior to referral.

38.	**Extra attention spelling elementary school (yes/no)**	Qualitative information on school history of the child, provided by the parents.	No	This variable was not informative as all children in the study received extra support at school prior to referral.
39.	**Remedial teaching spelling (yes/no)**	Qualitative information on school history of the child, provided by the parents.	No	This variable was not informative as all children in the study received extra support at school prior to referral.
40.	**Extra attention math elementary school (yes/no)**	Qualitative information on school history of the child, provided by the parents.	No	As the focus was on dyslexia (i.e. reading and spelling problems), we did not include information related to math.
41.	**Remedial teaching math (yes/no)**	Qualitative information on school history of the child, provided by the parents.	No	As the focus was on dyslexia (i.e. reading and spelling problems), we did not include information related to math.
42.	**School change elementary school (yes/no)**		No	Used by the clinician to judge whether the child received adequate reading instruction. If not, the child was not accepted for clinical care (based on definition and health care protocol) and therefore this variable was not informative for this study.
43.	**Number of school changes**		No	Used by the clinician to judge whether the child received adequate reading instruction. If not, the child was not accepted for clinical care (based on definition and health care protocol) and therefore this variable was not informative for this study.

44.	**Absenteeism elementary school (yes/no)**		No	Used by the clinician to judge whether the child received adequate reading instruction. If not, the child was not accepted for clinical care (based on definition and health care protocol) and therefore this variable was not informative for this study.
45.	**Absenteeism reason**	Reason of absenteeism	No	Used by the clinician to judge whether the child received adequate reading instruction. If not, the child was not accepted for clinical care (based on definition and health care protocol) and therefore this variable was not informative for this study.
46.	**Start Dutch education (grade)**	Indicated in which grade the child started Dutch education.	No	Used by the clinician to judge whether the child received adequate reading instruction. If not, the child was not accepted for clinical care (based on definition and health care protocol) and therefore this variable was not informative for this study.

47.	**Grade retention (yes/no)**	Indicated whether the child had to retake a grade or not.	Yes	To examine relation between grade retention and reading variables.
48.	**Grade retention (in which grade)**	In which grade the child had to retake a grade.	Yes	Children in kindergarten often have to retake a grade when they are not ready for elementary school, therefore only children that had to retake a grade from Grade 3 onwards received a score of 1 on this binary variable.
49.	**Language(s) spoken at home**		Yes	Included as a binary measure for multilingualism (with only one language = 0, more than one language = 1)
50.	**Date anamnesis**	Date of the intake (anamnesis)	No	Administrative variable. Not relevant for the network analysis.
51.	**Forgetfulness (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	We did not expect this binary variable of parents' judgement to provide additional information above the scores of memory tests we included in our analysis.
52.	**Attentional problems (yes/no)**	Qualitative information provided by the parents during anamnesis.	Yes	Although this was a binary variable based on parents' overall judgement, we nonetheless decided to include it in the analysis given high co-occurrence ADHD and dyslexia.

53.	**Stomach pain (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between stomach pain and reading disabilities.
54.	**Headache (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between headaches and reading disabilities.
55.	**Stuttering (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between stuttering and reading disabilities.
56.	**Anxiety (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	As this was a binary variable based on parents’ overall judgement, we decided that this was too general and unreliable to be included in the analysis.
57.	**Insomnia (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	As this was a binary variable based on parents’ overall judgement, we decided that this was too general and unreliable to be included in the analysis.
58.	**Aggressive behavior (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	As this was a binary variable based on parents’ overall judgement, we decided that this was too general and unreliable to be included in the analysis.

59.	**Motivational problems at school (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	As this was a binary variable based on parents' overall judgement, we decided that this was too general and unreliable to be included in the analysis.
60.	**Social behavior problems (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	As this was a binary variable based on parents' overall judgement, we decided that this was too general and unreliable to be included in the analysis.
61.	**Perinatal problems (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	As this was a binary variable based on parents' overall judgement, we decided that this was too general and unreliable to be included in the analysis.
62.	**Gestational age**		Yes	Measure of prematurity, as preterm delivery influences reading development.
63.	**Neonatal incubation (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between neonatal incubation and reading disabilities.
64.	**Duration neonatal incubation**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between neonatal incubation and reading disabilities.

65.	**Birth weight**		Yes	Measure of prematurity as preterm delivery influences reading development.
66.	**Hospitalization during first year (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between hospitalization during first year and reading disabilities.
67.	**Duration hospitalization**		No	No indications present in the relevant literature to expect an association between hospitalization during first year and reading disabilities.
68.	**Sitting and walking (early/average/late)**	Qualitative information provided by the parents during anamnesis.	No	As this variable was based on parents' overall judgement, we decided that this was too general and unreliable to be included in the analysis.
69.	**Talking (early/average/late)**	Qualitative information provided by the parents during anamnesis.	No	As this variable was based on parents' overall judgement, we decided that this was too general and unreliable to be included in the analysis.
70.	**Family risk dyslexia (yes/no)**		Yes	Family risk important factor in predicting dyslexia.

71.	**Strabismus (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	We did not expect this variable to provide extra information above the cognitive measure for visual perception we included in the network (see variable 158). Furthermore, children with a profound visual handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.
72.	**Depth vision problems (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	We did not expect this variable to provide extra information above the cognitive measure for visual perception we included in the network (see variable 158). Furthermore, children with a profound visual handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.
73.	**Amblyopia (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	We did not expect this variable to provide extra information above the cognitive measure for visual perception we included in the network (see variable 158). Furthermore, children with a profound visual handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.

74.	**Amblyopia year detection**	Qualitative information provided by the parents during anamnesis.	No	We did not expect this variable to provide extra information above the cognitive measure for visual perception we included in the network (see variable 158). Furthermore, children with a profound visual handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.
75.	**Amblyopia correction (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	We did not expect this variable to provide extra information above the cognitive measure for visual perception we included in the network (see variable 158). Furthermore, children with a profound visual handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.
76.	**Blurry vision (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	We did not expect this variable to provide extra information above the cognitive measure for visual perception we included in the network (see variable 158). Furthermore, children with a profound visual handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.

77.	**Blurry vision year detection**	Qualitative information provided by the parents during anamnesis.	No	We did not expect this variable to provide extra information above the cognitive measure for visual perception we included in the network (see variable 158). Furthermore, children with a profound visual handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.
78.	**Blurry vision correction (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	We did not expect this variable to provide extra information above the cognitive measure for visual perception we included in the network (see variable 158). Furthermore, children with a profound visual handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.
79.	**Other vision problems (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	We did not expect this variable to provide extra information above the cognitive measure for visual perception we included in the network (see variable 158). Furthermore, children with a profound visual handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.

80.	**Hearing loss (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between hearing problems and reading disabilities. Furthermore, children with a profound hearing handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.
81.	**Hearing loss year detection**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between hearing problems and reading disabilities. Furthermore, children with a profound hearing handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.
82.	**Hearing loss correction (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between hearing problems and reading disabilities. Furthermore, children with a profound hearing handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.

83.	**Middle ear infection (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between middle ear infections and reading disabilities. Furthermore, children with a profound hearing handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.
84.	**Middle ear infection year detection**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between middle ear infections and reading disabilities. Furthermore, children with a profound hearing handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.
85.	**Ear tubes (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between having ear tubes and reading disabilities. Furthermore, children with a profound hearing handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.

86.	**Ear tubes year**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between having ear tubes and reading disabilities. Furthermore, children with a profound hearing handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.
87.	**Other hearing problems (yes/no)**	Qualitative information provided by the parents during anamnesis.	No	No indications present in the relevant literature to expect an association between hearing problems and reading disabilities. Furthermore, children with a profound hearing handicap will not receive the dyslexia diagnosis according to the DSM-5 definition.
88.	**School dyslexia judgment (yes/no)**		No	Administrative variable related to the reasons for referral by the school of the child.
89.	**Student tracking system (Dutch: LVS) recent reading test**	System to assess reading progress in elementary school.	No	Variables 89-124 refer to information concerning support at Tier 2 and 3 provided by schools when referring a child to the clinic. As schools use different procedures.

				and instruments, these data were neither provided by schools, nor entered into the database, in a standardized manner. Therefore, these variables were not included in the analysis.
90.	**Date student tracking system recent reading test**		No	See reason variable 89.
91.	**Student tracking system recent spelling test**	System to assess spelling progress in elementary school.	No	See reason variable 89.
92.	**Date student tracking system recent spelling test**		No	See reason variable 89.
93.	**Student tracking system math test**	System to assess math progress in elementary school.	No	See reason variable 89.
94.	**Date student tracking system math test**		No	See reason variable 89.
95.	**Pretest reading**	Measures reading level prior to remedial teaching period 1.	No	See reason variable 89.
96.	**Date pretest reading**		No	See reason variable 89.

97.	**Pretest spelling**	Measures spelling level prior to remedial teaching period 1.	No	See reason variable 89.
98.	**Date pretest spelling**		No	See reason variable 89.
99.	**Interimtest reading**	Measures reading level between remedial teaching period 1 and 2.	No	See reason variable 89.
100.	**Date interimtest reading**		No	See reason variable 89.
101.	**Interimtest spelling**	Measures spelling level between remedial teaching period 1 and 2.	No	See reason variable 89.
102.	**Date interimtest spelling**		No	See reason variable 89.
103.	**Posttest reading**	Measures reading level after remedial teaching period 2.	No	See reason variable 89.
104.	**Date posttest reading**		No	See reason variable 89.
105.	**Posttest spelling**	Measures spelling level after remedial teaching period 2.	No	See reason variable 89.
106.	**Date posttest spelling**		No	See reason variable 89.
107.	**Remedial teaching period 1 (reading/spelling/both)**	Categorical variable that indicates for which domain remedial teaching was needed.	No	See reason variable 89.
108.	**Grade at which remedial teaching period 1 started**		No	See reason variable 89.

109.	**Remedial teaching period 1 start and end date**	No	See reason variable 89.
110.	**Remedial teaching period 2 (reading/spelling/both)**	No	See reason variable 89.
111.	**Grade at which remedial teaching period 2 started**	No	See reason variable 89.
112.	**Remedial teaching period 2 start and end date**	No	See reason variable 89.
113.	**Remedial teaching period 3 (reading/spelling/both)**	No	See reason variable 89.
114.	**Grade at which remedial teaching period 3 started**	No	See reason variable 89.
115.	**Remedial teaching period 3 start and end date**	No	See reason variable 89.

116.	**Remedial teaching period 4 (reading/spelling/both)**		No	See reason variable 89.
117.	**Remedial teaching period 4 starting at grade**		No	See reason variable 89.
118.	**Remedial teaching period 4 start and end date**		No	See reason variable 89.
119.	**Remedial teaching Tier 3 (reading/spelling/both)**	School-based remedial teaching which intensively addressed specific needs of the child regarding reading/spelling.	No	See reason variable 89.
120.	**Remedial teaching Tier 3 duration in weeks**		No	See reason variable 89.
121.	**Remedial teaching Tier 3 minutes/week**		No	See reason variable 89.
122.	**Remedial teaching Tier 3 (individual/group)**	Indicated whether children received remedial teaching individually or in group.	No	See reason variable 89.
123.	**Profession remedial teacher**		No	See reason variable 89.
124.	**Remedial teaching Tier 3 method**	Indicated which method was used for remedial teaching.	No	See reason variable 89.

125.	**IWAL reading text**	Measured text reading skill (accuracy and speed).	No	This task was only administered to a subgroup of our sample (Grade 4 and older).
126.	**AVI-2009**	Old version of a reading test measuring the child's text reading level.	No	This task was only administered to a subgroup of our sample.
127.	**AVI-new**	New version of a reading test measuring the child's text reading level.	No	This task was only administered to a subgroup of our sample.
128.	**BRUS (One-minute test)**	Measured word reading, in terms of the number of words read correctly within one minute.	Yes	Measure of reading fluency used to assess intervention progress as this variable had more records than other reading fluency measures.
129.	**IWAL-Brus tach**	Computerized version of one-minute test in which words are shortly presented on the computer screen (10ms, 100ms, 300ms or 500ms).	No	This is an experimental task that was not systematically administered for the majority of clients and correlated strongly with the included word reading tasks.
130.	**IWAL Dictation (words)**	Children needed to write down 40 one-syllable words which were orally presented in short sentences to assess spelling skills.	No	The task was not systematically administered to children in Grade 4 and older as it measured the basics of Dutch spelling that are normally mastered in Grade 1 and 2.

131.	**IWAL Dictation (sentences)**	Children needed to write down 14 orally presented sentences to assess spelling skills.	No	The task was not systematically administered to children in Grade 4 and older as it measured the basics of Dutch spelling that are normally mastered in Grade 1 and 2.
132.	**IWAL Standard Dictation**	Children needed to write down 19 orally presented sentences that contained familiar words to all elementary school children	No	This test was only administered to a subgroup of our sample.
133.	**IWAL picture dictation**	Spelling task	No	This test was not part of the standard diagnostic protocol, and therefore only incidentally administered.
134.	**IWAL vocabulary**	Receptive Vocabulary task	No	We included the Peabody Picture Vocabulary Test as measure of receptive vocabulary, as the IWAL vocabulary task was only administered to a subgroup of our sample.
135.	**PI dictation**	Spelling task	No	This task was not part of the standard diagnostic protocol, and therefore only incidentally administered.
136.	**IWAL auditory analysis**	Participants had to say monosyllabic words phone by phone; phoneme awareness task.	No	We included 3DM Phoneme Deletion as an index for phoneme awareness as this task

				was not part of the standard diagnostic protocol and therefore only incidentally administered.
137.	**IWAL auditory interference test**	Two groups of monosyllabic words were orally presented to produce inter-list interference. After presentation of the two groups, participants needed to repeat the first group and after that the second or first the second group and after that the first.	No	This task was not part of the standard diagnostic protocol, and therefore only incidentally administered (replaced as part of the standard protocol by 3DM test battery in 2009).
138.	**Leidse Diagnostische Test (LDT) [Knox Cube Test]**	Children needed to imitate a tapping pattern on four cubes showed by the test administrator.		This test was not part of the standard diagnostic protocol, and therefore only incidentally administered.
139.	**LDT Sentence Imitation**	Children had to listen to a story and had to reproduce the story sentences one by one immediately after hearing them.	No	This test was not part of the standard diagnostic protocol, and therefore only incidentally administered.
140.	**LDT Story Questions**	Children had to answer questions based on the story that has been told in the Sentence Imitation subtest of the LDT.	No	This test was not part of the standard diagnostic protocol, and therefore only incidentally administered (only for ages 4-8 and therefore only 20 records).

141.	**CB&WL Rapid automatized naming**	Rapid automatized naming of letters, digits, objects, and colors.	No	3DM version of rapid automatized naming was used because this version contained more missing values.
142.	**Taaltests voor Kinderen (TVK) vocabulary**	Children needed to name 60 pictures.	No	This test was only administered to younger children.
143.	**TVK Word Closure**	Children needed to say a word in which one or more phonemes were deleted and replaced by a short silence.	No	This test was not part of the standard diagnostic protocol, and therefore only incidentally administered (replaced as part of the standard protocol by 3DM test battery in 2009).
144.	**TVK Word forms Production**	Each item comprised 2 sentences, of which the second sentence was not complete. Children needed to complete the second sentence with a word based on the first sentence (i.e. past participle form, diminutive or comparative). Measured word morphology.	No	This test was not part of the standard diagnostic protocol, and therefore only incidentally administered.
145.	**Peabody receptive vocabulary**		Yes	Measure of receptive vocabulary.

146.	**3DM reading test (high, low and pseudowords)**		Yes	Measure of word reading accuracy and fluency for high frequent, low frequent and pseudowords.
147.	**3DM spelling test**		Yes	Measures accuracy and speed of spelling knowledge.
148.	**3DM letter-speech sound identification**		Yes	Measures letter-speech sound associations, which is considered important in reading acquisition.
149.	**3DM phoneme deletion**		Yes	Proxy of phoneme awareness.
150.	**3DM rapid naming (letters and digits)**		Yes	Rapid naming test with most records.
151.	**ADIT-C**	Dutch speech discrimination test for children (Auditieve Discriminatie Test). Tests peripheral auditory perception.	No	No indications present in the relevant literature to expect an association between peripheral auditory perception and reading disabilities.
152.	**HADIT-C**	Same as ADIT-C but only half of the trials.	No	No indications present in the relevant literature to expect an association between peripheral auditory perception and reading disabilities.
153.	**ASI identical figures**	Visual perception test	No	This task was only administered to adults.

154.	**HSI identical figures**	Visual perception test	No	The task was only administered to children in secondary education.
155.	**ASI silhouettes**	Visual perception test	No	The task was only administered to adults.
156.	**ASI picture memory**	Visual memory test	No	The task was only administered to adults.
157.	**HSI silhouettes**	Visual perception test	No	The task was only administered to children in secondary education.
158.	**GSO perception**	Visual perception test	Yes	As a measure of visual perception.
159.	**WISC-III**	Intelligence test	Yes	Included as intelligence estimation as this was the only intelligence test that was administered to the full age range of our sample.
160.	**WAIS**	Intelligence test	No	WISC-III was chosen as intelligence estimation (This test was only administered to older children).
161.	**SSON**	Intelligence test	No	WISC-III was chosen as intelligence estimation.
162.	**Intelligentieschatter**	Intelligence estimation, based on four subtests of WISC-III.	No	WISC-III was chosen as intelligence estimation
163.	**Revisie Amsterdamse Kinder Intelligentie Test (RAKIT)**	Intelligence test	No	WISC-III was chosen as intelligence estimation

164.	**Groninger Intelligence Test**	Intelligence test	No	WISC-III was chosen as intelligence estimation (This test was only administered to older children).
165.	**Raven progressive matrices**	Intelligence test	No	WISC-III was chosen as intelligence estimation (Raven was only incidentally administered).
166.	**Raven colored matrices**	Intelligence test	No	WISC-III was chosen as intelligence estimation (Raven was only incidentally administered).
167.	**12 WT**	Dutch version of Rey’s Auditory Verbal Learning Test	No	Less records than 15 WT, not possible to combine two tests due to different number of items.
168.	**15 WT**	Dutch version of Rey’s Auditory Verbal Learning Test	Yes	Measure of short- and long-term verbal memory, had more records than 12 WT.
169.	**REY visual design learning test**	Children had to memorize 15 geometric stimulus cards in 5 successive trials.	No	This test was not part of the standard diagnostic protocol, and therefore only incidentally administered (Only 17 records).
170.	**Wechsler Memory Scale visual memory reproduction**	Visual memory test	No	This task was only administered to older children (only 112 records).

171.	**IWAL visual recognition**	Visual memory test	No	This test was not part of the standard diagnostic protocol, and therefore only incidentally administered (only 19 records).
172.	**IWAL finger agnosia**	Test for finger agnosia	No	No indications present in the relevant literature to expect an association between finger agnosia and reading disabilities.
173.	**Zelf-Beoordelings-Vragenlijst voor kinderen (ZBV-K)**	Dutch version of State-Trait-Anxiety Inventory for children (STAI-C)	No	This test was not part of the standard diagnostic protocol, and therefore only incidentally administered (only 352 records).
174.	**Competentie Belevingsschaal voor kinderen (CBSK)**	Self-perception profile for children, measured self-concept in children.	No	This test was not part of the standard diagnostic protocol, and therefore only incidentally administered (only 385 records).
175.	**Schoolvragenlijst (SVL) [School Attitude Questionnaire]**	Questionnaire that measured socio-emotional functioning and attitudes toward school.	No	This test was not part of the standard diagnostic protocol, and therefore only incidentally administered (only 22 records).
176.	**Prestatie Motivatie Test-Kinderen (PMT-K) [Achievement Motivation Test for Children]**	Questionnaire that measured achievement motivation, negative and positive test anxiety, and social desirability	No	This test was not part of the standard diagnostic protocol, and therefore only incidentally administered (only 80 records).

Figure S1

General Networks Depicting the Relationship Between Variables in the Framework of Reading Disabilities Estimated With a Tuning Parameter γ = 0 (Left Panel) And γ = 0.25 (Right Panel)

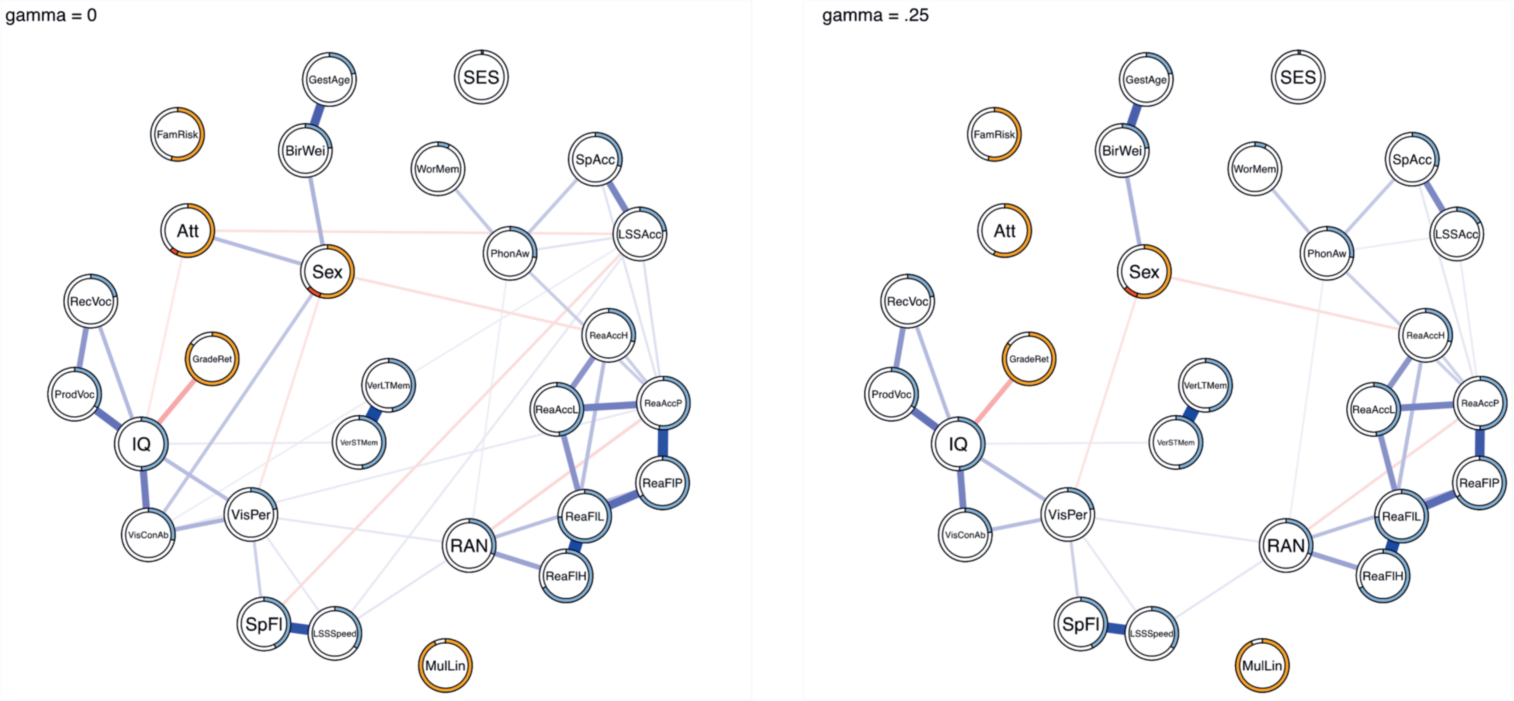
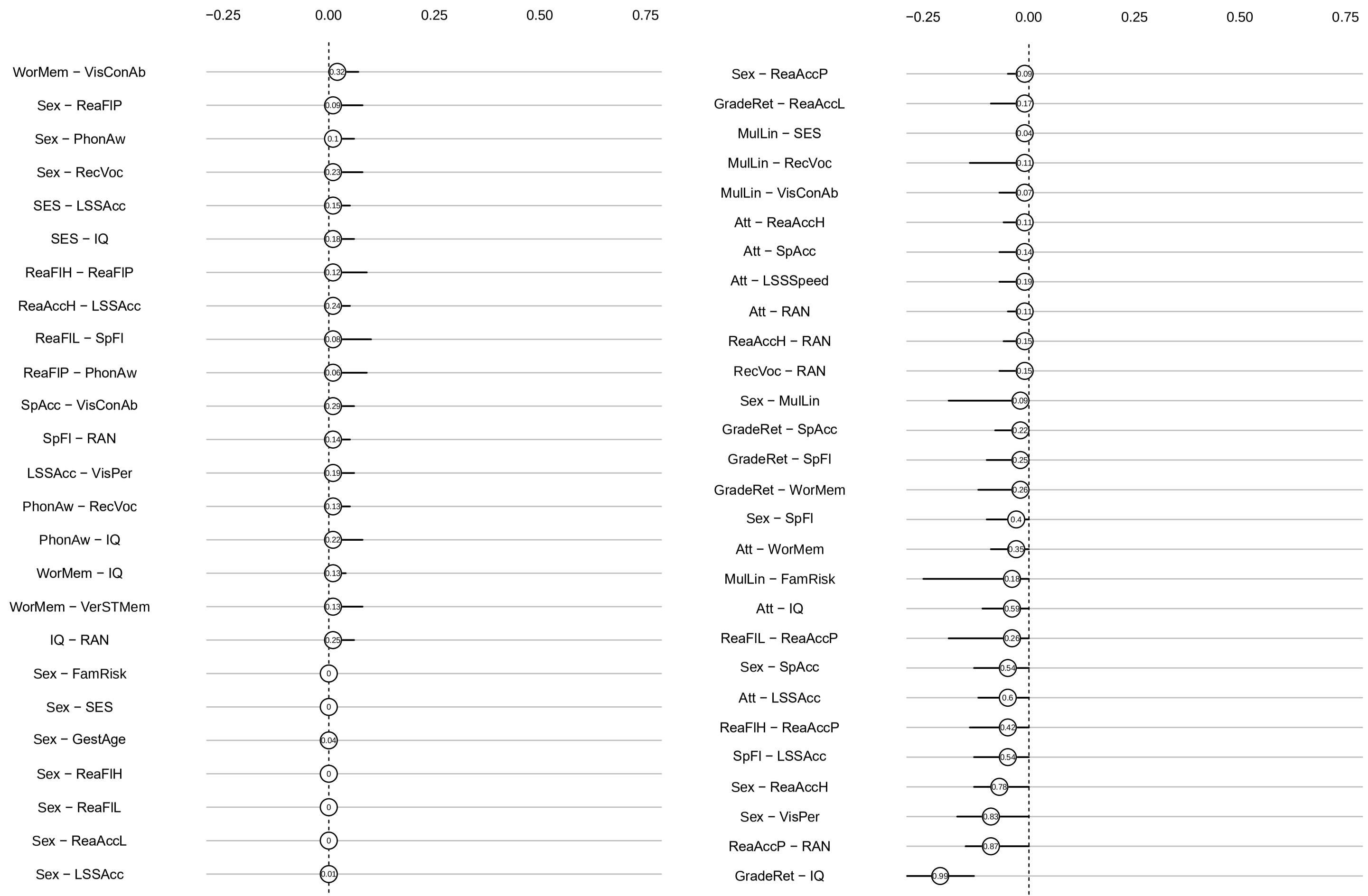


Figure S2

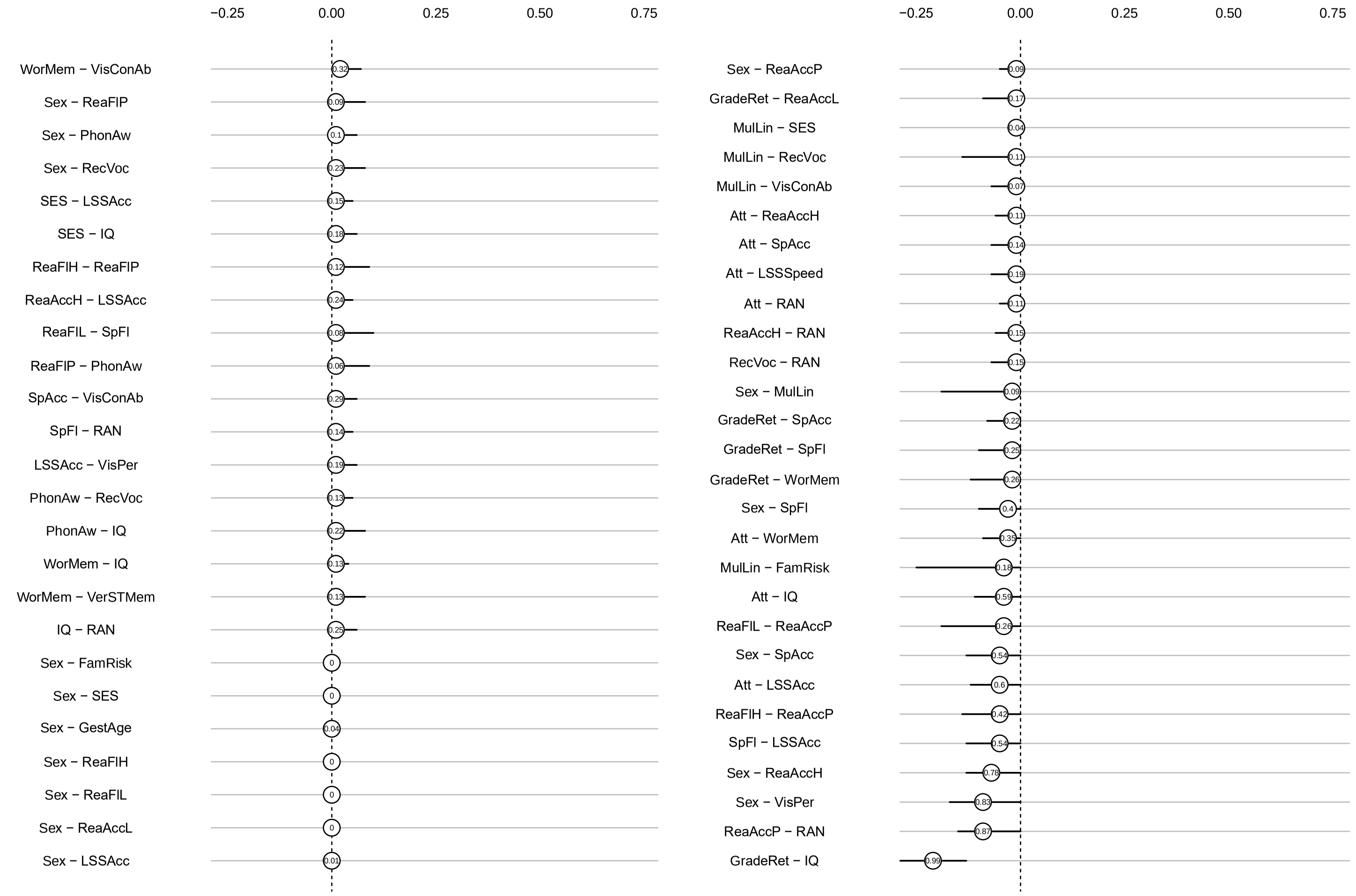
Bootstrapped Sampling Distribution of Edge Weights of the General Network



Note. Values indicate proportion of estimates whose absolute values were larger than zero. Black horizontal lines represent the 5% and 95% quantiles of the sampling distribution. For example, the edge weight between VerSTMem and VerLTMem was larger than zero in 100% of the bootstrap samples and its 5% and 95% quantiles lie around .55 and .65. Note that edges are ordered by the arithmetic mean of the sampling distribution in decreasing order and most edges that had a mean of (close to) zero were excluded for sake of brevity.

Figure S2 (continued)

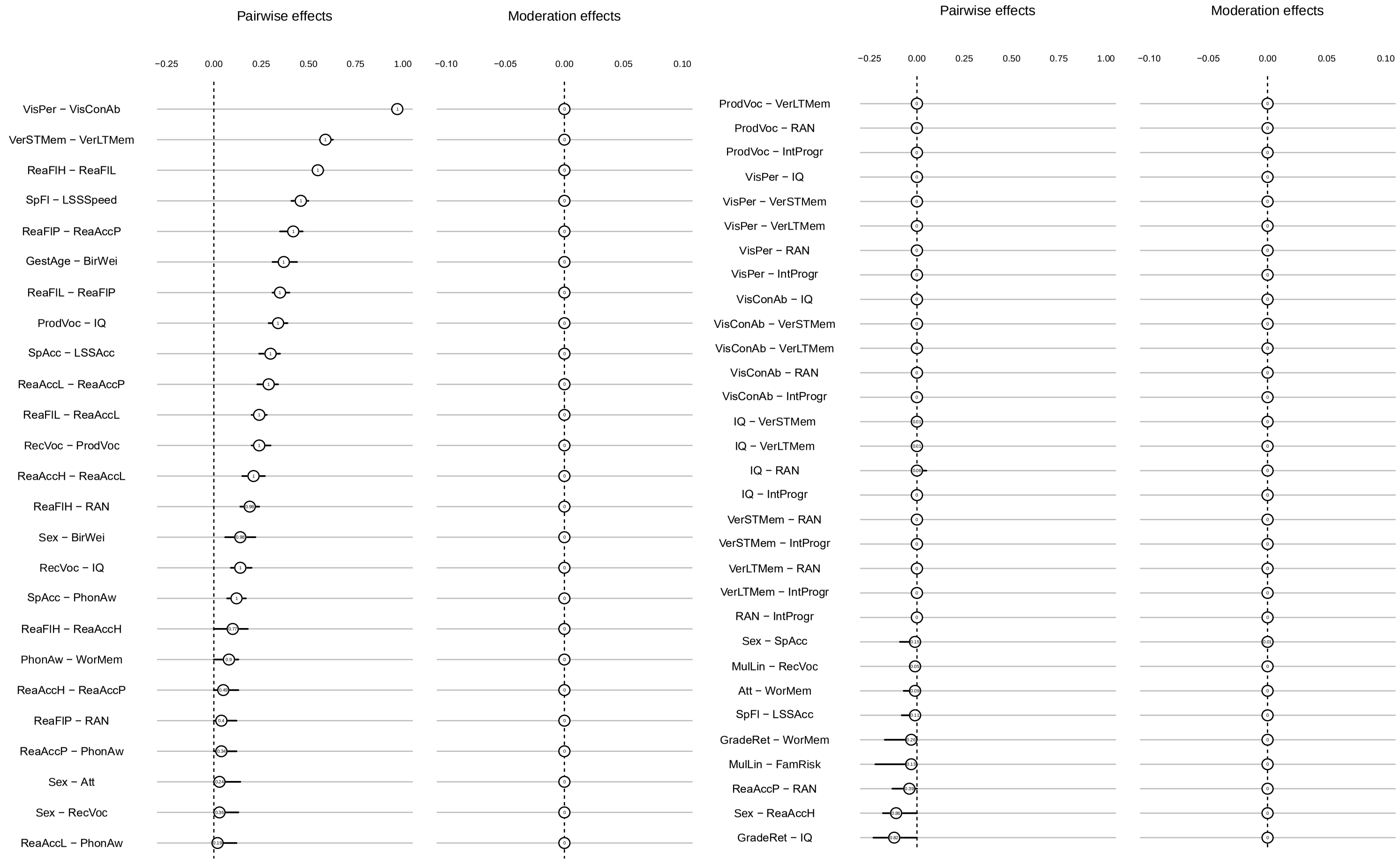
Bootstrapped Sampling Distribution of Edge Weights of the General Network



Note. Values indicate proportion of estimates whose absolute values were larger than zero. Black horizontal lines represent the 5% and 95% quantiles of the sampling distribution. For example, the edge weight between VerSTMem and VerLTMem was larger than zero in 100% of the bootstrap samples and its 5% and 95% quantiles lie around .55 and .65. Note that edges are ordered by the arithmetic mean of the sampling distribution in decreasing order and most edges that had a mean of (close to) zero were excluded for sake of brevity.

Figure S3

Bootstrapped Sampling Distribution of Edge Weights of the Intervention Network



Note. Pairwise effects are displayed in the first column and moderation effects in the second column. Bootstrapped sampling distribution of edge weights of the intervention network with in the first column pairwise effects and in the second column moderation effects. Values indicate proportion of estimates whose absolute values were larger than zero. Black horizontal lines represent the 5% and 95% quantiles of the sampling distribution. Moderation effects refer to the effect of the variable Intervention Progress on all pairwise interactions (e.g. moderation effect of Intervention progress on Visual Perception-Visuo-constructional abilities). Note that edges are ordered by the arithmetic mean of the sampling distribution in decreasing order and most edges that had a mean of (close to) zero were excluded for sake of brevity.

Appendix B

Appendix to Chapter 3

Table S1

Repeated Measures (M)ANOVA Results for Accuracy and Reaction Time with Condition (Goal-directed vs Implicit instructions) as Between-Subject Variable and Time (Bin 1 vs Bin 2 vs Bin 3 vs Bin 4) as Within-Subjects Variable

	Condition	**Time**	**Condition x Time**
MANOVA			
Block 1	$F(2,100) = 3.116$, $p = .049$, $\eta_p^2 = .059$	$F(6,606) = 18.462$, **$p < .001$**, $\eta_p^2 = .309$	$F(6,606) = 3.782$, **$p = .001$**, $\eta_p^2 = .072$
Block 2	$F(2,100) = 2.188$, $p = .117$, $\eta_p^2 = .042$	$F(6,606) = 12.002$, **$p < .001$**, $\eta_p^2 = .212$	$F(6,606) = 3.103$, **$p = .005$**, $\eta_p^2 = .060$
Block 3	$F(2,100) = 3.282$, $p = .042$, $\eta_p^2 = .062$	$F(6,606) = 1.386$, $p = .218$, $\eta_p^2 = .027$	$F(6,606) = 1.597$, $p = .145$, $\eta_p^2 = .031$
Accuracy			
Block 1	$F(1,101) = 5.922$, **$p = .017$**, $\eta_p^2 = .055$	$F(3,303) = 29.313$, **$p < .001$**, $\eta_p^2 = .225$	$F(3,303) = 6.176$, **$p < .001$**, $\eta_p^2 = .045$
Block 2	$F(1,101) = 3.587$, $p = .061$, $\eta_p^2 = .034$	$F(3,303) = 24.014$, **$p < .001$**, $\eta_p^2 = .192$	$F(3,303) = 6.166$, **$p < .001$**, $\eta_p^2 = .058$
Block 3	$F(1,101) = 6.626$, **$p = .012$**, $\eta_p^2 = .062$	$F(3,303) = 0.525$, $p = .666$, $\eta_p^2 = .005$	$F(3,303) = 1.781$, $p = .151$, $\eta_p^2 = .017$
Reaction time			
Block 1	$F(1,101) = 0.189$, $p = .665$, $\eta_p^2 =.002$	$F(3,303) = 17.393$, **$p < .001$**, $\eta_p^2 = .147$	$F(3,303) = 1.433$, $p = .233$, $\eta_p^2 = .014$
Block 2	$F(1,101) = 1.374$, $p = .244$, $\eta_p^2 =.013$	$F(3,303) = 2.868$, $p = .034$, $\eta_p^2 = .028$	$F(3,303) = 0.200$, $p = .896$, $\eta_p^2 = .002$
Block 3	$F(1,101) = 0.077$, $p = .782$, $\eta_p^2 =.001$	$F(3,303) = 2.331$, $p = .074$, $\eta_p^2 = .023$	$F(3,303) = 1.482$, $p = .219$, $\eta_p^2 = .014$

Note. The MANOVA was conducted on accuracy and reaction time together. P-values that remained significant after FDR correction are shown in bold. The Time variable was obtained by averaging reaction time and accuracy across 4 bins of 18 trials for Block 1 and 2, and of 14 trials for Block 3. Children learned four out of eight L-SS pairs in the first block, four new L-SS pairs in the second block, and all eight L-SS pairs together in the third block.

Table S2

Factorial (M)ANOVA Results for Accuracy and Reaction Time with Condition as Between-Subject Variable and Congruency and Day as Within-Subjects Variables

	MANOVA	**Accuracy**	**Reaction Time**
Condition	F(2,85) = 4.445, **p = .015**, η_p^2 = .095	F(1,86) = 8.975, **p = .004**, η_p^2 = .094	F(1,86) = 0.253, p = .616, η_p^2 = .003
Day	F(2,85) = 19.804, **p < .001**, η_p^2 = .318	F(1,86) = 35.57, **p < .001**, η_p^2 = .294	F(1,86) = 5.687, p = .019, η_p^2 = .062
Congruency	F(2,85) = 17.008, **p < .001**, η_p^2 = .286	F(1,86) = 5.690, p = .058, η_p^2 = .041	F(1,86) = 32.086, **p < .001**, η_p^2 = .272
Condition x Day	F(2,85) = 1.526, p = .223, η_p^2 = .035	F(1,86) = 0.46, p = .500, η_p^2 = .005	F(1,86) = 2.511, p = .117, η_p^2 = .028
Condition x Congruency	F(2,85) = 2.437, p = .094, η_p^2 = .054	F(1,86) = .272, p = .604, η_p^2 = .003	F(1,86) = 4.384, p = .039, η_p^2 = .049
Day x Congruency	F(2,85) = 2.726, p = .071, η_p^2 = .060	F(1,86) = 0.000, p = 1.000, η_p^2 = .000	F(1,86) = 5.478, p = .022, η_p^2 = .060
Condition x Day x Congruency	F(2,85) = 1.639, p = .200, η_p^2 = .037	F(1,86) = .902, p = .345, η_p^2 = .010	F(1,86) = 1.842, p = .178, η_p^2 = .021

Note. The MANOVA was conducted on accuracy and reaction time together. P-values that remained significant after FDR correction are shown in bold.

Table S3

Factorial (M)ANOVA Results for Symbol Knowledge, OMT Real Words and OMT Pseudowords with Condition as Between-Subject Variable and Day as Within-Subjects Variable

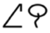
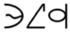
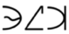
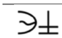
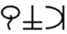
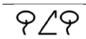
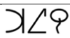
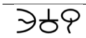
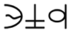
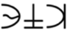
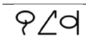
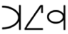
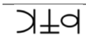
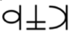
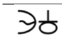
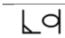
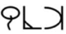
	Condition	**Day**	**Condition x Time**
MANOVA	$F(3,103) = 7.395$, **$p < .001$**, $\eta_p^2 = .177$	$F(3,103) = 16.680$, **$p < .001$**, $\eta_p^2 = .327$	$F(3,103) = 3.913$, **$p = .011$**, $\eta_p^2 = .102$
Symbol Knowledge	$F(1,105) = 19.672$, **$p < .001$**, $\eta_p^2 = .158$	$F(1,105) = 7.017$, **$p = .009$**, $\eta_p^2 = .063$	$F(1,105) = 0.103$, $p = .749$, $\eta_p^2 = .001$
OMT real words	$F(1,105) = 12.008$, **$p < .001$**, $\eta_p^2 = .103$	$F(1,105) = 48.686$, **$p < .001$**, $\eta_p^2 = .317$	$F(1,105) = 9.888$, **$p = .002$**, $\eta_p^2 = .086$
OMT pseudowords	$F(1,105) = 7.730$, **$p = .006$**, $\eta_p^2 = .069$	$F(1,105) = 25.559$, **$p < .006$**, $\eta_p^2 = .196$	$F(1,105) = 6.718$, **$p = .011$**, $\eta_p^2 = .060$
SUBGROUP SYMBOL KNOWLEDGE ≥ 5			
OMT real words	$F(1,64) = 0.272$, $p = .604$, $\eta_p^2 = .004$	$F(1,64) = 47.562$, **$p < .001$**, $\eta_p^2 = .426$	$F(1,64) = 3.706$, $p = .059$, $\eta_p^2 = .055$
OMT pseudowords	$F(1,64) = 0.042$, $p = .839$, $\eta_p^2 = .001$	$F(1,64) = 23.687$, **$p < .001$**, $\eta_p^2 = .270$	$F(1,64) = 2.385$, $p = .127$, $\eta_p^2 = .036$

Note. The MANOVA was conducted on all three dependent variables together. OMT = One-minute test. P-values that remained significant after FDR correction are shown in bold.

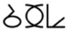
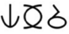
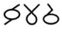
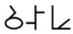
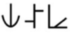
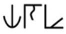
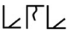
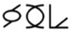
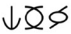
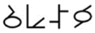
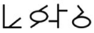
Table S4

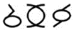
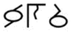
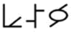
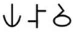
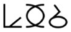
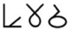
List of Artificial Words Resulting in Real Dutch Words or Pseudowords Used in Version A and B of the Artificial Learning Task with its Correct Answer and English Translation

VERSION A

Artificial Dutch	Correct answer	English translation	Artificial pseudo	Correct answer
	of	or		ot
	fout	wrong		foun
	zon	sun		zof
	zei	said		teif
	tot	until		fot
	zout	salt		zoun
	tent	tent		nent
	zijn	are		zijf
	ton	barrel		fon
	fijn	nice		nijf
	net	net		fet
	zou	would		oun
	en	and		et
	zet	put		tef

VERSION B

Artificial Dutch	Correct answer	English translation	Artificial pseudo	Correct answer
	is	is		ir
	soep	soup		woes
	reus	giant		seup
	sap	juice		wap
	wip	seesaw		pip
	roep	yell		woer
	spar	fir		pras

	roer	steering wheel		soer
	sip	glum		ris
	pa	dad		par
	oer	very old		weus
	was	was		oes
	poes	cat		peus
	as	ash		ap

Note. The two lists were presented on two different paper sheets. Children needed to decode as many words as possible within one minute.

Table S5

Correlations Between the Original and Alternative Version of the Artificial Learning Task to Examine Test-Retest Reliability

	Block 1				**Block 2**				**Block 3**			
	Bin1	Bin2	Bin3	Bin4	Bin1	Bin2	Bin3	Bin4	Bin1	Bin2	Bin3	Bin4
Accuracy (%)	.10	.39***	.35***	.45***	.31***	.41***	.37***	.45***	.35***	.48***	.25**	.32***
Reaction time	.29**	.42***	.44***	.33***	.46***	.43***	.29**	.44***	.40***	.39***	.42***	.33**

Note. Accuracy (percentage correct) and reaction time during the two versions of the learning task (Version A and B) were averaged across 4 bins of 18 trials for Block 1 and 2, and of 14 trials for Block 3 to map the learning curves. Correlations were computed between all corresponding bins of the two versions (e.g., Bin 1 of Version A with Bin 1 of Version B) to examine the test-retest reliability. ***p<.001, **p < .01

Appendix C

Appendix to Chapter 4

Table S1

Repeated Measures ANOVA Results for Accuracy and Reaction Time with Condition and Group as Between-Subject Variables and Time as Within-Subject Variable

	Accuracy	**Reaction Time**
Block 1		
Condition (GD vs implicit)	$F(1,119) = 0.15, p = .70, \eta_p^2 < .01$	$F(1,119) = 0.24, p = .63, \eta_p^2 < .01$
Group (DR vs TR)	$F(1,119) = 0.11, p = .74, \eta_p^2 < .01$	$F(1,119) = 1.92, p = .17, \eta_p^2 = .02$
Age (younger vs older)	$F(1,119) = 6.25, p = .01, \eta_p^2 = .05$	$F(1,119) = 0.07, p = .79, \eta_p^2 < .01$
Condition x Group	$F(1,119) = 0.28, p = .60, \eta_p^2 < .01$	$F(1,119) = 1.00, p = .32, \eta_p^2 < .01$
Group x Age	$F(1,119) = 0.21, p = .65, \eta_p^2 < .01$	$F(1,119) = 0.13, p = .72, \eta_p^2 < .01$
Condition x Age	$F(1,119) = 0.05, p = .83, \eta_p^2 < .01$	$F(1,119) = 1.41, p = .24, \eta_p^2 = .01$
Group x Condition x Age	$F(1,119) = 0.24, p = .63, \eta_p^2 < .01$	$F(1,119) = 0.90, p = .35, \eta_p^2 < .01$
Time	$F(3,357) = 35.78$, $\boldsymbol{p < .001}$, $\eta_p^2 = .23$	$F(3,357) = 17.32$, $\boldsymbol{p < .001,}$ $\eta_p^2 = .13$
Condition x Time	$F(3,357) = 0.39, p = .76, \eta_p^2 < .01$	$F(3,357) = 1.41, p = .24, \eta_p^2 = .01$
Group x Time	$F(3,357) = 0.43, p = .73, \eta_p^2 < .01$	$F(3,357) = 0.43, p = .73, \eta_p^2 < .01$
Time x Age	$F(3,357) = 2.60, p = .05, \eta_p^2 = .02$	$F(3,357) = 1.07, p = .36, \eta_p^2 < .01$
Condition x Group x Time	$F(3,357) = 1.09, p = .36, \eta_p^2 < .01$	$F(3,357) = 4.88$, $\boldsymbol{p < .01,}$ $\eta_p^2 = .04$
Time x Group x Age	$F(3,357) = 0.13, p = .94, \eta_p^2 < .01$	$F(3,357) = 2.04, p = .11, \eta_p^2 = .02$
Time x Condition x Age	$F(3,357) = 0.51, p = .68, \eta_p^2 < .01$	$F(3,357) = 0.33, p = .81, \eta_p^2 < .01$
Time x Condition x Group x Age	$F(3,357) = 0.45, p = .72, \eta_p^2 < .01$	$F(3,357) = 1.33, p = .26, \eta_p^2 = .01$
Block 2		

Condition (GD vs implicit)	$F(1,119) = 0.69, p = .41, \eta_p^2 < .01$	$F(1,119) = 1.51, p = .22, \eta_p^2 = .01$
Group (DR vs TR)	$F(1,119) = 0.23, p = .63, \eta_p^2 < .01$	$F(1,119) = 0.002, p = .96, \eta_p^2 < .01$.
Age	$F(1,119) = 10.95, \boldsymbol{p = .001}, \eta_p^2 = .08$	$F(1,119) = 0.82, p = .37, \eta_p^2 < .01$
Condition x Group	$F(1,119) = 0.46, p = .51, \eta_p^2 < .01$	$F(1,119) = 0.01, p = .94, \eta_p^2 < .01$
Group x Age	$F(1,119) = 0.32, p = .57, \eta_p^2 < .01$	$F(1,119) = 0.38, p = .54, \eta_p^2 < .01$
Condition x Age	$F(1,119) = 0.93, p = .34, \eta_p^2 < .01$	$F(1,119) = 1.94, p = .17, \eta_p^2 = .02$
Group x Condition x Age	$F(1,119) = 0.01, p = .90, \eta_p^2 < .01$	$F(1,119) = 0.61, p = .44, \eta_p^2 < .01$
Time	$F(3,357) = 37.51, \boldsymbol{p < .001}, \eta_p^2 = .24$	$F(3,357) = 2.67, p = .05, \eta_p^2 = .02$
Condition x Time	$F(3,357) = 0.91, p = .44, \eta_p^2 < .01$	$F(3,357) = 0.42, p = .74, \eta_p^2 < .01$
Group x Time	$F(3,357) = 1.77, p = .15, \eta_p^2 = .01$	$F(3,357) = 0.45, p = .72, \eta_p^2 < .01$
Time x Age	$F(3,357) = 0.53, p = .66, \eta_p^2 < .01$	$F(3,357) = 0.46, p = .71, \eta_p^2 < .01$
Condition x Group x Time	$F(3,357) = 2.18, p = .09, \eta_p^2 = .02$	$F(3,357) = 4.00, p < .01, \eta_p^2 = .03$
Time x Group x Age	$F(3,357) = 0.69, p = .56, \eta_p^2 < .01$	$F(3,357) = 1.01, p = .39, \eta_p^2 < .01$
Time x Condition x Age	$F(3,357) = 2.65, p = .05, \eta_p^2 = .02$	$F(3,357) = 0.78, p = .51, \eta_p^2 < .01$
Time x Condition x Group x Age	$F(3,357) = 0.26, p = .85, \eta_p^2 < .01$	$F(3,357) = 1.43, p = .23, \eta_p^2 = .01$
Block 3		
Condition (GD vs implicit)	$F(1,119) = 0.31, p = .58 , \eta_p^2 < .01$	$F(1,119) = 0.001, p = .97, \eta_p^2 < .01$
Group (DR vs TR)	$F(1,119) = 0.05, p = .82, \eta_p^2 < .01$.	$F(1,119) = 0.07, p = .80, \eta_p^2 < .01$
Age	$F(1,119) = 17.64, \boldsymbol{p < .001}, \eta_p^2 = .13$	$F(1,119) = 0.62, p = .43, \eta_p^2 < .01$
Condition x Group	$F(1,119) = 0.34, p = .56, \eta_p^2 < .01$	$F(1,119) = 0.63, p = .43, \eta_p^2 < .01$
Group x Age	$F(1,119) = 0.00, p = .98, \eta_p^2 < .01$	$F(1,119) = 0.12, p = .73, \eta_p^2 < .01$

Condition x Age	$F(1,119) = 0.02, p = .88, \eta_p^2 < .01$	$F(1,119) = 3.71, p = .06, \eta_p^2 = .03$
Group x Condition x Age	$F(1,119) = 1.31, p = .26, \eta_p^2 = .01$	$F(1,119) = 0.01, p = .91, \eta_p^2 < .01$
Time	$F(3,357) = 5.58, \boldsymbol{p} \mathbf{< .001}, \eta_p^2 = .04$	$F(3,357) = 2.16, p = .09, \eta_p^2 = .02$
Condition x Time	$F(3,357) = 0.79, p = .50, \eta_p^2 < .01$	$F(3,357) = 1.06, p = .37, \eta_p^2 < .01$
Group x Time	$F(3,357) = 1.26, p = .29, \eta_p^2 = .01$	$F(3,357) = 0.81, p = .49, \eta_p^2 < .01$
Time x Age	$F(3,357) = 1.76, p = .15, \eta_p^2 = .01$	$F(3,357) = 1.38, p = .25, \eta_p^2 = .01$
Condition x Group x Time	$F(3,357) = 0.71, p = .55, \eta_p^2 < .01$	$F(3,357) = 5.79, \boldsymbol{p} \mathbf{< .001,} \eta_p^2 = .05$
Time x Group x Age	$F(3,357) = 0.74, p = .53, \eta_p^2 < .01$	$F(3,357) = 1.62, p = .18, \eta_p^2 = .01$
Time x Condition x Age	$F(3,357) = 1.23, p = .30, \eta_p^2 = .01$	$F(3,357) = 0.56, p = .64, \eta_p^2 < .01$
Time x Condition x Group x Age	$F(3,357) = 0.57, p = .64, \eta_p^2 < .01$	$F(3,357) = 1.56, p = .20, \eta_p^2 = .01$

Note. P-values that remained significant at significance level .05 after FDR correction for multiple comparisons are shown in bold

Table S2

Factorial ANOVA Results for Accuracy and Reaction Time with Condition and Group as Between-Subjects Variables and Congruency as Within-Subjects Variable

	Accuracy	**Reaction Time**
Condition	$F(1,119) = 1.70, p = .20, \eta_p^2 = .01$	$F(1,119) = 0.55, p = .46, \eta_p^2 = .01$
Group	$F(1,119) = 0.72, p = .40, \eta_p^2 < .01$	$F(1,119) = 0.02, p = .89, \eta_p^2 = .01$
Age	$F(1,119) = 18.70$, **$p < .001$**, $\eta_p^2 = .14$	$F(1,119) = 0.43, p = .52, \eta_p^2 < .01$
Condition x Group	$F(1,119) = 0.00, p = .95, \eta_p^2 < .01$	$F(1,119) = 1.94, p = .17, \eta_p^2 = .02$
Condition x Age	$F(1,119) = 0.19, p = .66, \eta_p^2 < .01$	$F(1,119) = 0.83, p = .37, \eta_p^2 < .01$
Group x Age	$F(1,119) = 0.87, p = .35, \eta_p^2 < .01$	$F(1,119) = 0.49, p = .49, \eta_p^2 < .01$
Condition x Group x Age	$F(1,119) = 0.08, p = .77, \eta_p^2 < .01$	$F(1,119) = 0.24, p = .62, \eta_p^2 < .01$
Congruency	$F(1,119) = 0.51, p = .48, \eta_p^2 < .01$	$F(1,119) = 30.66$, **$p < .001$**, $\eta_p^2 = .20$
Condition x Congruency	$F(1,119) = 1.09, p = .30, \eta_p^2 = .01$	$F(1,119) = 0.49, p = .49, \eta_p^2 < .01$
Group x Congruency	$F(1,119) = 0.60, p = .44, \eta_p^2 = .01$	$F(1,119) = 0.50, p = .48, \eta_p^2 < .01$
Congruency x Age	$F(1,119) = 0.31, p = .58, \eta_p^2 < .01$	$F(1,119) = 3.53, p = .06, \eta_p^2 = .03$
Condition x Group x Congruency	$F(1,119) = 2.35, p = .13, \eta_p^2 = .02$	$F(1,119) = 0.06, p = .81, \eta_p^2 < .01$
Condition x Congruency x Age	$F(1,119) = 1.33, p = .25, \eta_p^2 = .01$	$F(1,119) = 0.03, p = .86, \eta_p^2 < .01$
Group x Congruency x Age	$F(1,119) = 0.21, p = .73, \eta_p^2 < .01$	$F(1,119) = 2.65, p = .11, \eta_p^2 = .01$
Condition x Group x Congruency x Age	$F(1,119) = 0.99, p = .32, \eta_p^2 < .01$	$F(1,119) = 2.01, p = .16, \eta_p^2 = .02$

Note. P-values that remained significant at significance level .05 after FDR correction for multiple comparisons are shown in bold.

Table S3

Component Characteristics

	Unrotated solution			Rotated solution		
	Eigenvalue	**Proportion var.**	**Cumulative**	**SS Loadings**	**Proportion var.**	**Cumulative**
Component 1	5.675	0.183	0.183	4.993	0.161	0.161
Component 2	5.057	0.163	0.346	4.464	0.144	0.305
Component 3	4.096	0.132	0.478	3.537	0.114	0.419
Component 4	3.340	0.108	0.586	3.436	0.111	0.530
Component 5	1.945	0.063	0.649	2.567	0.083	0.613
Component 6	1.536	0.050	0.698	2.142	0.069	0.682
Component 7	1.277	0.041	0.740	1.789	0.058	0.740

Note. SS = Sum Square.

Table S4

Component Loadings of Principal Component Analysis

	RC1	RC2	RC3	RC4	RC5	RC6	RC7	Uniqueness
Peak-to-peak FRN negative early amplitude		**0.888**						0.164
Peak-to-peak FRN negative late amplitude		**0.843**						0.187
Peak-to-peak FRN positive late amplitude		**0.824**						0.250
Peak-to-peak FRN positive early amplitude		**0.820**						0.221
P2 negative late amplitude		**0.769**						0.255
P2 negative early amplitude		**0.744**						0.157
P2 positive early amplitude		**0.667**						0.155
P2 positive late amplitude		**0.626**						0.157
FRN positive early amplitude	**0.898**							0.176
FRN negative early amplitude	**0.856**							0.248
FRN negative late amplitude	**0.760**							0.338
FRN positive late amplitude	**0.730**							0.359
AVL			**0.993**					0.006
AVL-impulsivity			**0.940**					0.099
AVL-hyperactivity			**0.875**					0.230
AVL-inattention			**0.873**					0.199
OMT real words				**0.926**				0.119
OMT pseudowords				**0.916**				0.142

Symbol knowledge	**0.907**				**0.169**
Letter-speech sound integration	**0.820**				0.254
LN negative late amplitude				**0.816**	0.311
LN positive early amplitude				**0.766**	0.297
LN positive late amplitude				**0.738**	0.337
LN negative early amplitude				**0.659**	0.400
Word reading Dutch		**0.890**			0.166
Rapid automatized naming		**-0.840**			0.315
Phoneme awareness		**0.722**			0.360
Peak-to-peak FRN positive early latency			**0.744**		0.375
Peak-to-peak FRN negative early latency			**0.640**		0.481
Peak-to-peak FRN negative late latency			**0.611**		0.583
Peak-to-peak FRN positive late latency			**0.589**		0.563

Appendix D

Appendix to Chapter 5

Principal component analysis. A principal component analysis with a varimax rotation and Kaiser normalization was conducted to obtain factor scores for letter-speech sound integration, based on following tasks: letter-speech sound discrimination (speed and accuracy), letter-speech sound identification (speed and accuracy) and computerized spelling (speed and accuracy). The factor analysis showed, as expected, clear convergence to two factors on all three timepoints, which we will refer to as letter-speech sound accuracy and letter-speech sound speed. Table S1 shows an overview of the variables and corresponding factor loadings.

Table S1

Variables of Interest and Corresponding Factor Loadings on two Factors

		Factor 1	Factor 2
T0	L-SS discrimination-accuracy	-0.10	**0.80**
	L-SS identification-accuracy	-0.03	**0.82**
	Spelling-accuracy	-0.03	**0.78**
	L-SS discrimination-speed	**0.79**	-0.12
	L-SS identification-speed	**0.86**	0.10
	Spelling-speed	**0.83**	-0.14
T1	L-SS discrimination-accuracy	-0.09	**0.80**
	L-SS identification-accuracy	-0.05	**0.77**
	Spelling-accuracy	0.03	**0.74**
	L-SS discrimination-speed	**0.80**	-0.04
	L-SS identification-speed	**0.85**	0.10
	Spelling-speed	**0.81**	-0.16
T2	L-SS discrimination-accuracy	-0.06	**0.82**
	L-SS identification-accuracy	-0.07	**0.77**
	Spelling-accuracy	0.07	**0.75**
	L-SS discrimination-speed	**0.80**	-0.02
	L-SS identification-speed	**0.87**	0.06
	Spelling-speed	**0.82**	-0.11

Note. L-SS = letter-speech sound. The different measures were assessed at three timepoints: before the intervention (T0), halfway through the intervention (T1), and at the end of the intervention (T2).

Table S2

Complete Overview of all Estimated Path Coefficients

			Raw est.	**S.E.**	**C.R.**	***p***	**Stand. est.**
T0 L-SS S	→	T1 L-SS S	0.515	0.014	36.510	< .001	0.515
T0 L-SS A	→	T1 L-SS S	0.091	0.014	6.464	< .001	0.091
T0 L-SS A	→	T1 L-SS A	0.531	0.014	37.780	< .001	0.531
T0 L-SS S	→	T1 L-SS A	-0.071	0.016	-4.329	< .001	-0.071
T1 L-SS S	→	T1 L-SS A	-0.012	0.016	-0.738	0.461	-0.012
T1 L-SS S	→	T2 L-SS S	0.599	0.015	41.000	< .001	0.601
T0 L-SS A	→	T2 L-SS S	0.014	0.015	0.933	0.351	0.014
T0 L-SS S	→	T2 L-SS S	0.104	0.015	7.096	< .001	0.104
T1 L-SS A	→	T2 L-SS S	0.146	0.015	9.929	< .001	0.147
T0 WRF	→	T1 WRF	0.853	0.014	61.317	< .001	0.722
T0 PS	→	T1 WRF	-0.045	0.008	-5.958	0.000	-0.072
T0 L-SS A	→	T2 L-SS A	0.182	0.015	11.852	< .001	0.183
T1 L-SS A	→	T2 L-SS A	0.539	0.016	34.639	< .001	0.541
T0 L-SS S	→	T2 L-SS A	-0.021	0.015	-1.382	0.167	-0.021
T0 L-SS S	→	T1 WRF	-1.216	0.093	-13.031	< .001	-0.182
T0 L-SS A	→	T1 WRF	-0.950	0.091	-10.494	< .001	-0.142
T1 L-SS S	→	T2 L-SS A	0.096	0.019	5.201	< .001	0.097
T1 L-SS S	→	T1 WRF	1.725	0.087	19.924	< .001	0.258
T1 L-SS A	→	T1 WRF	0.973	0.087	11.136	< .001	0.146
T2 L-SS S	→	T2 L-SS A	-0.160	0.018	-9.036	< .001	-0.160
T0 WRF	→	T2 WRF	0.222	0.017	12.889	< .001	0.166
T0 PS	→	T2 WRF	-0.018	0.007	-2.739	0.006	-0.025
T0 L-SS S	→	T2 WRF	-0.278	0.083	-3.336	< .001	-0.037
T0 L-SS A	→	T2 WRF	-0.541	0.081	-6.659	< .001	-0.072
T1 L-SS S	→	T2 WRF	-0.793	0.096	-8.284	< .001	-0.105
T1 L-SS A	→	T2 WRF	-0.215	0.091	-2.364	0.018	-0.028
T2 L-SS S	→	T2 WRF	1.441	0.090	15.963	< .001	0.190
T2 L-SS A	→	T2 WRF	0.833	0.086	9.688	< .001	0.110
T1 WRF	→	T2 WRF	0.812	0.014	56.665	< .001	0.720

Note. L-SS A = letter-speech sound accuracy; L-SS S = letter-speech sound speed; WRF = word reading fluency; PS = processing speed. The different measures were assessed at three timepoints: before the intervention (T0), halfway through the intervention (T1), and at the end of the intervention (T2). Raw est. = raw estimate; S.E. = standard error; C.R. = critical ratio; *p* = two-tailed p-value; Stand. est. = standardized estimate.

Appendix E

Appendix to Chapter 7

Table S1

Post Hoc Comparisons Between Conditions

			95% CI for Mean Difference					95% CI for Cohen's d		
		Mean Difference	**Lower**	**Upper**	**SE**	**t**	**Cohen's d**	**Lower**	**Upper**	**p_{tukey}**
AV+	AV	-5.823	-12.756	1.111	2.927	-1.989	-0.393	-0.875	0.089	0.119
	CC	3.467	-3.885	10.819	3.104	1.117	0.234	-0.275	0.742	0.505
AV	CC	9.290	2.034	16.545	3.063	3.033	0.627	0.118	1.135	0.008 **

Note. AV = Audiovisual only condition. AV+ = Audiovisual condition with motoric component. CC = Mathematical control condition. P-value and confidence intervals are adjusted for comparing a family of 3 estimates (confidence intervals corrected using the Tukey method). ** $p < .01$

Table S2

Correlation Matrix Between Outcome Measures and In-Game Measures

	LSS pre	**LSS post**	**Reading pre**	**Reading post**	**Game accuracy**	**Game level**
LSS pre	—					
LSS post	0.78 ***	—				
Reading pre	0.42 ***	0.47 ***	—			
Reading post	0.44 ***	0.44 ***	0.91 ***	—		
Game accuracy	0.30 **	0.42 ***	0.29 **	0.31 **	—	
Game level	0.38 ***	0.51 ***	0.39 ***	0.32 **	0.32 ***	—

Note. L-SS = letter-speech sound knowledge. * $p < .05$, ** $p < .01$, *** $p < .001$

References

Aarts, H., Custers, R., & Marien, H. (2008). Preparing and motivating behavior outside of awareness. *science*, *319*(5870), 1639–1639. https://doi.org/10.1126/science.1150432

Allor, J. H. (2002). The relationships of phonemic awareness and rapid naming to reading development. *Learning Disability Quarterly*, *25*(1), 47–57. https://doi.org/10.2307/1511190

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5®*. American psychiatric association.

Amora, K. K., Tretow, A., Verwimp, C., Tijms, J., Leppänen, P. H. T., & Csépe, V. (2022). Typical and atypical development of visual expertise for print as indexed by the visual word N1 (N170w): A Systematic Review. *Frontiers in Neuroscience*, *16*, 898800. https://doi.org/10.3389/fnins.2022.898800

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*(2), 71–82. https://doi.org/10.1076/chin.8.2.71.8724

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

Aravena, S., Snellings, P., Tijms, J., & van der Molen, M. W. (2013). A lab-controlled simulation of a letter–speech sound binding deficit in dyslexia. *Journal of Experimental Child Psychology*, *115*(4), 691–707. https://doi.org/10.1016/j.jecp.2013.03.009

Aravena, S., Tijms, J., Snellings, P., & van der Molen, M. W. (2018). Predicting individual differences in reading and spelling skill with artificial script–based letter–speech sound training. *Journal of Learning Disabilities*, *51*(6), 552–564. https://doi.org/10.1177/0022219417715407

Arbel, Y., & Fox, A. B. (2021). Electrophysiological examination of feedback-based learning in 8–11-year-old children. *Frontiers in Psychology*, *12*, 640270. https://doi.org/10.3389/fpsyg.2021.640270

Arbuthnott, K., & Frank, J. (2000). Trail Making Test, Part B as a measure of executive control: validation using a set-switching paradigm. *Journal of Clinical and Experimental*

Neuropsychology, *22*(4), 518–528. https://doi.org/10.1076/1380-3395(200008)22:4;1-0;FT518

Arnold, H. L., Plante, E., & Vance, R. (2022). Translating enhanced conversational recast to a telepractice setting. *Language, Speech & Hearing Services in Schools*, *53*(2), 275–289. https://doi.org/10.1044/2021_LSHSS-21-00051

Arrington, C. N., Kulesz, P. A., Francis, D. J., Fletcher, J. M., & Barnes, M. A. (2014). The contribution of attentional control and working memory to reading comprehension and decoding. *Scientific Studies of Reading*, *18*(5), 325–346. https://doi.org/10.1080/10888438.2014.902461

Astle, D. E., & Fletcher-Watson, S. (2020). Beyond the core-deficit hypothesis in developmental disorders. *Current Directions in Psychological Science*, 1–7. https://doi.org/10.1177/0963721420925518

Banfield, J. F., Wyland, C. L., MacRae, C. N., Munte, T. F., & Heatherton, T. F. (2004). The cognitive neuroscience of self-regulation. In R.F. Baumeister & K.D. Vohs (Eds.) *The Handbook of Self Regulation* (pp. 62–83). New York: Guilford Press.

Baron, I. S. (2004). *Neuropsychological Evaluation of the Child: Domains, Methods, & Case Studies*. New York: Oxford University Press.

Baron, J., & Treiman, R. (1980). Some problems in the study of differences in cognitive processes. *Memory & Cognition*, *8*(4), 313–321. https://doi.org/10.3758/BF03198270

Barron, K. E., & Harackiewicz, J. M. (2001). Achievement goals and optimal motivation: Testing multiple goal models. *Journal of Personality and Social Psychology*, *80*, 706–722. https://doi.org/10.1037/0022-3514.80.5.706

Bazen, L., de Bree, E. H., van den Boer, M., & de Jong, P. F. (2022). Perceived negative consequences of dyslexia: The influence of person and environmental factors. *Annals of Dyslexia*. https://doi.org/10.1007/s11881-022-00274-0

Becker, N., Vasconcelos, M., Oliveira, V., Santos, F. C. D., Bizarro, L., Almeida, R. M. M. D., Salles, J. F. D., & Carvalho, M. R. S. (2017). Genetic and environmental risk factors for developmental dyslexia in children: Systematic review of the last decade. *Developmental Neuropsychology*, *42*(7–8), 423–445. https://doi.org/10.1080/87565641.2017.1374960

Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *The European Journal of Neuroscience*, *27*(7), 1823–1835. https://doi.org/10.1111/j.1460-9568.2008.06138.x

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300. JSTOR. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berends, I. E., & Reitsma, P. (2006). Remediation of fluency: Word specific or generalised training effects? *Reading and Writing*, *19*(2), 221–234. https://doi.org/10.1007/s11145-005-5259-3

Berlin, R. (1887). *Eine besondere Art der Wortblindheit (Dyslexie)*. Wiesbaden: Bergmann.

Bishop, D. V. M. (2006). Dyslexia: What's the problem? *Developmental Science*, *9*(3), 256–257. https://doi.org/10.1111/j.1467-7687.2006.00484.x

Bitan, T., & Booth, J. R. (2012). Offline improvement in learning to read a novel orthography depends on direct letter instruction. *Cognitive Science*, *36*(5), 896–918. https://doi.org/10.1111/j.1551-6709.2012.01234.x

Blau, V., Reithler, J., van Atteveldt, N., Seitz, J., Gerretsen, P., Goebel, R., & Blomert, L. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: A functional magnetic resonance imaging study of dyslexic children. *Brain*, *133*(3), 868–879. https://doi.org/10.1093/brain/awp308

Blau, V., van Atteveldt, N., Ekkebus, M., Goebel, R., & Blomert, L. (2009). Reduced neural integration of letters and speech sounds links phonological and reading deficits in adult dyslexia. *Current Biology*, *19*(12), 503–508. https://doi.org/10.1016/j.cub.2009.06.008

Blomert, L. (2011). The neural signature of orthographic–phonological binding in successful and failing reading development. *NeuroImage*, *57*(3), 695–703. https://doi.org/10.1016/j.neuroimage.2010.11.003

Blomert, L., & Vaessen, A. (2009). *3DM differential diagnostics for dyslexia: Cognitive analysis of reading and spelling.* Amsterdam: Boom Test Publishers.

Boets, B., Smedt, B., Cleuren, L., Vandewalle, E., Wouters, J., & Ghesquière, P. (2010). Towards a further characterization of phonological and literacy problems in Dutch-speaking children with dyslexia. *British Journal of Developmental Psychology*, *28*(1), 5–31. https://doi.org/10.1348/026151010X485223

Boets, B., Wouters, J., van Wieringen, A., De Smedt, B., & Ghesquière, P. (2008). Modelling relations between sensory processing, speech perception, orthographic and phonological ability, and literacy achievement. *Brain and Language*, *106*(1), 29–40. https://doi.org/10.1016/j.bandl.2007.12.004

Bonifacci, P., Storti, M., Tobia, V., & Suardi, A. (2016). Specific Learning Disorders: A look inside children's and parents' psychological well-being and relationships. *Journal of Learning Disabilities*, *49*(5), 532–545. https://doi.org/10.1177/0022219414566681

Bonte, M., Correia, J. M., Keetels, M., Vroomen, J., & Formisano, E. (2017). Reading-induced shifts of perceptual speech representations in auditory cortex. *Scientific Reports*, *7*(1), 1. https://doi.org/10.1038/s41598-017-05356-3

Booth, J. N., Boyle, J. M. E., & Kelly, S. W. (2010). Do tasks make a difference? Accounting for heterogeneity of performance of children with reading difficulties on tasks of executive function: Findings from a meta-analysis. *British Journal of Developmental Psychology*, *28*(1), 133–176. https://doi.org/10.1348/026151009X485432

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13. https://doi.org/10.1002/wps.20375

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*(1), 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Bosse, M.-L. (2015). Learning to read and spell: How children acquire word orthographic knowledge. *Child Development Perspectives*, *9*(4), 222–226. https://doi.org/10.1111/cdep.12133

Bowers, J. S. (2020). Reconsidering the evidence that systematic phonics is more effective than alternative methods of reading instruction. *Educational Psychology Review*, *32*(3), 681–705. https://doi.org/10.1007/s10648-019-09515-y

Brem, S., Bach, S., Kucian, K., Kujala, J. V., Guttorm, T. K., Martin, E., Lyytinen, H., Brandeis, D., & Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter–speech sound correspondences. *Proceedings of the National Academy of Sciences*, *107*(17), 7939–7944. https://doi.org/10.1073/pnas.0904402107

Brem, S., Lang-Dullenkopf, A., Maurer, U., Halder, P., Bucher, K., & Brandeis, D. (2005). Neurophysiological signs of rapidly emerging visual expertise for symbol strings: *NeuroReport*, *16*(1), 45–48. https://doi.org/10.1097/00001756-200501190-00011

Breznitz, Z. (1987). Increasing first graders' reading accuracy and comprehension by accelerating their reading rates. *Journal of Educational Psychology*, *79*(3), 236–242.

Breznitz, Z. (2002). Asynchrony of visual-orthographic and auditory-phonological word recognition processes: An underlying factor in dyslexia. *Reading and Writing*, *15*, 15–42.

Breznitz, Z. (2006). *Fluency in Reading: Synchronization of Processes*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Bringmann, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., Wigman, J. T. W., & Snippe, E. (2019). What do centrality measures measure in psychological networks? *Journal of Abnormal Psychology*, *128*(8), 892–903. https://doi.org/10.1037/abn0000446

Brosnan, M., Demetre, J., Hamill, S., Robson, K., Shepherd, H., & Cody, G. (2002). Executive functioning in adults and children with developmental dyslexia. *Neuropsychologia*, *40*(12), 2144–2155. https://doi.org/10.1016/S0028-3932(02)00046-5

Bruner, J. (1961). The act of discovery. *Harvard Educational Review*, *31*, 21–32. https://doi.org/10.4324/9780203088609-13

Brunner, C., Delorme, A., & Makeig, S. (2013). *Eeglab – an Open Source Matlab Toolbox for Electrophysiological Research*. *58*(SI-1-Track-G). https://doi.org/10.1515/bmt-2013-4182

Brus, B. T., & Voeten, M. J. M. (1973). *Een Minuut Test: Verantwoording en handleiding [One Minute Decoding Test, Form A and B: Rationale and manual].* Berkhout.

Brysbaert, M. (2001). Prelexical phonological coding of visual words in Dutch: Automatic after all. *Memory & Cognition*, *29*(5), 765–773. https://doi.org/10.3758/BF03200479

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., Salas, N., Schöffelová, M., Defior, S., Mikulajová, M., Seidlová-Málková, G., & Hulme, C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science*, *23*(6), 678–686. https://doi.org/10.1177/0956797611434536

Cartwright, K. B. (2012). Insights from cognitive neuroscience: The importance of executive function for early reading development and education. *Early Education & Development*, *23*(1), 24–36. https://doi.org/10.1080/10409289.2011.615025

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Catts, H. W., Petscher, Y., Schatschneider, C., Bridges, M. S., & Mendoza, K. (2009). Floor effects associated with universal screening and their impact on early identification of reading disabilities. *Journal of Learning Disabilities*, *42*(2), 163–176. https://doi.org/10.1177/0022219408326219

Chein, J. M., & Schneider, W. (2012). The brain's learning and control architecture. *Current Directions in Psychological Science*, *21*(2), 78–84. https://doi.org/10.1177/0963721411434977

Conrad, N. J., & Levy, B. A. (2011). Training letter and orthographic pattern recognition in children with slow naming speed. *Reading and Writing*, *24*(1), 91–115. https://doi.org/10.1007/s11145-009-9202-x

Cooper, D. H., Roth, F. P., Speece, D. L., & Schatschneider, C. (2002). The contribution of oral language skills to the development of phonological awareness. *Applied Psycholinguistics*, *23*(3), 399–416. https://doi.org/10.1017/S0142716402003053

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. J. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, *54*, 13–29. https://doi.org/10.1016/j.jrp.2014.07.003

Craig, F., Operto, F. F., De Giacomo, A., Margari, L., Frolli, A., Conson, M., Ivagnes, S., Monaco, M., & Margari, F. (2016). Parenting stress among parents of children with Neurodevelopmental Disorders. *Psychiatry Research*, *242*, 121–129. https://doi.org/10.1016/j.psychres.2016.05.016

Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, *33*(2–3), 137–150. https://doi.org/10.1017/S0140525X09991567

Crone, E. A., Jennings, J. R., & Van der Molen, M. W. (2004). Developmental change in feedback processing as reflected by phasic heart rate changes. *Developmental Psychology*, *40*(6), 1228–1238. https://doi.org/10.1037/0012-1649.40.6.1228

Cuthbert, B. N., & Insel, T. R. (2010). Toward new approaches to psychotic disorders: The NIMH Research Domain Criteria project. *Schizophrenia Bulletin*, *36*(6), 1061–1062. https://doi.org/10.1093/schbul/sbq108

Davydov, V. V. (1995). The state of research on learning activity. *Journal of Russian & East European Psychology*, *33*(4), 55–70. https://doi.org/10.2753/RPO1061-0405330455

de Beer, J., Engels, J., Heerkens, Y., & van der Klink, J. (2014). Factors influencing work participation of adults with developmental dyslexia: A systematic review. *BMC Public Health*, *14*(1), 77. https://doi.org/10.1186/1471-2458-14-77

de Graaff, S., Bosman, A. M. T., Hasselman, F., & Verhoeven, L. (2009). Benefits of systematic phonics instruction. *Scientific Studies of Reading*, *13*(4), 318–333. https://doi.org/10.1080/10888430903001308

De Groot, B. J. A., Van den Bos, K. P., & Van der Meulen, B. F. (2015). *Fonemische Analyse Test (FAT-R). Gereviseerde versie*. Amsterdam: Pearson.

de Jong, P. F., & van der Leij, A. (2003). Developmental changes in the manifestation of a phonological deficit in dyslexic children learning to read a regular orthography. *Journal of Educational Psychology*, *95*(1), 22–40. https://doi.org/10.1037/0022-0663.95.1.22

de Oliveira, D. G., da Silva, P. B., Dias, N. M., Seabra, A. G., & Macedo, E. C. (2014). Reading component skills in dyslexia: Word recognition, comprehension and processing speed. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01339

Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience*, *16*(4), 234–244. https://doi.org/10.1038/nrn3924

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00781

Dijkstra, T., Frauenfelder, U., H., & Schreuder, R. (1993). Bidirectional grapheme-phoneme activation in a bimodal detection task. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(5), 931–950. https://psycnet.apa.org/doi/10.1037/0096-1523.19.5.931

Dilnot, J., Hamilton, L., Maughan, B., & Snowling, M. J. (2017). Child and environmental risk factors predicting readiness for learning in children at high risk of dyslexia. *Development and Psychopathology*, *29*(1), 235–244. https://doi.org/10.1017/S0954579416000134

Dion, J.-S., & Restrepo, G. (2016). A systematic review of the literature linking neural correlates of feedback processing to learning. *Zeitschrift Für Psychologie*, *224*(4), 247–256. https://doi.org/10.1027/2151-2604/a000260

Edgar, S., Carr, S. E., Connaughton, J., & Celenza, A. (2019). Student motivation to learn: Is self-belief the key to transition and first year performance in an undergraduate health professions program? *BMC Medical Education*, *19*, 111. https://doi.org/10.1186/s12909-019-1539-5

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading*, *18*, 116–125. https://doi.org/10.1111/j.1467-9817.1995.tb00077.x

Ehri, L. C. (1997). Learning to read and learning to spell are one and the same, almost. In *Learning to spell: Research, theory, and practice across languages.* (pp. 237–269). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

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

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the national reading panel's meta-analysis. *Review of Educational Research*, *71*(3), 393–447. https://doi.org/10.3102/00346543071003393

Elbro, C. (2010). Dyslexia as disability or handicap: When does vocabulary matter? *Journal of Learning Disabilities*, *43*(5), 469–478. https://doi.org/10.1177/0022219409357349

Elliott, J. G., & Grigorenko, E. L. (2014). *The Dyslexia Debate*. Cambridge, UK: Cambridge University Press.

Eppinger, B., Mock, B., & Kray, J. (2009). Developmental differences in learning and error processing: Evidence from ERPs. *Psychophysiology*, *46*(5), 1043–1053. https://doi.org/10.1111/j.1469-8986.2009.00838.x

Epskamp, S., Borsboom, D., & Fried, E. I. (2017). Estimating psychological networks and their accuracy: A tutorial paper. *ArXiv:1604.08462 [Stat]*. http://arxiv.org/abs/1604.08462

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*(4). https://doi.org/10.18637/jss.v048.i04

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, *23*(4), 617–634. https://doi.org/10.1037/met0000167

Epskamp, S., van Borkulo, C. D., van der Veen, D. C., Servaas, M. N., Isvoranu, A.-M., Riese, H., & Cramer, A. O. J. (2018). Personalized network modeling in psychopathology: the importance of contemporaneous and temporal connections. *Clinical Psychological Science*, *6*(3), 416–427. https://doi.org/10.1177/2167702617744325

Erhel, S., & Jamet, E. (2016). The effects of goal-oriented instructions in digital game-based learning. *Interactive Learning Environments*, *24*(8), 1744–1757. https://doi.org/10.1080/10494820.2015.1041409

Fair, D. A., Cohen, A. L., Power, J. D., Dosenbach, N. U. F., Church, J. A., Miezin, F. M., Schlaggar, B. L., & Petersen, S. E. (2009). Functional brain networks develop from a "local to distributed" organization. *PLoS Computational Biology*, *5*(5), e1000381. https://doi.org/10.1371/journal.pcbi.1000381

Fälth, L., Nilvius, C., & Anvegård, E. (2015). Intensive reading with reading lists—An intervention study. *Creative Education*, *06*(22), 2403–2409. https://doi.org/10.4236/ce.2015.622246

Farah, R., Ionta, S., & Horowitz-Kraus, T. (2021). Neuro-behavioral correlates of executive dysfunctions in dyslexia over development from childhood to adulthood. *Frontiers in Psychology*, *12*, 708863. https://doi.org/10.3389/fpsyg.2021.708863

Farasat, A., Nikolaev, A., Srihari, S. N., & Blair, R. H. (2015). Probabilistic graphical models in modern social network analysis. *Social Network Analysis and Mining*, *5*(1), 62. https://doi.org/10.1007/s13278-015-0289-6

Feder, K. P., & Majnemer, A. (2007). Handwriting development, competency, and intervention. *Developmental Medicine & Child Neurology*, *49*(4), 312–317. https://doi.org/10.1111/j.1469-8749.2007.00312.x

Ferdinand, N. K., Becker, A. M. W., Kray, J., & Gehring, W. J. (2016). Feedback processing in children and adolescents: Is there a sensitivity for processing rewarding feedback? *Neuropsychologia*, *82*, 31–38. https://doi.org/10.1016/j.neuropsychologia.2016.01.007

Ferdinand, N. K., & Kray, J. (2014). Developmental changes in performance monitoring: How electrophysiological data can enhance our understanding of error and feedback processing in childhood and adolescence. *Behavioural Brain Research*, *263*, 122–132. https://doi.org/10.1016/j.bbr.2014.01.029

Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K. E., Michaels, R., & Shaywitz, S. E. (2015). Achievement gap in reading is present as early as first grade and persists through adolescence. *The Journal of Pediatrics*, *167*(5), 1121-1125.e2. https://doi.org/10.1016/j.jpeds.2015.07.045

Fraga González, G., Smit, D. J. A., Molen, M. J. W., Tijms, J., Geus, E. J. C., & Molen, M. W. (2019). Probability learning and feedback processing in dyslexia: A performance and heart rate analysis. *Psychophysiology*, *56*(12). https://doi.org/10.1111/psyp.13460

Fraga González, G., Žarić, G., Tijms, J., Bonte, M., Blomert, L., Leppänen, P., & van der Molen, M. W. (2016). Responsivity to dyslexia training indexed by the N170 amplitude of the brain potential elicited by word reading. *Brain and Cognition*, *106*, 42–54. https://doi.org/10.1016/j.bandc.2016.05.001

Fraga González, G., Žarić, G., Tijms, J., Bonte, M., Blomert, L., & van der Molen, M. W. (2015). A randomized controlled trial on the beneficial effects of training letter-speech sound integration on reading fluency in children with dyslexia. *PLOS ONE*, *10*(12), e0143914. https://doi.org/10.1371/journal.pone.0143914

Francis, A. L., & Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 349–366. https://doi.org/10.1037/0096-1523.28.2.349

Francis, D. A., Caruana, N., Hudson, J. L., & McArthur, G. M. (2019). The association between poor reading and internalising problems: A systematic review and meta-analysis. *Clinical Psychology Review*, *67*, 45–60. https://doi.org/10.1016/j.cpr.2018.09.002

Fried, E. I., von Stockert, S., Haslbeck, J. M. B., Lamers, F., Schoevers, R. A., & Penninx, B. W. J. H. (2019). Using network analysis to examine links between individual depressive symptoms, inflammatory markers, and covariates. *Psychological Medicine*, 1–9. https://doi.org/10.1017/S0033291719002770

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science*, *18*(3), 233–239. https://doi.org/10.1111/j.1467-9280.2007.01882.x

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: true issues and false trails. *Psychological Bulletin*, *123*(1), 71–99. https://doi.org/10.1037/0033-2909.123.1.71

Froyen, D., Bonte, M., van Atteveldt, N., & Blomert, L. (2009). The long road to automation: Neurocognitive development of letter–speech sound processing. *Journal of Cognitive Neuroscience*, *21*(3), 567–580. https://doi.org/10.1162/jocn.2009.21061

Froyen, D., Van Atteveldt, N., Bonte, M., & Blomert, L. (2008). Cross-modal enhancement of the MMN to speech-sounds indicates early and automatic integration of letters and speech-sounds. *Neuroscience Letters*, *430*(1), 23–28. https://doi.org/10.1016/j.neulet.2007.10.014

Froyen, D., Willems, G., & Blomert, L. (2011). Evidence for a specific cross-modal association deficit in dyslexia: An electrophysiological study of letter-speech sound processing: Cross-modal association deficit in dyslexia. *Developmental Science*, *14*(4), 635–648. https://doi.org/10.1111/j.1467-7687.2010.01007.x

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, *21*(11), 1129–1164. https://doi.org/10.1002/spe.4380211102

Fugate, M. H. (1997). Letter training and its effect on the development of beginning reading skills. *School Psychology Quarterly*, *12*(2), 170–192. https://doi.org/10.1037/h0088957

Furlong, L. M., & Serry, T. A. (2022). An exploratory study of speech-language pathologists' clinical practice in the literacy domain: Comparing onsite practices with telepractice services during COVID-19. *International Journal of Speech-Language Pathology*, 1–13. https://doi.org/10.1080/17549507.2022.2030410

Furlong, L. M., Serry, T., Bridgman, K., & Erickson, S. (2021). An evidence-based synthesis of instructional reading and spelling procedures using telepractice: A rapid review in the context of COVID-19. *International Journal of Language & Communication Disorders*, *56*(3), 456–472. https://doi.org/10.1111/1460-6984.12619

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), e89900. https://doi.org/10.1371/journal.pone.0089900

Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A.-M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology*, *93*(3), 265–281. https://doi.org/10.1016/j.jecp.2005.08.003

Geelhoed, J. W., & Reitsma, P. (1999). *PI-dictee*. Lisse, The Netherlands: Swets & Zeitlinger.

Gellert, A. S., & Elbro, C. (2017). Does a dynamic test of phonological awareness predict early reading difficulties?: A longitudinal study from kindergarten through Grade 1. *Journal of Learning Disabilities*, *50*(3), 227–237. https://doi.org/10.1177/0022219415609185

Georgiou, G. K., Torppa, M., Manolitsis, G., Lyytinen, H., & Parrila, R. (2012). Longitudinal predictors of reading and spelling across languages varying in orthographic consistency. *Reading and Writing*, *25*(2), 321–346. https://doi.org/10.1007/s11145-010-9271-x

Gerber, A. J., Peterson, B. S., Giedd, J. N., Lalonde, F. M., Celano, M. J., White, S. L., Wallace, G. L., Lee, N. R., & Lenroot, R. K. (2009). Anatomical brain magnetic resonance imaging of typically developing children and adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(5), 465–470. https://doi.org/10.1097/CHI.0b013e31819f2715

Geurts, H. M., Grasman, R. P. P. P., Verté, S., Oosterlaan, J., Roeyers, H., van Kammen, S. M., & Sergeant, J. A. (2008). Intra-individual variability in ADHD, autism spectrum disorders and Tourette's syndrome. *Neuropsychologia*, *46*(13), 3030–3041. https://doi.org/10.1016/j.neuropsychologia.2008.06.013

Ghazalpour, A., Doss, S., Zhang, B., Wang, S., Plaisier, C., Castellanos, R., Brozell, A., Schadt, E. E., Drake, T. A., Lusis, A. J., & Horvath, S. (2006). Integrating genetic and network analysis to characterize genes related to mouse weight. *PLoS Genetics*, *2*(8), 11.

Godefroy, O., Cabaret, M., Petit-Chenal, V., Pruvo, J.-P., & Rousseaux, M. (1999). Control functions of the frontal lobes. modularity of the central-supervisory system? *Cortex*, *35*(1), 1–20. https://doi.org/10.1016/S0010-9452(08)70782-2

Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, *6*(4), 126–135. https://doi.org/10.1080/19388076709556976

Gough, P. B., Juel, C., & Griffith, P. L. (1992). Reading, spelling, and the orthographic cipher. In *Reading Acquisition* (pp. 35-48). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Graesser, A. C., Chipman, P., Leeming, F., & Biedenbach, S. (2009). Deep learning and emotion in serious games. In U. Ritterfeld, M. Cody, & P. Vorderer (Eds.), *Serious games: Mechanisms and effects* (pp. 81–100). New York: Routledge, Taylor & Francis.

Groen, Y., Wijers, A. A., Mulder, L. J. M., Minderaa, R. B., & Althaus, M. (2007). Physiological correlates of learning by performance feedback in children: A study of EEG event-related potentials and evoked heart rate. *Biological Psychology*, *76*(3), 174–187. https://doi.org/10.1016/j.biopsycho.2007.07.006

Grogan-Johnson, S., Schmidt, A. M., Schenker, J., Alvares, R., Rowan, L. E., & Taylor, J. (2013). A comparison of speech sound intervention delivered by telepractice and side-by-side service delivery models. *Communication Disorders Quarterly*, *34*(4), 210–220. https://doi.org/10.1177/1525740113484965

Guan, C. Q., Liu, Y., Chan, D. H. L., Ye, F., & Perfetti, C. A. (2011). Writing strengthens orthography and alphabetic-coding strengthens phonology in learning to read Chinese. *Journal of Educational Psychology*, *103*(3), 509–522. https://doi.org/10.1037/a0023730

Guan, C. Q., Smolen, E. R., Meng, W., & Booth, J. R. (2021). Effect of handwriting on visual word recognition in Chinese bilingual children and adults. *Frontiers in Psychology*, *12*, 628160. https://doi.org/10.3389/fpsyg.2021.628160

Guerra, G. (2022). *The contribution of auditory attention to reading processes of school-age children with and without dyslexia* [Doctoral Thesis, Maastricht University]. https://doi.org/10.26481/dis.20220322gg

Guerra, G., Tijms, J., Tierney, A., Vaessen, A., Dick, F., & Bonte, M. (In revision). *Auditory attention influences trajectories of symbol-speech sound learning in children with and without dyslexia.*

Guerra, G., Tijms, J., Vaessen, A., Tierney, A., Dick, F., & Bonte, M. (2021). Loudness and intelligibility of irrelevant background speech differentially hinder children's short story reading. *Mind, Brain and Education*, *15*(1), 77–87. https://doi.org/10.1111/mbe.12264

Hamel, R., & Schmittmann, V. D. (2006). The 20-Minute version as a predictor of the Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement*, *66*(6), 1039–1046. https://doi.org/10.1177/0013164406288169

Hamilton, L. G., Hayiou-Thomas, M. E., Hulme, C., & Snowling, M. J. (2016). The home literacy environment as a predictor of the early literacy development of children at family-risk of dyslexia. *Scientific Studies of Reading*, *20*(5), 401–419. https://doi.org/10.1080/10888438.2016.1213266

Hämmerer, D., Li, S.-C., Müller, V., & Lindenberger, U. (2011). Life span differences in electrophysiological correlates of monitoring gains and losses during probabilistic reinforcement learning. *Journal of Cognitive Neuroscience*, *23*(3), 579–592. https://doi.org/10.1162/jocn.2010.21475

Harlaar, N., Spinath, F. M., Dale, P. S., & Plomin, R. (2005). Genetic influences on early word recognition abilities and disabilities: A study of 7-year-old twins. *Journal of Child Psychology and Psychiatry*, *46*(4), 373–384. https://doi.org/10.1111/j.1469-7610.2004.00358.x

Harris, L. N., Creed, B., Perfetti, C. A., & Rickles, B. B. (2022). The role of word knowledge in error detection: A challenge to the broken error monitor account of dyslexia. *Annals of Dyslexia*, *72*(2), 384–402. https://doi.org/10.1007/s11881-021-00248-8

Haslbeck, J. M. B., Borsboom, D., & Waldorp, L. (2020). Moderated Network Models. *ArXiv:1807.02877 [Stat]*. http://arxiv.org/abs/1807.02877

Haslbeck, J. M. B., & Fried, E. I. (2017). How predictable are symptoms in psychopathological networks? A reanalysis of 18 published datasets. *Psychological Medicine*, *47*(16), 2767–2776. https://doi.org/10.1017/S0033291717001258

Haslbeck, J. M. B., & Waldorp, L. J. (2018). How well do network models predict observations? On the importance of predictability in network models. *Behavior Research Methods*, *50*(2), 853–861. https://doi.org/10.3758/s13428-017-0910-x

Haslbeck, J. M. B., & Waldorp, L. J. (2020). mgm: Estimating time-varying mixed graphical models in high-dimensional data. *ArXiv:1510.06871 [Stat]*. http://arxiv.org/abs/1510.06871

Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *NeuroImage*, *84*, 159–168. https://doi.org/10.1016/j.neuroimage.2013.08.028

Hawkins, E., Gathercole, S., Astle, D., The CALM Team, & Holmes, J. (2016). Language problems and ADHD symptoms: How specific are the links? *Brain Sciences*, *6*(4), 50. https://doi.org/10.3390/brainsci6040050

Hendren, R. L., Haft, S. L., Black, J. M., White, N. C., & Hoeft, F. (2018). Recognizing psychiatric comorbidity with reading disorders. *Frontiers in Psychiatry*, *9*, 101. https://doi.org/10.3389/fpsyt.2018.00101

Herdman, A. T., Fujioka, T., Chau, W., Ross, B., Pantev, C., & Picton, T. W. (2006). Cortical oscillations related to processing congruent and incongruent grapheme–phoneme pairs. *Neuroscience Letters*, *399*(1–2), 61–66. https://doi.org/10.1016/j.neulet.2006.01.069

Holmes, J., Guy, J., Kievit, R. A., Bryant, A., Mareva, S., CALM Team, & Gathercole, S. E. (2021). Cognitive dimensions of learning in children with problems in attention, learning, and memory. *Journal of Educational Psychology*, *113*(7), 1454–1480. https://doi.org/10.1037/edu0000644

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037/0033-295X.109.4.679

Horbach, J., Scharke, W., Cröll, J., Heim, S., & Günther, T. (2015). Kindergarteners' performance in a sound–symbol paradigm predicts early reading. *Journal of Experimental Child Psychology*, *139*, 256–264. https://doi.org/10.1016/j.jecp.2015.06.007

Horbach, J., Weber, K., Opolony, F., Scharke, W., Radach, R., Heim, S., & Günther, T. (2018). Performance in sound-symbol learning predicts reading performance 3 years later. *Frontiers in Psychology*, *9*, 1716. https://doi.org/10.3389/fpsyg.2018.01716

Horowitz-Kraus, T. (2011). Does development affect the error-related negativity of impaired and skilled readers?: An ERP study. *Developmental Neuropsychology*, *36*(7), 914–932. https://doi.org/10.1080/87565641.2011.606415

Horowitz-Kraus, T. (2014). Pinpointing the deficit in executive functions in adolescents with dyslexia performing the Wisconsin Card Sorting Test: An ERP Study. *Journal of Learning Disabilities*, *47*(3), 208–223. https://doi.org/10.1177/0022219412453084

Horowitz-Kraus, T. (2016a). Can the error-monitoring system differentiate ADHD from ADHD with reading disability? Reading and executive dysfunction as reflected in error monitoring. *Journal of Attention Disorders*, *20*(10), 889–902. https://doi.org/10.1177/1087054713488440

Horowitz-Kraus, T. (2016b). Improvement of the error-detection mechanism in adults with dyslexia following reading acceleration training: Error detection in dyslexia. *Dyslexia*, *22*(2), 173–189. https://doi.org/10.1002/dys.1523

Horowitz-Kraus, T., & Breznitz, Z. (2008). An error-detection mechanism in reading among dyslexic and regular readers – An ERP study. *Clinical Neurophysiology*, *119*(10), 2238–2246. https://doi.org/10.1016/j.clinph.2008.06.009

Horowitz-Kraus, T., & Breznitz, Z. (2013). Compensated dyslexics have a more efficient error detection system than noncompensated dyslexics. *Journal of Child Neurology*, *28*(10), 1266–1276. https://doi.org/10.1177/0883073812460917

Horowitz-Kraus, T., DiCesare, C., & Kiefer, A. W. (2018). Longer fixation times during reading are correlated with decreased connectivity in cognitive-control brain regions during rest in children: Resting state and eye movements when children read. *Mind, Brain, and Education*, *12*(1), 49–60. https://doi.org/10.1111/mbe.12168

Huang, L.-Y. (2014). Learning to read with the whole language approach: The teacher's view. *English Language Teaching*, *7*(5), p71. https://doi.org/10.5539/elt.v7n5p71

Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). Reading fluency assessment and instruction: What, why, and how? *The Reading Teacher*, *58*(8), 702–715.

Huettig, F., Lachmann, T., Reis, A., & Petersson, K. M. (2018). Distinguishing cause from effect – many deficits associated with developmental dyslexia may be a consequence of reduced and suboptimal reading experience. *Language, Cognition and Neuroscience*, *33*(3), 333–350. https://doi.org/10.1080/23273798.2017.1348528

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017–2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010

Hulme, C., & Snowling, M. J. (2013). *Developmental Disorders of Language Learning and Cognition*. New Jersey: Wiley-Blackwell.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research Domain Criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, *167*(7), 748–751. https://doi.org/10.1176/appi.ajp.2010.09091379

Iversen, S. D., & Dunnett, S. B. (1990). Functional organization of striatum as studied with neural grafts. *Neuropsychologia*, *28*(6), 601–626. https://doi.org/10.1016/0028-3932(90)90037-O

James, K. H. (2017). The importance of handwriting experience on the development of the literate brain. *Current Directions in Psychological Science*, *26*(6), 502–508. https://doi.org/10.1177/0963721417709821

James, K. H., & Atwood, T. P. (2009). The role of sensorimotor learning in the perception of letter-like forms: Tracking the causes of neural specialization for letters. *Cognitive Neuropsychology*, *26*(1), 91–110. https://doi.org/10.1080/02643290802425914

James, K. H., & Engelhardt, L. (2012). The effects of handwriting experience on functional brain development in pre-literate children. *Trends in Neuroscience and Education*, *1*(1), 32–42. https://doi.org/10.1016/j.tine.2012.08.001

Jaramillo-Alcázar, A., Venegas, E., Criollo-C, S., & Luján-Mora, S. (2021). An approach to accessible serious games for people with dyslexia. *Sustainability*, *13*(5), 2507. https://doi.org/10.3390/su13052507

Jiang, H., Fei, X., Liu, H., Roeder, K., Lafferty, J., Wasserman, L., Li, X., & Zhao, and T. (2020). *huge: High-Dimensional Undirected Graph Estimation* (1.3.4) [Computer software]. https://cran.r-project.org/web/packages/huge/index.html

Johnston, R. S., & Thompson, G. B. (1989). Is dependence on phonological information in children’s reading a product of instructional approach? *Journal of Experimental Child Psychology*, *48*(1), 131–145. https://doi.org/10.1016/0022-0965(89)90044-1

Jones, M., Ashby, J., & Branigan, H. (2012). Dyslexia and fluency: Parafoveal and foveal influences on rapid automatized naming. *Journal of Experimental Psychology. Human Perception and Performance*, *39*. https://doi.org/10.1037/a0029710

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*(3), 213–233. https://doi.org/10.1007/s11065-007-9040-z

Karipidis, I. I., Pleisch, G., Brandeis, D., Roth, A., Röthlisberger, M., Schneebeli, M., Walitza, S., & Brem, S. (2018). Simulating reading acquisition: The link between reading outcome and multimodal brain signatures of letter–speech sound learning in prereaders. *Scientific Reports (Nature Publisher Group)*, *8*, 1–13. https://doi.org/10.1038/s41598-018-24909-8

Karipidis, I., Pleisch, G., Di Pietro, S. V., Fraga-González, G., & Brem, S. (2021). Developmental trajectories of letter and speech sound integration during reading acquisition. *Frontiers in Psychology*, *12*, 750491. https://doi.org/10.3389/fpsyg.2021.750491

Karipidis, I., Pleisch, G., Röthlisberger, M., Hofstetter, C., Dornbierer, D., Stämpfli, P., & Brem, S. (2017). Neural initialization of audiovisual integration in prereaders at varying risk for developmental dyslexia. *Human Brain Mapping*, *38*(2), 1038–1055. https://doi.org/10.1002/hbm.23437

Kartal, G., & Terziyan, T. (2016). Development and evaluation of game-like phonological awareness software for kindergarteners: JerenAli. *Journal of Educational Computing Research*, *53*(4), 519–539. https://doi.org/10.1177/0735633115608397

Kazdin, A. E. (2004). Evidence-based treatments: Challenges and priorities for practice and research. *Child and Adolescent Psychiatric Clinics of North America*, *13*(4), 923–940. https://doi.org/10.1016/j.chc.2004.04.002

Kema, G. N., & Kema-van Leggelo, M. K. G. (1987). *G.S.O. 4-6: Handleiding*. Lisse, The Netherlands: Swets & Zeitlinger.

Khanolainen, D. (2020). Longitudinal effects of the home learning environment and parental difficulties on reading and math development across Grades 1–9. *Frontiers in Psychology*, *11*, 20.

Kim, J., Gilbert, J., Yu, Q., & Gale, C. (2021). Measures matter: A meta-analysis of the effects of educational apps on preschool to Grade 3 children's literacy and math skills. *AERA Open*, *7*, 23328584211004184. https://doi.org/10.1177/23328584211004183

Kim, Y.-S., Otaiba, S. A., Puranik, C., Folsom, J. S., & Gruelich, L. (2014). The contributions of vocabulary and letter writing automaticity to word reading and spelling for kindergartners. *Reading and Writing*, *27*(2), 237–253. https://doi.org/10.1007/s11145-013-9440-9

Kirby, J. R., Georgiou, G. K., Martinussen, R., Parrila, R., Bowers, P., & Landerl, K. (2010). Naming speed and reading: From prediction to instruction. *Reading Research Quarterly*, *45*(3), 341–362. https://doi.org/10.1598/RRQ.45.3.4

Klinzing, J. G., Niethard, N., & Born, J. (2019). Mechanisms of systems memory consolidation during sleep. *Nature Neuroscience*, *22*(10), 1598–1610. https://doi.org/10.1038/s41593-019-0467-3

Knoop-van Campen, C. A. N., Segers, E., & Verhoeven, L. (2018). How phonological awareness mediates the relation between working memory and word reading efficiency in children with dyslexia. *Dyslexia*, *24*(2), 156–169. https://doi.org/10.1002/dys.1583

Kohli, A., & Kaur, M. (2006). Wisconsin Card Sorting Test: Normative data and experience. *Indian Journal of Psychiatry*, *48*(3), 181–184. https://doi.org/10.4103/0019-5545.31582

Kohnen, S., Banales, E., & McArthur, G. (2021). Videoconferencing interventions for children with reading and spelling difficulties: A pilot study. *Telemedicine and E-Health*, *27*(5), 537–543. https://doi.org/10.1089/tmj.2020.0061

Kort, W., Compaan, E. L., Bleichrodt, N., Resing, W. C. M., Schittekatte, M., Bosman, M., Vermeir, G., & Verhaeghe, P. (2002). *Nederlandse bewerking van de Wechsler Intelligence Scale for Children III (WISC-III). Test en handleiding.* NIP Dienstencentrum.

Kovachy, V. N., Adams, J. N., Tamaresis, J. S., & Feldman, H. M. (2015). Reading abilities in school-aged preterm children: A review and meta-analysis. *Developmental Medicine and Child Neurology*, *57*(5), 410–419. https://doi.org/10.1111/dmcn.12652

Krahenbuhl, K. S. (2016). Student-centered education and constructivism: Challenges, concerns, and clarity for teachers. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *89*(3), 97–105. https://doi.org/10.1080/00098655.2016.1191311

Krämer, N., Schäfer, J., & Boulesteix, A.-L. (2009). Regularized estimation of large-scale gene association networks using graphical Gaussian models. *BMC Bioinformatics*, *10*, 384. https://doi.org/10.1186/1471-2105-10-384

Kraus, D., & Horowitz-Kraus, T. (2014). The effect of learning on feedback-related potentials in adolescents with dyslexia: An EEG-ERP study. *PLoS ONE*, *9*(6), e100486. https://doi.org/10.1371/journal.pone.0100486

Kronschnabel, J., Brem, S., Maurer, U., & Brandeis, D. (2014). The level of audiovisual print–speech integration deficits in dyslexia. *Neuropsychologia*, *62*, 245–261. https://doi.org/10.1016/j.neuropsychologia.2014.07.024

Kwak, S. G., & Kim, J. H. (2017). Central limit theorem: The cornerstone of modern statistics. *Korean Journal of Anesthesiology*, *70*(2), 144. https://doi.org/10.4097/kjae.2017.70.2.144

Kyllonen, P. C., Tirre, W. C., & Christal, R. E. (1991). Knowledge and processing speed as determinants of associative learning. *Journal of Experimental Psychology: General*, *120*, 57–79. https://doi.org/10.1037/0096-3445.120.1.57

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*(2), 293–323. https://doi.org/10.1016/0010-0285(74)90015-2

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

Landi, N., Frost, S. J., Mencl, W. E., Preston, J. L., Jacobsen, L. K., Lee, M., Yrigollen, C., Pugh, K. R., & Grigorenko, E. L. (2013). The COMT Val/Met polymorphism is associated with reading-related skills and consistent patterns of functional neural activation. *Developmental Science*, *16*(1), 13–23. https://doi.org/10.1111/j.1467-7687.2012.01180.x

Landsbergen, F. (2021). *Het dikke alfabetboek*. Tielt, The Netherlands: Lannoo.

Lassault, J., Sprenger-Charolles, L., Albrand, J.-P., Alavoine, E., Richardson, U., Lyytinen, H., & Ziegler, J. C. (2022). Testing the effects of GraphoGame against a computer-assisted math intervention in primary school. *Scientific Studies of Reading*, *26*(6), 1–20. https://doi.org/10.1080/10888438.2022.2052884

Law, J. M., De Vos, A., Vanderauwera, J., Wouters, J., Ghesquière, P., & Vandermosten, M. (2018). Grapheme-phoneme learning in an unknown orthography: a study in typical reading and dyslexic children. *Frontiers in Psychology*, *9*, 1393. https://doi.org/10.3389/fpsyg.2018.01393

Livingston, E. M., Siegel, L. S., & Ribary, U. (2018). Developmental dyslexia: Emotional impact and consequences. *Australian Journal of Learning Difficulties*, *23*(2), 107–135. https://doi.org/10.1080/19404158.2018.1479975

Luft, C. D. B. (2014). Learning from feedback: The neural mechanisms of feedback processing facilitating better performance. *Behavioural Brain Research*, *261*, 356–368. https://doi.org/10.1016/j.bbr.2013.12.043

Lukie, C. N., Montazer-Hojat, S., & Holroyd, C. B. (2014). Developmental changes in the reward positivity: An electrophysiological trajectory of reward processing. *Developmental Cognitive Neuroscience*, *9*, 191–199. https://doi.org/10.1016/j.dcn.2014.04.003

Lyytinen, H., Erskine, J., Kujala, J., Ojanen, E., & Richardson, U. (2009). In search of a science-based application: A learning tool for reading acquisition. *Scandinavian Journal of Psychology*, *50*(6), 668–675. https://doi.org/10.1111/j.1467-9450.2009.00791.x

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163–203. https://doi.org/10.1037/0033-2909.109.2.163

Majerus, S., & Cowan, N. (2016). The nature of verbal short-term impairment in dyslexia: the importance of serial order. *Frontiers in Psychology*, *7*, 1522. https://doi.org/10.3389/fpsyg.2016.01522

Makeig, S., Jung, T.-P., Bell, A. J., Ghahremani, D., & Sejnowski, T. J. (1997). Blind separation of auditory event-related brain responses into independent components. *Proceedings of the National Academy of Sciences*, *94*(20), 10979–10984. https://doi.org/10.1073/pnas.94.20.10979

Margari, L., Buttiglione, M., Craig, F., Cristella, A., de Giambattista, C., Matera, E., Operto, F., & Simone, M. (2013). Neuropsychopathological comorbidities in learning disorders. *BMC Neurology*, *13*(1), 198–198. https://doi.org/10.1186/1471-2377-13-198

Mascheretti, S., Bureau, A., Battaglia, M., Simone, D., Quadrelli, E., Croteau, J., Cellino, M. R., Giorda, R., Beri, S., Maziade, M., & Marino, C. (2013). An assessment of gene-by-environment interactions in developmental dyslexia-related phenotypes: Gene-by-environment interaction in dyslexia. *Genes, Brain and Behavior*, *12*(1), 47–55. https://doi.org/10.1111/gbb.12000

Matteucci, M. C., Scalone, L., Tomasetto, C., Cavrini, G., & Selleri, P. (2019). Health-related quality of life and psychological wellbeing of children with Specific Learning Disorders and their mothers. *Research in Developmental Disabilities*, *87*, 43–53. https://doi.org/10.1016/j.ridd.2019.02.003

Maughan, B., Rowe, R., Loeber, R., & Stouthamer-Loeber, M. (2003). Reading problems and depressed mood. *Journal of Abnormal Child Psychology*, *31*(2), 219–229. https://doi.org/10.1023/A:1022534527021

Maurer, U., Brem, S., Bucher, K., & Brandeis, D. (2005). Emerging neurophysiological specialization for letter strings. *Journal of Cognitive Neuroscience*, *17*(10), 1532–1552. https://doi.org/10.1162/089892905774597218

Mausbach, B. T., Moore, R., Roesch, S., Cardenas, V., & Patterson, T. L. (2010). The relationship between homework compliance and therapy outcomes: An updated meta-analysis. *Cognitive Therapy and Research*, *34*(5), 429–438. https://doi.org/10.1007/s10608-010-9297-z

Mayer, R. E. (2011). Multimedia learning and games. *Computer Games and Instruction*, 281–305. Charlotte, NC: Scopus.

Mazza, S., Gerbier, E., Gustin, M.-P., Kasikci, Z., Koenig, O., Toppino, T. C., & Magnin, M. (2016). Relearn faster and retain longer: along with practice, sleep makes perfect.

Psychological Science, *27*(10), 1321–1330. https://doi.org/10.1177/0956797616659930

McGrath, L. M., & Stoodley, C. J. (2019). Are there shared neural correlates between dyslexia and ADHD? A meta-analysis of voxel-based morphometry studies. *Journal of Neurodevelopmental Disorders*, *11*(1), 31. https://doi.org/10.1186/s11689-019-9287-8

McTigue, E. M., Solheim, O. J., Zimmer, W. K., & Uppstad, P. H. (2020). Critically reviewing GraphoGame across the world: Recommendations and cautions for research and implementation of computer-assisted instruction for word-reading acquisition. *Reading Research Quarterly*, *55*(1), 45–73. https://doi.org/10.1002/rrq.256

Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*(2), 322–352. https://doi.org/10.1037/a0026744

Menon, V., & D’Esposito, M. (2022). The role of PFC networks in cognitive control and executive function. *Neuropsychopharmacology*, *47*(1), 1. https://doi.org/10.1038/s41386-021-01152-w

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., Elbourne, D., Egger, M., & Altman, D. G. (2010). CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *BMJ (Clinical Research Ed.)*, *340*, c869. https://doi.org/10.1136/bmj.c869

Moll, K., Kunze, S., Neuhoff, N., Bruder, J., & Schulte-Körne, G. (2014). Specific learning disorder: Prevalence and gender differences. *PLoS ONE*, *9*(7), e103537. https://doi.org/10.1371/journal.pone.0103537

Moll, K., & Landerl, K. (2009). Double dissociation between reading and spelling deficits. *Scientific Studies of Reading*, *13*(5), 359–382. https://doi.org/10.1080/10888430903162878

Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., Streiftau, S., Lyytinen, H., Leppänen, P. H. T., Lohvansuu, K., Tóth, D., Honbolygó, F., Csépe, V., Bogliotti, C., Iannuzzi, S., Démonet, J.-F., Longeras, E., Valdois, S., George, F., … Landerl, K. (2014). Cognitive mechanisms underlying reading and spelling development in five

European orthographies. *Learning and Instruction*, *29*, 65–77. https://doi.org/10.1016/j.learninstruc.2013.09.003

Mommers, C., Verhoeven, L., & van der Linden, S. (1994). *Handleiding Veilig leren lezen. Structuur methode voor het aanvankelijk leesonderwijs*. Tilburg, The Netherlands: Zwijsen B.V.

Montessori, M. (1912). *The Montessori method*. New York, US: Frederick Stokes Company.

Mugnaini, D., Lassi, S., La Malfa, G., & Albertini, G. (2009). Internalizing correlates of dyslexia. *World Journal of Pediatrics*, *5*(4), 255–264. https://doi.org/10.1007/s12519-009-0049-7

Muijselaar, M. M. L., Swart, N. M., Steenbeek-Planting, E. G., Droop, M., Verhoeven, L., & de Jong, P. F. (2017). Developmental relations between reading comprehension and reading strategies. *Scientific Studies of Reading*, *21*(3), 194–209. https://doi.org/10.1080/10888438.2017.1278763

Muis, K. R., Ranellucci, J., Trevors, G., & Duffy, M. C. (2015). The effects of technology-mediated immediate feedback on kindergarten students' attitudes, emotions, engagement and learning outcomes during literacy skills development. *Learning and Instruction*, *38*, 1–13. https://doi.org/10.1016/j.learninstruc.2015.02.001

Mulder, F., Timman, Y., & Verhallen, S. (2009). *Handreiking bij de Basiswoordenlijst Amsterdamse Kleuters (BAK)*. Amsterdam: ITTA.

Nagarajan, S., Mahncke, H., Salz, T., Tallal, P., Roberts, T., & Merzenich, M. M. (1999). Cortical auditory signal processing in poor readers. *Proceedings of the National Academy of Sciences*, *96*(11), 6483–6488. https://doi.org/10.1073/pnas.96.11.6483

Nagler, T., Korinth, S. P., Linkersdörfer, J., Lonnemann, J., Rump, B., Hasselhorn, M., & Lindberg, S. (2015). Text-fading based training leads to transfer effects on children's sentence reading fluency. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00119

Narciss, S., Sosnovsky, S., Schnaubert, L., Andrès, E., Eichelmann, A., Goguadze, G., & Melis, E. (2014). Exploring feedback and student characteristics relevant for personalizing feedback strategies. *Computers & Education*, *71*, 56–76. https://doi.org/10.1016/j.compedu.2013.09.011

Nash, H. M., Gooch, D., Hulme, C., Mahajan, Y., McArthur, G., Steinmetzger, K., & Snowling, M. J. (2017). Are the literacy difficulties that characterize developmental dyslexia associated with a failure to integrate letters and speech sounds? *Developmental Science*, *20*(4), e12423. https://doi.org/10.1111/desc.12423

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769)*. National Institute of Child Health & Development. https:// www1.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf

Neumann, M. M. (2018). Using tablets and apps to enhance emergent literacy skills in young children. *Early Childhood Research Quarterly*, *42*, 239–246. https://doi.org/10.1016/j.ecresq.2017.10.006

Norton, E. S., & Wolf, M. (2012). Rapid Automatized Naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology*, *63*(1), 427–452. https://doi.org/10.1146/annurev-psych-120710-100431

Nyhus, E., & Barceló, F. (2009). The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: A critical update. *Brain and Cognition*, *71*(3), 437–451. https://doi.org/10.1016/j.bandc.2009.03.005

O'Brien, G., & Yeatman, J. D. (2020). Bridging sensory and language theories of dyslexia: Toward a multifactorial model. *Developmental Science*, *24*(3). https://doi.org/10.1111/desc.13039

O'Connor, R. E., White, A., & Swanson, H. L. (2007). Repeated reading versus continuous reading: influences on reading fluency and comprehension. *Exceptional Children*, *74*(1), 31–46. https://doi.org/10.1177/001440290707400102

Ozernov-Palchik, O., & Gaab, N. (2016). Tackling the 'dyslexia paradox': Reading brain and behavior for early markers of developmental dyslexia: Tackling the 'dyslexia paradox.' *Wiley Interdisciplinary Reviews: Cognitive Science*, *7*(2), 156–176. https://doi.org/10.1002/wcs.1383

Ozernov-Palchik, O., Yu, X., Wang, Y., & Gaab, N. (2016). Lessons to be learned: How a comprehensive neurobiological framework of atypical reading development can inform educational practice. *Current Opinion in Behavioral Sciences*, *10*, 45–58. https://doi.org/10.1016/j.cobeha.2016.05.006

Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, *6*(4), 351–371. https://doi.org/10.1007/BF02213420

Paige, D. D. (2011). Testing the acceleration hypothesis: Fluency outcomes utilizing still-versus accelerated-text in sixth-grade students with reading disabilities. *Literacy Research and Instruction*, *50*(4), 294–312. https://doi.org/10.1080/19388071.2010.518661

Parkin, A. J., & Streete, S. (1988). Implicit and explicit memory in young children and adults. *British Journal of Psychology*, *79*(3), 361. https://doi.org/10.1111/j.2044-8295.1988.tb02295.x

Partanen, M., & Siegel, L. S. (2014). Long-term outcome of the early identification and intervention of reading disabilities. *Reading and Writing*, *27*(4), 665–684. https://doi.org/10.1007/s11145-013-9472-1

Patel, P., Torppa, M., Aro, M., Richardson, U., & Lyytinen, H. (2021). Assessing the effectiveness of a game-based phonics intervention for first and second grade English language learners in India: A randomized controlled trial. *Journal of Computer Assisted Learning*, *38*(1), jcal.12592. https://doi.org/10.1111/jcal.12592

Pecini, C., Spoglianti, S., Bonetti, S., Di Lieto, M. C., Guaran, F., Martinelli, A., Gasperini, F., Cristofani, P., Casalini, C., Mazzotti, S., Salvadorini, R., Bargagna, S., Palladino, P., Cismondo, D., Verga, A., Zorzi, C., Brizzolara, D., Vio, C., & Chilosi, A. M. (2019). Training RAN or reading? A telerehabilitation study on developmental dyslexia. *Dyslexia*, dys.1619. https://doi.org/10.1002/dys.1619

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*, 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, *101*(2), 385–413. https://doi.org/10.1016/j.cognition.2006.04.008

Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development*, *72*(3), 816–833. https://doi.org/10.1111/1467-8624.00317

Pennington, B. F., Santerre-Lemmon, L., Rosenberg, J., MacDonald, B., Boada, R., Friend, A., Leopold, D. R., Samuelsson, S., Byrne, B., Willcutt, E. G., & Olson, R. K. (2012). Individual prediction of dyslexia by single versus multiple deficit models. *Journal of Abnormal Psychology*, *121*(1), 212–224. https://doi.org/10.1037/a0025823

Perry, C., Zorzi, M., & Ziegler, J. C. (2019). Understanding dyslexia through personalized large-scale computational models. *Psychological Science*, *30*(3), 386–395. https://doi.org/10.1177/0956797618823540

Peters, S., Braams, B. R., Raijmakers, M. E. J., Koolschijn, P. C. M. P., & Crone, E. A. (2014). The neural coding of feedback learning across child and adolescent development. *Journal of Cognitive Neuroscience*, *26*(8), 1705–1720. https://doi.org/10.1162/jocn_a_00594

Peterson, R. L., & Pennington, B. F. (2015). Developmental Dyslexia. *Annual Review of Clinical Psychology*, *11*(1), 283–307. https://doi.org/10.1146/annurev-clinpsy-032814-112842

Pfost, M., Blatter, K., Artelt, C., Stanat, P., & Schneider, W. (2019). Effects of training phonological awareness on children's reading skills. *Journal of Applied Developmental Psychology*, *65*, 101067. https://doi.org/10.1016/j.appdev.2019.101067

Pleisch, G., Karipidis, I., Brauchli, C., Röthlisberger, M., Hofstetter, C., Stämpfli, P., Walitza, S., & Brem, S. (2019). Emerging neural specialization of the ventral occipitotemporal cortex to characters through phonological association learning in preschool children. *NeuroImage*, *189*, 813–831. https://doi.org/10.1016/j.neuroimage.2019.01.046

Poljac, E., Simon, S., Ringlever, L., Kalcik, D., Groen, W. B., Buitelaar, J. K., & Bekkering, H. (2010). Impaired task switching performance in children with dyslexia but not in children with autism. *Quarterly Journal of Experimental Psychology*, *63*(2), 401–416. https://doi.org/10.1080/17470210902990803

Premeti, A., Bucci, M. P., & Isel, F. (2022). Evidence from ERP and eye movements as markers of language dysfunction in dyslexia. *Brain Sciences*, *12*(1), 1. https://doi.org/10.3390/brainsci12010073

Prensky, M. (2003). Digital game-based learning. *Computers in Entertainment*, *1*(1), 21. https://doi.org/10.1145/950566.950596

Preston, J. L., Molfese, P. J., Frost, S. J., Mencl, W. E., Fulbright, R. K., Hoeft, F., Landi, N., Shankweiler, D., & Pugh, K. R. (2016). Print-speech convergence predicts future reading outcomes in early readers. *Psychological Science*, *27*(1), 75–84. https://doi.org/10.1177/0956797615611921

Quinn, J. M., Wagner, R. K., Petscher, Y., & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development*, *86*(1), 159–175. https://doi.org/10.1111/cdev.12292

R Core Team. (2019). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. https://www.R-project.org

Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *Quarterly Journal of Experimental Psychology*, *61*(1), 129–141. https://doi.org/10.1080/17470210701508822

Rastle, K., Lally, C., Davis, M. H., & Taylor, J. S. H. (2021). The dramatic impact of explicit instruction on learning to read in a new writing system. *Psychological Science*, 095679762096879. https://doi.org/10.1177/0956797620968790

Raven, J. C., Court, J. H., & Raven, J. (1984). *Manual for Raven's progressive matrices and vocabulary scales*. London: Lewis.

Raven, J. C., Court, J. H., & Raven, J. E. (1998). *Manual for Raven's progressive matrices and vocabulary scales*. Bloomington, MN: Pearson.

Raven, J. C., & Raven, J. (2020). *Raven's 2 Progressive Matrices | Diagnostische versie*. https://www.pearsonclinical.nl/ravens2

Reitsma, P., & Verhoeven, L. (2016). *Acquisition of Reading in Dutch*. Dordrecht, The Netherlands: Floris Publications.

Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.

Reynolds, R. E. (2000). Attentional resource emancipation: toward understanding the interaction of word identification and comprehension processes in reading. *Scientific Studies of Reading*, *4*(3), 169–195. https://doi.org/10.1207/S1532799XSSR0403_1

Richardson, U., & Lyytinen, H. (2014). The GraphoGame method: The theoretical and methodological background of the technology-enhanced learning environment for learning to read. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, *10*(1), 39–60. https://doi.org/10.17011/ht/urn.201405281859

Ridderinkhof, K. R., van den Wildenberg, W. P. M., Segalowitz, S. J., & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition*, *56*(2), 129–140. https://doi.org/10.1016/j.bandc.2004.09.016

Roberts, J., Hodgson, R., & Dolan, P. (2011). "It's driving her mad": Gender differences in the effects of commuting on psychological health. *Journal of Health Economics*, *30*(5), 1064–1076. https://doi.org/10.1016/j.jhealeco.2011.07.006

Romanovska, L., & Bonte, M. (2021). How learning to read changes the listening brain. *Frontiers in Psychology*, *12*, 726882. https://doi.org/10.3389/fpsyg.2021.726882

Romanovska, L., Janssen, R., & Bonte, M. (2021). Cortical responses to letters and ambiguous speech vary with reading skills in dyslexic and typically reading children. *NeuroImage: Clinical*, *30*, 102588. https://doi.org/10.1016/j.nicl.2021.102588

Roos, L. E., Pears, K., Bruce, J., Kim, H. K., & Fisher, P. A. (2015). Impulsivity and the association between the feedback-related negativity and performance on an inhibitory control task in young at-risk children: Feedback processing in at-risk children. *Psychophysiology*, *52*(5), 704–713. https://doi.org/10.1111/psyp.12389

Rose, J. (2006). *Independent review of the teaching of early reading: Final report. UK: Department for Education and Skills*. London, UK: TSO.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

RStudio Team. (2022). *RStudio: Integrated Development for R.* RStudio, Inc. http://www.rstudio.com/

Saan, R. J., & Deelman, B. G. (1986). *De 15-woordentest A en B (Een voorlopige handleiding)*. Afdeling Neuropsychologie, AZG.

Saine, N. L., Lerkkanen, M.-K., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2011). Computer-assisted remedial reading intervention for school beginners at risk for reading disability: Computer-assisted reading intervention. *Child Development*, *82*(3), 1013–1028. https://doi.org/10.1111/j.1467-8624.2011.01580.x

Samuels, S. J. (2006). Toward a model of reading fluency. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 24–46). Newark, US: International Reading Association.

Sanislow, C. A., Ferrante, M., Pacheco, J., Rudorfer, M. V., & Morris, S. E. (2019). Advancing translational research using NIMH research domain criteria and computational methods. *Neuron*, *101*(5), 779–782. https://doi.org/10.1016/j.neuron.2019.02.024

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

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: phonological awareness and some other promising predictors. In P. Accardo, A. Capute, & B. Shapiro (Eds.), *Specific Reading Disability: A View of the Spectrum* (pp. 75–119). Timonium, US: York Press.

Schaerlaekens, A., Kohnstamm, D., & Lejaegere, M. (1999). *Streeflijst woordenschat voor zesjarigen [Word frequency list for Dutch 6-year-olds]*. Lisse, The Netherlands: Swets & Zeitlinger.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2). 23-74.

Schlichting, L. (2005). *Peabody Picture Vocabulary Test-III-NL*. Amsterdam, The Netherlands: Harcourt Assessment B.V.

Schmalz, X., Marinus, E., Coltheart, M., & Castles, A. (2015). Getting to the bottom of orthographic depth. *Psychonomic Bulletin & Review*, *22*(6), 1614–1629. https://doi.org/10.3758/s13423-015-0835-2

Schneider, W., & Chein, J. M. (2003). Controlled & automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science*, *27*(3), 525–559. https://doi.org/10.1207/s15516709cog2703_8

Scholte, E. M., & Van der Ploeg, J. D. (2005). *Handleiding ADHD-Vragenlijst (AVL) [Manual ADHD-Questionnaire (AVL)]*. Houten, The Netherlands: Bohn Stafleu Van Loghum.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461. https://doi.org/10.1214/aos/1176344136

Schwertman, N. C., Owens, M. A., & Adnan, R. (2004). A simple more general boxplot method for identifying outliers. *Computational Statistics & Data Analysis*, *47*(1), 165–174. https://doi.org/10.1016/j.csda.2003.10.012

Segers, E., Damhuis, C. M. P., Van De Sande, E., & Verhoeven, L. (2016). Role of executive functioning and home environment in early reading development. *Learning and Individual Differences*, *49*, 251–259. https://doi.org/10.1016/j.lindif.2016.07.004

Sesma, H. W., Mahone, E. M., Levine, T., Eason, S. H., & Cutting, L. E. (2009). The contribution of executive skills to reading comprehension. *Child Neuropsychology*, *15*(3), 232–246. https://doi.org/10.1080/09297040802220029

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143–174. https://doi.org/10.1348/000712603321661859

Shapiro, L. R., & Solity, J. (2008). Delivering phonological and phonics training within whole-class teaching. *British Journal of Educational Psychology*, *78*(4), 597–620. https://doi.org/10.1348/000709908X293850

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218. https://doi.org/10.1016/0010-0277(94)00645-2

Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine*, *326*(3), 145–150. https://doi.org/10.1056/NEJM199201163260301

Shaywitz, S. E., Morris, R., & Shaywitz, B. A. (2008). The education of dyslexic children from childhood to young adulthood. *Annual Review of Psychology*, *59*(1), 451–475. https://doi.org/10.1146/annurev.psych.59.103006.093633

Shreffler, J., & Huecker, M. R. (2022). Type I and Type II errors and statistical power. In *StatPearls [Internet]*. StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK557530/

Skiada, R., Soroniati, E., Gardeli, A., & Zissis, D. (2014). EasyLexia: A mobile application for children with learning difficulties. *Procedia Computer Science*, *27*, 218-228. https://doi.org/10.1016/j.procs.2014.02.025

Smart, D., Youssef, G. J., Sanson, A., Prior, M., Toumbourou, J. W., & Olsson, C. A. (2017). Consequences of childhood reading difficulties and behaviour problems for educational achievement and employment in early adulthood. *British Journal of Educational Psychology*, *87*(2), 288–308. https://doi.org/10.1111/bjep.12150

Smith-Spark, J. H., Henry, L. A., Messer, D. J., Edvardsdottir, E., & Zięcik, A. P. (2016). Executive functions in adults with developmental dyslexia. *Research in Developmental Disabilities*, *53–54*, 323–341. https://doi.org/10.1016/j.ridd.2016.03.001

Smulders, S. F. A., Soetens, E., & van der Molen, M. W. (2016). What happens when children encounter an error? *Brain and Cognition*, *104*, 34–47. https://doi.org/10.1016/j.bandc.2016.02.004

Snellings, P., van der Leij, A., de Jong, P. F., & Blok, H. (2009). Enhancing the reading fluency and comprehension of children with reading disabilities in an orthographically transparent language. *Journal of Learning Disabilities*, *42*(4), 291–305. https://doi.org/10.1177/0022219408331038

Snowling, M. J. (1998). Dyslexia as a phonological deficit: Evidence and implications. *Child Psychology and Psychiatry Review*, *3*(1), 4–11. https://doi.org/10.1017/S1360641797001366

Snowling, M. J. (2000). *Dyslexia, Second edition*. Malden: Blackwell Publishers Inc.

Snowling, M. J. (2013). Early identification and interventions for dyslexia: A contemporary view: Early identification and interventions for dyslexia: a contemporary view. *Journal of Research in Special Educational Needs*, *13*(1), 7–14. https://doi.org/10.1111/j.1471-3802.2012.01262.x

Snowling, M. J., Gallagher, A., & Frith, U. (2003). Family risk of dyslexia is continuous: individual differences in the precursors of reading skill. *Child Development*, *74*(2), 358–373. https://doi.org/10.1111/1467-8624.7402003

Snowling, M. J., & Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological Bulletin*, *142*(5), 498–545. https://doi.org/10.1037/bul0000037

Stafford, T., & Vaci, N. (2022). Maximizing the potential of digital games for understanding skill acquisition. *Current Directions in Psychological Science*, *31*(1), 49–55. https://doi.org/10.1177/09637214211057841

Stanovich, K. E. (1996). Toward a more inclusive definition of dyslexia. *Dyslexia*, *2*(3), 154–166. https://doi.org/10.1002/(SICI)1099-0909(199611)2:3<154::AID-DYS63>3.0.CO;2-B

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, *9*(4), 255–266. http://dx.doi.org/10.1038/nrn2331

Stein, J. (2001). The magnocellular theory of developmental dyslexia. *Dyslexia (Chichester, England)*, *7*(1), 12–36. https://doi.org/10.1002/dys.186

Stein, J., & Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends in Neurosciences*, *20*(4), 147–152. https://doi.org/10.1016/S0166-2236(96)01005-3

Straatemeier, M. (2014). *Math Garden: A new educational and scientific instrument.* https://dare.uva.nl/search?identifier=67ab8977-5e36-4f9f-962d-be294f2b8c09

Struiksma, C., van der Leij, A., & Stoel, R. (2009). Response to fluency-oriented intervention of Dutch poor readers. *Learning and Individual Differences*, *19*(4), 541–548. https://doi.org/10.1016/j.lindif.2009.07.005

Swanson, H. L., Xinhua Zheng, & Jerman, O. (2009). Working memory, short-term memory, and reading disabilities: A selective meta-analysis of the literature. *Journal of Learning Disabilities*, *42*(3), 260–287. https://doi.org/10.1177/0022219409331958

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, *14*(9), 400–410. https://doi.org/10.1016/j.tics.2010.06.008

Tamnes, C. K., Walhovd, K. B., Torstveit, M., Sells, V. T., & Fjell, A. M. (2013). Performance monitoring in children and adolescents: A review of developmental changes in the error-related negativity and brain maturation. *Developmental Cognitive Neuroscience*, *6*, 1–13. https://doi.org/10.1016/j.dcn.2013.05.001

Taylor, O. D., Armfield, N. R., Dodrill, P., & Smith, A. C. (2014). A review of the efficacy and effectiveness of using telehealth for paediatric speech and language assessment. *Journal* *of Telemedicine and Telecare*, *20*(7), 405–412. https://doi.org/10.1177/1357633X14552388

ten Braak, D., Kleemans, T., Størksen, I., Verhoeven, L., & Segers, E. (2018). Domain-specific effects of attentional and behavioral control in early literacy and numeracy development. *Learning and Individual Differences*, *68*, 61–71. https://doi.org/10.1016/j.lindif.2018.10.001

ten Oever, S., Romei, V., van Atteveldt, N., Soto-Faraco, S., Murray, M. M., & Matusz, P. J. (2016). The COGs (context, object, and goals) in multisensory processing. *Experimental Brain Research*, *234*(5), 1307–1323. https://doi.org/10.1007/s00221-016-4590-z

Thomas, M. S. C. (2020). Developmental disorders: Few specific disorders and no specific brain regions. *Current Biology*, *30*(7), R304–R306. https://doi.org/10.1016/j.cub.2020.02.019

Thompson, P. A., Hulme, C., Nash, H. M., Gooch, D., Hayiou-Thomas, E., & Snowling, M. J. (2015). Developmental dyslexia: Predicting individual risk. *Journal of Child Psychology and Psychiatry*, *56*(9), 976–987. https://doi.org/10.1111/jcpp.12412

Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Tijms, J. (2007). The development of reading accuracy and reading rate during treatment of dyslexia. *Educational Psychology*, *27*(2), 273–294. https://doi.org/10.1080/01443410601066800

Tijms, J. (2011). Effectiveness of computer-based treatment for dyslexia in a clinical care setting: Outcomes and moderators. *Educational Psychology*, *31*(7), 873–896. https://doi.org/10.1080/01443410.2011.621403

Tijms, J., de Bree, E. H., Bonte, M., de Jong, P. F., Loykens, E., & Reij, R. (2021). *Protocol Dyslexie Diagnostiek en Behandeling 3.0*. Nederlands Kwaliteitsinstituut Dyslexie (NKD)

Tilanus, E. A. T., Segers, E., & Verhoeven, L. (2016). Responsiveness to intervention in children with dyslexia. *Dyslexia*, *22*(3), 214–232. https://doi.org/10.1002/dys.1533

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology*, *40*(1), 7–26. https://doi.org/10.1016/S0022-4405(01)00092-9

Torgesen, J. K. (2005). Recent discoveries from research on remedial interventions for children with dyslexia. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 521–537). Oxford: Blackwell Publishing.

Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities*, *34*(1), 33–58.

Torppa, M., Lyytinen, P., Erskine, J., Eklund, K., & Lyytinen, H. (2010). Language development, literacy skills, and predictive connections to reading in Finnish children with and without familial risk for dyslexia. *Journal of Learning Disabilities*, *43*(4), 308–321. https://doi.org/10.1177/0022219410369096

Torppa, M., Poikkeus, A.-M., Laakso, M.-L., Tolvanen, A., Leskinen, E., Leppanen, P. H. T., Puolakanaho, A., & Lyytinen, H. (2007). Modeling the early paths of phonological awareness and factors supporting its development in children with and without familial risk of dyslexia. *Scientific Studies of Reading*, *11*(2), 73–103. https://doi.org/10.1080/10888430709336554

Vaessen, A., & Blomert, L. (2010). Long-term cognitive dynamics of fluent reading development. *Journal of Experimental Child Psychology*, *105*(3), 213–231. https://doi.org/10.1016/j.jecp.2009.11.005

van Atteveldt, N., Formisano, E., Blomert, L., & Goebel, R. (2006). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cerebral Cortex*, *17*(4), 962–974. https://doi.org/10.1093/cercor/bhl007

van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron*, *43*(2), 271–282. https://doi.org/10.1016/j.neuron.2004.06.025

van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2007). Top–down task effects overrule automatic multisensory responses to letter–sound pairs in auditory association cortex. *NeuroImage,* *36*(4), 1345-1360. https://doi.org/10.1016/j.neuroimage.2007.03.065

van Bergen, E., de Jong, P. F., Maassen, B., & van der Leij, A. (2014). The effect of parents' literacy skills and children's preliteracy skills on the risk of dyslexia. *Journal of Abnormal Child Psychology*, *42*(7), 1187–1200. https://doi.org/10.1007/s10802-014-9858-9 van Bergen, E., van der Leij, A., & de Jong, P. F. (2014). The intergenerational multiple deficit model and the case of dyslexia. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00346

Van Breukelen, G. J. P. (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology*, *59*(9), 920–925. https://doi.org/10.1016/j.jclinepi.2006.02.007

van de Ven, M., de Leeuw, L., van Weerdenburg, M., & Steenbeek-Planting, E. g. (2017). Early reading intervention by means of a multicomponent reading game. *Journal of Computer Assisted Learning*, *33*(4), 320–333. https://doi.org/10.1111/jcal.12181

Van De Voorde, S., Roeyers, H., & Wiersema, J. R. (2010). Error monitoring in children with ADHD or reading disorder: An event-related potential study. *Biological Psychology*, *84*(2), 176–185. https://doi.org/10.1016/j.biopsycho.2010.01.011

van der Helden, J., Boksem, M. A. S., & Blom, J. H. G. (2010). The importance of failure: Feedback-related negativity predicts motor learning efficiency. *Cerebral Cortex*, *20*(7), 1596–1603. https://doi.org/10.1093/cercor/bhp224

van Uittert, A., Verhoeven, L., & Segers, E. (2021). Responsiveness to a game-based intervention to enhance reading efficiency in first graders. *Journal of Computer Assisted Learning*, *38*(1), 178–191. https://doi.org/10.1111/jcal.12599

Vanden Bempt, F., Economou, M., Van Herck, S., Vanderauwera, J., Glatz, T., Vandermosten, M., Wouters, J., & Ghesquière, P. (2021). Digital game-based phonics instruction promotes print knowledge in pre-readers at cognitive risk for dyslexia. *Frontiers in Psychology*, *12*, 720548. https://doi.org/10.3389/fpsyg.2021.720548

Vander Stappen, C., Dricot, L., & Van Reybroeck, M. (2020). RAN training in dyslexia: Behavioral and brain correlates. *Neuropsychologia*, *146*, 107566. https://doi.org/10.1016/j.neuropsychologia.2020.107566

Varvara, P., Varuzza, C., Sorrentino, A. C. P., Vicari, S., & Menghini, D. (2014). Executive functions in developmental dyslexia. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00120

Vaughn, S., Wanzek, J., Murray, C. S., Scammacca, N., Linan-Thompson, S., & Woodruff, A. L. (2009). Response to early reading intervention examining higher and lower responders. *Exceptional Children*, *75*(2), 165–183. https://doi.org/10.1177/001440290907500203

Verhagen, W., Aarnoutse, C., & van Leeuwe, J. (2008). Phonological awareness and naming speed in the prediction of dutch children's word recognition. *Scientific Studies of Reading*, *12*(4), 301–324. https://doi.org/10.1080/10888430802132030

Verhoeven, L. (1991). *Drie-minuten-toets [Word decoding test]*. Arnhem, The Netherlands: CITO.

Verwimp, C., Snellings, P., Wiers, R. W., & Tijms, J. (2023). Goal-directedness enhances letter-speech sound learning and consolidation in an unknown orthography. *Child Development*, *n/a*(n/a). https://doi.org/10.1111/cdev.13901

Verwimp, C., Tijms, J., Snellings, P., Haslbeck, J. M. B., & Wiers, R. W. (2021). A network approach to dyslexia: Mapping the reading network. *Development and Psychopathology*, 1–15. https://doi.org/10.1017/S0954579421000365

Vidal, C., Content, A., & Chetail, F. (2017). BACS: The Brussels Artificial Character Sets for studies in cognitive psychology and neuroscience. *Behavior Research Methods*, *49*(6), 2093–2112. https://doi.org/10.3758/s13428-016-0844-8

Vuilleumier, P. (2015). Affective and motivational control of vision. *Current Opinion in Neurology*, *28*(1), 29. https://doi.org/10.1097/WCO.0000000000000159

Wang, F., Karipidis, I., Pleisch, G., Fraga-González, G., & Brem, S. (2020). Development of print-speech integration in the brain of beginning readers with varying reading skills. *Frontiers in Human Neuroscience*, *14*, 289. https://doi.org/10.3389/fnhum.2020.00289

Wang, H., Nation, K., Gaskell, M. G., Robidoux, S., Weighall, A., & Castles, A. (2022). Nap effects on preschool children's learning of letter-sound mappings. *Child Development*, cdev.13753. https://doi.org/10.1111/cdev.13753

Welsh, M. C., Friedman, S. L., & Spieker, S. J. (2006). Executive functions in developing children: Current conceptualizations and questions for the future. In *Blackwell Handbook of Early Childhood Development* (pp. 167–187). Malden, MA: John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470757703.ch9

Wiley, R. W., & Rapp, B. (2021). The effects of handwriting experience on literacy learning. *Psychological Science*, *32*(7), 1086–1103. https://doi.org/10.1177/0956797621993111

Willcutt, E. G., Betjemann, R. S., McGrath, L. M., Chhabildas, N. A., Olson, R. K., DeFries, J. C., & Pennington, B. F. (2010). Etiology and neuropsychology of comorbidity between RD and ADHD: The case for multiple-deficit models. *Cortex*, *46*(10), 1345–1361. https://doi.org/10.1016/j.cortex.2010.06.009

Willcutt, E. G., & Pennington, B. F. (2000). Psychiatric comorbidity in children and adolescents with reading disability. *Journal of Child Psychology and Psychiatry*, *41*(8), 1039–1048. https://doi.org/10.1111/1469-7610.00691

Wilmot, A., Hasking, P., Leitão, S., Hill, E., & Boyes, M. (2023). Understanding mental health in developmental dyslexia: A scoping review. *International Journal of Environmental Research and Public Health*, *20*(2), 1653. https://doi.org/10.3390/ijerph20021653

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. https://doi.org/10.1017/S0142716400009620

Wimmer, H., & Mayringer, H. (2002). Dysfluent reading in the absence of spelling difficulties: A specific disability in regular orthographies. *Journal of Educational Psychology*, *94*(2), 272–277. https://doi.org/10.1037/0022-0663.94.2.272

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*(2), 249–265. https://doi.org/10.1037/a0031311

Wouters, P., & van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education*, *60*(1), 412–425. https://doi.org/10.1016/j.compedu.2012.07.018

Wright, C., Conlon, E. G., & Wright, M. (2011). Voice over the internet protocol as a medium for delivering reading intervention: Evidence from a single case. *SAGE Open*, *1*(3), 215824401142815. https://doi.org/10.1177/2158244011428159

Xia, Z., Yang, T., Cui, X., Hoeft, F., Liu, H., Zhang, X., Liu, X., & Shu, H. (2022). Atypical relationships between neurofunctional features of print-sound integration and reading abilities in Chinese children with dyslexia. *Frontiers in Psychology*, *12*, 748644. https://doi.org/10.3389/fpsyg.2021.748644

Xiao, P., Zhu, K., Liu, Q., Xie, X., Jiang, Q., Feng, Y., Wu, X., Tang, J., & Song, R. (2022). Association between developmental dyslexia and anxiety/depressive symptoms among children in China: The chain mediating of time spent on homework and stress. *Journal of Affective Disorders*, *297*, 495–501. https://doi.org/10.1016/j.jad.2021.10.120

Xu, W., Kolozsvari, O. B., Monto, S. P., & Hämäläinen, J. A. (2018). Brain responses to letters and speech sounds and their correlations with cognitive skills related to reading in children. *Frontiers in Human Neuroscience*, *12*. https://doi.org/10.3389/fnhum.2018.00304

Xue, G., Chen, C., Jin, Z., & Dong, Q. (2006). Language experience shapes fusiform activation when processing a logographic artificial language: An fMRI training study. *NeuroImage*, *31*(3), 1315–1326. https://doi.org/10.1016/j.neuroimage.2005.11.055

Yap, R., & van der Leij, A. (1993). Word processing in dyslexics: An automatic decoding deficit? *Reading and Writing*, *5*(3), 261–279. https://doi.org/10.1007/BF01027391

Yildirim, O., & Surer, E. (2021). Developing adaptive serious games for children with specific learning difficulties: A two-phase usability and technology acceptance study. *JMIR Serious Games*, *9*(2), e25997. https://doi.org/10.2196/25997

Yoncheva, Y. N., Blau, V. C., Maurer, U., & McCandliss, B. D. (2010). Attentional focus during learning impacts N170 ERP responses to an artificial script. *Developmental Neuropsychology*, *35*(4), 423–445. https://doi.org/10.1080/87565641.2010.480918

Žarić, G., Fraga González, G., Tijms, J., van der Molen, M. W., Blomert, L., & Bonte, M. (2014). Reduced neural integration of letters and speech sounds in dyslexic children scales with individual differences in reading fluency. *PLoS ONE*, *9*(10), e110337. https://doi.org/10.1371/journal.pone.0110337

Žarić, G., Fraga González, G., Tijms, J., van der Molen, M. W., Blomert, L., & Bonte, M. (2015). Crossmodal deficit in dyslexic children: Practice affects the neural timing of letter-speech sound integration. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00369

Zeguers, M. H. T., Snellings, P., Huizenga, H. M., & van der Molen, M. W. (2014). Time course analyses of orthographic and phonological priming effects during word recognition in a transparent orthography. *The Quarterly Journal of Experimental Psychology*, *67*(10), 1925–1943. https://doi.org/10.1080/17470218.2013.879192

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, *11*(1), 37–63. https://doi.org/10.1016/S0885-2014(96)90027-1

Zemlak, J. L., Wilson, P., VanGraafeiland, B., & Rodney, T. (2021). Telehealth and the Psychiatric Mental Health Nurse Practitioner: Beyond the COVID-19 Pandemic. *Journal of the American Psychiatric Nurses Association*, 10783903211045120. https://doi.org/10.1177/10783903211045119

Ziegler, J. C., Perry, C., & Zorzi, M. (2020). Learning to read and dyslexia: from theory to intervention through personalized computational models. *Current Directions in Psychological Science*, *29*(3), 293–300. https://doi.org/10.1177/0963721420915873

List of publications

Part of this dissertation

Verwimp, C., Lemsom, J., Bruinsma, S., Snellings, P., Wiers, R. W., & Tijms, J. (To be submitted). Reciprocal associations between letter-speech sound integration and word reading fluency during dyslexia intervention.

Author contributions: CV, SB, PS and JT: conceptualization. SB: data curation. CV and SB: data analysis. CV: writing—original draft. JL, SB, PS, RW, and JT: writing—review and editing. PS, RW, and JT: supervision.

Verwimp, C., Snellings, P., Wiers, R. W., & Tijms, J. (2023). Goal-directedness enhances letter-speech sound learning and consolidation in an unknown orthography. *Child Development.* https://doi.org/10.1111/cdev.13901

Author contributions: CV and JT: conceptualization. CV: data collection, data analysis, data curation, writing—original draft. PS, RW, and JT: writing—review and editing, supervision.

Verwimp, C., Snellings, P., Wiers, R. W., Tijms, J., & Bonte, M. (To be submitted). Behavioral and neural correlates of cognitive control during artificial letter-speech sound learning in typical and dyslexic readers.

Author contributions: CV and JT: conceptualization. CV: data collection, data analysis, data curation, writing—original draft. MB: data preprocessing and analysis—review and supervision. PS, RW, JT, MB: writing—review and editing, supervision.

Verwimp, C., Snellings, P., Wiers, R. W., & Tijms, J. (2023). A randomised proof-of-concept trial on the effectiveness of a game-based training of phoneme-grapheme correspondences in pre-readers. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12821

Author contributions: CV and JT: conceptualization. CV: data collection, data analysis, data curation, writing—original draft. PS, RW, JT: writing—review and editing, supervision.

Verwimp, C., Tijms, J., Snellings, P., Haslbeck, J. M. B., & Wiers, R. W. (2021). A network approach to dyslexia: Mapping the reading network. *Development and Psychopathology*, 1–15. https://doi.org/10.1017/S0954579421000365

Author contributions: CV and JT: conceptualization. CV: analysis, data curation, writing—original draft. JH: analysis—review and supervision. JT, PS, JH, RW: writing—review and editing. JT, PS, RW: supervision.

Verwimp, C., Vaessen, A., Snellings, P., Wiers, R. W., & Tijms, J. (To be submitted). The COVID generation: Online dyslexia treatment equally effective as face-to-face treatment in a Dutch sample.

Author contributions: CV, AV and JT: conceptualization. AV: data curation, CV: analysis, writing—original draft. AV, PS, RW, JT: writing—review and editing, supervision.

Other publications

Amora, K. K., Tretow, A., **Verwimp, C.** (shared first authorship), Tijms, J., Leppänen, P. H. T., & Csépe, V. (2022). Typical and atypical development of visual expertise for print as indexed by the visual word N1 (N170w): A systematic review. *Frontiers in Neuroscience*, *16*, 898800. https://doi.org/10.3389/fnins.2022.898800

Khanolainen, D. **Verwimp, C.**, Tolvanen, A., Salminen, J., & Torppa, M. (Under review). Developing and validating an abbreviated adult reading history questionnaire in the Finnish and Dutch contexts

Verwimp, C., Bempt, F. V., Kellens, S., Economou, M., Vandermosten, M., Wouters, J., ... & Vanderauwera, J. (2020). Pre-literacy heterogeneity in Dutch-speaking kindergartners: Latent profile analysis. *Annals of Dyslexia*, *70*(3), 275-294. https://doi.org/10.1007/s11881-020-00207-9

Verwimp, C., Meiboom, N., Tijms, J. (Submitted). Treatment of reading and spelling problems in children with co-occurring attention-deficit– hyperactivity disorder and dyslexia.

Summary

In a world where literacy is more important than ever, being unable to read comes with many challenges. Although reading might seem straightforward, a significant percentage of the population struggles persistently to develop fluent reading skills. Individuals with dyslexia are at risk for adverse long-term academic, economic, and psychosocial consequences. Therefore, it is highly relevant to examine the underlying causes of dyslexia and how these can be matched to specialized interventions. Although researchers have tried to assign the variety of symptoms associated with dyslexia to one single underlying cause for decades, the focus has now shifted toward probabilistic, dimensional models in which a disorder results from complex interactions between symptoms. However, most studies still examine factors in isolation and have stringent inclusion criteria, overestimating the homogeneity within individuals who experience reading difficulties and leaving it unclear how factors within and across diagnostic boundaries interact. The overarching aims of this dissertation were to provide more insight into the multidimensional etiology of developmental dyslexia, more insight into neurocognitive mechanisms that contribute to the development of integrated and automatized letter-speech sound correspondences needed for fluent reading and how these can be translated into clinical practice, and more insight into how co-occurring attentional difficulties alter the reading development.

In *Chapter 2*, we mapped the complexity of links between cognitive, environmental, and demographic factors related to reading using network analysis. Our network revealed two interconnected subparts: one that included reading fluency and accuracy measures, and another one that consisted of intelligence-related measures. Interestingly, phoneme awareness was linked to the precise and accurate processing of letter-speech sound (L-SS) mappings, whereas rapid automatized naming was more closely associated with the automatic integration of visual and speech information. Our findings suggest that various factors contribute to typical and atypical reading development, each influencing different aspects of the reading process. Consequently, our research challenges the commonly held belief that a single underlying deficit causes dyslexia. By utilizing the network approach to psychopathology, we have demonstrated how we can examine complex interactions within the reading network, which can inform more individualized interventions in the future.

In *Chapter 3*, we shed light on attentional processes involved in developing and integrating cross-modal letter-speech sound correspondences. One hundred and seven children were instructed to learn eight new letter-speech sound correspondences that consisted of eight unknown characters that were linked to Dutch speech sounds. The learning task was either preceded by instructions that revealed the goal of the task and directed children toward that goal or instructions that told them that the goal would become evident during the task itself. Assessments of artificial word-reading ability and symbol knowledge were conducted immediately following the learning session and on the subsequent day to examine the effect of sleep. Children who received goal-directed instructions learned the new script faster and more efficiently and had better learning outcomes than their implicitly instructed peers. Our results highlight that directing children toward the goal can promote L-SS learning and consolidation, providing insights into how top-down control impacts the initial reading skill acquisition phase.

To get more insight into how top-down control impacts reading acquisition in a clinical sample, the same artificial learning task was used to examine learning differences between children with dyslexia and typical readers, as reported in *Chapter 4*. Importantly, children with co-occurring attentional difficulties were included in both groups. We examined the interaction between behavioral and neural cognitive control processes, i.e., goal-directed behavior, post-error slowing, and neural performance monitoring, as well as how behavioral ADHD measures were related to learning outcomes. Our results revealed substantial inter-individual differences in learning outcomes, but we did not find any consistent significant group differences between instruction conditions (goal-directed vs. implicit instruction) or reading groups (children with dyslexia vs. typical readers) on a behavioral or neural level. However, our results suggest that, on average, children with dyslexia followed a slightly different learning trajectory compared to their typical reading peers, which was also influenced by instruction manipulation. This suggests that while some individuals have deficits in letter-speech sound mapping and/or feedback processing, we did not find evidence to assume this was true in all dyslexic learners. This aligns with the multiple deficit model that proposed that no single deficit is sufficient to cause dyslexia, but multiple interacting factors might increase the risk. Interestingly, our behavioral attentional measures were associated with learning outcomes of the artificial learning task and reading performance in the native language, suggesting that attentional difficulties may impact both reading performance and acquisition of alphabetic knowledge. Associations between feedback-related negativity amplitudes and behavioral measures did not survive correction for multiple comparisons. The fact that learning outcomes were related to reading ability within the native language validates using such an artificial learning paradigm. This paradigm allows us to further examine factors that contribute to the transition from letter-speech sound integration to consolidated audiovisual objects and how these relate to individual learning indices.

The second section of this thesis discussed mechanisms involved in reading intervention, shedding light on the significant reported inter-individual differences in intervention response with the ultimate aim to ameliorate future interventions for struggling readers. In *Chapter 5*, we created a detailed window concerning the reciprocal associations between the development of letter-speech sound integration and word reading fluency within a Dutch dyslexia treatment program. A path analysis revealed that the L-SS training program significantly improved L-SS integration and word reading fluency in children with dyslexia, with more automatized L-SS associations leading to more fluent word reading. Moreover, especially children with weaker L-SS integration at the beginning of the intervention benefitted more during the intervention, suggesting that children whose reading difficulties were not associated with poor L-SS might need a slightly different approach to target reading problems. The intervention's second half was crucial in achieving the full potential for word reading fluency growth. During the intervention, processing speed only had a small negative relationship with word reading fluency, and age and IQ did not contribute significantly to the model. This study provide valuable insights into the mechanisms of change within a specialized dyslexia intervention, so we can more optimally adapt interventions to individual needs.

In *Chapter 6*, we examined whether online specialized reading treatment is as effective as the more conventional face-to-face treatment in improving children's reading and spelling skills, as interventions massively moved to online platforms because of the global COVID-19 pandemic. Using Bayesian methods, 254 children who received treatment as usual were compared to 162 children who received online treatment. The results showed that children in the online treatment group had slightly fewer treatment sessions but made similar progress in their reading and spelling abilities, after controlling for the number of sessions, when compared to those who received the face-to-face treatment. Interestingly, the conditions did not differ in the percentage of homework sessions they completed, suggesting that moving interventions online does not affect treatment adherence in this context. This study provided support to the small body of literature that supports the equivalence of online provided dyslexia treatments compared to traditional face-to-face methods in a Dutch clinical sample. Future studies should examine whether some children benefit more from online treatment than others and whether our results can be generalized to other treatment methods and languages.

Last, in line with the digitalization of treatment programs, *Chapter 7* discusses the results of a short, intensive, randomized controlled trial in which we examined the effectiveness of the game-based intervention tool 'KlankKr8'. The standard version of the game was developed to boost letter-speech sound (L-SS) correspondences in a motivational game environment. In an extended version of the game, we added a motoric handwriting component, given the importance of handwriting in audiovisual integration. Math Garden was used as a control condition that did not involve any reading exercises. We found that pre-readers who played the standard version of the game improved in L-SS knowledge after three weeks of intervention, but this knowledge did not transfer to better word reading ability. Children who played the motoric version of the game did not perform significantly better than children in the control condition but also not significantly worse than children in the standard condition. These findings suggest that this game could be a useful preventive intervention for at-risk children in kindergarten to improve L-SS knowledge before learning to read, but further research is necessary to explore whether more extended intervention periods lead to a transfer from improved letter-speech sound knowledge to improved reading skills and to an effect of handwriting practice on the integration of letters and speech sounds.

Samenvatting

In een wereld waar geletterdheid belangrijker is dan ooit, brengt het niet kunnen lezen veel uitdagingen met zich mee. Hoewel lezen eenvoudig lijkt, worstelt een aanzienlijk percentage van de bevolking met het ontwikkelen van vloeiende leesvaardigheden. Personen met dyslexie lopen het risico op nadelige academische, economische en psychosociale gevolgen op lange termijn. Daarom is het relevant om de onderliggende oorzaken van dyslexie te onderzoeken en hoe deze kunnen worden gekoppeld aan gepersonaliseerde interventies. Hoewel onderzoekers al tientallen jaren proberen de verscheidenheid aan symptomen gerelateerd aan dyslexie aan één enkele onderliggende oorzaak toe te wijzen, is de focus nu verschoven naar probabilistische, dimensionale modellen waarin een stoornis het gevolg is van complexe interacties tussen symptomen. De meeste studies onderzoeken echter afzonderlijke factoren en hanteren strenge inclusiecriteria, waardoor de homogeniteit binnen individuen met dyslexie sterk wordt overschat en het onduidelijk blijft hoe factoren binnen en tussen verschillende diagnoses met elkaar interageren. De doelen van dit proefschrift waren om meer inzicht te verschaffen in de multidimensionale etiologie van dyslexie, meer inzicht te krijgen in neurocognitieve mechanismen die bijdragen aan de ontwikkeling van geautomatiseerde letter-klankkoppelingen die nodig zijn om vloeiend te kunnen lezen en hoe deze kunnen worden vertaald naar klinische studies, en meer inzicht in hoe comorbide aandachtsproblemen de leesontwikkeling beïnvloeden.

In *Hoofdstuk 2* brachten we de complexe associaties tussen cognitieve, omgevings- en demografische factoren die gerelateerd zijn aan lezen in kaart met behulp van een netwerkanalyse. We vonden twee onderling verbonden subonderdelen in het netwerk: een die bestond uit maten van leesvloeiendheid en -accuraatheid, en een andere die bestond uit intelligentie-gerelateerde maten. Foneembewustzijn was geassocieerd met de nauwkeurige verwerking van letter-klankkoppelingen, terwijl benoemsnelheid geassocieerd was met de automatische integratie van visuele en auditieve informatie. Onze bevindingen suggereren dat verschillende factoren bijdragen aan de typische en atypische leesontwikkeling, die elk verschillende aspecten van het leesproces beïnvloeden. Deze bevindingen trekken de algemeen aanvaarde ‘core deficit’ overtuiging in twijfel, wat impliceert dat dyslexie veroorzaakt wordt door één enkel onderliggend tekort. Door gebruik te maken van de netwerkbenadering hebben we aangetoond hoe we complexe interacties binnen het leesnetwerk kunnen onderzoeken, wat in de toekomst kan leiden tot meer geïndividualiseerde interventies.

In *Hoofdstuk 3* onderzochten we aandachtsprocessen die betrokken zijn bij het ontwikkelen en integreren van crossmodale letter-klankkoppelingen. Honderdenzeven kinderen kregen de opdracht om acht nieuwe letter-klankkoppelingen te leren die bestonden uit acht onbekende symbolen die gekoppeld waren aan Nederlandse spraakklanken. Kinderen kregen ofwel instructies die het doel van de taak onthulden en de kinderen aanspoorden dat doel te bereiken, ofwel instructies die hen vertelden dat het doel tijdens de taak zelf duidelijk zou worden. Leesvermogen in het artificiële schrift en symboolkennis werden gemeten meteen na de leertaak, alsook op de volgende dag om het effect van slaap te onderzoeken. Kinderen die doelgerichte instructies kregen leerden het nieuwe schrift sneller en efficiënter en hadden betere leeruitkomsten dan hun impliciet geïnstrueerde leeftijdsgenoten. Onze resultaten tonen aan dat het onthullen van het leerdoel en kinderen aansporen dit doel te bereiken het leren en consolideren van letter-klankkoppelingen kan bevorderen, waardoor inzicht wordt verkregen in de manier waarop *top-down* processen de initiële fase van de leesontwikkeling beïnvloeden.

Om meer inzicht te krijgen in hoe *top-down* controle de leesontwikkeling in een klinische steekproef beïnvloedt werd dezelfde leertaak gebruikt om leerverschillen tussen kinderen met dyslexie en typische lezers te onderzoeken, zoals beschreven in *Hoofdstuk 4*. Kinderen met comorbide aandachtsproblemen werden geïncludeerd in beide groepen. We onderzochten de interactie tussen gedragsmatige en neurale cognitieve controleprocessen: *goal-directed behavior*, *post-error slowing* en neurale *performance monitoring*, evenals hoe gedragsmatige ADHD maten verband hielden met leeruitkomsten. We vonden substantiële interindividuele verschillen in leeruitkomsten, maar geen consistente significante groepsverschillen tussen instructiecondities (doelgericht vs. impliciete instructie) of kinderen met dyslexie vs. typische lezers op gedrags- of neuraal niveau. Onze resultaten suggereren echter dat kinderen met dyslexie gemiddeld gezien een iets ander leertraject volgden dan hun leesgenoten zonder dyslexie, dat ook werd beïnvloed door de instructies die ze kregen. Dit suggereert dat hoewel sommige individuen tekorten hebben in het associëren van letters en klanken en/of het verwerken van feedback, we niet kunnen aannemen dat dit bij alle kinderen met dyslexie het geval is. Dit komt overeen met het *multiple deficit model*, dat inhoudt dat geen enkel tekort voldoende is om dyslexie te veroorzaken, maar dat meerdere op elkaar inwerkende factoren het risico kunnen vergroten. Verder vonden we dat onze gedragsmatige aandachtsmaten geassocieerd waren met de leeruitkomsten van de artificiële leertaak en leesprestaties in de moedertaal, wat suggereert dat aandachtsproblemen zowel de leesprestaties als de verwerving van alfabetische kennis kunnen beïnvloeden. Associaties tussen *feedback-related negativity* amplitudes en gedragsmaten overleefden de correctie voor *multiple comparisons* niet. Het feit dat kinderen hun leerruitkomsten geassocieerd waren met leesvaardigheid in hun moedertaal valideert het gebruik van dergelijke leerparadigma om factoren die bijdragen aan de overgang van letter-klankintegratie naar geconsolideerde audiovisuele objecten verder te onderzoeken.

In het tweede deel van dit proefschrift worden leesinterventie mechanismen besproken, waarbij we trachtten licht te werpen op de significante gerapporteerde interindividuele verschillen in interventierespons met als uiteindelijke doel toekomstige interventies voor zwakke lezers te verbeteren. In *Hoofdstuk 5* onderzochten we de wederkerige associaties tussen de ontwikkeling van letter-spraakintegratie en vloeiend woordlezen binnen een Nederlands behandelprogramma voor dyslexie. Uit een padanalyse bleek dat het trainingsprogramma de letter-klankintegratie en het vloeiend woordlezen bij kinderen met dyslexie aanzienlijk verbeterde, waarbij meer geautomatiseerde letter-klankassociaties leidden tot vloeiender woordlezen. Bovendien profiteerden vooral kinderen met een zwakkere letter-klankintegratie aan het begin van de interventie meer tijdens de interventie. Dit suggereert dat kinderen van wie de leesproblemen niet geassocieerd waren met slechte letter-klankintegratie mogelijk een iets andere benadering nodig hebben om leesproblemen aan te pakken. De tweede helft van de interventie was cruciaal voor het bereiken van het volledige potentieel voor de groei van de woordleesvaardigheid. Tijdens de interventie had de verwerkingssnelheid slechts een kleine negatieve relatie met het vloeiend lezen van woorden, en leeftijd en IQ droegen niet significant bij aan het model. Deze studie biedt waardevolle inzichten in de mechanismen die bijdragen aan verbetering gedurende een gespecialiseerde dyslexie-interventie, zodat we uiteindelijk interventies beter kunnen aanpassen aan individuele behoeften.

In *Hoofdstuk 6* onderzochten we of online leesbehandeling even effectief is als de meer conventionele *face-to-face* methode bij het verbeteren van de lees- en spellingvaardigheden, aangezien interventies vanwege de wereldwijde COVID-19 pandemie massaal naar online platforms zijn verschoven. Met behulp van Bayesiaanse methoden werden 254 kinderen die de gebruikelijke behandeling kregen vergeleken met 162 kinderen die online behandeling kregen. De resultaten toonden aan dat kinderen in de online behandelgroep iets minder behandelsessies kregen, maar vergelijkbare vooruitgang boekten in hun lees- en spellingsvaardigheden (na controle voor het aantal sessies) in vergelijking met kinderen die de *face-to-face* behandeling kregen. Interessant genoeg verschilden de twee condities niet in het percentage huiswerksessies die ze voltooiden. Dit suggereert dat het verplaatsen van interventies naar online platforms de therapietrouw in deze context niet beïnvloedt. Deze studie bood ondersteuning aan de kleine hoeveelheid literatuur die de equivalentie van online aangeboden dyslexiebehandelingen ondersteunt in vergelijking met traditionele *face-to-face* methoden in een Nederlandse klinische steekproef. Toekomstige studies moeten onderzoeken of sommige kinderen meer baat hebben bij online behandeling dan anderen en of onze resultaten kunnen worden gegeneraliseerd naar andere behandelmethoden en talen.

Ten slotte bespreekt *Hoofdstuk 7*, in lijn met de digitalisering van behandelprogramma's, de resultaten van een korte, intensieve, *randomized controlled trial* waarin we de effectiviteit van het spel 'KlankKr8' onderzochten. De standaardversie van het spel werd ontwikkeld om de correspondenties tussen letters en spraakgeluid (L-S) te versterken in een motiverende spelomgeving. In een aangepaste versie van het spel hebben we een motorische handschriftcomponent toegevoegd, gegeven het belang van handschrift bij audiovisuele integratie. Rekentuin werd gebruikt als controleconditie waarbij geen leesgerelateerde vaardigheden geoefend werden. Onze resultaten toonden dat kleuters die de standaardversie van het spel speelden na drie weken hun L-S kennis verbeterden, maar dat deze kennis niet vertaald werd naar betere woordleesvaardigheden. Kinderen die de motorische versie van het spel speelden presteerden niet significant beter dan kinderen in de controleconditie, maar ook niet significant slechter dan kinderen in de standaardconditie. Deze bevindingen suggereren dat dit spel een nuttige preventieve interventie zou kunnen zijn voor kleuters die risico hebben op dyslexie om de L-S kennis een boost te geven voordat ze leren lezen, maar verder onderzoek is nodig om te onderzoeken of langere interventieperiodes leiden tot een vertaling van verbeterde L-S kennis naar betere leesprestaties en tot een effect van handschrifttraining op de integratie van letters en spraakklanken.

Dankwoord

Het is zover, ik heb de finishlijn gehaald. Wat voor lange tijd ver en onbereikbaar leek, is eindelijk aangebroken. Het is een ongelofelijke reis geweest, met veel hoogtes maar ook veel laagtes. Wat begon als een project waarin we elke 6 maanden zouden reizen om samen te komen met ons Europees consortium veranderde al snel in een 2 jaar durende lockdown. Weinig contact met collega's, uitgestelde dataverzameling, en vele plannen A die veranderden in plannen B en C. Wanneer ik dacht dat het niet erger kon werd onze testcomputer gestolen (uiteraard zonder recente back-up) en besloten ze het lab te renoveren. Ik kan je vertellen, geen goed idee om kinderen met dyslexie en ADHD te testen terwijl de kamer ernaast afgebroken wordt. Maar ik heb het gehaald. Ongeveer 4 jaar later ligt hier mijn proefschrift op tafel, en dit heb ik te danken aan een aantal mensen die me tijdens deze reis vergezeld hebben. Daarom neem ik graag de tijd om wat mensen te bedanken die elk op hun eigen manier hebben bijgedragen.

Allereerst wil ik mijn promotorenteam bedanken. Reinout, Patrick, en Jurgen, wat hebben jullie mij van in het begin al met open armen ontvangen. Onze eerste kennismaking gebeurde – al in het pre-coronatijdperk – online. Wanneer vrienden vroegen "En, hoe ging dat gesprek dan?" antwoordde ik steevast met "Goh ja, het voelde wel goed, maar dat heb je met alle Nederlanders zeker?" Enkele maanden later volgde het verdict: ik mocht bij jullie beginnen! Tijdens de eerste maanden probeerden we elkaar regelmatig te spreken. Waar we het in het begin vooral over praktische zaken hadden evolueerde al snel tot updates over lopende projecten en gepubliceerde artikelen. Reinout, wat was het fijn om jou er bij te hebben! Jouw academische wijsheid is buitengewoon. Dankzij jou konden we ons onderzoek steeds vanuit een breder perspectief bekijken. Waar als leesonderzoeker sommige keuzes vanzelfsprekend leken kwam jij steeds met kritische vragen die ons onderzoek naar een beter niveau tilde. Patrick, ook jou ben ik ontzettend dankbaar voor jouw kritische blik en goede adviezen. Jouw kennis over dyslexie en meertaligheid is enorm. En jou kon ik tenminste wel blij maken met een 'tas koffie', waar ik bij andere collega's vragend aangekeken werd. Jurgen, jou wil ik in het bijzonder bedanken. Jij nam me gedurende 4 jaar onder de arm en leerde me alle kneepjes van het vak. Jouw deur kon niet harder openstaan aangezien we het merendeel van de tijd een kamer deelden. Dat ik van tijd tot tijd de notitieblaadjes weer een keer terug moest duwen naar jouw kant van de tafel nam ik er graag bij ☺ Bedankt om steeds te luisteren naar onzekerheden en frustraties en waar nodig een helpende hand te bieden. Bedankt voor de vele constructieve feedback en inspirerende woorden. Dankjewel om steeds je rustige zelf te zijn wanneer mijn hoofd soms explodeerde. Merci! Wie hier ook niet mag ontbreken is Milene Bonte van de Universiteit Maastricht. Milene, jouw hulp bij de EEG studie was onmisbaar. Dankjewel om me steeds met open armen in Maastricht of Heemstede te willen ontvangen na mijn paniekerige mailtjes over de data. Jouw enthousiasme voor onderzoek werkt aanstekelijk.

Daarnaast wil ik graag mijn promotiecommissie bedanken die de tijd heeft genomen om mijn werk te lezen en te beoordelen. Thank you Prof. dr. Silvia Brem, Dr. Brenda Jansen, Prof. dr. Richard Ridderinkhof, Prof. dr. Judith Rispens, and Prof. dr. Minna Torppa for accepting the invitation to be members of my PhD examination committee. Thank you for taking time to evaluate my dissertation and providing me with thoughtful and challenging questions during the defense.

I was lucky to be part of the Neo-PRISM-C consortium. Professor Papadopoulos, Timos, thank you for setting up this project and convincing all partners with your enthusiasm to collaborate with you. Without you, none of this would have been possible. Thank you to all PIs and ESRs who were part of the consortium for sharing excellent training sessions, fruitful discussions, many coffee breaks, and the most delicious food at the most wonderful places. Although we saw each other less than planned due to the pandemic, meeting so many brilliant minds was an absolute honor. But I have to admit, Finnish karaoke would have been the cherry on the cake.

A special thank you goes to Ariane and Kathleen. Countless hours of writing a review together coexisted with daily rowing sessions, sauna moments with lake dips, and wine evenings. Ariane, waiting for each other to go to the hotel lobby as hotel room neighbors when we just met turned into sharing Air BnB’s on work trips, drinking Glühwein or Klöki together on Christmas markets, and even enjoying wedding parties. No other person is so much into research and cycling as you are and can talk about it with so much enthusiasm. Thank you for all the nice moments we had together in different countries, I hope visiting each other in Tampere or Antwerp becomes a tradition. Kathleen, thank you for always being so open and joyful. Every moment with you is fun. I will always be your ‘Belgian twin’.

Niemand kan natuurlijk onderzoek doen zonder medewerking van participanten. In de afgelopen 3 jaren testten we maar liefst 176 kleuters, 237 schoolkinderen, en 42 volwassenen (did someone say pandemic?). Duizend maal dank aan alle participanten maar ook ouders, leerkrachten en scholen die keer op keer engagement toonden om deel te nemen aan onze studies. Dat ik in staat was om zoveel data te verzamelen getuigt natuurlijk van heel veel helpende handen. Eline, Sjoerdje, Lea, Jesse, Hannah, Esmee, Isabel, Luuk, Kaylee, Lizette en Julia, dankjewel voor het brengen van leven in de brouwerij op de uni en jullie onschatbare bijdrage aan mijn proefschrift. Jesse, jij in het bijzonder. Bedankt voor de gezellige overwerkavonden die gevuld waren met random google zoekacties en soms eindigden in spontane etentjes.

Prof. dr. Pol Ghesquière, tijdens mijn bachelor psychologie in Leuven zorgde jij er al voor dat mijn interesse in leerstoornissen aangewakkerd werd. Ik zou bijna durven zeggen dat ik hier vandaag niet zou staan moest ik toen niet zo'n enthousiaste professor gehad hebben. Lieve Jolijn, ook jij speelt zonder het te beseffen een grote rol in mijn academische carrière. Tijdens eindeloze weekend MRI sessies vroeg jij of ik er ooit al aan gedacht had om een PhD te doen. Sinds dat moment gingen we samen actief op zoek naar interessante posities, waaronder deze PhD plek. Jij moedigde me aan om te solliciteren, ook al voldeed ik misschien niet volledig aan alle vereisten. Door in mij te geloven geloofde ik ook in mezelf. Merci voor jouw eindeloze enthousiasme en vertrouwen, ik hoop dat we in de toekomst nog vaak mogen samenwerken.

Er zijn ook enkele personen die al langer meegaan dan de duur van dit project. Lieve familie, mama, papa, Johnny, Paula, al van kleins af aan steunen jullie mij in alles wat ik doe. Dankjewel om steeds in mij te geloven, al snappen jullie niet altijd goed waar mijn ideeën altijd vandaan komen. Studeren in Leuven, verhuizen naar Amsterdam, een marathon lopen in Hamburg, altijd verzin ik wel iets nieuws. Al was het met drie van de vier (plus)ouders als leerkracht niet altijd even makkelijk uit te leggen waarom je in hemelsnaam zou moeten onderzoeken of kinderen 'luisteren als je ze niks uitlegt'. Dankjewel om me steeds uit te dagen te verantwoorden waarom mijn onderzoek wél nuttig was. Dimi, bedankt om mijn zelf-geprogrammeerde computertaakjes uit te testen en meteen te hulp te schieten bij Word lay-out frustraties.

Lieve schoonfamilie, jullie ook bedankt voor de interesse in mijn doctoraat. Lieve, er zijn weinig mensen die zoveel gevatte vragen hebben gesteld over mijn onderzoek als jij. Jouw interesse in andermans leven en jouw 'wil om te weten' stel ik zeer op prijs. Hans en Meike, hoewel ik eerder bescheiden ben over gepubliceerde artikelen zijn jullie steeds degenen die vragen om een kopie (al denk ik niet dat ik jullie er ooit eentje heb gegeven). Dankjewel voor de vele gezellige momenten in Hamburg of SPO. Ik ben blij dat ik deel mag zijn van Vanhooren XL.

Liefste vriendinnetjes van het middelbaar, liefste Tafelliers. Met jullie ben ik samen volwassenen geworden, gegroeid. Toen ik vier jaar geleden aankondigde dat ik een PhD ging doen kreeg ik de reactie 'Jij? Een PhD? Dat had ik nu nooit gedacht!'. Maar kijk, hier sta ik nu, met de meesten van jullie in de zaal als supporter. Wat we als snotneuzen nooit hadden kunnen bedenken zit er nu bij de meeste onder ons allemaal aan te komen: werken, verhuizen, trouwen en kinderen. Ik had me geen betere vrienden kunnen bedenken om al deze levensfases mee te delen. Dankjewel om altijd te vragen hoe het gaat, de wijntjes na frustrerende werkdagen, en ‘de eerste zondag van de maand’ die weliswaar bijna nooit op de eerste zondag van de maand gebeurt. Het zijn deze kleine dingen die mijn leven in balans houden. Lieve Femke en Bjorn, ook jullie kan ik niet genoeg bedanken. Dankjewel voor de leukste weekendjes in Amsterdam, Herentals, en Rekkem. Voor jullie is het nooit te ver, als er maar kroketten of bitterballen te eten zijn. Merci om altijd jullie lieve en gastvrije zelf te zijn.

Last but not least. Lieve Jakob, koekie, liefje, mijn mannetje. Vele koosnaampjes zijn gepasseerd de afgelopen jaren maar jij bleef steeds mijn trouwste supporter langs de zijlijn. Waar we als studenten planden om lekker te gaan samenwonen draaide anders uit toen ik deze PhD positie in Amsterdam kreeg. Uitkijken naar de weekends samen in Amsterdam of Herentals verzachtte de pijn van niet steeds bij elkaar te zijn. Bedankt om steeds te luisteren naar mijn frustraties over hoe een statistische test net niet significant was (hoewel ik niet meer op 1 hand kan tellen hoe vaak ik moest uitleggen wat een p-waarde is), of waarom kinderen niet gewoon kunnen stilzitten tijdens hersenonderzoek. Dankjewel om steeds te geloven in mij en me keer op keer te motiveren om door te gaan. Maar vooral bedankt om steeds jezelf te zijn. Laten we elkaar levenslang inspireren.

The project has come to an end, but there is one thing that will be keeping us connected for ever: ‘You want to learn, you want to concur, you want to go further, MERAKI’ (professor Papadopoulos, 2023).

Cara
Oktober 2023