

## Regular Article

# A network approach to dyslexia: Mapping the reading network

Cara Verwimp<sup>1,2,3</sup> , Jurgen Tijms<sup>1,2,3</sup>, Patrick Snellings<sup>1,2</sup>, Jonas M. B. Haslbeck<sup>4</sup> and Reinout W. Wiers<sup>1</sup>

<sup>1</sup>Department of Developmental Psychology, University of Amsterdam, Amsterdam, The Netherlands; <sup>2</sup>Rudolf Berlin Center, Amsterdam, The Netherlands;

<sup>3</sup>RID, Amsterdam, The Netherlands and <sup>4</sup>Department of Psychological Methods, University of Amsterdam, Amsterdam, The Netherlands

### Abstract

Research on the etiology of dyslexia typically uses an approach based on a single core deficit, 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 related to dyslexia by employing a network analysis using a large clinical database of 1,257 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 (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.

**Keywords:** dyslexia, multiple-deficit model, network analysis, reading disability

(Received 20 November 2020; revised 2 February 2021; accepted 4 April 2021; First Published online 27 July 2021)

Developmental dyslexia (henceforth dyslexia) is a neurodevelopmental disorder characterized by a failure to develop accurate and fluent reading (Ozernov-Palchik, Yu, Wang, & Gaab, 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, feelings of shame) and academic consequences (Hendren, Haft, Black, White, & Hoefft, 2018; Livingston, Siegel, & Ribary, 2018; Mugnaini, Lassi, La Malfa, & Albertini, 2009). Effective reading interventions are thus needed to prevent such long-term consequences.

Current interventions mainly focus on phonics instruction, comprising repeated reading practice and phoneme awareness (PA) training, along with decoding strategies (Fraga González et al., 2015; Galuschka, Ise, Krick, & Schulte-Körne, 2014; Scammacca, Roberts, Vaughn, & Stuebing, 2015). However, the meta-analysis of Galuschka et al. (2014) showed that dyslexia remediation only moderately improved foundational reading skills (mean effect size of .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 responses, with a substantial number of children with dyslexia that remain 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 interventions.

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 inter-individual 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 (Insel et al., 2010; Sanislow, Ferrante, Pacheco, Rudorfer, & Morris, 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

**Author for Correspondence:** Cara Verwimp; E-mail: [c.t.verwimp@uva.nl](mailto:c.t.verwimp@uva.nl)

**Cite this article:** Verwimp C, Tijms J, Snellings P, Haslbeck JMB, Wiers RW (2023). A network approach to dyslexia: Mapping the reading network. *Development and Psychopathology* 35: 1011–1025. <https://doi.org/10.1017/S0954579421000365>

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.



environmental), operate probabilistically by increasing (or decreasing) the risk that one develops a reading disorder (Pennington, 2006; van Bergen, van der Leij, & de Jong, 2014).

In the current study, we aimed to provide a more detailed window on the multidimensional character of dyslexia, drawing 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 PA and rapid automatized naming (RAN) are the two most frequently reported problems in dyslexia (Caravolas *et al.*, 2012; Georgiou, Torppa, Manolitsis, Lyytinen, & Parrila, 2012). PA refers to the ability to segment and manipulate 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; it is found to be especially associated with reading fluency (Araújo, Reis, Petersson, & Fátima, 2015). As learning to read involves matching visual symbols (*i.e.*, letters) to corresponding speech sounds, the formation of letter–speech sound (L–SS) mappings has received growing interest in recent years (*e.g.*, Blomert, 2011; Žarić *et al.*, 2014). It has been shown that a deficit in developing automatized L–SS associations results in dysfluent reading (Aravena, Snellings, Tijms, & van der Molen, 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, Zheng, & Jerman, 2009). Children with dyslexia have been 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, Lyytinen, Erskine, Eklund, & Lyytinen, 2010). Slow vocabulary development and poor productive language skills appear to be predictive for later reading disabilities (Snowling, Gallagher, & Frith, 2003). Others suggest that the origin of dyslexia might be related to a deficit in visual processing. More specifically, dyslexics have 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, Spinath, Dale, & Plomin, 2005; Torppa *et al.*, 2010), indicating the importance of genetic risk factors. Several environmental factors, such as socioeconomic status (SES), have also been reported to contribute to unique variance in the development of reading (*e.g.*, Becker *et al.*, 2017; Kovachy, Adams, Tamaresis, & Feldman, 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, Hamilton, Maughan, & Snowling, 2017; Hamilton, Hayiou-Thomas, Hulme, & Snowling, 2016). In addition, prematurity has been associated with developing deficits in reading abilities at a later age (Kovachy *et al.*, 2015).

The recent literature thus suggests a variety of factors that contribute to the risk of developing dyslexia. Although the literature on dyslexia is extensive, it remains unclear how variations in combinations of risk and protective factors contribute to reading

development and how this combination influences intervention outcomes, since evidence mostly originates from parsimonious single-deficit studies that examine one component (Astle & Fletcher-Watson, 2020). In those studies, by employing an arbitrary threshold value, children with diverse symptoms are routinely classified as either dyslexic or typically developing readers. 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. In 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, Nikolaev, Srihari, & Blair, 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, thus 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 that represent conditional dependencies between pairs of nodes after controlling for all other nodes in the network (Epskamp, Borsboom, & Fried, 2018). 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 the analysis of individual factors concurrently, which may reveal patterns that are currently neglected. The most important symptoms in the network and their interrelations can be identified, along with 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, Perry, & Zorzi, 2020).

The present study thus aimed to elucidate the complex relations between risk factors and symptoms in dyslexia. In addition, we aimed 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 (a) how numerous variables (including reading, spelling, several language and memory processes, SES, and general abilities) relate to each other in the

framework of reading disabilities and (b) whether the amount of progress one made during reading intervention influenced other relations in the network. In total, 28 reading-related variables, including cognitive, demographic, and environmental measures, were used to estimate a general network structure in 1,257 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 the 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 were collected at a nationwide, clinical center for learning disabilities in the Netherlands (RID) over the period 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 6 to 33 years, many diverging tests were used based on participant 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 the children were native Dutch speakers that had been referred to the clinical center for diagnostic screening for dyslexia because of severe and persistent reading disabilities at school (i.e., below the tenth percentile on standard reading measures or below the tenth percentile on spelling in combination with a score below the sixteenth percentile on reading) and who resisted extra remedial support at school prior to referral. The data set 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 (<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 the sake of brevity, this is reported in Table S1 of the Supplementary Material. 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, RAN, and L–SS speed) were reversed, such that for all variables higher scores represented better performance. Finally, as not all the 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 is provided in Table 1.

### Cognitive assessments

#### Reading and spelling measures

**Word reading.** Reading was measured with the computerized reading task from the Differential Diagnosis Dyslexia Maastricht Battery (3DM) (Blomert & Vaessen, 2009). This word reading task comprised three different levels – high-frequency words, low-frequency words, and pseudowords. Each level contained 75 words that were displayed on five 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 s 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 the analyses (i.e., for high-frequency, low-frequency, and pseudowords).

**Spelling.** Spelling was assessed with the computerized spelling task from the 3DM (Blomert & Vaessen, 2009). In this task, a word was presented both aurally and visually, with 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 (percentage correct) and fluency (seconds/item) were included in the analyses.

#### Other cognitive measures

**L–SS mapping.** L–SS mapping was measured with the L–SS 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 (percentage correct) and response time (seconds/item) were included in the analyses as L–SS accuracy and L–SS speed, respectively.

**PA.** PA 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 provide the resulting word. The score was computed as the percentage of correct responses.

**RAN.** RAN 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.** Receptive vocabulary and productive vocabulary were respectively 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 (WISC-III) (Kort et al., 2002). In the PPVT, a word was aurally presented and the child was asked to choose the correct image from four alternatives. In the WISC test, the child was asked needed to describe the meanings of words of increasing complexity. Both scores were defined as the

**Table 1.** Overview of included variables

Abbreviation	Meaning in the network	Task	Variable type
ReaFIH	Reading fluency in high-frequency words	3DM (reading task)	Continuous
ReaFIL	Reading fluency in low-frequency words	3DM (reading task)	Continuous
ReaFIP	Reading fluency in pseudowords	3DM (reading task)	Continuous
ReaAccH	Reading accuracy in high-frequency words	3DM (reading task)	Continuous
ReaAccL	Reading accuracy in low-frequency words	3DM (reading task)	Continuous
ReaAccP	Reading accuracy in pseudowords	3DM (reading task)	Continuous
SpFI	Spelling fluency	3DM (spelling task)	Continuous
SpAcc	Spelling accuracy	3DM (spelling task)	Continuous
LSSSpeed	L-SS speed	3DM (L-SS identification task)	Continuous
LSSAcc	L-SS accuracy	3DM (L-SS identification task)	Continuous
PhonAw	PA	3DM (phoneme deletion task)	Continuous
RAN	RAN	3DM (RAN 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 Words Test	Continuous
VerLTMem	Verbal long-term memory	15 Words Test (recall)	Continuous
VisConAb	Visuo-constructional abilities	WISC-III (block design)	Continuous
VisPer	Visual perception	GSO test	
IQ	Intelligence	WISC-III (similarities, information, comprehension, arithmetic, picture completion, picture arrangement, object assembly, substitution)	Continuous
Att	Attention	—	Binary
BirWei	Birth weight	—	Continuous
GestAge	Gestational age	—	Continuous
Sex	Gender (0 = girl, 1 = boy)	—	Binary
FamRisk	Family risk	—	Binary
MuLin	Multilingualism	—	Binary
GradeRet	Grade retention (after third grade)	—	Binary
SES	Income as proxy of SES	—	Continuous
IntProgr	Intervention progress – difference in word reading fluency scores prior to and after intervention	One-minute test	Continuous

3DM = Differential Diagnosis Dyslexia Maastricht Battery; GSO = Groninger School Onderzoek; L-SS = letter–speech sound; PA = phoneme awareness; PPVT = Peabody Picture Vocabulary Task  
 RAN = rapid automatized naming; SES = socioeconomic background; WISC-III = Wechsler Intelligence Scale for Children, third edition

number of correct items, with a maximum score of 240 for the PPVT and a maximum score of 70 for the WISC vocabulary subtest.

*Working memory.* Working memory was measured with the digit span reversed subtest from the WISC-III (Kort *et al.*, 2002). Children were asked to memorize a sequence of digits and then repeat them in 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 asked to recall as many words as possible from an aurally presented list containing 15 high-frequency Dutch nouns, comprising five learning blocks. After a 20-min delay, the children were asked to recall as many words as possible from the list they learned before. Short-term memory was calculated as the number of words recalled during the fifth learning block and long-term memory was calculated as the number of words recalled during “delayed recall”, both with 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 were asked 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, with time bonus points awarded for 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) test (Kema & Kema-van Leggelo, 1987). For each item, four abstract figures were presented – two identical ones and two distractors. The children were asked to match the two identical figures, solving as many tasks as possible within a 5-min time limit. 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 have led to spurious relations in the network. As a result, IQ was defined as the sum of the norm scores of the subtests: similarities, information, comprehension, arithmetic, picture completion, picture arrangement, object assembly, and substitution.

**Attention.** Attention was included as a binary variable. It reflected the judgement of 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 completed by parents. SES was calculated as the mean disposable income per household, based on postal code, and was 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 L–SS mappings were taught explicitly and consequently repeated intensively in order to obtain automatized L–SS associations. Regular L–SS correspondences were trained and, subsequently, irregular L–SS mappings were taught with increasing difficulty (i.e., first short vowels, then long vowels, then diphthongs). Children were asked to pronounce the displayed vowels and were then 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 the reader to Tijms (2011) or Fraga González et al. (2015).

#### *Intervention data*

Intervention data were used to examine whether intervention progress moderated the network. Here, we considered 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. 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-Minuuut Test; Brus & Voeten, 1973). In this test, children were asked to read as many words as possible within 1 min, from 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.

#### *Statistical analysis*

This data set 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: (a) estimate the edges, (b) calculate node predictability, (c) visualize the network, (d) assess node centrality, and (e) 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 data set 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 least absolute shrinkage and selection operator regularization (LASSO) (Tibshirani, 1996). The strength of this penalty is controlled by a parameter  $\lambda$ , selected with the extended Bayesian information criterion (EBIC) (Haslbeck & Waldorp, 2020). The EBIC itself has a tuning parameter  $\gamma$ , which in practice controls the trade-off between sensitivity and precision, with a default of 0.25 (Haslbeck, Borsboom, & Waldorp, 2019). Typically, 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, LASSO has a low likelihood of false positives (Krämer, Schäfer, & Boulesteix, 2009). 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 were expected to be small by nature and therefore more likely to be shrunk to zero. Therefore, to allow for recovering small edges in the model – particularly those between different domains – the  $\gamma$  value was set to 0 for the purpose of this study (note that  $\gamma = 0$  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 non-zero (Haslbeck & Waldorp, 2020). All continuous variables were scaled to a mean of 0 and standard deviation (SD) of 1 in

order to avoid penalization of a given parameter depending on the *SD* of the associated variable (Haslbeck *et al.*, 2019).

Node predictability – how well a node can be predicted by nodes it shares an edge with – was calculated for all the 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 next step involved visualizing the estimated edges and the node predictability using the R-package *qgraph* (version 1.6.5) (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 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 – respectively, these quantify 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 (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). To investigate the robustness of our network obtained using  $\gamma = 0$ , the same estimation procedure was repeated with the default value of  $\gamma = .25$  and estimations of the two networks were compared (see Supplementary Material Figure S1). We ran 100 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 Supplementary Material Figure S2).

For the second research question, we aimed to examine whether the network of children that made progress during the reading interventions differed from those that did not. One way to do this was to split the data set into two subgroups – responders to intervention and non-responders. However, this required an arbitrary cutoff 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.*, 2019). Therefore, rather than dichotomizing our sample into 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.*, 2019). 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 was estimated using the R-package *mgm*, with intervention progress as the moderator (Haslbeck *et al.*, 2019). Node predictability was calculated for all the nodes in the network (Haslbeck & Waldorp, 2018) and the reliability of estimates was examined (for full details, see Haslbeck *et al.*, 2019).

## 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 1,257 children (696 boys), aged between 85 and 158 months ( $M = 114.00$ ,  $SD = 14.42$ ) with scores on 28 variables for the general network. Selection of only children who had 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 provided in Table 2.

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.

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

RAN was positively connected to reading fluency in high-frequency words and pseudowords, but not to reading fluency in low-frequency words. Regarding the reading accuracy variables, only a negative association between rapid naming and reading accuracy in pseudowords emerged, meaning 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 PA, visual perception and L–SS speed.

L–SS speed was found to have a large association with spelling fluency, which means that the children who were more fluent in identifying correct L–SS associations also were more fluent in spelling. Children that were more accurate in identifying L–SS associations were more accurate in spelling, but on average needed more time to spell words. These results are indicated by, respectively, the positive edge between L–SS accuracy and spelling accuracy and the negative edge between L–SS accuracy and spelling fluency. Simultaneously, L–SS accuracy was positively associated with reading accuracy in pseudowords, suggesting that children

**Table 2.** Descriptive statistics for the total sample and the intervention subsample

	Total sample (N = 1,257)		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
ReaFIH	31.85 (10.53)	2.00–60.00	29.95 (10.35)	2.00–53.00
ReaFIL	24.04 (9.94)	1.00–48.00	22.09 (9.41)	2.00–48.00
ReaFIP	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)	1,050–8,450	3472.00 (680.18)	1,060–8,450
GestAge	39.57 (2.10)	25.00–50.00	39.55 (2.16)	25.00–50.00
SES	2857.00 (1017.37)	1,100–10,000	2834.00 (951.35)	1,300–10,000
IntProgr	–	–	9.84 (6.62)	–20.00–33.00
Intervention period (months)	–	–	7.41 (1.78)	3.50–16.23

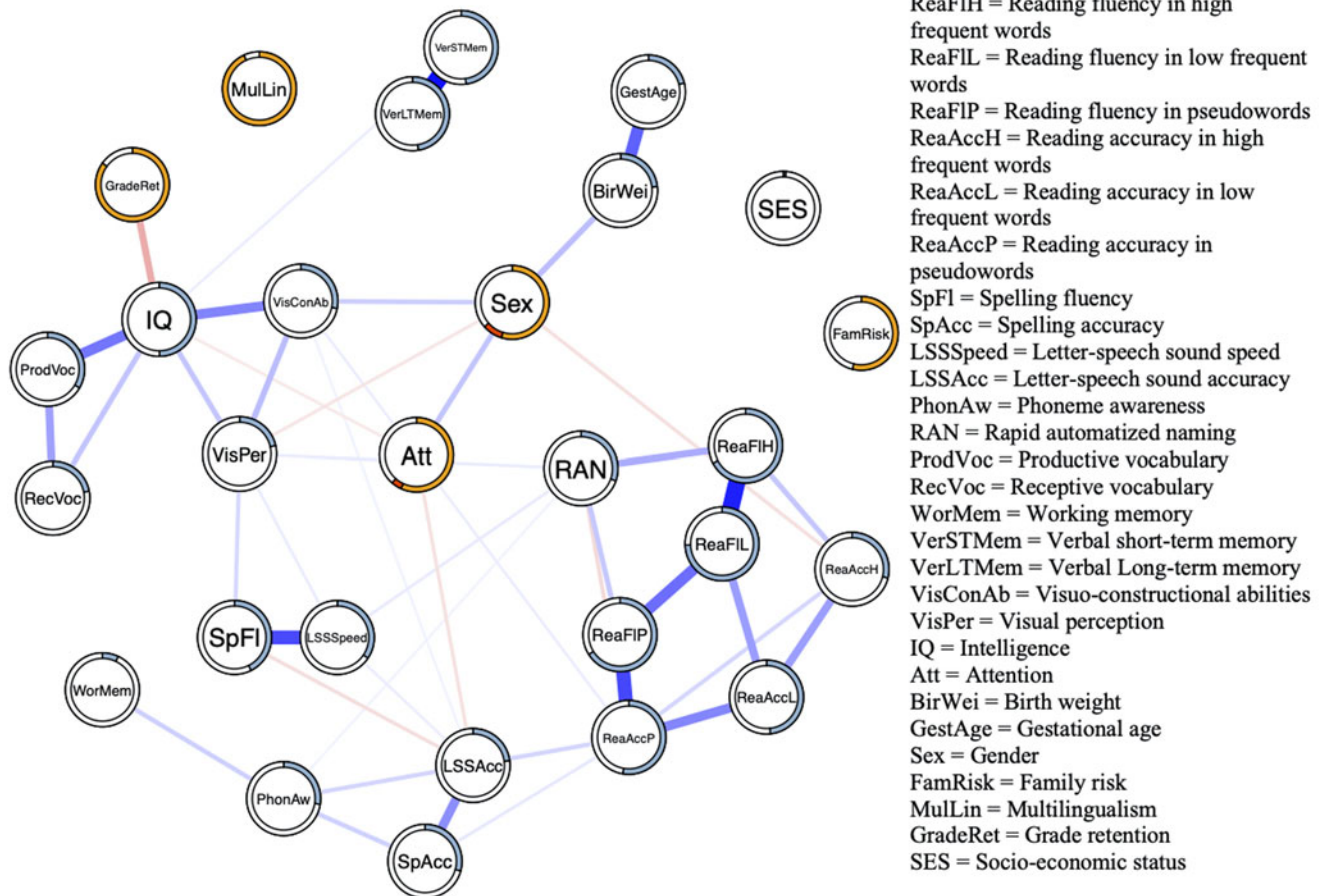
Note: Children in the intervention sample were selected from the total sample. The intervention period depended on the number of sessions. It is therefore possible that children differed in the intervention period but received an equal number of intervention sessions.

who were better in accurately identifying corresponding letters and speech sounds were more accurate in decoding pseudowords and vice versa.

On the left-hand 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, L-SS 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 directly 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, were also not directly related to reading accuracy or fluency. In contrast, a negative association between gender and reading accuracy of high-frequency words

appeared – girls were, on average, more accurate in reading high-frequency words. Gender was also related to attention, which in turn was directly related to L-SS accuracy. This latter association suggests that children with attentional problems are, on average, less accurate in identifying correct L-SS correspondences. Finally, some variables were disconnected from the network: SES, 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 (Figure 1). The proportions of explained variance and accuracy are reported for the continuous and binary variables, respectively. The highest predictability was observed for reading fluency measures (74.50%, 67.30%, and 66.10% for low-frequency words, high-frequency words, and pseudowords respectively). The lowest estimates were observed for family risk, multilingualism, and grade retention (all zero). In addition, low predictability estimates were found for gender (0.88%), attention (0.46%), SES (0.12%), and working memory (0.71%). Low predictability



**Figure 1.** The general network for children with reading difficulties. 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.

implies that all other nodes together share nearly no variance with these variables.

Re-estimating the models with the more conservative  $\gamma = 0.25$  led to nearly identical models (correlations of adjacency matrices  $r > .98$ ). However, some associations disappeared (these were all small ones). In the more conservative model, attention was no longer associated with gender, intelligence, and L-SS accuracy. In addition, the association between visuo-constructional abilities and L-SS accuracy disappeared, and the latter was no longer associated with spelling fluency (see Supplementary Material Figure S1). Finally, the positive relation between L-SS 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, it should be noted that the meaning of closeness and betweenness is questionable in a psychological context given the more complex assumptions and the rather complicated interpretation of these two centrality measures (Bringmann *et al.*, 2019). We therefore mainly focused our interpretations on node strength. The centrality plot

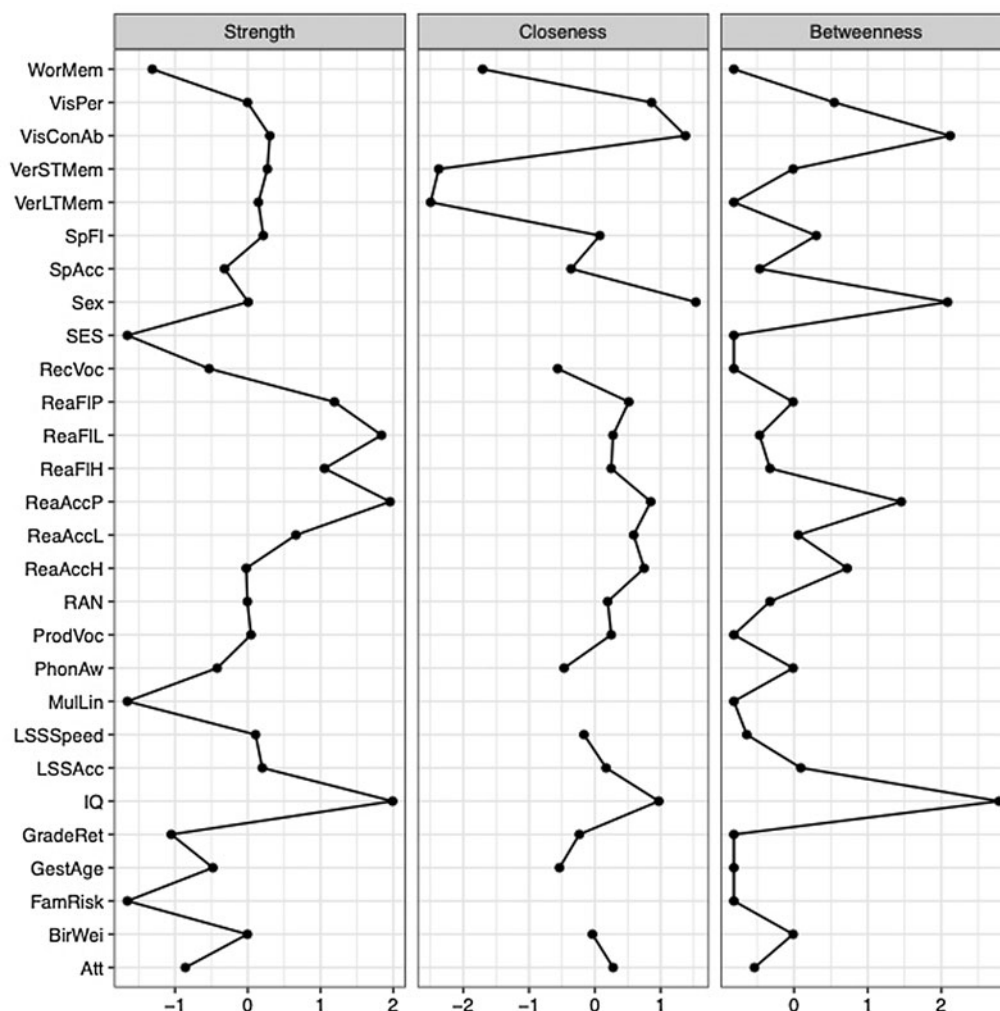
for the general network is depicted in Figure 2, suggesting that nodes differed substantially in their centrality estimates. Reading fluency in low-frequency words, reading accuracy of pseudowords, and intelligence had the highest strength centrality score, whereas visual perception, visuo-constructional 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. 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 of the Supplementary Material. Most edges were estimated reliably as they were included in all or nearly all of the 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.

### Intervention network

Children within the intervention sample were, on average, more fluent in reading after they attended the intervention program





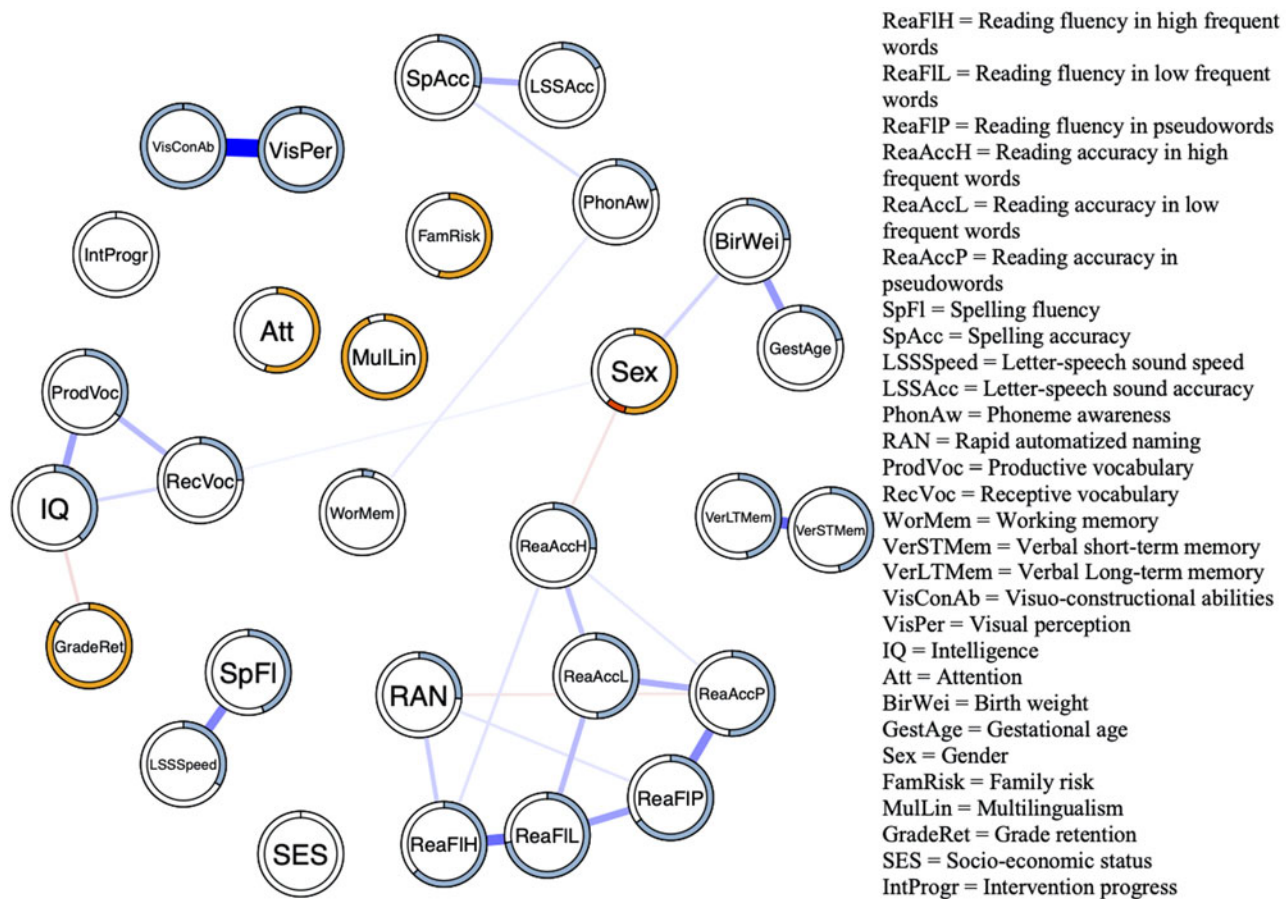
**Figure 2.** Centrality indices for the general network (shown as standardized z-scores).

(intervention progress:  $M = 9.84$ ,  $SD = 6.62$ ). However, note that some children also read less fluently after reading intervention. Those who were severely hampered in fluent reading prior to intervention made the biggest progress. Including intervention progress as a moderator led to the network depicted in Figure 3. Compared with the general network (Figure 1), a large number of associations attenuated or disappeared. On the right-hand side of the network, the subpart comprising reading accuracy and fluency variables remained. In contrast to the general network, RAN was no longer associated with PA and L-SS speed. Moreover, L-SS and spelling fluency were no longer connected to the other variables in the network. Simultaneously, although better spelling accuracy was still associated with better PA and better L-SS accuracy, none of these three variables appeared to be associated with better reading performance. Further, 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 the moderator.

The bootstrapped models evaluating the edge weight accuracy of the intervention network are shown in Figure S3 of the Supplementary Material. Again, most estimated edges were included in all or nearly all of the 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.

## Discussion

Typically, research on the etiology of dyslexia uses an approach based on a single core deficit (Astle & Fletcher-Watson, 2020). However, given the high number of children with dyslexia that remain hampered in fluent reading after intervention and the large inter-individual variability in intervention responses, one needs to take the multidimensional character of dyslexia into account. To further the understanding of how multiple factors interact in the context of reading disorders, this study explored links between cognitive, environmental, and demographic variables related to dyslexia by employing a network analysis using a large clinical database of 1,257 elementary school children (55.37% boys). To our knowledge, this study is the first to



**Figure 3.** Network displaying the relationships between variables in the framework of reading disabilities. Only children that received reading intervention were included in this sample. Intervention progress was included as a moderator. Positive associations are represented as blue edges and negative associations are represented as red edges in the network.

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 on the right-hand side of the network, and one comprising intelligence-related measures on the left-hand side of the network. Interestingly, our network showed that 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 L–SS 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, Aarnoutse, & van Leeuwe, 2008). Furthermore, RAN was functionally related to the automated convergence of visual and speech information as well as to visual processing, which are important in more advanced phases toward reading fluency (Vaessen & Blomert,

2010). In the next section, some findings are discussed in more detail.

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

The phonological theory of dyslexia claims that PA – the ability to focus on and manipulate individual sounds in spoken words – is crucial for the establishment and automatization of L–SS correspondences, which in turn underlie accurate and fluent word recognition (Pennington *et al.*, 2012). In our network, a moderate direct association between PA and L–SS accuracy was found, but no direct relation was found between PA and reading fluency. As mentioned earlier, our results suggest that PA supports decoding, 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 the 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-frequency words tend to become familiar more quickly than low-frequency words and pseudowords, leading to an earlier cognitive shift for high-frequency than for low-frequency words and pseudowords. Therefore, as our sample contained children of different ages, ranging from beginning to more experienced readers, the associations between PA and reading measures in high- and low-frequency words might be attenuated whereas these relations might be found when only beginning readers were included in the sample.

Previous research suggests that verbal memory is subsumed under PA, as both involve phonological processing. However, no association between verbal memory and PA emerged in the current network. In contrast, a relation between PA and working memory did appear. We found that the relation between working memory and reading was mediated by PA, which aligns with a recent study by Knoop-van Campen, Segers, and Verhoeven (2018), which showed that working memory predicted reading efficiency via its relationship with PA. It is postulated that this is particularly true for more difficult PA 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 as to whether reading and spelling skills are aspects of the same phenomenon or that they 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, Neuhoff, Bruder, & Schulte-Körne, 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 L-SS 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 have reported a correlation between preschool vocabulary and phonological awareness as well as later reading abilities (Cooper, Roth, Speece, & Schatschneider, 2002; Kim, Otaiba, Puranik, Folsom, & Grulich, 2014), as an increase in vocabulary size may lead to better PA. These findings suggest that vocabulary is especially important in early reading development. Moreover, vocabulary knowledge seems to be especially related to reading comprehension (Quinn, Wagner, Petscher, & Lopez, 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 PA 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 increase, or decrease, in nodes with which the targeted node shares an edge (Borsboom, 2017). According to the relations in our network, targeting PA and L-SS mappings mainly led 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 L-SS 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 for developing more tailor-made intervention programs.

The relatively high comparable node predictability values for reading fluency measures in all three levels (i.e., high-frequency words, low-frequency words, 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 sharing an edge with a node 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 interventions should include rapid naming training. However, although there is an abundance of evidence to show that RAN is one of the best predictors of reading fluency (Araújo *et al.*, 2015), studies have failed to find a reliable impact on either RAN or reading measures after training RAN as an isolated component (Conrad & Levy, 2011; Fugate, 1997). In the framework of our network, RAN – or more generally the under time-pressure mapping of print and speech information – should be trained in the context of 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; Pfof, Blatter, Artelt, Stanat, & Schneider, 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, Dricot, & Van Reybroeck, 2020). Another closely related variable was reading pseudowords. Interestingly, it was previously proposed to include pseudowords in remedial instruction in order to acquire better fluent reading skills (Fälth, Nilvius, & Anvegård, 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 did 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 comprised 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 with 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 PA and L-SS 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 had 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 were set to zero. 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 versus multiple deficit?*

The prevailing single-deficit theory has led to discussions in the field about the core deficit in dyslexia (e.g., PA vs. RAN or phonological deficit vs. L-SS integration deficit). Although evidence has shown that a single mechanism cannot be considered as the “core” of dyslexia (e.g., O’Brien & Yeatman, 2020), most evidence still originates from single-deficit studies that examine one component (Astle & Fletcher-Watson, 2020); therefore, relations between factors, especially those between different domains (e.g., cognitive, environmental), remain largely understood. 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 (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 was 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 PA, whereas this variable should ideally be conceptualized much more broadly. Moreover, the strong association between variables may have been due to the measures being part of the same test battery – 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, in network science, there are no clear guidelines about whether to combine items in one node instead of keeping them separately. In this study, we thought it reasonable to model the three levels of reading fluency and accuracy (i.e., high-frequency words, low-frequency words, 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 made it unlikely to detect moderation effects as these tend to be quite small in observational studies in general (Haslbeck et al., 2019). 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., 2019).

Third, as our study was conducted in a clinical sample, our sample only included children with reading problems. Therefore, no strong claims can be made as to 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 would allow examination of 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 the relationships between variables change over time. In 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. In addition, other studies 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 obtain 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.

**Supplementary Material.** The supplementary material for this article can be found at <https://doi.org/10.1017/S0954579421000365>.

**Authors' contributions.** CV and JT were involved in the conception and design of the study. CV performed the analyses and wrote the first version of the manuscript. JH gave substantial feedback on the analyses. All authors gave substantial feedback on the manuscript and approved the final version.

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

**Conflicts of Interest.** None.

## References

- Allor, J. H. (2002). The relationships of phonemic awareness and rapid naming to reading development. *Learning Disability Quarterly*, 25, 47–57. doi:10.2307/1511190
- Araújo, S., Reis, A., Petersson, K. M., & Faisca, L. (2015). Rapid automatized naming and reading performance: A meta-analysis. *Journal of Educational Psychology*, 107, 868–883. doi: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, 691–707. doi:10.1016/j.jecp.2013.03.009
- Astle, D. E., & Fletcher-Watson, S. (2020). Beyond the core-deficit hypothesis in developmental disorders. *Current Directions in Psychological Science*, 29, 1–7. doi:10.1177/0963721420925518
- Becker, N., Vasconcelos, M., Oliveira, V., Santos, F. C. D., Bizarro, L., Almeida, R. M. M. 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, 423–445. doi:10.1080/87565641.2017.1374960
- Blomert, L. (2011). The neural signature of orthographic–phonological binding in successful and failing reading development. *NeuroImage*, 57, 695–703. doi: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.
- Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, 16, 5–13. doi: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, 91–121. doi:10.1146/annurev-clinpsy-050212-185608
- Bringmann, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., ... Snippe, E. (2019). What do centrality measures measure in psychological networks? *Journal of Abnormal Psychology*, 128, 892–903. doi:10.1037/abn0000446
- Brus, B. T., & Voeten, M. J. M. (1973). *Een-Minut Test*. Nijmegen: Berkhout.
- Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., ... Hulme, C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science*, 23, 678–686. doi:10.1177/0956797611434536
- Conrad, N. J., & Levy, B. A. (2011). Training letter and orthographic pattern recognition in children with slow naming speed. *Reading and Writing*, 24, 91–115. doi: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, 399–416. doi:10.1017/S0142716402003053
- Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Möttus, 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. doi:10.1016/j.jrp.2014.07.003
- Cuthbert, B. N., & Insel, T. R. (2010). Toward new approaches to psychotic disorders: The NIMH Research Domain Criteria project. *Schizophrenia Bulletin*, 36, 1061–1062. doi:10.1093/schbul/sbq108

- 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, 235–244. doi:10.1017/S0954579416000134
- Ehri, L. C. (1997). Learning to read and learning to spell are one and the same, almost. In C. A. Perfetti, L. Rieben & M. Fayol (Eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 237–269). Mahwah, NJ: Lawrence Erlbaum Associates.
- Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, 50, 195–212. doi:10.3758/s13428-017-0862-1
- 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), 1–18. doi:10.18637/jss.v048.i04.
- Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, 23, 617–634. doi:10.1037/met0000167
- Fälth, L., Nilvius, C., & Anvegård, E. (2015). Intensive reading with reading lists: An intervention study. *Creative Education*, 6, 2403–2409. doi:10.4236/ce.2015.622246
- 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, 62. doi:10.1007/s13278-015-0289-6
- 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, 1121–1125.e2. doi:10.1016/j.jpeds.2015.07.045
- 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, e0143914. doi:10.1371/journal.pone.0143914
- 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*, 50, 2682–2690. doi:10.1017/S0033291719002770
- Fruchterman, T. M., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, 21, 1129–1164. doi:10.1002/spe.4380211102
- Fugate, M. H. (1997). Letter training and its effect on the development of beginning reading skills. *School Psychology Quarterly*, 12, 170–192. doi:10.1037/h0088957
- 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, e89900. doi:10.1371/journal.pone.0089900
- 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, 227–237. doi: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, 321–346. doi:10.1007/s11145-010-9271-x
- Ghazalpour, A., Doss, S., Zhang, B., Wang, S., Plaisier, C., Castellanos, R., ... Horvath, S. (2006). Integrating genetic and network analysis to characterize genes related to mouse weight. *PLoS Genetics*, 2, 11. doi:10.1371/journal.pgen.0020130
- 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, 401–419. doi:10.1080/10888438.2016.1213266
- 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, 373–384. doi:10.1111/j.1469-7610.2004.00358.x
- Haslbeck, J., Borsboom, D., & Waldorp, L. (2019). Moderated network models. *Multivariate Behavioral Research*, 1–32. doi:10.1080/00273171.2019.1677207
- Haslbeck, J. M. B., & Fried, E. I. (2017). How predictable are symptoms in psychopathological networks? A reanalysis of 18 published datasets. *Psychological Medicine*, 47, 2767–2776. doi: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, 853–861. doi: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. *Journal of Statistical Software*, 93(8), 1–46. doi:10.18637/jss.v093.i08
- Hendren, R. L., Haft, S. L., Black, J. M., White, N. C., & Hoefft, F. (2018). Recognizing psychiatric comorbidity with reading disorders. *Frontiers in Psychiatry*, 9, 101. doi:10.3389/fpsy.2018.00101
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... Wang, P. (2010). Research domain criteria (RDoc): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, 167, 748–751. doi:10.1176/appi.ajp.2010.09091379
- Jiang, H., Fei, X., Liu, H., Roeder, K., Lafferty, J., Wasserman, L., ... Zhao, T. (2020). huge: High-Dimensional Undirected Graph Estimation. R package version 1.3.4. Retrieved from <https://CRAN.R-project.org/package=huge>
- Kema, G. N., & Kema-van Leggelo, M. K. G. (1987). *G.S.O. 4-6: Handleiding*. Lisse: Swets en Zeitlinger.
- Kim, Y.-S., Otaiba, S. A., Puranik, C., Folsom, J. S., & Grulich, L. (2014). The contributions of vocabulary and letter writing automaticity to word reading and spelling for kindergartners. *Reading and Writing*, 27, 237–253. doi: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, 341–362. doi:10.1598/RRQ.45.3.4
- 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, 156–169. doi:10.1002/dys.1583
- Kort, W., Compaan, E. L., Bleichrodt, N., Resing, W. C. M., Schittekatte, M., Bosmans, M., ... Verhaeghe, P. (2002). *Wechsler Intelligence Scale for Children III (WISC-III). Test en handleiding*. Amsterdam: 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 & Child Neurology*, 57, 410–419. doi:10.1111/dmcn.12652
- Krämer, N., Schäfer, J., & Boulesteix, A. (2009). Regularized estimation of large-scale gene association networks using graphical Gaussian models. *BMC Bioinformatics*, 10, 384. doi:10.1186/1471-2105-10-384
- Livingston, E. M., Siegel, L. S., & Ribary, U. (2018). Developmental dyslexia: Emotional impact and consequences. *Australian Journal of Learning Difficulties*, 23, 107–135. doi:10.1080/19404158.2018.1479975
- Mascheretti, S., Bureau, A., Battaglia, M., Simone, D., Quadrelli, E., Croteau, J., ... 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, 47–55. doi:10.1111/gbb.12000
- Moll, K., Kunze, S., Neuhoff, N., Bruder, J., & Schulte-Körne, G. (2014). Specific learning disorder: Prevalence and gender differences. *PLoS One*, 9, e103537. doi:10.1371/journal.pone.0103537
- Moll, K., & Landerl, K. (2009). Double dissociation between reading and spelling deficits. *Scientific Studies of Reading*, 13, 359–382. doi:10.1080/10888430903162878
- Mugnaini, D., Lassi, S., La Malfa, G., & Albertini, G. (2009). Internalizing correlates of dyslexia. *World Journal of Pediatrics*, 5, 255–264. doi:10.1007/s12519-009-0049-7
- O'Brien, G., & Yeatman, J. D. (2020). Bridging sensory and language theories of dyslexia: Toward a multifactorial model. *Developmental Science*, doi:10.1111/desc.13039
- 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. doi:10.1016/j.cobeha.2016.05.006
- Pecini, C., Spoglianti, S., Bonetti, S., Di Lieto, M. C., Guarani, F., Martinelli, A., ... Chilosi, A. M. (2019). Training RAN or reading? A telerehabilitation study on developmental dyslexia. *Dyslexia*, 25, 318–331. doi:10.1002/dys.1619

- Pennington, B. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, *101*, 385–413. doi: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*, 816–833. doi:10.1111/1467-8624.00317
- Pennington, B. F., Santerre-Lemmon, L., Rosenberg, J., MacDonald, B., Boada, R., Friend, A., ... Olson, R. K. (2012). Individual prediction of dyslexia by single versus multiple deficit models. *Journal of Abnormal Psychology*, *121*, 212–224. doi:10.1037/a0025823
- 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. doi:10.1016/j.appdev.2019.101067
- Preston, J. L., Molfese, P. J., Frost, S. J., Mencl, W. E., Fulbright, R. K., Hoefft, F., ... Pugh, K. R. (2016). Print-speech convergence predicts future reading outcomes in early readers. *Psychological Science*, *27*, 75–84. doi: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*, 159–175. doi:10.1111/cdev.12292
- Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *Quarterly Journal of Experimental Psychology*, *61*, 129–141. doi:10.1080/17470210701508822
- R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.
- Saan, R. J., & Deelman, B. G. (1986). *De 15-woordentest A en B (Een voorlopige handleiding)*. Groningen: Afdeling Neuropsychologie, AZG.
- 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*, 779–782. doi: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*, 369–390. doi: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: York Press.
- Schlichting, L. (2005). *Peabody Picture Vocabulary Test-III-NL*. Amsterdam: Harcourt Assessment B.V.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464. doi:10.1214/aos/1176344136
- 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*, 7–14. doi: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*, 358–373. doi:10.1111/1467-8624.7402003
- Stanovich, K. E. (1996). Toward a more inclusive definition of dyslexia. *Dyslexia*, *2*, 154–166. doi:10.1002/(SICI)1099-0909(199611)2:3<154::AID-DYS63>3.0.CO;2-B
- Stein, J., & Walsh, V. (1997). To see but not to read; The magnocellular theory of dyslexia. *Trends in Neurosciences*, *20*, 147–152. doi:10.1016/S0166-2236(96)01005-3
- Swanson, H. L., Zheng, X., & Jerman, O. (2009). Working memory, short-term memory, and reading disabilities: A selective meta-analysis of the literature. *Journal of Learning Disabilities*, *42*, 260–287. doi:10.1177/0022219409331958
- Thomas, M. S. C. (2020). Developmental disorders: Few specific disorders and no specific brain regions. *Current Biology*, *30*, R304–R306. doi:10.1016/j.cub.2020.02.019
- Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society Series B (Methodological)*, *58*, 267–288. doi: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*, 273–294. doi:10.1080/01443410601066800
- Tijms, J. (2011). Effectiveness of computer-based treatment for dyslexia in a clinical care setting: Outcomes and moderators. *Educational Psychology*, *31*, 873–896. doi:10.1080/01443410.2011.621403
- Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology*, *40*, 7–26. doi:10.1016/s0022-4405(01)00092-9
- 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*, 33–58. doi:10.1177/002221940103400104
- 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*, 308–321. doi:10.1177/0022219410369096
- Vaessen, A., & Blomert, L. (2010). Long-term cognitive dynamics of fluent reading development. *Journal of Experimental Child Psychology*, *105*, 213–231. doi:10.1016/j.jecp.2009.11.005
- 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*, 297–309. doi:10.3389/fnhum.2014.00346
- Vander Stappen, C., Dricot, L., & Van Reybroeck, M. (2020). RAN training in dyslexia: Behavioral and brain correlates. *Neuropsychologia*, *146*, 107566. doi:10.1016/j.neuropsychologia.2020.107566
- 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*, 301–324. doi:10.1080/10888430802132030
- Wimmer, H., & Mayringer, H. (2002). Dysfluent reading in the absence of spelling difficulties: A specific disability in regular orthographies. *Journal of Educational Psychology*, *94*, 272–277. doi:10.1037/0022-0663.94.2.272
- 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*, e110337. doi:10.1371/journal.pone.0110337
- 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*, 293–300. doi:10.1177/0963721420915873