Check for
updates

Orthographic depth and developmental dyslexia: a meta-analytic study

Desiré Carioti[1] · Marta Franca Masia[1] · Simona Travellini[1,2] · Manuela Berlingeri[1,2,3]

Abstract

Cross-cultural studies have suggested that reading deficits in developmental dyslexia (DD) can be moderated by orthographic depth. To further explore this issue and assess the moderating role of orthographic depth in the developmental cognitive trajectories of dyslexic and typical readers, we systematically reviewed 113 studies on DD that were published from 2013 to 2018 and selected 79 in which participants received an official DD diagnosis. Each study was classified according to orthographic depth (deep vs. shallow) and participant age (children vs. adults). We assessed the difference between DD and control groups' performance in reading tasks and in a wide range of cognitive domains associated with reading (phonological awareness (PA), rapid automatized naming (RAN), short-term working memory (WM), and nonverbal reasoning), including age and orthographies as moderators. We found an age-by-orthography interaction effect in word reading accuracy and a significant effect of age in pseudoword reading accuracy, but we found no effect of age and orthographic depth on the fluency parameters. These results suggest that reading speed is a reliable index for discriminating between DD and control groups across European orthographies from childhood to adulthood. A similar pattern of results emerged for PA, RAN, and short-term/WM. Our findings are discussed in relation to their impact on clinical practice while considering the orthographic depth and developmental level.

Keywords Developmental dyslexia · Orthographic depth · Phonological awareness · RAN · Reading skills · Working memory

Developmental dyslexia (DD) "is a specific learning disability that is neurological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor

✉ Manuela Berlingeri
manuela.berlingeri@uniurb.it

[1] DISTUM, Department of Humanities, University of Urbino Carlo Bo, Urbino, Italy

[2] Center of Clinical Developmental Neuropsychology, ASUR Marche, Area Vasta 1, Pesaro, Italy

[3] NeuroMi, Milan Center for Neuroscience, Milan, Italy

⌂ Springer

spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede the growth of vocabulary and background knowledge." (International Dyslexia Association, 2002).

Although DD has been extensively studied in the past decades, the debate about its causes has continued because of its multifaceted manifestations (Parrila et al., 2020).

Different theoretical frameworks tried to account for causes of DD by assuming deficit in some domain-general cognitive areas as Phonological Awareness (Bishop & Snowling, 2004; Bradley & Bryant, 1978; Démonet et al., 2004; Gabrieli, 2009; Peterson & Pennington, 2015; Vellutino et al., 2004; Snowling, 1981; Stanovich, 1988; Vellutino, 1979) and Working Memory (see Perceptual Anchoring theory; Banai & Ahissar, 2004, 2010), or by attributing to DD some specific neurobiological dysfunctions in the Magnocellular system (Stein & Walsh, 1997; see Stein, 2019 for a review), or in the cerebellar system (Nicolson et al., 2001; Nicolson & Fawcett, 1990, 2011, 2005; Nicolson et al., 1999). These neurobiological alterations would be at the basis of behavioral deficits observed in DD, like deficits in phonological awareness, visuo-ocular motion, contrast sensitivity (Stein, 2019), and process automation (Nicolson et al., 2001).

In particular, the phonological theory attributes to DD a primary deficit in perceiving and manipulating single linguistic unit sounds (the phonemes) and in accurately relating them to their correspondent graphical representation (the graphemes). Based on empirical studies, this deficit seems to prevail in the spectrum of observed DD behavioral deficits (Ramus et al., 2003; Danelli et al., 2017; Reid et al., 2007), also at a cross-linguistic level (Landerl et al., 2013; Paulesu et al., 2001; Ziegler et al., 2010).

Moreover, according to the double deficit hypothesis (Wolf & Bowers, 1999), phonological awareness would represent just one of the two main cognitive markers of DD. The other one would be the Rapid Automatized Naming (RAN), considered as a phonological task by some authors (Bowey et al., 2005; Clarke et al., 2005; Torgesen et al., 1997), but conceived as independent from phonology by Wolf and Bowers (1999). The idea of co-occurrence of phonological and rapid naming deficit found support in several studies on children and adults (see Parrila et al., 2020 for a review) and in cross-linguistic investigations (Landerl et al., 2013; Ziegler et al., 2010). This is quite surprising if we consider that each language has its proper system of phonological rules. In this perspective, phonological awareness and also RAN (if we focused on the lexical retrieval component) are two linguistic constructs that should be as specific as language and reading. Nevertheless, based on the literature (Landerl et al., 2013; Ziegler et al., 2010; Araújo & Faísca, 2019), they seem to be impaired in DD, although their language-specificity.

In line with the multifaceted spectrum of behavioral deficits that emerged by experimental studies and the wide number of causal theoretical frameworks, Pennington (2006) proposed a multiple cognitive deficit model to describe DD. Some comparative studies (Ramus et al., 2003; Danelli et al., 2017) supported this vision highlighting the number of different cognitive and perceptual deficits associated with dyslexia in different DD patients, and suggesting that they could be moderated by some further variables related to both stages of reading acquisition and language-specific issues.

Currently, DD is widely accepted to manifest differently as a function of age and orthography; a growing body of evidence (Borleffs et al., 2018) has shown that language-specific

orthographic codes influence reading difficulties as well as reading acquisition (Seymour, 2005; Seymour et al., 2003; Ziegler & Goswami, 2005; Ziegler et al., 2003).

Nevertheless, to account for cross-linguistic differences was initially hard, as the first age of research on reading in general and on DD, in particular, was characterized by an "Anglo-Saxon bias," namely, by an extensive volume of evidence from English-speaking countries (Share, 2008; Ziegler et al., 2003). Indeed, among papers on dyslexia between 1990 and 2010, the proportion of those that involve English readers exceeded 50% before 2000, whereas the number of studies in other orthographies increased only in the last 20 years (see Fig. 1). Consequently, most empirical studies and reading models were widely conceived and more suitable for Anglophone participants. Thus, from the late 1990s, the need to revisit all these models in light of orthographic depth became compelling even for neuroimaging issues (Devoto et al., under review; Martin et al., 2016; Martin et al., 2015; Richlan, 2014; Richlan et al., 2011).

To investigate the language-specific and universal manifestations of DD, in the last 20 years, some cross-linguistic comparisons have provided a clearer picture of the core cognitive deficits associated with reading disability in childhood and adulthood. Similarly, this meta-analytic study will compare the reading and cognitive skills of non-impaired and dyslexic readers in orthographies characterized by different degrees of consistency. This approach will provide further insights into the nature of reading acquisition differences and the universal character of some common cognitive alterations already indicated as PA and RAN, which are core DD deficits (Georgiou et al., 2008; Landerl et al., 2013; Parrila et al., 2020; Ziegler et al., 2010).
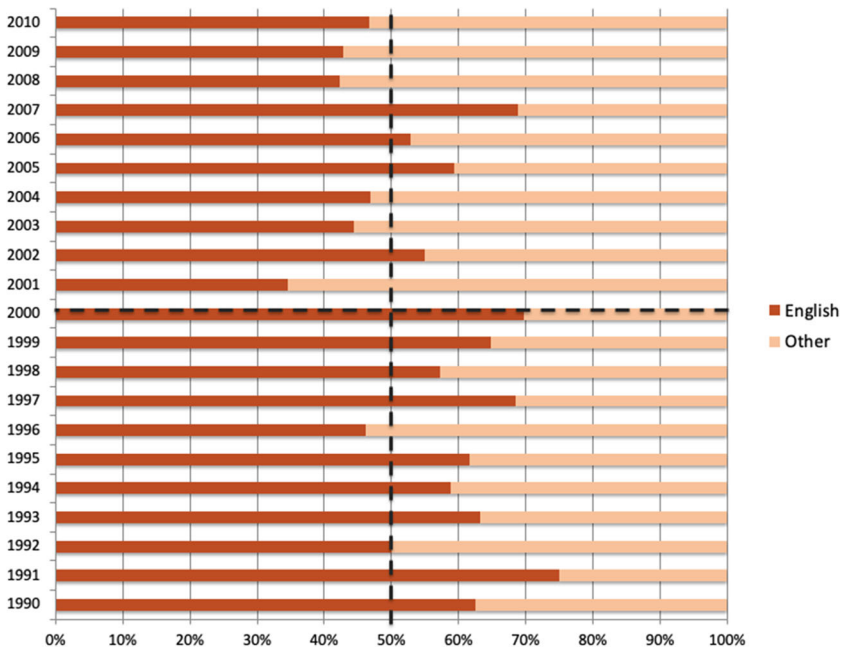
Fig. 1 Proportion of English studies published between 1990 and 2010 whose titles included the word "dyslexia." Before 2000, with only one exception, the percentage of dyslexia studies that include English readers exceeded 50%. Only in the last 20 years, the number of studies published in other languages and orthographies increased, but the percentage of those for English readers always exceeded 30%. This represents the "Anglo-Saxon bias" that characterized the research on the reading process and developmental dyslexia

The following section will briefly define (i) developmental trends in DD, (ii) how orthographic consistency influences reading and reading acquisition, and (iii) how orthographic depth shapes the manifestation of DD to introduce the main issues that this meta-analytic study will address.

Dyslexia in children and adults

As previously highlighted, DD is a developmental disorder that compromises reading efficiency during its acquisition and significantly affects reading proficiency in the life-long period (Scarborough, 1984). The above-mentioned phonological deficit reported by dyslexic children is still found in adults with DD, also when compensation for reading proficiency emerged (Bruck, 1992; Danelli et al., 2017; Elbro et al., 1994; Fostick & Revah, 2018; Olofsson, 2002; Pennington et al., 1990; Ramus et al., 2003; Reid et al., 2007; Scarborough, 1984; Shaywitz et al., 1999; Snowling, 1995; Undheim, 2009).

Indeed, as highlighted by Danelli et al. (2017), magnocellular and motor/cerebellar dysfunctions observed in children are less relevant in adult dyslexic readers (Ramus et al., 2003), while failures in reading fluency and phonological awareness may still be evident (Nergård-Nilssen & Hulme, 2014; see Reis et al., 2020 for a review).

Moreover, according to Eloranta et al. (2019), adults with DD can still exhibit lower proficiency in rapid naming, working memory, and verbal skills. In particular, the rapid naming performance would predict compensational outcomes in adult reading (Eloranta et al., 2019). In a recent meta-analysis by Reis et al. (2020), adult DD readers showed lower performances in several cognitive measures (phonological awareness, working memory, RAN, full IQ, verbal IQ, non-verbal IQ, etc.), although the most severe DD symptoms emerged for reading and writing skills.

These findings dovetail with claims of both phonological theory (Bishop & Snowling, 2004; Bradley & Bryant, 1978; Démonet et al., 2004; Gabrieli, 2009; Peterson & Pennington, 2015; Vellutino et al., 2004; Snowling, 1981; Stanovich, 1988; Vellutino, 1979) and double deficit hypothesis (Wolf & Bowers, 1999), suggesting that phonological awareness and RAN are the more reliable markers of DD, regardless of age, together with some others cognitive aspects as working memory and verbal skills. Reis et al. (2020) showed that orthographic consistency could play a role in moderating DD manifestations in adult readers of deep and shallow orthographies. In line with this, it is even more relevant to consider both developmental and cross-linguistic perspectives when approaching DD.

Orthographic consistency and reading

Katz and Frost (1992) asserted that a highly stable orthographic system where single-unit sounds and their assembly remain unchanged between words is defined as a "shallow orthography" (see, for example, the Italian word "casa," where the spelling pattern of the four graphemes corresponds to the pronunciation of the four distinct phonemes [/ˈkasa/]). By contrast, in some orthographies, certain letters have no phonemic representation (for example, "w" in the English word "whole" [/həʊl/] or "t" in the French word "chat" [/ʃa/]), and some vowels change sounds as a function of the letters they are linked to (for example, the "u" in "ugly" and "huge" in English). Moreover, some decoding rules are generally inconsistent and

less stable; these orthographies are referred to as "deep" because the grapheme-to-phoneme correspondence is more "opaque" (Katz & Frost, 1992).

As mentioned above, the different rules that must be implemented across languages and orthographies to obtain a correct and fluent reading process cannot, by definition, allow the application of identical acquisition mechanisms and the manifestation of the same impairment patterns in reading disabilities (Ziegler & Goswami, 2005).

For what concerns accuracy, literacy acquisition studies revealed that readers of shallow orthographies exhibit faster word and nonword reading performance, reaching 80–90% accuracy at the end of the first grade, while readers of deep orthographies suffer from delays in accurate decoding, specifically for nonword reading (Aro & Wimmer, 2003; Frith et al., 1998; Seymour et al., 2003; Seymour, 2005).

According to Seymour et al. (2003), these cross-linguistic differences can be attributed not only to orthographic depth but also to syllabic complexity, a factor that seems to affect in particular fluency and accuracy of nonword reading, that is, the so-called *phonological recoding*. To advance the body of knowledge on language-specific challenges in reading acquisition, Ziegler and Goswami (2005) proposed the *Grain Size Theory*, suggesting that beginner readers encounter three problems: the *availability* of pre-reading phonological and linguistic unit representation, the *consistency* between letters and sounds, and the *granularity* of the orthography.

These works all challenged the validity of the reading models devised for English participants and especially the notion that alphabetic languages maintain a universal switch from a primary alphabetic reading strategy to automatic sight-word reading (Aro, 2004; Ehri, 1987, 1991, 1995, 2005; Frith, 1985).

Children learning to read in shallow orthographies seem mostly dependent on grapheme-to-phoneme conversion to decode words, whereas children learning deep orthographies would benefit more from orthographic whole-word analysis because of the orthographic system's irregularity (Aro, 2004). Parallel with these predictions, concerning reading fluency, some studies (Ellis & Hooper, 2001; Ellis et al., 2004) reported that younger readers of shallow orthographies have longer latencies when reading longer words while English readers showed no word length effect. Conversely, the latter were more prone to incur substitution errors while readers of shallow orthographies committed nonword reading pronunciation errors more frequently (Ellis et al., 2004). These results suggest that readers of shallow orthographies implement single-phoneme mapping and grapheme-to-phoneme strategies to decode words, while those of deep orthographies depend less on phonological recoding to acquire orthographic representations and lexical storage. This is inconsistent with the idea that the complete acquisition of the alphabetic stage of reading (Ehri & McCormick, 1998; Frith, 1985) is crucial to achieving orthographic analysis; thus, authors such as Castles and Coltheart 2004; Castles et al., 2003) questioned even the idea that phoneme awareness is a cause, rather than a result, of reading acquisition and proficiency in general. Nevertheless, the mutual relation between PA and reading proficiency is widely accepted (Gottardo et al., 2016; Hulme et al., 2005; Perfetti et al., 1987).

Dyslexia in deep and shallow orthographies: universal and language-specific features

Cross-linguistic studies that explore the language-specific and universal characteristics of DD remain limited; the main challenge to an empirical cross-linguistic comparison of reading is the different methods for evaluating fluency. For example, word and pseudoword reading tasks

usually test sight-word reading and phonological recoding, but a different reading fluency measure can be conceptualized via reading assessment in different orthographies; indeed, deep orthographies usually provide an accuracy-based fluency measure that scores participant performance in terms of the number of accurate words read in 45 seconds or 1 minute (e.g., the English Test of Word Reading Efficiency (TOWRE) test; Torgesen et al., 2012). Conversely, shallow orthographies, wherein the ceiling level is more often achieved (Araújo et al., 2015), usually measure fluency using speed (seconds or syllables/seconds) and thus generate one index for accuracy and a separate one for fluency (e.g., the Italian DDE-2 Test; Sartori et al., 2007). Moreover, the overall accuracy level is more often considered by deep orthographies (Share, 2008; Sprenger-Charolles et al., 2011). These methodological differences complicate the direct comparison of performance between, for example, Italian and English readers without adopting ad hoc reading tasks that "parcel out" the specific effects of different procedures and language-specific psycholinguistic variables.

Landerl et al.'s (1997) pioneer study compared reading performance between English and German dyslexic children using 192 words and 192 ad-hoc pseudowords that were similar in spelling and pronunciation and identical in meaning. The authors found that English dyslexics were less accurate and slower in reading pseudowords and were more impaired in reading low-frequency words than their German counterparts. These results indicate a more severe behavioral manifestation of DD in the English context that, as the authors argue, would be "triggered by the key orthographic feature distinguishing German and English orthography, namely the difference in the consistency of grapheme-phoneme relations for vowels" (Landerl et al., 1997, p. 328). Similarly, Ziegler et al. (2003) observed differences between German and English dyslexic readers in overall accuracy but a common impairment in reading speed in all tasks. Notably, both groups reported lower reading fluency levels when compared to age- and level-matched controls but did not perform differently from level-matched controls in terms of accuracy; Ziegler et al. (2003) interpreted this as a delay in accurate reading decoding acquisition. Nevertheless, considering the remarkable difficulty in accurate pseudoword decoding also shown by typical English readers, the authors described this finding as a language-specific feature (Ziegler et al., 2003). This observation is consistent with Paulesu et al.'s (2000) observed behavioral disadvantage for proficient adult readers of deep orthographies and thus seems to represent a language-specific characteristic that remains stable across age. Interestingly, beyond the accuracy results, Ziegler et al. (2003) interpreted the speed-level cross-linguistic deficit as a core DD feature. In particular, DD would be characterized by a "specific nonword reading deficit and a phonological decoding mechanism that operates extremely slowly and serially" (p. 188).

The pattern shown by developmental dyslexic readers in different orthographies was consistent with that reported by typical readers (Landerl et al., 1997; Ziegler et al., 2003). Accordingly, English developmental dyslexic readers reported a lower accuracy than Germans not because of a more severe deficit but because of a generally tougher achievement of reading proficiency in deep orthographies that, in turn, would further exacerbate behavioral symptoms. Within this framework, the first universal behavioral sign of DD would be a reading speed deficit, while, according to the authors, a phonological deficit would represent a cognitive underpinning. Paulesu et al. (2001) reported similar results: despite cross-linguistic differences in reading performance, with Italian readers committing fewer errors for word and nonword reading compared to English and French readers, dyslexic adult readers showed consistent impairment across countries compared to control adult readers. These conclusions are further supported by the meta-analytic results of Reis et al. (2020). Indeed, the most reliable cross-

linguistic differences emerged for reading accuracy, in which DD readers of shallow orthographies reported smaller effect sizes when compared to their age-matched controls.

All these findings, when considered with results of behavioral studies on readers of deep (mostly English) orthographies (e.g., Elbro et al., 1994; Griffiths & Frith, 2002; Kemp et al., 2009; Ramus et al., 2003) and more transparent orthographies such as Finnish, Italian, and Spanish (Burani et al., 2006; Laasonen et al., 2012; Leinonen et al., 2001; Suárez-Coalla & Cuetos, 2015), let emerge a clear picture. Dyslexic readers seem to be characterized by a phonological deficit associated with difficulties in implementing an orthographic reading strategy (Ziegler et al., 2008).

Although a reading speed deficit was mainly observed among adult Italian (Burani et al., 2006) and Spanish (Suárez-Coalla & Cuetos, 2015) dyslexic readers, this cannot be considered a language-specific feature of dyslexia because of reading assessment differences and because an accuracy-based reading evaluation is predominant in deep orthographies (Sprenger-Charolles et al., 2011).

Landerl et al.'s (2013) result further supported the hypothesis that dyslexic readers may manifest universal phonological and automation deficits (expressed in terms both of rapid naming and reading fluency; as suggested by Ziegler et al., 2003) since childhood. The authors compared dyslexics and controls from five different orthographies using Seymour et al.'s (2003) five-level classification. They observed that the best predictors of reading deficit were PA and RAN tasks while intelligence quotient (IQ) and short-term/WM played a minor role and that the digit span score predicted the reading process in more inconsistent orthographies (Landerl et al., 2013). Moll et al. (2014) replicated these results on typical readers, finding that PA and RAN were strong predictors of second-graders' reading skills, with RAN predicting speed and PA accounting for accuracy. These results are also in line with Ziegler et al. (2010), suggesting that PA and RAN are the most reliable reading skill predictors thus far. Once again, this behavioral evidence seems to support the phonological theory (Démonet et al., 2004; Gabrieli, 2009; Peterson & Pennington, 2015; Vellutino et al., 2004; Bishop & Snowling, 2004; Stanovich, 1988; Bradley & Bryant, 1978; Vellutino, 1979; Snowling, 1981) and the double deficit hypothesis (Wolf & Bowers, 1999).

Research objectives

Several authors (Brislin, 1976; He & Van de Vijver, 2013; Peña, 2007) acknowledged several methodological issues with cross-cultural data collection, most of which were associated with sampling and difficulties in finding adequate cross-linguistic tasks (Landerl et al., 1997; Seymour, 2005). However, the substantial literature on the behavioral and cognitive markers of DD provides a more reliable picture of how orthographic depth affects reading skills development if a meta-analytic approach is employed.

A meta-analytic approach represents a robust method for exploring the universal characteristics of DD by considering the effect of both age and orthographic depth and determining the reliable signatures of DD from findings that are not systematically replicated across studies. This approach enables a more detailed "cognitive characterization" of DD by describing, beyond its universal signs, cognitive characteristics moderated by an interaction between developmental and orthographic factors.

Given this phenomenon's complexity, we expect different scenarios regarding how DD influences reading accuracy or fluency measures:

- between-groups differences, namely, differences between developmental dyslexic readers and age-matched controls irrespective of age and orthographies, which will help us identify universal cognitive DD markers;
- between-groups differences moderated by age, namely, differences between children and adults with and without DD, which will show variations in the developmental trajectories of reading acquisition across orthographies;
- between-groups differences moderated by orthography, namely, differences between controls and developmental dyslexic readers representing linguistic-specific effects regardless of age;
- between-groups differences moderated by age and orthographic depth, namely, the behavioral differences between developmental dyslexic readers and controls specifically characterizing these groups' developmental trajectories in a specific linguistic context.

Additionally, we consider other cognitive dimensions, such as RAN, PA, nonverbal reasoning, and short-term/WM, to test whether systematic cognitive differences exist between the two groups and whether they are moderated by our factors of interest.

Finally, this meta-analytic approach will help determine whether the two available reading fluency measures—time-unlimited (TU) and time-limited (TL) tasks—are associated with the same pattern of findings.

Methods

Literature search

To compare the reading skills of people with and without DD, from a wide range of European orthographies, we used the US National Library of Medicine's PubMed online database to search the following keywords: "dyslexia AND test," "dyslexia AND behavioral/behavioural tasks," and the more general "dyslexia."

Due to the high amount of literature about DD, we decided to limit the period meta-analyzed to 5 years. Accordingly, only papers published between 2013 and 2018 were considered. 2013 was considered as a good starting point as it corresponds to the publication year of DSM-5 (American Psychiatric Association, 2013) and, thus, to the updating of diagnostic criteria for learning disorders at an international level, at least for what concerns children.

The first query was run in January 2017; the pool of selected papers was further updated in January 2019.

Initially, we considered 1,994 papers for inclusion, of which 291 were found using "dyslexia AND test," 59 using "dyslexia AND behavioral/behavioural tasks," and 1,644 using "dyslexia." After removing duplicates ($N = 633$), an initial title inspection reduced the selection to 595 studies. Each abstract was reviewed for inclusion criteria and relevance.

Inclusion and exclusion criteria

We defined inclusion and exclusion criteria based on The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher et al., 2009) and specific requirements imposed by our research questions.

Inclusion criteria limited the selection to studies wherein DD was the main topic. In particular, applied inclusion criteria are reported in follows:

- all participants had official DD diagnoses by a professional neuropsychologist, speech therapist, or other authority in compliance with each country's provisions;
- an age-matched control group was present;
- the orthographies involved were all European and rated as "deep" or "shallow" according to Seymour et al.'s (2003) classification;
- the results reported clear neuropsychological and cognitive data;
- raw data, with means and standard deviations, were reported or could eventually be calculated based on information about the type of reported scores (T points, C points, standard scores, etc.);
- the type of measure was clearly expressed as, for example, "accuracy," "proportion of accurate answers," "total reading time," or "words read in a minute."

However, we excluded the following: studies wherein comorbidity with other pathologies was present, even from a neurodevelopmental point of view; studies wherein participants were preschoolers or children with a familial risk of dyslexia; studies conducted in countries where a non-European language is spoken; studies involving poor readers or investigating psychological but not DD-associated cognitive aspects (e.g., quality of life, school satisfaction, and performance anxiety); single-case papers; editorials; reviews and other meta-analyses.

Literature search results

Of the 595 papers selected based on their titles, 113 met the inclusion criteria. Further, we removed 21 studies whose participants read in Portuguese, Dutch, and Swedish: as this study aimed to compare languages with higher transparency and opaqueness, orthographies at the third level of orthographic depth (the medium level per Seymour et al., 2003) were considered confusing and were excluded. From the remaining 92 papers, we extracted behavioral indexes of interest for word reading (accuracy and fluency), non-lexical decoding (accuracy and fluency), RAN, phonologic awareness, short-term/WM, and nonverbal reasoning. Of these, 13 were excluded after a quality check of the methods and reported data, narrowing the sample to 79. Figure 2 shows the selection process.

Coding procedures

Different datasets were created for each study dimension based on the information from the papers. The general cognitive dimensions meta-analyzed were word reading, non-lexical decoding, PA, RAN, short-term/WM, and nonverbal reasoning. The first two dimensions helped quantify the reading-associated behavioral deficit in DD across country and age; the last four were used to explore DD-associated cognitive markers and identify universal aspects of this neurodevelopmental deficit. Each dimension was explored by one or more independent meta-analyses depending on the nature of the reported data. For reading measures, for example, we considered two parameters (accuracy and fluency), but we had to address the heterogeneity of adopted measures in each country. In particular, as described in the introduction, reading fluency data are often heterogeneous. Fluency parameters can be measured as total reading time by tests, such as the Spanish PROLEC-R (Cuetos et al., 2007), the Italian
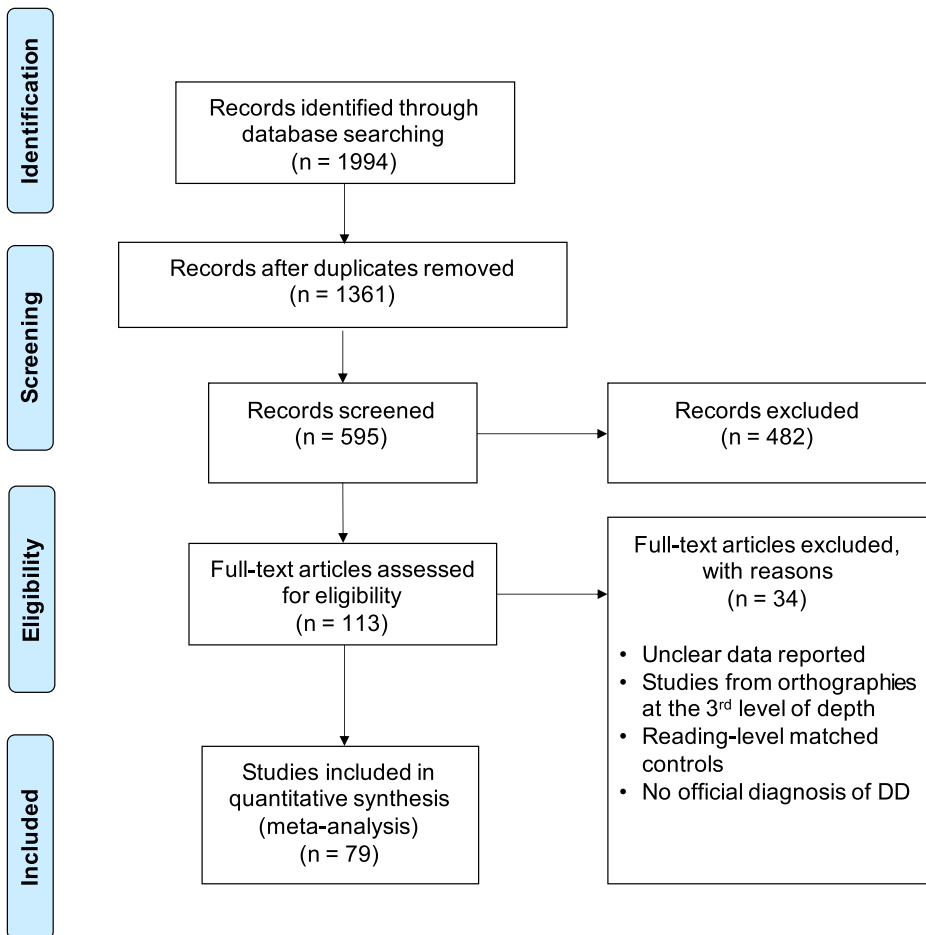
Fig. 2 The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram (Moher et al., 2009) shows the entire literature recognition and paper selection process. Studies were excluded if they contained incomplete methodological information or met the exclusion criteria

DDE-2 battery (Sartori et al., 2007), the French Belec (Mousty & Leybaert, 1999) and ECLA +16 (Gola-Asmussen et al., 2010), or the number of words accurately read in a given time, as in the widely used One Minute Test (Brus & Voeten, 1973), Lecture en Une Minute (Khomsi, 1999), the Dyslexia Screening Test-Junior (Fawcett & Nicolson, 2004), the Ein-Minuten Leseflüssigkeitstest (Willburger & Landerl, 2009), and the TOWRE (Torgesen et al., 2012). The following sections report these two modalities as TU and TL. Figure 3 depicts the heterogeneity of these measures.

Apart from reading fluency, we reported the tests and measures meta-analyzed for each cognitive dimension:

1. *word reading* (including reading high- and low-frequency as well as regular and irregular words, all mainly based on lexical identification), for which three independent meta-analyses for accuracy (total score, percentage, and proportion), TU (sec, msec), and TL (number or percentage of words read) were implemented;
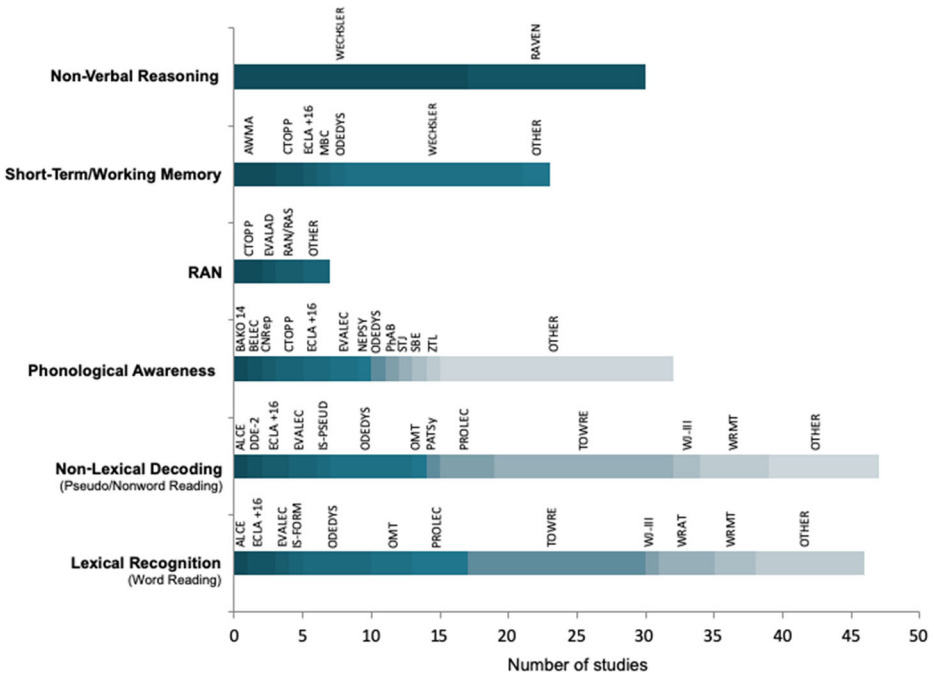
Fig. 3 Heterogeneity of tasks and standardized tests available in European countries for each cognitive dimension typically included in the clinical assessment of developmental dyslexia

2. *non-lexical decoding* (including nonword and pseudowords reading tests requiring a phonological recoding strategy), for which three independent meta-analyses for accuracy (total score, percentage, and proportion), TU (sec, msec), and TL (number or percentage of words read) were implemented;

3. *phonological awareness*, conceptualized as either *phonological manipulation* (including tasks such as phoneme deletion and elision, segmentation and synthesis, and spoonerisms) or *nonword repetition*, for which all the data in the two meta-analyses concerned accuracy;

4. *RAN*, for which the meta-analysis reported tests on the rapid naming of, mainly, objects, colors, and numbers. Performances were reported in terms of speed (sec, msec);

5. *short-term/WM*, for which the overall meta-analysis primarily reported scores for forward and backward digit span or quotient measures by the WM index of Wechsler's battery;

6. *nonverbal reasoning*, for which the meta-analysis reported scores from the performance IQ or matrix reasoning subtests of Wechsler's tests, Raven's matrices (standard progressive matrices, Raven, 1958; colored progressive matrices, Raven, 2003; Raven, 1956), or other nonverbal reasoning tests as the Cattel Culture Fair Intelligence Test (CCFIT; Cattell, 1940) and the Naglieri Nonverbal Ability Test (Naglieri, 2003).

Orthographic depth classification of European orthographies. Only European orthographies were considered in the present study. In particular, languages and orthographies included were classified as shallow or deep, referring to Seymour's classification (2003). Here European orthographies were rated at a different depth level based on their grapheme-to-phoneme consistency and rated as "complex" or "simple" based on their syllabic

complexity. Seymour's classification is reported in Table 1. As widely supported in literature (see Schmalz et al., 2015 for a review) and according to Seymour's classification, English represents the deeper and more complex level.

Factor and variable coding. From each selected paper, we extracted cognitive dimensions, specific tasks within each dimension, and the type of raw measure associated with each task. For example, from Cantiani et al. (2015), we extracted the "phonological awareness" dimension, the "phonological manipulation" task, and the corresponding "accuracy" measure. Each task's mean raw score and standard deviation were reported for both the dyslexia group (DG) and the control group (CG), with each group's mean age (expressed in years). Subsequently, we re-coded age into a dichotomous variable: children (8–13 years) and adults (19–32 years). Moreover, for each measure, we also included whether the participants' countries adopted a shallow or deep orthography according to Seymour et al.'s (2003) classification of European orthographies (see Table 1). Specifically, the shallow level included only measures extracted from studies in which the orthography was at level 1 or 2; the deep level corresponded to levels 4 and 5. See Fig. 4 and Table S1 for more details about orthographies of papers included.

## Statistical analysis

Effect size calculation

The analyses were conducted in R (R Core Team, 2015) with the "metafor" package (Viechtbauer, 2010). The effect size of each measure, namely, standardized differences between DG and CG, was calculated using the *escalc* function. If multiple measures using the same cognitive dimension were reported within a single paper (e.g., high- and low-frequency word reading accuracy in the same participant samples), then a single effect size was computed using the *agg* function for the aggregation of dependent effect sizes in the R package "MAd" (Del Re & Hoyt, 2014).

Data analysis

First, we identified outliers by exploring the boxplots and calculating the I^2 index (Higgins et al., 2003) using the *rma* function and computing a random effect model (REML). Second, we ran a second REML to check whether between-studies heterogeneity was reduced and decide

Table 1 Seymour's classification of European Orthographies (Seymour et al., 2003, p. 146)

		Orthographic Depth				
		Shallow 1st level	2nd level	3rd level	4th level	Deep 5th level
Syllabic Complexity	*Simple*	Finnish	Greek Italian Spanish	Portuguese	French	English
	Complex		German Norwegian Icelandic	Dutch Swedish	Danish	
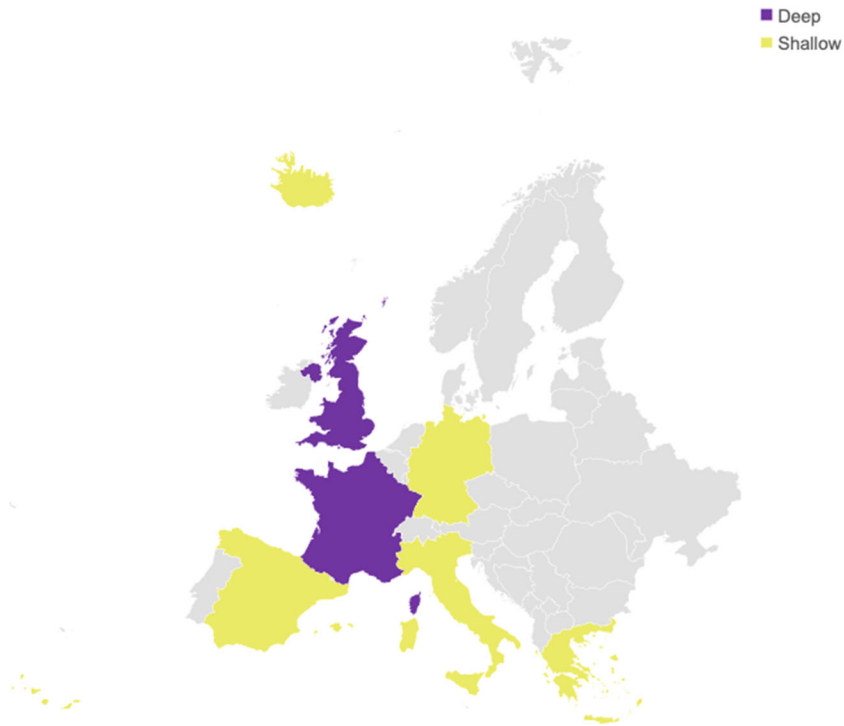
Fig. 4 Representation of European countries involved in the meta-analysis. Countries in purple adopt a deep orthography; those in yellow represent countries that implement a shallow orthography. Sixteen further studies included were conducted in the US, and one additional research was done in Australia; all were categorized as "deep orthography"

whether to conduct a further REML including the categorical variables "age" and "orthography" as moderators or to run a fixed effect (FE) model without any further moderator (only when the *tau* value is 0). Finally, we selected the best-fit model based on the Aikake information criterion, a widely used model selection standard (Arnold, 2010).

Moreover, to check whether effect sizes of nonverbal reasoning would predict variations in effect sizes of word reading (accuracy, time unlimited, and limited), we run three further meta-regressions. Only studies that provide data for both nonverbal reasoning and word reading accuracy and fluency (TU and TL) were included in the meta-regressions.

Results

This meta-analytic study included 79 articles including 450 ESs comparing dyslexic with age-matched typical readers. The total sample size was 14,947 participants, 49% (mean sample size = 659.4; SD = 209.4) for DG and 51% (mean sample size = 682.6; SD = 206.12) for CG. In all meta-analyses the DG (mean $age_{(adults)}$ = 23.54, SD = 2.88; mean $age_{(children)}$ = 10.22, SD = 0.77) and CG (mean $age_{(adults)}$ = 23.69, SD = 3.76; mean $age_{(children)}$ = 9.91, SD = 0.72) groups were equivalent in age (as reported in Table S3 of supplementary materials).

Overall, for what concerns the orthographic depth, 67.4% of studies included were rated as "deep" and 32.5% as "shallow," providing further evidence in favor of the Anglo-Saxon bias.

In what follows, results of models run on each reading and cognitive dimension are reported. We reported only the best-fit model results for each dimension together with the ensuing heterogeneity level, the main effect, or the interaction effect (of moderators, in the case of REML) when significant. Moreover, tables report the number of studies for each moderator and the average effect sizes in each model.

Word reading

Accuracy

This study initially reported 59 observations for both DG and CG in the word reading accuracy dataset. Of these, 36 were dependent and were thus aggregated through the *agg* procedure, resulting in 41 total observations extracted from 38 studies (Casini et al. (2018) reported results from two experiments, Nittrouer and Lowenstein (2013) reported results from two different samples for both DG and CG, and Ruffino et al. (2014) reported data of two different samples of DD readers with and without phonological deficit).

The best-fit model for the data was the REML with "orthography" and "age" as moderators (see supplementary materials for model selection details). The residual heterogeneity test was significant ($QE_{(37)} = 56.5$, $p = 0.02$; $I^2 = 32.63\%$, $tau^2 = 0.07$, SE = 0.05), and the same can be said for the moderators' test ($QM_{(3)} = 39.3$, $p < 0.001$). The model, which stratified the studies based on the two moderators (see Table 2), revealed a significant difference between DG and CG ($Z = 14.34$, SE = 0.15, $p < 0.001$), and an age-by-orthography interaction ($Z = 2.24$, SE = 0.32, $p = 0.02$). In particular, as summarized in Table 3 and Fig. 5, children of shallow orthographies reported smaller effect size (1.53, 1.42–1.63) than those of deep (2.38, 2.33–2.42) and, accordingly, the increasing trend between children and adult was broader in deep (from 2.83 to 0.83) than in shallow orthographies (from 1.53 to 1.12).

Fluency—time unlimited (TU)

We obtained 19 observations from 19 studies for word reading TU by aggregating six dependent measures from the 25 initial observations in the dataset.

The REML without moderators was the best-fit model (see supplementary materials for more details) and revealed a significant between-group difference ($Z = -13.43$, SE = 0.10, $p < 0.001$; see Fig. 6, panel a, and Table 2 for more details) with a nonsignificant level of heterogeneity ($Q_{(18)} = 27.26$, $p = 0.07$, $I^2 = 37.44\%$, $tau^2 = 0.07$, SE = 0.06).

Fluency—time limited (TL)

First, we aggregated the dependent measures from Sigurdardottir et al. (2017) and Bonifacci et al. (2017) in single effect size, with the final dataset including 25 studies with 26 observations (Casini et al. (2018) and Halliday et al., (2014) reported measures from two different dyslexic and control samples). After exploring the boxplots, we rated Lallier et al. (2013c) and Bogon et al. (2014) as outliers and removed them. Thus, the first REML was implemented on 24 observations extracted from 23 papers.

The best-fit model for the data was the REML without moderators (see supplementary materials for details about model selection). The model revealed a nonsignificant level of heterogeneity ($Q_{(23)} =$

Table 2 Distribution of the observations in each reading and cognitive dimension considered, type of model run, and average Effect Sizes (ESs) extracted

	Orthographic Depth	Age	Records (k)	Participants (N)	Model	Average ES	min.	max.
Word Reading								
Accuracy	Shallow	Children	6	286	REML with moderators	1.53	1.42	1.63
		Adults	9	392		1.12	1.04	1.20
	Deep	Children	13	799		2.38	2.33	2.42
		Adults	13	461		0.83	0.75	0.91
		Tot.	41	1938				
Fluency (TU)	Shallow	Children	3	116	REML			
		Adults	4	186				
	Deep	Children	4	488				
		Adults	8	296				
		Tot.	19	1086				
Fluency (TL)	Shallow	Children	3	180	REML	-1.47	-1.50	-1.44
		Adults	2	130				
	Deep	Children	11	362				
		Adults	8	279		1.82	1.76	1.88
		Tot.	24	951				
Non-Lexical Decoding								
	Orthographic Depth	Age	Records (k)	Participants (N)	Model	Average ES	min.	max.
Accuracy	Shallow	Children	6	289	REML with moderators			
		Adults	9	392				
	Deep	Children	7	581				
		Adults	16	552				
	Regardless of Orthographic Depth	Children	13	870		1.97	1.94	2.01
		Adults	25	944		1.23	1.18	1.28
		Tot.	38	1814				
Fluency (TU)	Shallow	Children	5	228	REML			
		Adults	3	143				
	Deep	Children	4	485				
		Adults	8	306				
		Tot.	20	1202				
Fluency (TL)	Shallow	Children	4	204	REML with moderators	-1.40	-1.43	-1.37
		Adults	2	132				
	Deep	Children	8	263				
		Adults	7	245				

Table 2 (continued)

	Orthographic Depth	Age	Records (k)	Participants (N)	Model	Average ES	min.	max.
Phonological Manipulation		Tot.	21	844		2.01	1.95	2.08
Phonological Awareness					**Model**	**Average ES**	**min.**	**max.**
	Shallow	Children	4	195	REML			
		Adults	3	144				
	Deep	Children	13	836				
		Adults	17	556				
		Tot.	37	1731		1.19	1.16	1.21
Nonword repetition	Shallow	Children	6	372	REML			
		Adults	3	125				
	Deep	Children	8	292				
		Adults	5	163				
		Tot.	22	952		1.12	1.07	1.17
Other Cognitive Skills					**Model**	**Average ES**	**min.**	**max.**
RAN	Shallow	Children	2	99	REML			
		Adults	2	85				
	Deep	Children	8	570				
		Adults	15	490				
		Tot.	27	1244		-1.05	-1.11	-0.99
Short-Term/WM	Shallow	Children	5	222	REML			
		Adults	3	140				
	Deep	Children	8	503				
		Adults	14	431				
		Tot.	30	1296		0.84	0.79	0.88
Non-Verbal Reasoning	Shallow	Children	8	384	FE			
		Adults	9	380				
	Deep	Children	12	442				
		Adults	20	683				
		Tot.	49	1889		0.19	0.16	0.22

Table 3 Results of the REML with moderators run on word reading accuracy, non-lexical decoding accuracy and fluency (TL)

		estimate	se	z	p	CI. lb	CI. ub
Word Reading							
Accuracy		estimate	se	z	p	CI. lb	CI. ub
	Intercept	2.17	0.15	14.34	**< .001**	1.87	2.46
	Age	-1.19	0.20	-5.87	**< .001**	-1.58	-0.79
	Orthography	-0.57	0.24	-2.31	**0.02**	-1.05	-0.08
	Orthography Age*	0.72	0.32	2.24	**0.02**	0.09	1.35
Non-Lexical Decoding							
Accuracy		estimate	se	z	p	CI. lb	CI. ub
	Intercept	2.05	0.16	12.39	**< .001**	1.73	2.38
	Age	-0.77	0.20	-3.8	**< .001**	-1.18	-0.37
	Orthography	-0.03	0.25	-0.13	0.89	-0.52	0.45
	Orthography Age*	0.17	0.31	0.53	0.58	-0.44	0.79
Fluency (TL)	*Intercept*	2.69	0.25	10.61	**< .001**	2.19	3.18
	Age	-0.38	0.36	-1.06	0.28	-1.08	0.32
	Orthography	-0.66	0.41	-1.62	0.1	-1.47	013
	Orthography Age*	0.31	0.64	0.48	0.62	-0.94	1.56

Significant effects are highlighted in bold

32.11, $p = 0.09$, $I^2 = 32.23\%$, tau^2 = 0.07, SE = 0.07) and a significant between-group difference (Z=18.32, SE =0.1, $p < 0.001$; see Fig. 6, panel b, and Table 2 for more details).

Non-lexical decoding

Accuracy

The dataset initially included 40 observations, but after exploration of the boxplot, we identified 2 outliers (Asbjørnsen et al. (2014) and the group without a phonological deficit in Ruffino et al. (2014)) and removed them from the dataset. The final dataset included, thus, 38 observations.

The best-fit model included the moderators "age" and "orthography" (see the supplementary materials for more details on model selection).

The residual heterogeneity test was significant (QE$_{(34)}$ = 55.21, p = .01; I^2 = 38.9%, tau^2 = 0.07, SE = 0.05); an overall significant effect of the moderators was observed (QM$_{(3)}$ = 21.36, $p < 0.001$). Table 2 shows the stratification of the studies. In particular, we found a significant between-groups difference (Z = 12.39, SE = 0.16, $p < 0.001$), a main effect of age (Z = −3.8, SE = 0.20, $p < 0.001$), but no effects of orthography (Z = −0.13, SE = 0.25, p = 0.89) nor of interaction (Z = 0.53, SE = 0.31, p = 0.58) emerged (see Table 3 and Fig. 7). Accordingly, we observed an effect size of 1.97 (1.94–2.91) for children and 1.23 (1.18–1.28) for adults, regardless of orthography.

Fluency—time unlimited (TU)

We rated one study (Danelli et al., 2017) as an outlier through the boxplot exploration and removed it from the dataset. Moreover, 2 observations from the same study (Dole et al., 2014) were aggregated through the *agg* procedure, obtaining 20 observations from 19 studies. The REML without moderators revealed significant heterogeneity (Q$_{(19)}$=34.65, p = 0.01, I^2 = 45.1 %, tau^2 = 0.08, SE = 0.06) but, as reported in the supplementary materials, was also the best-fit model and revealed a significant between-group difference (Z= −15.76, SE =0.09, $p < 0.001$; see Fig. 8, panel a, and Table 2).
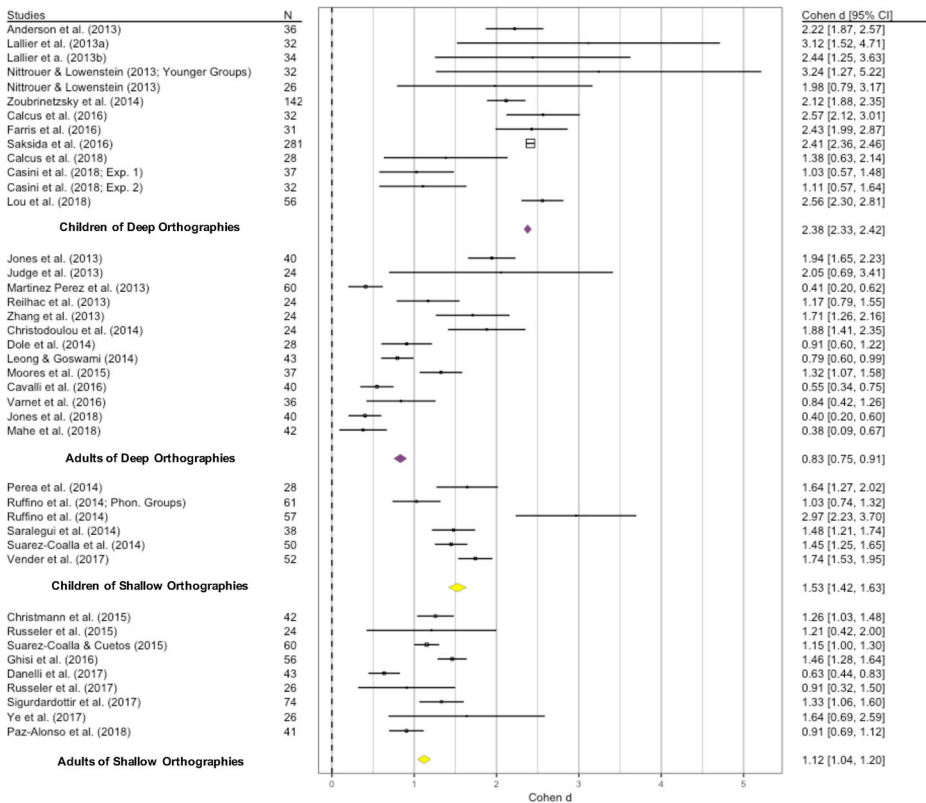
Fig. 5 Forest plot for the meta-analysis of the 41 word reading accuracy observations. The average effect size associated with each group (adults and children under deep and shallow orthographies) is represented by diamonds. The yellow diamonds indicate the average effect size for shallow orthographies; the purple ones refer to deep orthographies
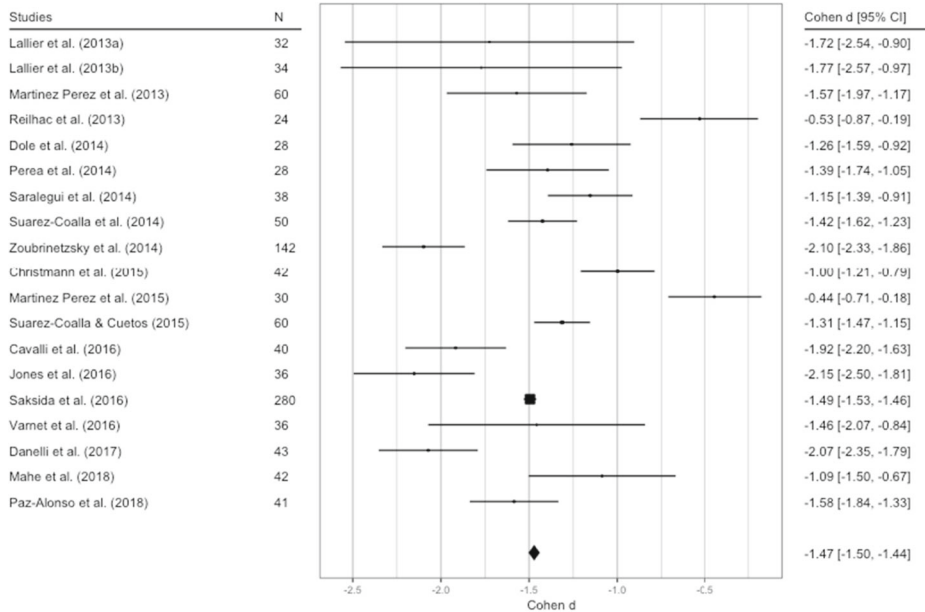
Fluency—time limited (TL)

The initial dataset had 21 observations from 20 studies (Halliday et al. (2014) reported data from a younger and an older group of children). The best-fit model included the two moderators (see supplementary materials) from which we obtained a significant residual heterogeneity ($QE_{(17)} = 40.38$, $p = 0.001$, $I^2 = 58.1\%$; $tau^2 = 0.25$, SE $= 0.15$) associated with a lack of effect of the two moderators ($QM_{(3)} = 3.75$, $p = 0.28$). Accordingly, only a significant difference between DG and CG emerged ($Z = 10.61$, SE $= 0.25$, $p < 0.001$). See Tables 1 and 2, and Fig. 8, panel b.
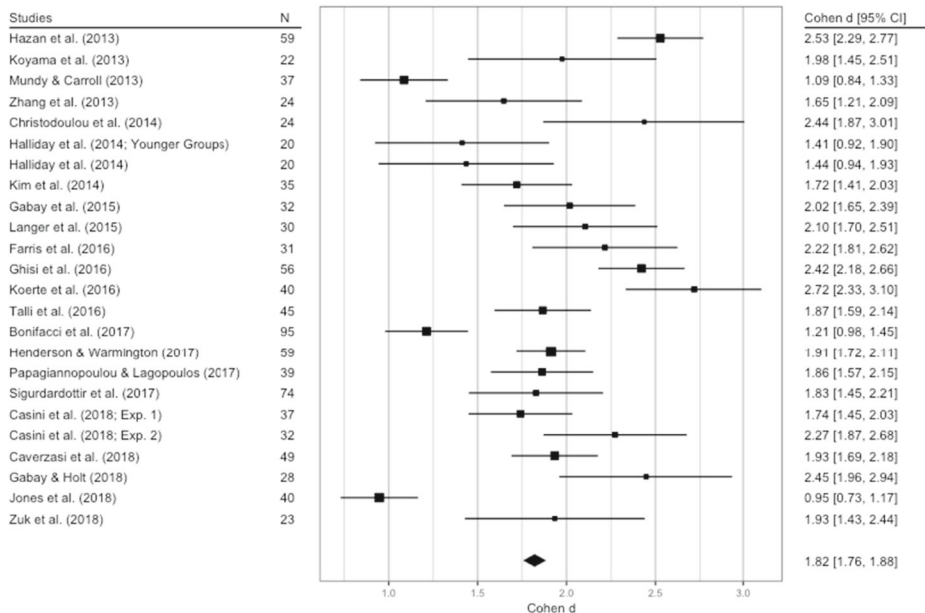
Phonological awareness

Phonological manipulation

We obtained 56 observations extracted from 36 studies (Casini et al. (2018), Ruffino et al. (2014), Altarelli et al., (2013, and Lallier et al., (2013c) reported results from two experiments in the same

a. Word Reading (TU)

b. Word Reading (TL)

Fig. 6 Forest plots of the word reading fluency meta-analyses. The average effect size for each fluency variable (either time limited (TL) or time unlimited (TU)) is represented by the black diamond. Panel **a**: effect size of studies wherein the recorded measure of fluency was TU. Panel **b**: effect size of studies where the recorded measure of fluency was TL, i.e., the number of correct words read in a limited time

work); after the agg procedure and the remotion of 3 outliers (Sumner et al., (2014); Anderson et al., (2013); Du & Kelly, (2013)), 37 observations from 33 studies remained.

Springer
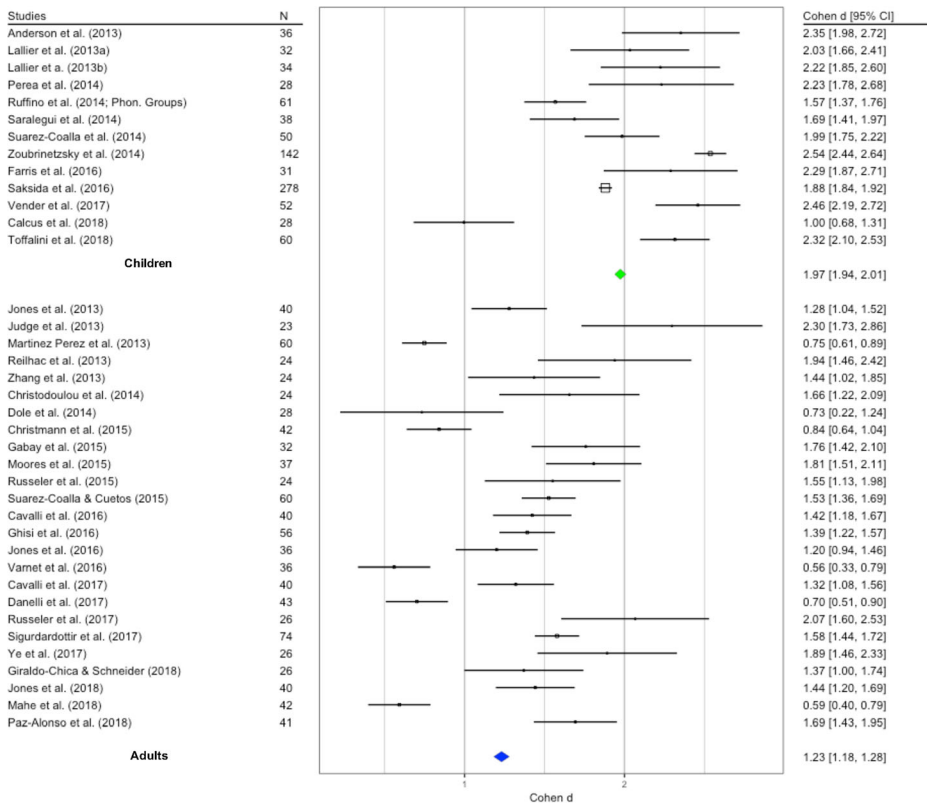
Fig. 7 Forest plot of the random effect model used on non-lexical decoding accuracy. The average effect size associated with each participant group (adults and children under deep and shallow orthographies) is represented by diamonds. The green diamond shows the average effect size for children; the blue one indicates adult readers
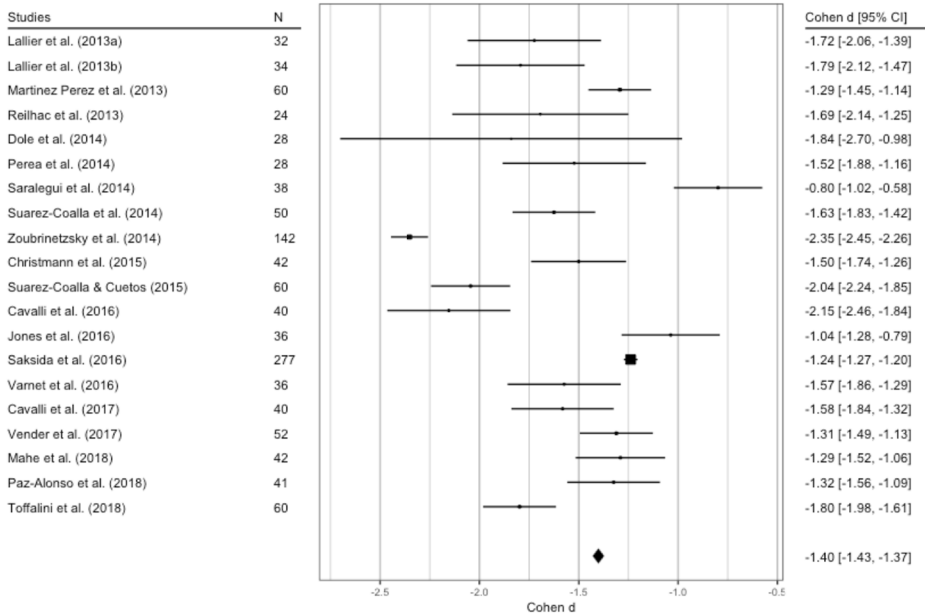
The best fitting model was the REML without moderators (see supplementary materials for more details). It reveals a nonsignificant level of heterogeneity ($Q_{(36)} = 42.29$, $p = 0.21$, $I^2 = 23.18\%$, $tau^2 = 0.03$; SE = 0.03) while a significant between-groups difference emerged ($Z = 16.40$, SE = 0.06, $p < 0.001$; see Fig. 9, panel a, and Table 2).
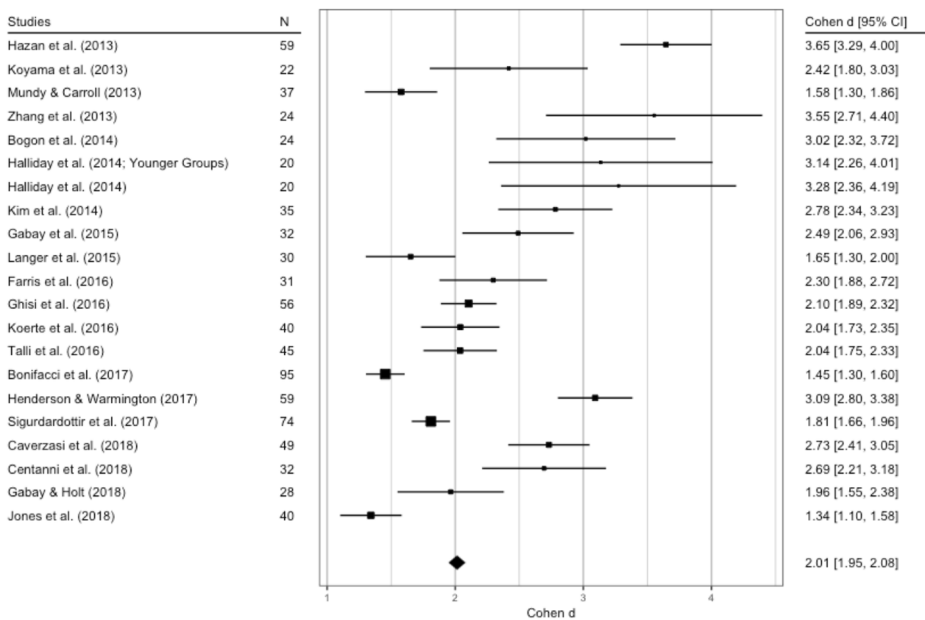
Nonword repetition

The initial dataset included 23 observations from 20 selected studies; Cantiani et al. (2015) particularly, reported 2 aggregated dependent measures, and the studies of Halliday et al. (2014) and Ruffino et al. (2014) reported data from two different samples. After the *agg* procedure, the resulting dataset included 22 observations.

The best fitting model was the REML without moderators (see supplementary materials for more details). It revealed a significant level of heterogeneity ($Q_{(21)} = 34.67$, $p = 0.03$, $I^2 = 39.7\%$, $tau^2 = 0.07$; SE = 0.06), and a significant between-groups difference ($Z = 12.94$, SE = 0.09, $p < 0.001$, see Fig. 9, panel b, and Table 2 for more details).

a. Non-Lexical Decoding (TU)

Studies	N	Cohen d [95% CI]
Lallier et al. (2013a)	32	-1.72 [-2.06, -1.39]
Lallier et al. (2013b)	34	-1.79 [-2.12, -1.47]
Martinez Perez et al. (2013)	60	-1.29 [-1.45, -1.14]
Reilhac et al. (2013)	24	-1.69 [-2.14, -1.25]
Dole et al. (2014)	28	-1.84 [-2.70, -0.98]
Perea et al. (2014)	28	-1.52 [-1.88, -1.16]
Saralegui et al. (2014)	38	-0.80 [-1.02, -0.58]
Suarez-Coalla et al. (2014)	50	-1.63 [-1.83, -1.42]
Zoubrinetzsky et al. (2014)	142	-2.35 [-2.45, -2.26]
Christmann et al. (2015)	42	-1.50 [-1.74, -1.26]
Suarez-Coalla & Cuetos (2015)	60	-2.04 [-2.24, -1.85]
Cavalli et al. (2016)	40	-2.15 [-2.46, -1.84]
Jones et al. (2016)	36	-1.04 [-1.28, -0.79]
Saksida et al. (2016)	277	-1.24 [-1.27, -1.20]
Varnet et al. (2016)	36	-1.57 [-1.86, -1.29]
Cavalli et al. (2017)	40	-1.58 [-1.84, -1.32]
Vender et al. (2017)	52	-1.31 [-1.49, -1.13]
Mahe et al. (2018)	42	-1.29 [-1.52, -1.06]
Paz-Alonso et al. (2018)	41	-1.32 [-1.56, -1.09]
Toffalini et al. (2018)	60	-1.80 [-1.98, -1.61]
		-1.40 [-1.43, -1.37]

Cohen d axis: -2.5, -2.0, -1.5, -1.0, -0.5

b. Non-Lexical Decoding (TL)

Studies	N	Cohen d [95% CI]
Hazan et al. (2013)	59	3.65 [3.29, 4.00]
Koyama et al. (2013)	22	2.42 [1.80, 3.03]
Mundy & Carroll (2013)	37	1.58 [1.30, 1.86]
Zhang et al. (2013)	24	3.55 [2.71, 4.40]
Bogon et al. (2014)	24	3.02 [2.32, 3.72]
Halliday et al. (2014; Younger Groups)	20	3.14 [2.26, 4.01]
Halliday et al. (2014)	20	3.28 [2.36, 4.19]
Kim et al. (2014)	35	2.78 [2.34, 3.23]
Gabay et al. (2015)	32	2.49 [2.06, 2.93]
Langer et al. (2015)	30	1.65 [1.30, 2.00]
Farris et al. (2016)	31	2.30 [1.88, 2.72]
Ghisi et al. (2016)	56	2.10 [1.89, 2.32]
Koerte et al. (2016)	40	2.04 [1.73, 2.35]
Talli et al. (2016)	45	2.04 [1.75, 2.33]
Bonifacci et al. (2017)	95	1.45 [1.30, 1.60]
Henderson & Warmington (2017)	59	3.09 [2.80, 3.38]
Sigurdardottir et al. (2017)	74	1.81 [1.66, 1.96]
Caverzasi et al. (2018)	49	2.73 [2.41, 3.05]
Centanni et al. (2018)	32	2.69 [2.21, 3.18]
Gabay & Holt (2018)	28	1.96 [1.55, 2.38]
Jones et al. (2018)	40	1.34 [1.10, 1.58]
		2.01 [1.95, 2.08]

Cohen d axis: 1, 2, 3, 4

Fig. 8 Forest plots of meta-analyses performed on non-lexical decoding fluency. The average effect size for each fluency variable (either time limited (TL) or time unlimited (TU)) is represented by the black diamond. Panel **a**: effect size of studies wherein the fluency measure was recorded as TU. Panel **b**: effect size of studies where the recorded measure of fluency was TL, i.e., the number of correct pseudo/nonwords read in a limited time. Notably, Halliday et al. (2014) reported results from different groups of dyslexics and controls

RAN

The initial dataset had 25 studies that reported 50 overall observations; after merging the effect sizes from the repeated measures, 27 data points extracted from 25 studies remained (Altarelli et al. (2013), Jones et al. (2013), and Lallier et al. (2013c) reported the data of 2 independent samples for both dyslexics and controls).

The REML without moderators showed the best fit to the data (see the supplementary materials) and revealed a nonsignificant heterogeneity ($Q_{(26)}$ = 31.21, p = 0.22; I^2 = 23.89%, tau^2 = 0.05, SE = 0.06). A significant between-group difference emerged (Z = −11.62, SE = 0.09, p < 0.001; see Fig. 10 panel a, and Table 2).

Short-term/WM

After aggregating 10 dependent measures, we obtained 30 observations from 25 studies (Casini et al. (2018), Altarelli et al. (2013), Jones et al. (2013), Lallier et al. (2013c), and Nittrouer et al. (2018) reported results from two different samples of children for both DG and CG).

The REML without moderators showed the best data fit (see the supplementary materials), and the heterogeneity level was nonsignificant ($Q_{(29)}$ = 32.27, p = 0.3, I^2 = 15.6%, tau^2 = 0.02, SE = 0.04). The DG and CG showed significantly different performances (Z = 11.6, SE = 0.07, p < 0.001; see Fig. 10, panel b, and Table 2).

Nonverbal reasoning

This dataset had 49 observations extracted from 47 papers (Lallier et al. (2013c) and Halliday et al. (2014) reported results from two different dyslexic and control samples).

The tau value of the first REML was null (tau^2 = 0; SE = 0.02); hence, we ran an FE model and obtained a better data fit (see the supplementary materials). The FE model did not show a nonsignificant level of heterogeneity ($Q_{(48)}$ = 48.79, p = 0.44) but showed a significant between-groups difference (Z = 4.46, SE = 0.04, p < 0.001; see Fig. 11 and Table 2).

Nonverbal reasoning and word reading

To understand whether systematic differences in nonverbal reasoning found in DG and CG would affect reading skills, we used a meta-regression approach. Effect sizes of nonverbal reasoning were tested as moderators on all indices of word reading (Accuracy, TU, TL).

In particular, 21 studies were included for word reading accuracy, and no effect of moderators was found ($QM_{(1)}$ = 2.44, p = 0.11), unless a significant degree of heterogeneity emerged by the REML ($QE_{(19)}$ = 41.1, p = 0.002, I^2 = 54.02%, tau^2=0.18, SE = 0.11).

A similar pattern of results emerged for both indices of word reading fluency (12 studies for TU, 15 studies for TL), in which no significant level of heterogeneity (TU: $QE_{(10)}$ = 16.51, p = 0.08, I^2= 41.11%, tau^2=0.1, SE = 0.11; TL: $QE_{(13)}$ = 20.42, p = 0.08, I^2= 28.63%, tau^2=0.07, SE = 0.09) and no significant effect of nonverbal reasoning (TU: $QM_{(1)}$ = 0.72, p = 0.39; TL: $QM_{(1)}$ = 2.98, p = 0.08) emerged.

a. Phonological Manipulation
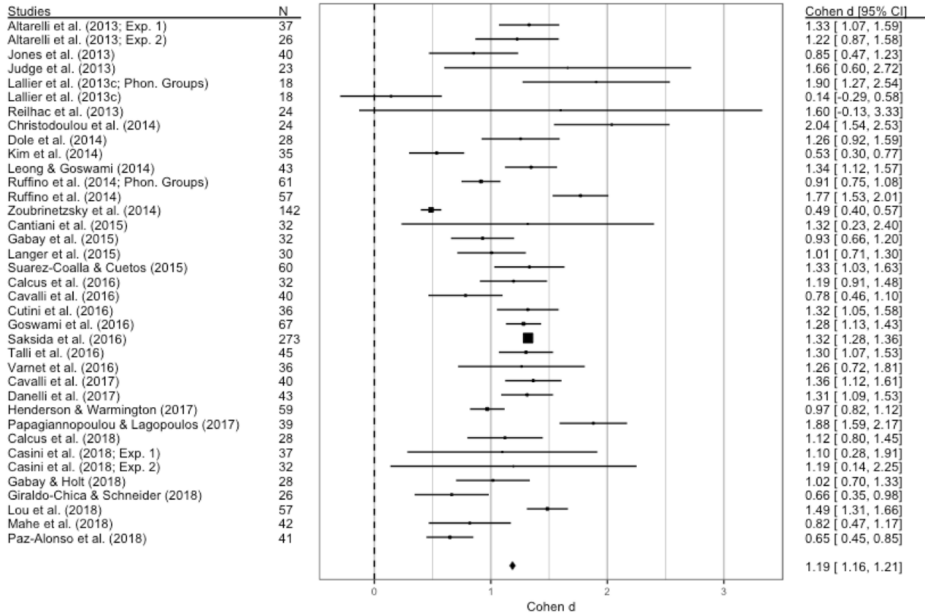
b. Nonword Repetition
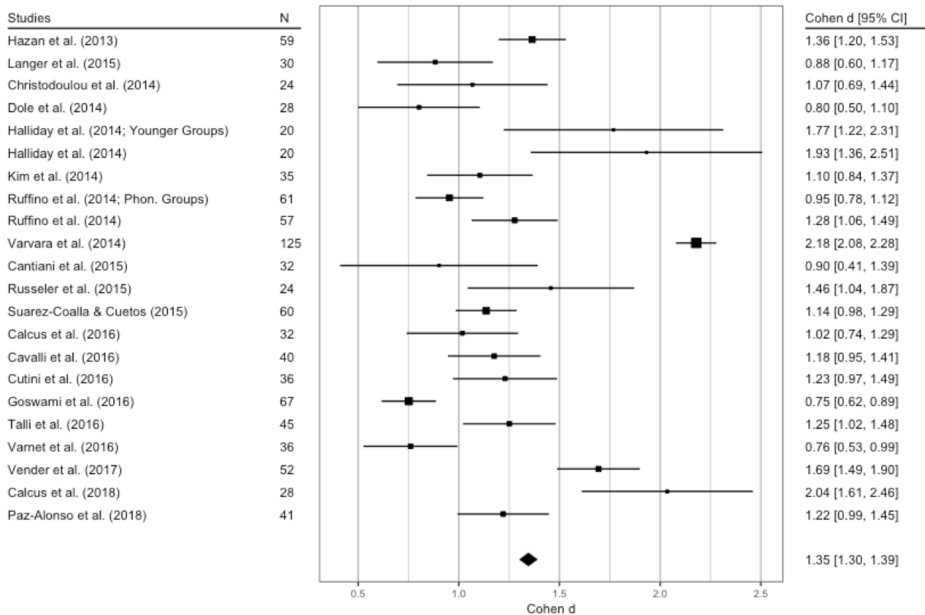
Fig. 9 Forest plots of meta-analyses conducted on phonological awareness. The average effect size for each model is represented by the black diamond. Panel **a**: effect size of the difference between typical and dyslexic readers in phonological manipulation tasks (such as phoneme deletion and elision, segmentation and synthesis, and spoonerisms). Panel **b**: effect size of the difference between typical and dyslexic readers in nonword repetition
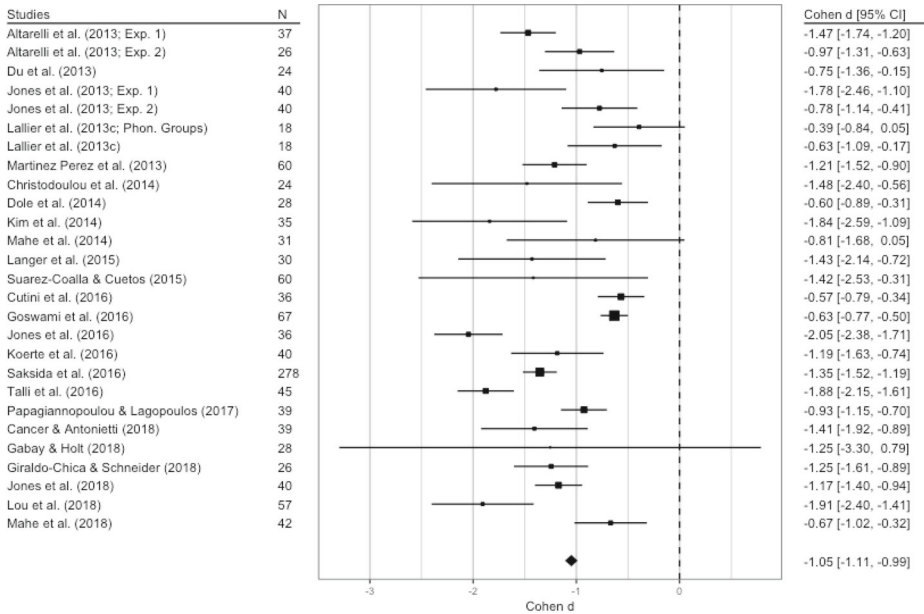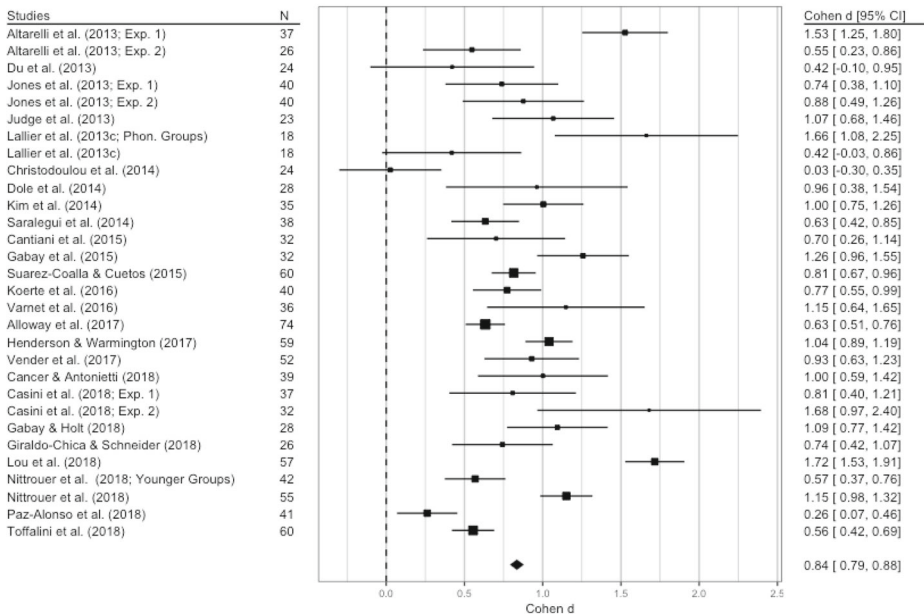
Fig. 10 Forest plots of meta-analyses performed on rapid automatized naming (panel **a**) and short-term working memory (panel **b**). The average effect size is indicated by the black diamond

Discussion

Dyslexia research has been prolific in the last 80 years; a search of the term "developmental dyslexia" in international repositories such as PubMed retrieves a list of 10,145 studies from 1946 to 2019 (as

Studies	N		Cohen d [95% CI]
Anderson et al. (2013)	36		0.61 [0.39, 0.84]
Du et al. (2013)	24		-0.51 [-0.85, -0.17]
Jones et al. (2013)	40		0.22 [0.02, 0.42]
Judge et al. (2013)	24		-0.78 [-1.13, -0.43]
Lallier et al. (2013c; Phon. Groups)	18		0.87 [0.40, 1.35]
Lallier et al. (2013c)	18		0.12 [-0.32, 0.55]
Martinez Perez et al. (2013)	60		0.17 [0.04, 0.30]
Mundy & Carroll (2013)	37		0.01 [-0.21, 0.22]
Palladino et al. (2013)	46		0.00 [-0.17, 0.17]
Bogon et al. (2014)	24		0.17 [-0.16, 0.49]
Christodoulou et al. (2014)	24		0.55 [0.21, 0.89]
Dole et al. (2014)	28		0.50 [0.21, 0.79]
Halliday et al. (2014; Younger Groups)	20		0.70 [0.28, 1.11]
Halliday et al. (2014)	20		0.48 [0.07, 0.88]
Huestegge et al. (2014)	42		0.00 [-0.19, 0.19]
Mahe et al. (2014)	31		0.15 [-0.11, 0.40]
Suarez-Coalla et al. (2014)	50		0.14 [-0.02, 0.30]
Biname et al. (2015)	52		0.16 [0.01, 0.31]
Cantiani et al. (2015)	32		0.58 [0.32, 0.84]
Christmann et al. (2015)	42		0.37 [0.18, 0.56]
Furnes & Norman (2015)	44		0.27 [0.09, 0.45]
Gabay et al. (2015)	32		0.40 [0.15, 0.65]
Langer et al. (2015)	30		0.83 [0.55, 1.12]
Martinez Perez et al. (2015)	30		0.32 [0.05, 0.58]
Russeler et al. (2015)	24		0.48 [0.15, 0.82]
Suarez-Coalla & Cuetos (2015)	60		-0.51 [-0.65, -0.38]
Calcus et al. (2016)	32		0.61 [0.36, 0.87]
Cavalli et al. (2016)	40		0.12 [-0.08, 0.32]
Cutini et al. (2016)	36		-0.54 [-0.77, -0.32]
Jones et al. (2016)	36		0.13 [-0.09, 0.35]
Palladino et al. (2016)	23		0.00 [-0.35, 0.35]
Varnet et al. (2016)	36		0.27 [0.05, 0.49]
Alloway et al. (2017)	74		0.13 [0.01, 0.25]
Bonifacci et al. (2017)	95		0.37 [0.24, 0.50]
Danelli et al. (2017)	43		0.11 [-0.08, 0.29]
Henderson & Warmington (2017)	59		0.14 [0.00, 0.27]
Russeler et al. (2017)	26		0.12 [-0.18, 0.42]
Sigurdardottir et al. (2017)	74		0.40 [0.29, 0.50]
Ye et al. (2017)	26		0.53 [0.22, 0.84]
Burani et al. (2018)	72		0.31 [0.20, 0.42]
Calcus et al. (2018)	28		0.24 [-0.05, 0.52]
Caverzasi et al. (2018)	58		0.27 [0.11, 0.42]
Centanni et al. (2018)	32		0.84 [0.56, 1.13]
Gabay & Holt (2018)	28		0.58 [0.29, 0.87]
Giraldo-Chica & Schneider (2018)	26		0.33 [0.02, 0.63]
Jones et al. (2018)	40		-0.45 [-0.66, -0.25]
Mahe et al. (2018)	42		0.48 [0.29, 0.68]
Paz-Alonso et al. (2018)	41		-0.18 [-0.37, 0.02]
Zuk et al. (2018)	24		0.68 [0.34, 1.03]
			0.19 [0.16, 0.22]

Fig. 11 Forest plot of the meta-analysis conducted for nonverbal reasoning. The average effect size is represented by the black diamond

of December 2019). With this meta-analysis, we summarized the most recent (2013–2018) behavioral findings on this topic considering both orthographic depth and a developmental perspective (operationalized as a dichotomous variable: children vs. adults). Of the many DD-associated behavioral features, we focused on two main aspects: behavioral performance in reading tasks and the main cognitive underpinnings of DD. At the cognitive level, we specifically explored PA, RAN, and short-term/WM, which are considered the main reading ability predictors (Landerl et al., 2013; Moll et al., 2014; Paulesu et al., 2001; Vaessen et al., 2010; Ziegler et al., 2010), and nonverbal reasoning, which is often used to assess the discrepancy criterion (Gustafson & Samuelsson, 1999; Goswami, 2001). Our aim was twofold: to identify DD's universal manifestations and to determine the features of DD that can be moderated by age, orthography, or an interaction of the two at the behavioral and cognitive levels.

We will first discuss the behavioral findings extracted by the reading tasks. Second, we will discuss the meta-analysis' results for the cognitive dimensions of interest and, third, conclude with a general discussion, the implications for clinical practice, limitations, and future directions.

Reading skills across ages and orthographies

The introduction stated that cognitive models of reading (Coltheart et al., 2001; Perry et al., 2010) and reading acquisition (Ehri, 1987, 1991, 1995, 2005, 2014; Ehri & McCormick, 1998;

Frith, 1985; Ziegler et al., 2020) are usually based on a dual-route: a sub-lexical route implying a grapheme-to-phoneme conversion (i.e., *phonological recoding*) and a lexical route based on orthographic whole-word analysis (i.e., lexical recognition). Accordingly, phonological recoding is now widely assumed to represent the first reading strategy to be developed; in a further step, the recurrent presentations of the same words would support the development of a whole-word lexical retrieval strategy (Ehri, 2005; Grainger & Ziegler, 2011; Ziegler & Goswami, 2006). Consequently, to address both the linguistic and developmental effects on reading skills while separating lexical processing from phonological recoding, we ran separate meta-analyses for words and for the process we defined as "non-lexical decoding" (nonword/ pseudoword reading), considering accuracy and fluency level (measured as either TL or TU).

Interestingly, the pattern of results was the same for the fluency indices from both word and pseudoword reading, but some intriguing differences were found for the accuracy indices. Indeed, word reading accuracy showed a systematic difference between dyslexic readers and age-matched controls that was moderated by an orthography-by-age interaction effect, while between-group differences at the fluency level were "universal," that is, not moderated by the orthographic depth and stable at the developmental level.

As stated previously, we expected between-group differences moderated by both age and orthographic depth to represent behavioral differences between dyslexics and controls that characterize both groups' developmental trajectories in a specific linguistic context. In particular, the significant orthography-by-age interaction effect found for word reading accuracy suggests that orthographic depth shapes the developmental trajectories of dyslexic and typical readers (Landerl et al., 1997; Ziegler et al., 2003). In deep orthographies, the average effect size between developmental dyslexic readers and age-matched controls starts from 2.4 to reach a mean Cohen's D of 0.8 in adulthood. Shallow orthographies show a smoother developmental trajectory slope (ranging from a mean Cohen's D of 1.5 in children to 1.1 in adulthood). These results suggest that while both typical and dyslexic readers in shallow orthographies showed a higher level of lexical decoding accuracy than children in deep orthographies, the language-specific effect in the behavioral manifestation of DD becomes less significant in adulthood. Indeed, what is more worthy to note here is that developmental trajectories are moderated by orthographic depth.

As already observed (Aro & Wimmer, 2003; Caravolas et al., 2013; Seymour et al., 2003), the acquisition of familiar-word reading seems more challenging for typical readers of deep orthographies than for their shallow counterparts. In this perspective, the gap reduction between DG and CG of deep orthographies from children to adulthood can be explained as a delayed maturation of reading efficacy by typical readers rather than as a sign of better compensation by dyslexic readers of shallow orthographies. In this sense, the differences observed in word reading accuracy between orthographies should be interpreted as a language-specific effect rather than as a feature of dyslexia.

A different pattern of results emerged for accuracy in phonological recoding, that is, non-lexical decoding. Here, the difference between deep and shallow orthographies did not emerge, while a significant difference between children and adult readers emerged: in children, the average effect size between age-matched controls and developmental dyslexics starts from 1.9 to reach1.2 in adulthood.

As Ziegler et al. (2003) suggested, a key finding of cross-linguistic studies in European languages is that children learning to read in a shallow orthography first rely on non-lexical decoding, whereas children in countries with a deep orthographic system, such as English-speaking countries, are forced to support grapheme-phoneme decoding with rhyming and

whole-word strategies because of the orthographic system's inconsistency. Nevertheless, our data suggest that DD would be characterized by an initial universal deficit in the development of phonological serial recoding during childhood, which would prevent the accurate reading of pseudowords. This deficit in phonological recoding would represent, according to the phonological theory (Bishop & Snowling, 2004; Bradley & Bryant, 1978; Démonet et al., 2004; Gabrieli, 2009; Peterson & Pennington, 2015; Vellutino et al., 2004; Snowling, 1981; Stanovich, 1988; Vellutino, 1979), a core deficit in DD, and, according to our results, is also maintained in adulthood, even though with a lesser degree of severity.

These results are in line with the idea of compensation in reading accuracy as a function of age and print exposure (Eloranta et al., 2019). In other words, reading fluency could represent a more stable marker of DD across ages than reading accuracy as, in this latter measure, a ceiling level can be more easily achieved in adulthood (Reis et al., 2020).

Indeed, more consistent findings were observed for reading speed of both word and pseudoword: no influence of orthography or age was detected for reading fluency expressed in terms of TL and TU.

According to our hypothesis, the between-group differences not moderated by age and orthography would represent the empirical manifestation of universal cognitive DD markers in line with Landerl et al. (1997) and Ziegler et al. (2003). This fluency deficit appears to similarly affect lexical decoding and phonological recoding and, according to our results, remains the most stable and reliable DD index in reading tasks. Moreover, our analysis of reading fluency needs a further methodological comment; namely, similar results emerged from the TL and TU measures of reading fluency, suggesting that the two methods are virtually identical and interchangeable. Nevertheless, the results on reading fluency, especially for models applied on time-limited measures, should be taken with *a grain of salt*: as shown in Table 2, for both word reading and Non-Lexical decoding, the number of studies in deep orthographies considerably exceed those in the shallow level. Further studies are needed to address whether this negative finding is replicated or if it represents a mere artifact owing to a lack of statistical power.

Notably, the meta-analytic approach allowed us to quantitatively support the role of reading fluency in assessing DD regardless of orthography. Beyond the above-mentioned methodological considerations, fluency measures suggest that children and adults with DD have a universal reading speed impairment. Interestingly, adults did not show any improvement in reading speed, contrary to reading accuracy, which suggests that while accuracy measures can capitalize on experience and years of reading practice, fluency measures (i.e., measures of the reading process' automation) remain impaired. Accordingly, once again, we can conclude that the automation deficit must be considered a more reliable marker of DD.

Cognitive predictors/markers of reading across orthographies

As described earlier, cross-linguistic studies have focused more on cognitive predictors/ markers[1] of reading skills than on reading per se. For example, Ziegler et al. (2010) tested

---

[1] With the expression predictor/marker, we want to refer to a relatively fine-grained clinical/methodological distinction. In statistical terms, the independent variables within a generalized linear model (GLM) can be referred to as "predictors." This term from the clinical and cognitive point of view may be, anyway, misleading. Here we adopted the following distinction: predictors are those cognitive skills measured before a complete reading mastering is achieved (in principle, before the third grade of school, at least) and that can predict the level of future reading performance. Markers are those cognitive skills that can distinguish typical from dyslexic readers once the learning process of reading should be achieved.

universal reading predictors for typical Finnish, Hungarian, Dutch, Portuguese, and French second-grade readers, whereas others (Landerl et al., 2013; Moll et al., 2014; Vaessen et al., 2010) focused on PA and RAN.

Ziegler et al. (2010) suggested that PA, in terms of phoneme deletion, was a direct predictor of reading accuracy and speed in all the tested orthographies except Finnish. The authors further argued that PA, as measured by phoneme deletion, would facilitate reading irrespective of a specific country's orthographic system, but the strength of the association between PA and reading would be moderated by the consistency of the grapheme-to-phoneme correspondence (Ziegler et al., 2010). This is consistent with our results on phonological manipulation (comprising phoneme deletion) and nonword repetition. Interestingly, although our meta-analytic approach did not allow for a direct test of the association between reading performance and phonological skills, we observed a universal PA deficit in dyslexic readers regardless of orthographic depth (consistent with Landerl et al., 2013). Nonetheless, since we observed no moderation effect of orthographic depth on PA task performance, we cannot completely support Ziegler's assumption (Ziegler et al., 2010).

Similarly, RAN was identified as a significant reading skill predictor in all orthographies (Ziegler et al., 2010) despite the smaller effect size of this relationship compared to the one between PA and reading skills (Araújo & Faísca, 2019; Araújo et al., 2015; Kirby et al., 2010; Norton & Wolf, 2012), suggesting that RAN is a significant reading predictor albeit secondary to PA. Landerl et al.'s (2013) cross-linguistic study compared the RAN of typical and dyslexic readers; RAN was a significant marker in all the orthographies, and this was the case for both alphanumeric and picture RAN. Similarly, our meta-analysis combined data from different types of RAN and replicated Landerl et al.'s (2013) results, yet a significant level of data heterogeneity remained. This could be because of the specific materials adopted in each study and the disproportionate number of studies on shallow versus deep orthographies. Nevertheless, Araújo et al.'s (2015) meta-analysis' results seem to support our findings and identify RAN as one of the best reading ability markers across orthographies.

Finally, notably, our results do not completely fit those of Vaessen et al. (2010). Vaessen et al. (2010) suggested that the relevance of PA and RAN in predicting reading skills would change across school grades, with PA becoming less relevant and RAN more significant over time. However, we had to exclude Hungarian, Dutch, and Portuguese studies from our meta-analyses because of their median-level orthographic depth (see Seymour's classification), and these were the exact orthographies investigated by Vaessen et al. (2010). Moreover, we could not further test this hypothesis because we pooled all the studies involving children without considering different grades.

All we can say is that adult dyslexic readers seem equally impaired in both PA and RAN as children with DD: as in the case of reading fluency, this suggests that the diagnosis of DD is supported at every age and in every orthographic system by poor phonological knowledge and automation deficit. Moreover, our results identified short-term/WM as a cognitive marker to distinguish between dyslexic readers and typical readers in both children and adults and all types of European orthographies. However, in line with Landerl et al. (2013), short-term/WM seems to play a minor role in the entire picture, as the average difference between DDs and typical readers was smaller than 1 SD and thus less relevant than the impairments observed for PA and RAN.

Nevertheless, since dyslexia is marked by a minor but stable disadvantage in short-term/WM regardless of age and orthography, we can also claim that poor performance in WM tasks characterizes the reading disability profile. Therefore, the impact of verbal WM on DD seems

overestimated, as verbal WM measures may inevitably be associated with phonological skills (see, for instance, tasks under "verbal working memory" in Peng et al.'s (2018) meta-analysis). WM and PA are considered both concurrent and combined markers of reading ability, at least in children (Knoop-van Campen et al., 2018) and represent two important components of tasks like nonword repetition and phonological manipulation. Accordingly, the conflicting results in the literature about the role of WM in DD (Gray et al., 2019; Maehler & Schuchardt, 2016; Menghini et al., 2011) could be explained by the need to separate the phonological component and consider the role of WM as is.

To conclude, our meta-analytic work seems to support both the double deficit hypothesis (Wolf & Bowers, 1999) and, more in general, to the phonological theory (Démonet et al., 2004; Vellutino et al., 2004; Bishop & Snowling, 2004).

Moreover, we included nonverbal reasoning measures in our meta-analytic procedure for two main reasons: to have a measure of intelligence (as this is a key feature of DD diagnosis) and to assess the reliability and consistency of the visuoperceptual deficit often identified as a possible DD marker (Skottun, 2000; Stein et al., 2000) across studies. Interestingly, our results suggest that while the average gap between DD and CG was 0.19 SD, a value that indicates normal performance in the nonverbal reasoning tests, we found a systematic between-group difference. This means that even within the "normal" performance range, from a psychological viewpoint, some DD-related factors as visuo-perceptual or working memory weaknesses can negatively affect the performance of dyslexic readers in reasoning tasks, through which a "lack of normality" emerges from a statistical perspective (Capitani, 1997). This should be considered when examining the "discrepancy" criterion. Nevertheless, the discrepancy criterion, when tested through a meta-regression, emerged as relevant since differences in nonverbal IQ scores of DG and CG had no significant influence in predicting their word reading outcomes (in both accuracy and fluency indices). Nonverbal reasoning, in these terms, cannot account for the word reading deficit observed for DD readers, further suggesting is independence from more general reasoning skills.

Finally, in terms of developmental level, all cognitive predictors of DD remain impaired across the entire life span regardless of orthographic level; here, PA, automation, and short-term/WM (even if with lesser strength) are all concurrent predictors of an efficient reading system.

General discussion and some clinical remarks

This meta-analytic study attempted to summarize the most recent literature on DD from a cross-linguistic perspective and compared impaired and non-impaired children and adult readers. As a result, we obtained some relevant cross-linguistic and developmental findings useful when approaching DD from both a research and a clinical standpoint.

First, our results suggest that fluency must be considered the most relevant parameter for DD diagnosis. The disadvantage in reading speed, in lexical recognition and phonological recoding, comprises the universal manifestation of reading deficits irrespective of age and orthographic depth. Thus, the adoption of time-limited approaches in reading tasks does not provide either inconsistent or less robust evidence. As the reading process is tedious for those with DD, a good compromise would be to adopt time-limited reading tasks to avoid upsetting dyslexic readers with long and difficult reading tasks with questionable reliability and clinical validity. Nevertheless, consistent with Sprenger-Charolles et al. (2011), our results warn against adopting a solely accuracy-based reading skill evaluation. Although accuracy is clearly

an important parameter, especially when assessing cross-linguistic differences in reading skills, it can easily peak because of orthographic transparency or deficit compensation (at least for this parameter) in adulthood.

Second, PA, RAN, and reading fluency are the most reliable DD cognitive markers. As suggested by numerous studies, phonology is a relevant component of the reading process associated with impairment in dyslexic readers irrespective of age and orthographic depth; consequently, a PA assessment should always be included in clinical evaluations. This suggestion reveals a relevant issue, namely, the significant lack of empirical studies that included in their cognitive batteries standardized PA measures, particularly for what concerns shallow orthographies. Moreover, regarding PA, RAN should be mandatory in clinical evaluations even though the high level of heterogeneity in our meta-analysis suggests the need to identify the optimal methodological procedures and materials for evaluating this crucial cognitive dimension.

For some short-term/WM tasks such as nonword repetition (e.g. the task of Gathercole and collagues (1994), widely used in english countries), PA is also an underpinning main component. Accordingly, to avoid the risk of systematically attributing a deficit to the wrong cognitive component and to develop a comprehensive idea of verbal short term memory ability, clinical practice should adopt digit—rather than nonword—repetition when assessing short-term memory.

Finally, the results recommend trusting the discrepancy criterion. Although we found a systematic disadvantage in nonverbal reasoning among people with DD, the pooled effect was minimal (0.19). Such a small between-group difference might depend on DD-related weakness in visuoperceptual and WM skills, rather than on a more general cognitive disadvantage. In other words, dyslexic readers' nonverbal reasoning was nevertheless in the normal range irrespectively by age. Therefore, nonverbal reasoning skills, as supported by the results of our meta-regressions, remains a valid criterion to distinguish DD from more generic intellectual impairments.

In conclusion, the results of our meta-analysis further support cross-linguistic studies on lexical reading and phonological recoding. Despite a fluency deficit in DD across European orthographies, the accuracy measures' impairment seems strongly linked to orthographic depth. Conversely, cognitive markers of reading ability, such as PA, RAN, and short-term/WM, confirm their roles in both children and adults regardless of orthography and thus represent the universal core deficits of DD.

These results offer valuable information to researchers and clinicians. On the one hand, the results provide new research perspectives regarding possible DD treatments and the development of new diagnostic tools. On the other hand, the above-mentioned clinical considerations could also be adopted in situations where differentiating DD from other reading deficits compounded by external environmental issues can be challenging, as often observed in children with multilingual backgrounds. Indeed, identifying and diagnosing DD in these specific multilingual profiles should be based primarily on the presence of a reading fluency deficit along with poor PA and RAN performance.

Limitations and future directions

To conclude, we would like to capitalize on some limitations of this meta-analytic study to suggest some future directions for the research agenda on DD. First, some of our results must be considered with caution because of studies' lopsidedness across ages for deep versus shallow orthographies, particularly in the case of time-limited reading fluency and RAN measures.

Although limited to studies published between 2013 and 2018, the present study was nevertheless useful in summarizing the status quo of literature on reading and DD and in highlighting some crucial points that, from a clinical perspective, represent fundamental measures to be investigated for DD diagnosis. However, more research would help clarify how orthographic knowledge shapes phonological recoding and, through an empirical comparison, whether adult readers in deep orthographies possess better word reading accuracy than readers in shallow orthographies.

We limited our analyses to an orthographic depth level evaluation without considering the complexity, the granularity, and the entropy level considered by other classifications and approaches (e.g., Grain Size Theory; Ziegler & Goswami, 2005). In addition, future meta-analyses could go beyond our dichotomous categorization (shallow vs. deep) by including studies that cover all five levels of Seymour's classification to provide more specific differences between orthographies.

Finally, a higher level of between-studies consistency should be reached when reporting WM-measures' results. Although short-term/WM is widely considered a separate process, we were restricted from making this distinction herein because the studies often reported forward and backward digit span indices as composite scores.

Nonetheless, through a meta-analytic approach, we attempted to address the difficulties in conducting a cross-linguistic comparison of a wide range of cognitive abilities along the life span. This meta-analysis afforded us a direct comparison between children and adults to describe the developmental trajectories of reading and DD in shallow and deep European orthographies.

References

Studies included in the meta-analyses are marked with an asterisk

*Alloway, T. P., Tewolde, F., Skipper, D., & Hijar, D. (2017). Can you spell dyslexia without SLI? Comparing the cognitive profiles of dyslexia and specific language impairment and their roles in learning. *Research in Developmental Disabilities*, *65*, 97–102. https://doi.org/10.1016/j.ridd.2017.04.013

*Altarelli, I., Monzalvo, K., Iannuzzi, S., Fluss, J., Billard, C., Ramus, F., & Dehaene-Lambertz, G. (2013). A functionally guided approach to the morphometry of occipitotemporal regions in developmental dyslexia: evidence for differential effects in boys and girls. *Journal of Neuroscience*, 33(27), 11296-11301. https://doi.org/10.1523/JNEUROSCI.5854-12.2013

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub.

*Anderson, A., Lin, C. Y., & Wang, M. (2013). Native and novel language prosodic sensitivity in English-speaking children with and without dyslexia. *Dyslexia*, *19*(2), 92–112. https://doi.org/10.1002/dys.1451

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

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.

Arnold, T. W. (2010). Uninformative parameters and model selection using Akaike's Information Criterion. *Journal of Wildlife Management, 74*(6), 1175–1178. https://doi.org/10.1111/j.1937-2817.2010.tb01236.x.

Aro, M. (2004). Learning to read: The effect of orthography. *Jyväskylän Yliopisto. ISSN*, 951-39-1722-3.

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics, 24*(4), 621–635. https://doi.org/10.1017/S0142716403000316.

*Asbjørnsen, A. E., Obrzut, J. E., & Oyler, J. D. (2014). A cross-cultural comparison of verbal learning and memory functions in reading disabled American and Norwegian adolescents. *Scandinavian Journal of Psychology*, *55*(2), 115–122. https://doi.org/10.1111/sjop.12108

Banai, K., & Ahissar, M. (2004). Poor frequency discrimination probes dyslexics with particularly impaired working memory. *Audiology Neurootology, 9*(6), 328–340. https://doi.org/10.1159/000081282.

Banai, K., & Ahissar, M. (2010). On the importance of anchoring and the consequences of its impairment in dyslexia. *Dyslexia, 16*(3), 240–257. https://doi.org/10.1002/dys.407.

*Binamé, F., Danzio, S., & Poncelet, M. (2015). Relative ease in creating detailed orthographic representations contrasted with severe difficulties to maintain them in long-term memory among dyslexic children. *Dyslexia*, *21*(4), 361–370. https://doi.org/10.1002/dys.1506

Bishop, D. V., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: same or different? *Psychological Bulletin, 130*(6), 858–886. https://doi.org/10.1037/0033-2909.130.6.858.

*Bogon, J., Finke, K., Schulte-Körne, G., Müller, H. J., Schneider, W. X., & Stenneken, P. (2014). Parameter-based assessment of disturbed and intact components of visual attention in children with developmental dyslexia. *Developmental Science, 17*(5), 697–713. https://doi.org/10.1111/desc.12150

*Bonifacci, P., Canducci, E., Gravagna, G., & Palladino, P. (2017). English as a foreign language in bilingual language-minority children, children with dyslexia and monolingual typical readers. *Dyslexia*, *23*(2), 181–206. https://doi.org/10.1002/dys.1553

Borleffs, E., Maassen, B. A. M., Lyytinen, H., & Zwarts, F. (2018). Cracking the code: The impact of orthographic transparency and morphological-syllabic complexity on reading and developmental dyslexia. *Frontiers in Psychology, 9*, 2534. https://doi.org/10.3389/fpsyg.2018.02534.

Bowey, J. A., McGuigan, M., & Ruschena, A. (2005). On the association between serial naming speed for letters and digits and word-reading skill: Towards a developmental account. *Journal of Research in Reading, 28*(4), 400–422. https://doi.org/10.1111/j.1467-9817.2005.00278.x.

Bradley, L., & Bryant, P. E. (1978). Difficulties in auditory organisation as a possible cause of reading backwardness. *Nature*, 271(5647), 746-747. doi :https://doi.org/10.1038/271746a0

Brislin, R. W. (1976). Comparative research methodology: Cross-cultural studies. *International Journal of Psychology, 11*(3), 215–229. https://doi.org/10.1080/00207597608247359.

Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental psychology, 28*(5), 874–886. https://doi.org/10.1037/0012-1649.28.5.874.

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

Burani, C., Barca, L., & Ellis, A. W. (2006). Orthographic complexity and word naming in Italian: Some words are more transparent than others. *Psychonomic Bulletin and Review, 13*(2), 346–352. https://doi.org/10.3758/bf03193855.

*Burani, C., Marcolini, S., Traficante, D., & Zoccolotti, P. (2018). Reading derived words by Italian children with and without dyslexia: The effect of root length. *Frontiers in Psychology*, *9*, 647. https://doi.org/10.3389/fpsyg.2018.00647

*Calcus, A., Lorenzi, C., Collet, G., Colin, C., & Kolinsky, R. (2016). Is there a relationship between speech identification in noise and categorical perception in children with dyslexia? *Journal of Speech, Language, and Hearing Research*, *59*(4), 835–852. https://doi.org/10.1044/2016_JSLHR-H-15-0076

*Calcus, A., Deltenre, P., Colin, C., & Kolinsky, R. (2018). Peripheral and central contribution to the difficulty of speech in noise perception in dyslexic children. *Developmental Science*, *21*(3), e12558. https://doi.org/10.1111/desc.12558

*Cancer, A., & Antonietti, A. (2018). Rapid automatized naming, verbal working memory, and rhythm discrimination as predictors of reading in Italian undergraduate students with and without dyslexia. *Brain Sciences*, 8(5). https://doi.org/10.3390/brainsci8050087

*Cantiani, C., Lorusso, M. L., Perego, P., Molteni, M., & Guasti, M. T. (2015). Developmental dyslexia with and without language impairment: ERPs reveal qualitative differences in morphosyntactic processing. *Developmental Neuropsychology*, 40(5), 291–312. https://doi.org/10.1080/87565641.2015.1072536

Capitani, E. (1997). Normative data and neuropsychological assessment. Common problems in clinical practice and research. *Neuropsychological Rehabilitation, 7*(4), 295–310. https://doi.org/10.1080/713755543.

Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science, 24*(8), 1398–1407. https://doi.org/10.1177/0956797612473122.

*Casini, L., Pech-Georgel, C., & Ziegler, J. C. (2018). It's about time: Revisiting temporal processing deficits in dyslexia. *Developmental Science*, 21(2). https://doi.org/10.1111/desc.12530

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition, 91*(1), 77–111. https://doi.org/10.1016/s0010-0277(03)00164-1.

Castles, A., Holmes, V. M., Neath, J., & Kinoshita, S. (2003). How does orthographic knowledge influence performance on phonological awareness tasks? *Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology, 56*(3), 445–467. https://doi.org/10.1080/02724980244000486.

Cattell, R. B. (1940). A culture-free intelligence test. *I. Journal of Educational Psychology, 31*(3), 161–179. https://doi.org/10.1037/h0059043.

*Cavalli, E., Casalis, S., El Ahmadi, A., Zira, M., Poracchia-George, F., & Colé, P. (2016). Vocabulary skills are well developed in university students with dyslexia: Evidence from multiple case studies. *Research in Developmental Disabilities*, 51–52, , 89–102. https://doi.org/10.1016/j.ridd.2016.01.006

*Cavalli, E., Colé, P., Pattamadilok, C., Badier, J. M., Zielinski, C., Chanoine, V., & Ziegler, J. C. (2017). Spatiotemporal reorganization of the reading network in adult dyslexia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 92, 204–221. https://doi.org/10.1016/j.cortex.2017.04.012

*Caverzasi, E., Mandelli, M. L., Hoeft, F., Watson, C., Meyer, M., Allen, I. E., Papinutto N., Wang C., Gandini Wheeler-Kingshott C. A.M., Marco E. J., Mukherjee P., Miller Z. A., Miller B. L., Hendren R., Shapiro K. A. Gorno-Tempini, M. L. (2018). Abnormal age-related cortical folding and neurite morphology in children with developmental dyslexia. *NeuroImage. Clinical*, 18, 814–821. https://doi.org/10.1016/j.nicl.2018.03.012

*Centanni, T. M., Pantazis, D., Truong, D. T., Gruen, J. R., Gabrieli, J. D., & Hogan, T. P. (2018). Increased variability of stimulus-driven cortical responses is associated with genetic variability in children with and without dyslexia. *Developmental cognitive neuroscience*, 34, 7-17. https://doi.org/10.1016/j.dcn.2018.05.008

*Christmann, C. A., Lachmann, T., & Steinbrink, C. (2015). Evidence for a general auditory processing deficit in developmental dyslexia from a discrimination paradigm using speech versus nonspeech sounds matched in complexity. *Journal of Speech, Language, and Hearing Research*, 58(1), 107–121. https://doi.org/10.1044/2014_JSLHR-L-14-0174

*Christodoulou, J. A., Del Tufo, S. N., Lymberis, J., Saxler, P. K., Ghosh, S. S., Triantafyllou, C., ... & Gabrieli, J. D. (2014). Brain bases of reading fluency in typical reading and impaired fluency in dyslexia. *PLOS ONE*, 9(7), e100552. https://doi.org/10.1371/journal.pone.0100552

Clarke, P., Hulme, C., & Snowling, M. (2005). Individual differences in RAN and reading: A response timing analysis. *Journal of Research in Reading, 28*(2), 73–86. https://doi.org/10.1111/j.1467-9817.2005.00255.x.

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

Cuetos, F., Rodríguez, B., Ruano, E., & Arribas, D. (2007). *Prolec-r. Evaluación de los procesos lectores–revisado*. Madrid: Texas Education Agency.

*Cutini, S., Szűcs, D., Mead, N., Huss, M., & Goswami, U. (2016). Atypical right hemisphere response to slow temporal modulations in children with developmental dyslexia. *Neuroimage*, 143, 40–49. https://doi.org/10.1016/j.neuroimage.2016.08.012

*Danelli, L., Berlingeri, M., Bottini, G., Borghese, N. A., Lucchese, M., Sberna, M., Price C. J. Paulesu, E. (2017). How many deficits in the same dyslexic brains? A behavioral and fMRI assessment of comorbidity in adult dyslexics. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 97, 125–142. https://doi.org/10.1016/j.cortex.2017.08.038

Del Re, A., & Hoyt, W. T. (2014). MAd: Meta-analysis with mean differences. *R package version 0.8–2*. URL http://cran.r-project.org/web/packages/MAd.

Démonet, J. F., Taylor, M. J., & Chaix, Y. (2004). Developmental dyslexia. *Lancet, 363*(9419), 1451–1460. https://doi.org/10.1016/S0140-6736(04)16106-0.

Devoto, F., Carioti, D., Danelli, L., & Berlingeri, M. (under review). A meta-analysis of functional neuroimaging studies on developmental dyslexia across European orthographies: the ADOD model. *Language, Cognition and Neuroscience.*

*Dole, M., Meunier, F., & Hoen, M. (2014). Functional correlates of the speech-in-noise perception impairment in dyslexia: An MRI study. *Neuropsychologia, 60*, 103–114. https://doi.org/10.1016/j.neuropsychologia.2014.05.016

*Du, W., & Kelly, S. W. (2013). Implicit sequence learning in dyslexia: A within-sequence comparison of first- and higher-order information. *Annals of Dyslexia, 63*(2), 154–170. https://doi.org/10.1007/s11881-012-0077-1

Ehri, L. C. (1987). Learning to read and spell words. *Journal of Reading Behavior, 19*(1), 5–31. https://doi.org/10.1080/10862968709547585.

Ehri, L. C. (1991). Development of the ability to read words. *Handbook of Reading Research, 2*, 383–417.

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18*(2), 116–125. https://doi.org/10.1111/j.1467-9817.1995.tb00077.x.

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. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading, 18*(1), 5–21. https://doi.org/10.1080/10888438.2013.819356.

Ehri, L. C., & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading and Writing Quarterly, 14*(2), 135–163. https://doi.org/10.1080/1057356980140202.

Elbro, C., Nielsen, I., & Petersen, D. K. (1994). Dyslexia in adults: Evidence for deficits in non-word reading and in the phonological representation of lexical items. *Annals of Dyslexia, 44*(1), 203–226. https://doi.org/10.1007/BF02648162.

Ellis, N. C., & Hooper, A. M. (2001). Why learning to read is easier in Welsh than in English: Orthographic transparency effects evinced with frequency-matched tests. *Applied Psycholinguistics, 22*(4), 571–599. https://doi.org/10.1017/S0142716401004052.

Ellis, N. C., Natsume, M., Stavropoulou, K., Hoxhallari, L., Van Daal, V. H. P., Polyzoe, N., et al. (2004). The effects of orthographic depth on learning to read alphabetic, syllabic, and logographic scripts. *Reading Research Quarterly, 39*(4), 438–468. https://doi.org/10.1598/RRQ.39.4.5.

Eloranta, A. K., Närhi, V. M., Eklund, K. M., Ahonen, T. P., & Aro, T. I. (2019). Resolving reading disability— Childhood predictors and adult-age outcomes. *Dyslexia, 25*(1), 20–37. https://doi.org/10.1002/dys.1605.

*Farris, E. A., Ring, J., Black, J., Lyon, G. R., & Odegard, T. N. (2016). Predicting growth in word level reading skills in children with developmental dyslexia using an object rhyming functional neuroimaging task. *Developmental Neuropsychology, 41*(3), 145–161. https://doi.org/10.1080/87565641.2016.1158264

Fawcett, A., & Nicolson, R. (2004). *The dyslexia screening test: Junior (DST-J)*. Harcourt Assessment.

Fostick, L., & Revah, H. (2018). Dyslexia as a multi-deficit disorder: Working memory and auditory temporal processing. *Acta psychologica, 183*, 19–28. https://doi.org/10.1016/j.actpsy.2017.12.010.

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

Frith, U. (1985). Beneath the surface of developmental dyslexia. *Surface Dyslexia, 32*(1), 301–330.

Frith, U., Wimmer, H., & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading, 2*(1), 31–54. https://doi.org/10.1207/s1532799xssr0201_2.

*Furnes, B., & Norman, E. (2015). Metacognition and reading: Comparing three forms of metacognition in normally developing readers and readers with dyslexia. *Dyslexia, 21*(3), 273-284. https://doi.org/10.1002/dys.1501

*Gabay, Y., & Holt, L. L. (2018). Short-term adaptation to sound statistics is unimpaired in developmental dyslexia. *PLOS ONE, 13*(6), e0198146. https://doi.org/10.1371/journal.pone.0198146

*Gabay, Y., Thiessen, E. D., & Holt, L. L. (2015). Impaired statistical learning in developmental dyslexia. *Journal of Speech, Language, and Hearing Research, 58*(3), 934–945. https://doi.org/10.1044/2015_JSLHR-L-14-0324

Gabrieli, J. D. (2009). Dyslexia: a new synergy between education and cognitive neuroscience. *Science, 325*(5938), 280–283. https://doi.org/10.1126/science.1171999.

Georgiou, G. K., Parrila, R., Kirby, J. R., & Stephenson, K. (2008). Rapid naming components and their relationship with phonological awareness, orthographic knowledge, speed of processing, and different reading outcomes. *Scientific Studies of Reading, 12*(4), 325–350. https://doi.org/10.1080/10888430802378518.

*Ghisi, M., Bottesi, G., Re, A. M., Cerea, S., & Mammarella, I. C. (2016). Socioemotional features and resilience in Italian university students with and without dyslexia. *Frontiers in psychology, 7*, 478. https://doi.org/10.3389/fpsyg.2016.00478

*Giraldo-Chica, M., & Schneider, K. A. (2018). Hemispheric asymmetries in the orientation and location of the lateral geniculate nucleus in dyslexia. *Dyslexia, 24*(2), 197–203. https://doi.org/10.1002/dys.1580

Gola-Asmussen, C., Lequette, C., Pouget, G., Rouyer, C., & Zorman, M. (2010). *ECLA-16+: Évaluation des compétences de lecture chez l'adulte de plus de 16 ans*. France: Grenoble: *CogniSciences*.

Goswami, U. (2001). Developmental dyslexia. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (Vol. 11). Elsevier.

*Goswami, U., Cumming, R., Chait, M., Huss, M., Mead, N., Wilson, A. M., Barnes L. Fosker, T. (2016). Perception of filtered speech by children with developmental dyslexia and children with specific language impairments. *Frontiers in Psychology*, 7, 791. https://doi.org/10.3389/fpsyg.2016.00791

Gottardo, A., Pasquarella, A., Chen, X., & Ramirez, G. (2016). The impact of language on the relationships between phonological awareness and word reading in different orthographies: A test of the psycholinguistic grain size theory in bilinguals. *Applied Psycholinguistics, 37*(5), 1083–1115. https://doi.org/10.1017/S0142716415000508.

Grainger, J., & Ziegler, J. C. (2011). A dual-route approach to orthographic processing. *Frontiers in Psychology, 2*, 54. https://doi.org/10.3389/fpsyg.2011.00054.

Gray, S., Fox, A. B., Green, S., Alt, M., Hogan, T. P., Petscher, Y., & Cowan, N. (2019). Working memory profiles of children with dyslexia, developmental language disorder, or both. *Journal of Speech, Language, and Hearing Research, 62*(6), 1839–1858. https://doi.org/10.1044/2019_JSLHR-L-18-0148.

Griffiths, S., & Frith, U. (2002). Evidence for an articulatory awareness deficit in adult dyslexics. *Dyslexia, 8*(1), 14–21. https://doi.org/10.1002/dys.201.

Gustafson, S., & Samuelsson, S. (1999). Intelligence and dyslexia: Implications for diagnosis and intervention. *Scandinavian Journal of Psychology, 40*(2), 127–134. https://doi.org/10.1111/1467-9450.00109.

*Halliday, L. F., Barry, J. G., Hardiman, M. J., & Bishop, D. V. (2014). Late, not early mismatch responses to changes in frequency are reduced or deviant in children with dyslexia: An event-related potential study. *Journal of Neurodevelopmental Disorders*, 6(1), 21. https://doi.org/10.1186/1866-1955-6-21

*Hazan, V., Messaoud-Galusi, S., & Rosen, S. (2013). The effect of talker and intonation variability on speech perception in noise in children with dyslexia. *Journal of Speech, Language, and Hearing Research*, 56(1), 44–62. https://doi.org/10.1044/1092-4388(2012/10-0107)

He, J., & Van de Vijver, F. J. (2013). Methodological issues in cross-cultural studies in educational psychology. *Advancing cross-cultural perspectives on educational psychology: A festschrift for Dennis McInerney*, 39–56.

*Henderson, L. M., & Warmington, M. (2017). A sequence learning impairment in dyslexia? It depends on the task. *Research in Developmental Disabilities*, 60, 198–210. https://doi.org/10.1016/j.ridd.2016.11.002

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557.

*Huestegge, L., Rohrßen, J., van Ermingen-Marbach, M., Pape-Neumann, J., & Heim, S. (2014). Devil in the details? Developmental dyslexia and visual long-term memory for details. *Frontiers in Psychology*, 5, 686. https://doi.org/10.3389/fpsyg.2014.00686

Hulme, C., Snowling, M., Caravolas, M., & Carroll, J. (2005). Phonological skills are (probably) one cause of success in learning to read: A comment on Castles and Coltheart. *Scientific Studies of Reading, 9*(4), 351–365. https://doi.org/10.1207/s1532799xssr0904_2.

International Dyslexia Association (2002). Definition of dyslexia. Retrieved from International Dyslexia Association: https://dyslexiaida.org/definition-of-dyslexia/.

*Jones, M. W., Branigan, H. P., Parra, M. A., & Logie, R. H. (2013). Cross-modal binding in developmental dyslexia. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 39(6), 1807–1822. https://doi.org/10.1037/a0033334

*Jones, M. W., Kuipers, J. R., & Thierry, G. (2016). ERPs reveal the time-course of aberrant visual-phonological binding in developmental dyslexia. *Frontiers in Human Neuroscience*, 10, 71. https://doi.org/10.3389/fnhum.2016.00071

*Jones, M. W., Kuipers, J. R., Nugent, S., Miley, A., & Oppenheim, G. (2018). Episodic traces and statistical regularities: Paired associate learning in typical and dyslexic readers. *Cognition*, 177, 214–225. https://doi.org/10.1016/j.cognition.2018.04.010

*Judge, J., Knox, P. C., & Caravolas, M. (2013). Spatial orienting of attention in dyslexic adults using directional and alphabetic cues. *Dyslexia*, 19(2), 55–75. https://doi.org/10.1002/dys.1452

Katz, L., & Frost, R. (1992). Chapter 4. The reading process is different for different orthographies: The orthographic depth hypothesis. *Advances in Psychology, 94*, 67–84. https://doi.org/10.1016/S0166-4115(08)62789-2.

Kemp, N., Parrila, R. K., & Kirby, J. R. (2009). Phonological and orthographic spelling in high-functioning adult dyslexics. *Dyslexia, 15*(2), 105–128. https://doi.org/10.1002/dys.364.

Khomsi, A. (1999). Epreuve d'évaluation de la compétence en lecture: Lecture de mots et compréhension-révisée. *Les Editions du Centre de Psychologie appliquée*.

*Kim, S., Lombardino, L. J., Cowles, W., & Altmann, L. J. (2014). Investigating graph comprehension in students with dyslexia: An eye tracking study. *Research in Developmental Disabilities*, *35*(7), 1609–1622. https://doi.org/10.1016/j.ridd.2014.03.043

Kirby, J. R., Georgiou, G. K., Martinussen, R., & Parrila, R. (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.

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.

*Koerte, I. K., Willems, A., Muehlmann, M., Moll, K., Cornell, S., Pixner, S., Steffinger D., Keeser D., Heinen F., Kubicki M., Shenton M. E., Ertl-Wagner B. Schulte-Körne, G. (2016). Mathematical abilities in dyslexic children: A diffusion tensor imaging study. *Brain Imaging and Behavior*, *10*(3), 781–791. https://doi.org/10.1007/s11682-015-9436-y

*Koyama, M. S., Di Martino, A., Kelly, C., Jutagir, D. R., Sunshine, J., Schwartz, S. J., … Milham, M. P. (2013). Cortical signatures of dyslexia and remediation: An intrinsic functional connectivity approach. *PLOS ONE*, *8*(2), e55454. https://doi.org/10.1371/journal.pone.0055454

Laasonen, M., Service, E., Lipsanen, J., & Virsu, V. (2012). Adult developmental dyslexia in a shallow orthography: Are there subgroups? *Reading and Writing, 25*(1), 71–108. https://doi.org/10.1007/s11145-010-9248-9.

*Lallier, M., Donnadieu, S., & Valdois, S. (2013a). Investigating the role of visual and auditory search in reading and developmental dyslexia. *Frontiers in Human Neuroscience*, *7*, 597. https://doi.org/10.3389/fnhum.2013.00597

*Lallier, M., Donnadieu, S., & Valdois, S. (2013b). Developmental dyslexia: exploring how much phonological and visual attention span disorders are linked to simultaneous auditory processing deficits. *Annals of dyslexia*, *63*(2), 97-116. https://doi.org/10.1007/s11881-012-0074-4

*Lallier, M., Thierry, G., Tainturier, MJ., (2013c). On the importance of considering individual profiles when investigating the role of auditory sequential deficits in developmental dyslexia. *Cognition*;126(1):121-127. https://doi.org/10.1016/j.cognition.2012.09.008

Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: A German–English comparison. *Cognition, 63*(3), 315–334. https://doi.org/10.1016/s0010-0277(97)00005-x.

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H., Lohvansuu, K., et al. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 54*(6), 686–694. https://doi.org/10.1111/jcpp.12029.

*Langer, N., Benjamin, C., Minas, J., & Gaab, N. (2015). The neural correlates of reading fluency deficits in children. Cerebral Cortex, 25(6), 1441-1453. https://doi.org/10.1093/cercor/bht330

Leinonen, S., Müller, K., Leppänen, P. H. T., Aro, M., Ahonen, T., & Lyytinen, H. (2001). Heterogeneity in adult dyslexic readers: Relating processing skills to the speed and accuracy of oral text reading. *Reading and Writing, 14*(3/4), 265–296. https://doi.org/10.1023/A:1011117620895.

*Leong, V., & Goswami, U. (2014). Impaired extraction of speech rhythm from temporal modulation patterns in speech in developmental dyslexia. *Frontiers in Human Neuroscience*, *8*, 96. https://doi.org/10.3389/fnhum.2014.00096

*Lou, C., Duan, X., Altarelli, I., Sweeney, J. A., Ramus, F., & Zhao, J. (2019). White matter network connectivity deficits in developmental dyslexia. *Human Brain Mapping*, *40*(2), 505–516. https://doi.org/10.1002/hbm.24390

Maehler, C., & Schuchardt, K. (2016). The importance of working memory for school achievement in primary school children with intellectual or learning disabilities. *Research in Developmental Disabilities, 58*, 1–8. https://doi.org/10.1016/j.ridd.2016.08.007.

*Mahé, G., Doignon-Camus, N., Dufour, A., & Bonnefond, A. (2014). Conflict control processing in adults with developmental dyslexia: An event related potentials study. *Clinical Neurophysiology*, *125*(1), 69–76. https://doi.org/10.1016/j.clinph.2013.06.005

*Mahé, G., Pont, C., Zesiger, P., & Laganaro, M. (2018). The electrophysiological correlates of developmental dyslexia: New insights from lexical decision and reading aloud in adults. *Neuropsychologia*, *121*, 19–27. https://doi.org/10.1016/j.neuropsychologia.2018.10.025

Martin, A., Schurz, M., Kronbichler, M., & Richlan, F. (2015). Reading in the brain of children and adults: A meta-analysis of 40 functional magnetic resonance imaging studies. *Human Brain Mapping, 36*(5), 1963–1981. https://doi.org/10.1002/hbm.22749.

Martin, A., Kronbichler, M., & Richlan, F. (2016). Dyslexic brain activation abnormalities in deep and shallow orthographies: A meta-analysis of 28 functional neuroimaging studies. *Human Brain Mapping, 37*(7), 2676–2699. https://doi.org/10.1002/hbm.23202.

*Martinez Perez, T., M. P, Steve, M., & Martine, P. (2013). Impaired short-term memory for order in adults with dyslexia. *Research in developmental disabilities*, 34(7), 2211-2223. https://doi.org/10.1016/j.ridd.2013.04.005

*Martinez Perez, T., Poncelet, M., Salmon, E., & Majerus, S. (2015). Functional alterations in order short-term memory networks in adults with dyslexia. *Developmental Neuropsychology, 40*(7–8), 407–429. https://doi.org/10.1080/87565641.2016.1153098

Menghini, D., Finzi, A., Carlesimo, G. A., & Vicari, S. (2011). Working memory impairment in children with developmental dyslexia: Is it just a phonological deficit? *Developmental Neuropsychology, 36*(2), 199–213. https://doi.org/10.1080/87565641.2010.549868.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & the PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Open Medicine, 3*(3), e123–e130. https://doi.org/10.1016/j.ijsu.2010.02.007.

Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., Streitau, 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., Soares-Boucaud, I., le Heuzey, M. F., Billard, C., O'Donovan, M., Hill, G., Williams, J., Brandeis, D., Maurer, U., Schulz, E., van der Mark, S., Müller-Myhsok, B., Schulte-Körne, G., & 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.

*Moores, E., Tsouknida, E., & Romani, C. (2015). Adults with dyslexia can use cues to orient and constrain attention but have a smaller and weaker attention spotlight. *Vision research*, *111*, 55-65. https://doi.org/10.1016/j.visres.2015.03.019

Mousty, P., & Leybaert, J. (1999). Evaluation des habiletés de lecture et d'orthographe au moyen de BELEC: Données longitudinales auprès d'enfants francophones testés en 2ème et 4ème années. *Revue européenne de psychologie appliquée, 49*(4), 325–342.

*Mundy, I. R., & Carroll, J. M. (2013). Spelling-stress regularity effects are intact in developmental dyslexia. *Quarterly Journal of Experimental Psychology*, *66*(4), 816–828. https://doi.org/10.1080/17470218.2012.719530

Naglieri, J. A. (2003). Naglieri nonverbal ability tests. In *Handbook of nonverbal assessment* (pp. 175–189). Springer.

Nergård-Nilssen, T., & Hulme, C. (2014). Developmental dyslexia in adults: Behavioral manifestations and cognitive correlates. *Dyslexia, 20*(3), 191–207. https://doi.org/10.1002/dys.1477.

Nicolson, R. I., & Fawcett, A. J. (1990). Automaticity: a new framework for dyslexia research? *Cognition, 35*(2), 159–182. https://doi.org/10.1016/0010-0277(90)90013-a.

Nicolson, R. I., & Fawcett, A. J. (2005). Developmental dyslexia, learning and the cerebellum. *Journal of Neural Transmission Supplementa*, (69), 19–36. https://doi.org/10.1007/3-211-31222-6_2.

Nicolson, R. I., & Fawcett, A. J. (2011). Dyslexia, dysgraphia, procedural learning and the cerebellum. *Cortex, 47*(1), 117–127. https://doi.org/10.1016/j.cortex.2009.08.016.

Nicolson, R. I., Fawcett, A. J., Berry, E. L., Jenkins, I. H., Dean, P., & Brooks, D. J. (1999). Association of abnormal cerebellar activation with motor learning difficulties in dyslexic adults. *Lancet, 353*(9165), 1662–1667. https://doi.org/10.1016/S0140-6736(98)09165-X.

Nicolson, R. I., Fawcett, A. J., & Dean, P. (2001). Developmental dyslexia: the cerebellar deficit hypothesis. *Trends in Neuroscience, 24*(9), 508–511. https://doi.org/10.1016/s0166-2236(00)01896-8.

*Nittrouer, S., & Lowenstein, J. H. (2013). Perceptual organization of speech signals by children with and without dyslexia. *Research in Developmental Disabilities*, *34*(8), 2304–2325. https://doi.org/10.1016/j.ridd.2013.04.018

*Nittrouer, S., Krieg, L. M., & Lowenstein, J. H. (2018). Speech recognition in noise by children with and without dyslexia: How is it related to reading? *Research in Developmental Disabilities*, *77*, 98–113. https://doi.org/10.1016/j.ridd.2018.04.014

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*, 427–452. https://doi.org/10.1146/annurev-psych-120710-100431.

Olofsson, Å. (2002). Twenty years of phonological deficits: Lundberg's sample revisited. *Dyslexia & literacy*, 151–162.

*Palladino, P., Bellagamba, I., Ferrari, M., & Cornoldi, C. (2013). Italian children with dyslexia are also poor in reading English words, but accurate in reading English pseudowords. *Dyslexia*, *19*(3), 165–177. https://doi.org/10.1002/dys.1456

*Palladino, P., Cismondo, D., Ferrari, M., Ballagamba, I., & Cornoldi, C. (2016). L2 spelling errors in Italian children with dyslexia. *Dyslexia*, *22*(2), 158–172. https://doi.org/10.1002/dys.1522

*Papagiannopoulou, E. A., & Lagopoulos, J. (2017). P300 event-related potentials in children with dyslexia. *Annals of Dyslexia*, *67*(1), 99–108. https://doi.org/10.1007/s11881-016-0122-6

Parrila, R., Georgiou, G. K., & Papadopoulos, T. C. (2020). Dyslexia in a consistent orthography: Evidence from reading-level match design. *Dyslexia, 26*(4), 343–358. https://doi.org/10.1002/dys.1650

Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., Cotelli, M., Cossu, G., Corte, F., Lorusso, M., Pesenti, S., Gallagher, A., Perani, D., Price, C., Frith, C. D., & Frith, U. (2000). A cultural effect on brain function. *Nature Neuroscience, 3*(1), 91–96. https://doi.org/10.1038/71163.

Paulesu, E., Démonet, J. F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., Cappa, S. F., Cossu, G., Habib, M., Frith, C. D., & Frith, U. (2001). Dyslexia: Cultural diversity and biological unity. *Science, 291*(5511), 2165–2167. https://doi.org/10.1126/science.1057179.

*Paz-Alonso, P. M., Oliver, M., Lerma-Usabiaga, G., Caballero-Gaudes, C., Quiñones, I., Suárez-Coalla, P., Duñabeitia J. A., Cuetos F. Carreiras, M. (2018). Neural correlates of phonological, orthographic and semantic reading processing in dyslexia. *NeuroImage. Clinical*, *20*, 433–447. https://doi.org/10.1016/j.nicl.2018.08.018

Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child development, 78*(4), 1255–1264. https://doi.org/10.1111/j.1467-8624.2007.01064.x.

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., Dardick, W., & Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin, 144*(1), 48–76. https://doi.org/10.1037/bul0000124.

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., Van Orden, G. C., Smith, S. D., Green, P. A., & Haith, M. M. (1990). Phonological processing skills and deficits in adult dyslexics. *Child development, 61*(6), 1753–1778. https://doi.org/10.1111/j.1467-8624.1990.tb03564.x.

*Perea, M., Jiménez, M., Suárez-Coalla, P., Fernández, N., Viña, C., & Cuetos, F. (2014). Ability for voice recognition is a marker for dyslexia in children. *Experimental Psychology*, *61*(6), 480–487. https://doi.org/10.1027/1618-3169/a000265

Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly (1982–)*, 283–319.

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist dual Process (CDP++) model. *Cognitive Psychology, 61*(2), 106–151. https://doi.org/10.1016/j.cogpsych.2010.04.001.

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

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain, 126*(4), 841–865. https://doi.org/10.1093/brain/awg076.

Raven, J. C. (1956). *Guide to using the coloured progressive matrices Sets A, Ab, B. (Revised Order, 1956)*. H. K. Lewis.

Raven, J. C. (1958). *Mill Hill vocabulary scale: MHV; Raven's Standard Progressive Matrices: SPM*. HK Lewis.

Raven, J. (2003). Raven progressive matrices. In *Handbook of nonverbal assessment* (pp. 223–237). Berlin: Springer.

R Core Team (2015). R: A language and environment for statistical computing.

Reid, A. A., Szczerbinski, M., Iskierka-Kasperek, E., & Hansen, P. (2007). Cognitive profiles of adult developmental dyslexics: theoretical implications. *Dyslexia, 13*(1), 1–24. https://doi.org/10.1002/dys.321.

*Reilhac, C., Peyrin, C., Démonet, J. F., & Valdois, S. (2013). Role of the superior parietal lobules in letter-identity processing within strings: FMRI evidence from skilled and dyslexic readers. *Neuropsychologia*, *51*(4), 601–612. https://doi.org/10.1016/j.neuropsychologia.2012.12.010

Reis, A., Araújo, S., Morais, I. S., & Faísca, L. (2020). Reading and reading-related skills in adults with dyslexia from different orthographic systems: a review and meta-analysis. *Annals of Dyslexia, 70*(3), 339–368. https://doi.org/10.1007/s11881-020-00205-x.

Richlan, F. (2014). Functional neuroanatomy of developmental dyslexia: The role of orthographic depth. *Frontiers in Human Neuroscience, 8*, 347. https://doi.org/10.3389/fnhum.2014.00347.

Richlan, F., Kronbichler, M., & Wimmer, H. (2011). Meta-analyzing brain dysfunctions in dyslexic children and adults. *Neuroimage, 56*(3), 1735–1742. https://doi.org/10.1016/j.neuroimage.2011.02.040.

*Ruffino, M., Gori, S., Boccardi, D., Molteni, M., & Facoetti, A. (2014). Spatial and temporal attention in developmental dyslexia. Frontiers in human neuroscience, 8, 331. https://doi.org/10.3389/fnhum.2014.00331

*Rüsseler, J., Gerth, I., Heldmann, M., & Münte, T. F. (2015). Audiovisual perception of natural speech is impaired in adult dyslexics: An ERP study. *Neuroscience*, *287*, 55–65. https://doi.org/10.1016/j.neuroscience.2014.12.023

*Rüsseler, J., Ye, Z., Gerth, I., Szycik, G. R., & Münte, T. F. (2018). Audio-visual speech perception in adult readers with dyslexia: An fMRI study. *Brain Imaging and Behavior*, *12*(2), 357–368. https://doi.org/10.1007/s11682-017-9694-y

*Saksida, A., Iannuzzi, S., Bogliotti, C., Chaix, Y., Démonet, J. F., Bricout, L., Billard C., Nguyen-Morel M.A., le Heuzey M.F., Soares-Boucaud I., George F., Ziegler J. C. Ramus, F. (2016). Phonological skills, visual attention span, and visual stress in developmental dyslexia. *Developmental Psychology*, *52*(10), 1503–1516. https://doi.org/10.1037/dev0000184

*Saralegui, I., Ontañón, J. M., Fernandez-Ruanova, B., Garcia-Zapirain, B., Basterra, A., & Sanz-Arigita, E. J. (2014). Reading networks in children with dyslexia compared to children with ocular motility disturbances revealed by fMRI. *Frontiers in Human Neuroscience*, *8*, 936. https://doi.org/10.3389/fnhum.2014.00936

Sartori, G., Job, R., & Tressoldi, P. (2007). DDE-2. *Bateria per la valutazione della dislessia e della disortografia evolutiva*, *2*.

Scarborough, H. S. (1984). Continuity between childhood dyslexia and adult reading. *British journal of psychology, 75*(3), 329–348. https://doi.org/10.1111/j.2044-8295.1984.tb01904.x.

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.

Seymour, P. H. K. (2005). Early reading development in European orthographies. In M. J. Snowling & C. Hulme (Eds.), *Blackwell handbooks of developmental psychology. The science of reading: A handbook* (pp. 296–315). Blackwell Publishing. https://doi.org/10.1002/9780470757642.ch16.

Seymour, P. H., 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.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134*(4), 584–615. https://doi.org/10.1037/0033-2909.134.4.584.

Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Shneider, A. E., Marchione, K. E., Stuebing, K. K., Francis, D. J., Pugh, K. R., & Shaywitz, B. A. (1999). Persistence of dyslexia: The Connecticut longitudinal study at adolescence. *Pediatrics, 104*(6), 1351–1359. https://doi.org/10.1542/peds.104.6.1351.

*Sigurdardottir, H. M., Danielsdottir, H. B., Gudmundsdottir, M., Hjartarson, K. H., Thorarinsdottir, E. A., & Kristjánsson, Á (2017). Problems with visual statistical learning in developmental dyslexia. *Scientific Reports*, *7*(1), 606. https://doi.org/10.1038/s41598-017-00554-5

Skottun, B. C. (2000). On the conflicting support for the magnocellular-deficit theory of dyslexiaResponse to Stein, Talcott and Walsh (2000). *Trends in Cognitive Sciences, 4*(6), 211–212. https://doi.org/10.1016/s1364-6613(00)01485-6.

Snowling, M. J. (1981). Phonemic deficits in developmental dyslexia. *Psychological Research, 43*(2), 219–234. https://doi.org/10.1007/BF00309831.

Snowling, M. J. (1995). Phonological processing and developmental dyslexia. *Journal of Research in Reading, 18*(2), 132–138. https://doi.org/10.1111/j.1467-9817.1995.tb00079.x.

Sprenger-Charolles, L., Siegel, L. S., Jiménez, J. E., & Ziegler, J. C. (2011). Prevalence and reliability of phonological, surface, and mixed profiles in dyslexia: A review of studies conducted in languages varying in orthographic depth. *Scientific Studies of Reading, 15*(6), 498–521. https://doi.org/10.1080/10888438.2010.524463.

Stanovich, K. E. (1988). The right and wrong places to look for the cognitive locus of reading disability. *Annals of Dyslexia, 38*(1), 154–177. https://doi.org/10.1007/BF02648254.

Stein, J. (2019). The current status of the magnocellular theory of developmental dyslexia. *Neuropsychologia, 130*, 66–77. https://doi.org/10.1016/j.neuropsychologia.2018.03.022.

Stein, J., & Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends in Neuroscience, 20*(4), 147–152. https://doi.org/10.1016/s0166-2236(96)01005-3.

Stein, J., Talcott, J., & Walsh, V. (2000). Controversy about the visual magnocellular deficit in developmental dyslexics. *Trends in Cognitive Sciences, 4*(6), 209–211. https://doi.org/10.1016/s1364-6613(00)01484-4.

*Suárez-Coalla, P., & Cuetos, F. (2015). Reading difficulties in Spanish adults with dyslexia. *Annals of Dyslexia*, *65*(1), 33–51. https://doi.org/10.1007/s11881-015-0101-3

*Suárez-Coalla, P., Avdyli, R., & Cuetos, F. (2014). Influence of context-sensitive rules on the formation of orthographic representations in Spanish dyslexic children. *Frontiers in Psychology*, *5*, 1409. https://doi.org/10.3389/fpsyg.2014.01409

*Sumner, E., Connelly, V., & Barnett, A. L. (2014). The influence of spelling ability on handwriting production: Children with and without dyslexia. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(5), 1441–1447. https://doi.org/10.1037/a0035785

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory, 2*(2), 103–127.

*Talli, I., Sprenger-Charolles, L., & Stavrakaki, S. (2016). Specific language impairment and developmental dyslexia: What are the boundaries? Data from Greek children. *Research in Developmental Disabilities*, *49–50*, 339–353. https://doi.org/10.1016/j.ridd.2015.12.014

*Toffalini, E., Marsura, M., Garcia, R. B., & Cornoldi, C. (2019). A cross-modal working memory binding span deficit in reading disability. *Journal of Learning Disabilities*, *52*(2), 99–108. https://doi.org/10.1177/0022219418786691

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Burgess, S., & Hecht, S. (1997). Contributions of phonological awareness and rapid automatic naming ability to the growth of word-reading skills in second-to fifth-grade children. *Scientific studies of reading, 1*(2), 161–185. https://doi.org/10.1207/s1532799xssr0102_4. Pearson Clinical Assessment.

Torgesen, J. K., Wagner, R., & Rashotte, C. (2012). *Test of word reading efficiency (TOWRE-2)*. Pearson Clinical Assessment.

Undheim, A. M. (2009). A thirteen-year follow-up study of young Norwegian adults with dyslexia in childhood: reading development and educational levels. *Dyslexia, 15*(4), 291–303. https://doi.org/10.1002/dys.384.

Vaessen, A., Bertrand, D., Tóth, D., Csépe, V., Faísca, L., Reis, A., & Blomert, L. (2010). Cognitive development of fluent word reading does not qualitatively differ between transparent and opaque orthographies. *Journal of Educational Psychology, 102*(4), 827–842. https://doi.org/10.1037/a0019465.

*Varnet, L., Meunier, F., Trollé, G., & Hoen, M. (2016). Direct viewing of dyslexics' compensatory strategies in speech in noise using auditory classification images. *PLOS ONE, 11*(4), e0153781. https://doi.org/10.1371/journal.pone.0153781

*Varvara, P., Varuzza, C., Sorrentino, A. C., Vicari, S., & Menghini, D. (2014). Executive functions in developmental dyslexia. *Frontiers in Human Neuroscience, 8*, 120. https://doi.org/10.3389/fnhum.2014.00120

Vellutino, F. R. (1979). *Dyslexia: Theory and research*. MIT Press.

*Vender, M., Mantione, F., Savazzi, S., Delfitto, D., & Melloni, C. (2017). Inflectional morphology and dyslexia: Italian children's performance in a nonword pluralization task. *Annals of Dyslexia, 67*(3), 401–426. https://doi.org/10.1007/s11881-017-0152-8

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48.

Willburger, E., & Landerl, K. (2009). Der ein-minuten leseflüssigkeitstest—ein verfahren zur diagnose der leistung im wort-und pseudowortlesen [1-minute Reading Fluency Test]. *Tests und Trends, 7*, 65–80.

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of educational psychology, 91*(3), 415–438. https://doi.org/10.1037/0022-0663.91.3.415.

*Ye, Z., Rüsseler, J., Gerth, I., & Münte, T. F. (2017). Audiovisual speech integration in the superior temporal region is dysfunctional in dyslexia. *Neuroscience, 356*, 1–10. https://doi.org/10.1016/j.neuroscience.2017.05.017

*Zhang, Y., Whitfield-Gabrieli, S., Christodoulou, J. A., & Gabrieli, J. D. (2013). Atypical balance between occipital and fronto-parietal activation for visual shape extraction in dyslexia. *PLOS ONE, 8*(6), e67331. https://doi.org/10.1371/journal.pone.0067331

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3.

Ziegler, J. C., & Goswami, U. (2006). Becoming literate in different languages: Similar problems, different solutions. *Developmental Science, 9*(5), 429–436. https://doi.org/10.1111/j.1467-7687.2006.00509.x.

Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology, 86*(3), 169–193. https://doi.org/10.1016/s0022-0965(03)00139-5.

Ziegler, J. C., Castel, C., Pech-Georgel, C., George, F., Alario, F. X., & Perry, C. (2008). Developmental dyslexia and the dual route model of reading: Simulating individual differences and subtypes. *Cognition, 107*(1), 151–178. https://doi.org/10.1016/j.cognition.2007.09.004.

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., Saine, N., Lyytinen, H., Vaessen, A., & Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science, 21*(4), 551–559. https://doi.org/10.1177/0956797610363406.

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.

*Zoubrinetzky, R., Bielle, F., & Valdois, S. (2014). New insights on developmental dyslexia subtypes: Heterogeneity of mixed reading profiles. *PLOS ONE, 9*(6), e99337. https://doi.org/10.1371/journal.pone.0099337

*Zuk, J., Perdue, M. V., Becker, B., Yu, X., Chang, M., Raschle, N. M., & Gaab, N. (2018). Neural correlates of phonological processing: Disrupted in children with dyslexia and enhanced in musically trained children. *Developmental Cognitive Neuroscience, 34*, 82–91. https://doi.org/10.1016/j.dcn.2018.07.001