A stitch in time...: Comparing late-identified, late-emerging and early-identified dyslexia

de Bree, E.H.; van den Boer, M.; Toering, B.; de Jong, P.F.

RESEARCH ARTICLE

WILEY

A stitch in time...: Comparing late-identified, late-emerging and early-identified dyslexia

Elise H. de Bree[1,2] 🆔 | Madelon van den Boer[1] 🆔 |
Boukje M. Toering[3] | Peter F. de Jong[1]

[1]Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, Netherlands

[2]Department of Education and Pedagogy, Utrecht University, Utrecht, Netherlands

[3]Marnix Academy for the Training of Primary School Teachers, Utrecht, Netherlands

Correspondence
Elise H. de Bree, Department of Education and Pedagogy, Utrecht University, P.O. Box 80140, Utrecht 3580 TC, Netherlands.
Email: e.h.debree@uu.nl

When dyslexia is diagnosed late, the question is whether this is due to late-*emerging* (LE) or late-*identified* (LI) problems. In a random selection of dyslexia-diagnosis case files we distinguished early-diagnosed (Grade 1–3, $n = 116$) and late-diagnosed (Grade 4–6) dyslexia. The late-diagnosed files were divided into LE ($n = 54$) and LI dyslexia ($n = 45$). The LE group consisted of children whose national-curriculum literacy outcomes did not warrant referral for dyslexia diagnosis in Grades 1–2; the LI group of children whose literacy outcomes did, but who were referred for diagnostic assessment after Grade 3. At the time of diagnosis, the percentage of poor performers on word-level literacy measures generally did not differ between the groups. Only the LE group contained fewer poor performers than the early-diagnosed and LI group on some word-reading measures. All groups showed similar distributions of phonological difficulties. There were no indications of compensation through vocabulary, memory or IQ in either late-diagnosed group. Our diagnosis-based study confirms and extends previous research-based studies on LE dyslexia. Moreover, it shows that LI dyslexia exists, which can be regarded as the existence of instructional

casualties. The findings speak to issues of identification, diagnosis and compensation and call for further efforts to improve the early identification of dyslexia.

Key messages

- Ideally, severe and persistent word-level reading difficulties (dyslexia) are detected as early as possible.
- Our findings confirm findings that late-emerging dyslexia exists, meaning that the severe and persistent poor performance surface later.
- We also found that late-identified dyslexia exists, meaning that children were referred later than their reading outcomes warranted.
- The late-emerging and late-identified groups do not show evidence of compensation of literacy abilities.
- Efforts should be made to avoid such instructional casualties (late-identified dyslexia) and to support the school literacy curriculum.

1 | INTRODUCTION

Dyslexia is a disorder characterized by consistently poor word reading and spelling performances that cannot be accounted for by general learning difficulties, sensory deficits, or inadequate teaching These word-level literacy deficits are caused by multiple risk factors (Peterson & Pennington, 2015). As these severe and persistent word reading and spelling difficulties generally surface in the first years of literacy instruction, children with dyslexia are usually identified and diagnosed during the first years of primary school. There is, however, also a group of children whose literacy difficulties are identified and diagnosed later, typically after Grade 3. This group forms approximately one-third of samples of children with literacy difficulties (Catts, Compton, Tomblin, & Bridges, 2012; Leach, Scarborough, & Rescorla, 2003; Lipka, Lesaux, & Siegel, 2006; Torppa, Eklund, van Bergen, & Lyytinen, 2015). An important question is why some children's difficulties are identified in the early stages of literacy instruction and those of others only later.

2 | EARLY AND LATE-DIAGNOSED DYSLEXIA

The issue of early and late-diagnosed literacy problems has been targeted in a few studies. The main question is whether the literacy problems of the late group truly developed at a later stage than those of the early group. Previous studies reported that this is the case (Catts et al., 2012; Leach et al., 2003; Lipka et al., 2006; Torppa et al., 2015). Once identified, the outcomes of the early and late groups are generally similar or show a step-wise pattern of performance with a no-literacy problem group outperforming the late group, and the late group outperforming the early group on some measures (Bazen, van den Boer, de Jong, & de Bree, 2020; Catts et al., 2012; Compton, Fuchs, Fuchs, Elleman, & Gilbert, 2008; Leach et al., 2003; Lipka et al., 2006; Torppa et al., 2015).

The previous studies into early and late-emerging (LE) dyslexia share some characteristics concerning their literacy assessment. First, some studies (initially) categorized the sample of late poor readers on poor word reading and reading comprehension difficulties (Catts et al., 2012; Compton et al., 2008; Etmanskie, Partanen, & Siegel, 2016). Whereas poor word reading is one of the defining characteristics of dyslexia, this is not the case for poor reading

comprehension. Poor word readers are not necessarily poor comprehenders (Bazen et al., 2020; Moojen et al., 2020). This means that it is important to separate word reading from reading comprehension. Moreover, both in studies designed to establish the prevalence of poor readers (Compton et al., 2008; Etmanskie et al., 2016) and in studies designed to compare early and late groups (Bazen et al., 2020; Leach et al., 2003; Lipka et al., 2006; Torppa et al., 2015) the numbers of poor word-readers only have been small. Larger samples of early and late poor word readers are needed to draw more firm conclusions concerning LE and late-identified (LI)dyslexia.

Furthermore, the previous early- and late studies included word-level reading measures, but only Torppa et al. (2015) and Leach et al. (2003) included measures of word reading accuracy and fluency in younger readers. Other studies only included accuracy (Catts et al., 2012; Etmanskie et al., 2016; Lipka et al., 2006) or fluency (Bazen et al., 2020). Ideally, both are assessed: In the initial stages of reading development, word recognition will be slow and error-prone, making accuracy an adequately sensitive measure. As reading acquisition progresses, the ease or speed with which words can be read (correctly) becomes a more sensitive indicator of word reading ability (van Viersen et al., 2018). Additionally, both word and pseudoword reading outcomes have been found to be informative (Bazen et al., 2020; Torppa et al., 2015) as both decoding (pseudowords) and sight word reading processes (words) can be impaired in poor readers (Peterson, Pennington, Olson, & Wadsworth, 2014). In order to further understand the difference between early and late literacy difficulties, a comprehensive assessment of word-level reading is needed.

Also, the spelling performance of the early and late groups has received little attention so far; only two studies included a standardized measure of spelling (Etmanskie et al., 2016; Lipka et al., 2006) and two a raw measure (Bazen et al., 2020; Leach et al., 2003). Given that poor spelling is taken to be a feature of dyslexia and is often part of a diagnosis of dyslexia (Lyon, Shaywitz, & Shaywitz, 2003; Peterson & Pennington, 2015), it is informative to include both word reading and spelling outcomes in an evaluation of early and late dyslexia.

3 | LATE-EMERGING OR LATE-IDENTIFIED DYSLEXIA?

On the basis of the previous studies, the consensus seems to be that late-diagnosed literacy difficulties are actually late-emerging and were not present early on. Nevertheless, there are also indications that "instructional casualties" (Lyon, 2002; Snow, 2016) occur (Barbiero et al., 2012), referring to students whose literacy problems were present early on, but were not identified. Potentially, the group with late-diagnosed literacy problems can be a merger of children with literacy problems that are late-emerging and children with early problems that were missed and were therefore late identified.

One way to look into the late group is by using formal diagnoses of dyslexia. Such diagnoses include a range of literacy and cognitive skills, and take into account the development and persistence of literacy problems over time. Formal diagnoses of late-diagnosed children allow a division between LE dyslexia, with persistent literacy deficits truly developing at a later age, and LI dyslexia, with persistent literacy deficits being present much earlier than the actual time of referral for diagnosis. Literacy outcomes at the time of diagnosis of the LI group can then be compared to the early as well as the LE group. Such a comparison could include text-level literacy outcomes in the early grades, as the results of this task might be used in the referral process. Indeed, a German study into information that teachers used in the identification of children that need reading interventions showed that teachers identified more students with severe reading difficulties on the text level than on the word level (Schmitterer & Brod, 2021).

Next to literacy outcomes, it can be assessed whether LI children differ from early-diagnosed and LE children in terms of risk factors. The key risk factor associated with dyslexia is a phonological deficit (Peterson & Pennington, 2015), which usually concerns skills related to grapheme-phoneme associations, phoneme awareness (PA), and rapid automatized naming (RAN) (de Jong & van der Leij, 2003; Moll et al., 2014; Norton & Wolf, 2012). Previous research has shown that the late group performs better than the early-identified group on these risk factors before the literacy difficulties surface (Lipka et al., 2006) and resembles the early group once the literacy difficulties have become apparent (Bazen et al., 2020; Leach et al., 2003). However, the pattern of findings for the cognitive risk factors is not entirely consistent: one study reported that the early and late groups showed (equally) low performance on RAN prior to the appearance of literacy difficulties (Torppa et al., 2015). Vice versa, PA performance of

the late group at or after the time of diagnosis is not always as poor as that of the early-identified group (Bazen et al., 2020; Torppa et al., 2015). By dividing the late-diagnosed group into LE and LI children we can shed further light on the presence and strength of a phonological deficit in the different groups.

Furthermore, a comparison can be made of performance on protective factors, reflecting potential strengths in more general cognitive abilities such as intelligence, memory, or language skills. These might reduce or mask a present literacy and/or phonological deficit (Haft, Myers, & Hoeft, 2016; van Viersen, de Bree, & de Jong, 2019; van Viersen, Kroesbergen, Slot, & de Bree, 2016). The studies that compared early and late groups on one or more potential protective factors, that is, non-verbal intelligence (Bazen et al., 2020; Leach et al., 2003), verbal intelligence (Leach et al., 2003; Torppa et al., 2015), verbal short-term memory (Bazen et al., 2020; Torppa et al., 2015), or working memory (Lipka et al., 2006) did not find significant differences between the early and late groups. Similarly, Torppa et al. (2015) and Bazen et al. (2020) found no differences in vocabulary between early and LI groups with dyslexia and normally reading participants. Although these findings suggest that intelligence, memory, and vocabulary do not function as protective factors that compensate for a literacy deficit of LI children during the early primary school years, it needs to be assessed whether this is the case for children with LE as well as LI dyslexia.

4 | PRESENT STUDY

In the current study, we used case files of formal diagnoses of dyslexia and divided these files into early-diagnosed (Grades 1–3) and late-diagnosed children (Grades 4–6) with dyslexia in the semi-transparent orthography Dutch. The case files of late-diagnosed children were divided into LE and LI groups. We addressed the following questions:

1. To what extent does the LI group differ from the early-diagnosed and LE group on word reading and spelling outcomes at the time of diagnosis and on text-level reading measures at the end of Grade 1?
2. To what extent does the LI group differ from the early-diagnosed and LE group on phonological measures at the time of diagnosis?
3. To what extent does the LI group differ from the early-diagnosed and LE group on potentially compensatory measures?

We had the following expectations:

1. All three groups were expected to show poor literacy outcomes at the time of diagnosis. A very tentative hypothesis, based on the finding by Schmitterer and Brod (2021), is that text-level reading outcomes of the LI group might be better than their word reading and spelling problems at the end of Grade 1.
2. Based on the hypothesis that LE dyslexia is characterized by a less severe phonological deficit (Bazen et al., 2020; Lipka et al., 2006; Torppa et al., 2015), children with LE dyslexia might perform better than early-diagnosed and LI children on tasks assessing a phonological deficit. For children with LI dyslexia, a phonological deficit similar to the early-identified group was expected.
3. Expectations concerning the protective factors were less clear. For the LE group, dyslexia could be LE due to the presence of protective factors. This would mean that children with LE dyslexia would display higher intelligence, memory, and vocabulary scores compared to early-diagnosed children. Empirical findings so far do not suggest this to be the case (Bazen et al., 2020; Leach et al., 2003; Lipka et al., 2006; Torppa et al., 2015). For the LI group, it could be the case that their performance on early protective factors (vocabulary) is higher than that of the early-emerging group. Schmitterer and Brod (2021) found that German teachers' evaluations of a student's need for reading interventions were partly based on children's vocabulary outcomes. The early-identified group might thus have lower vocabulary outcomes than the LI group. An alternative option could be that the LI group's performance on tasks related to protective factors is not high and therefore does not differ strongly from their literacy abilities. The absence of a clear discrepancy might move teachers' attention away from the literacy problems.

5 | METHOD

5.1 | Dyslexia diagnosis in the Netherlands

In the Netherlands, the diagnosis of dyslexia entails two steps. First, at school, a response-to-intervention approach is followed Fuchs & Fuchs, 2006; Scheltinga, van der Leij, & Struiksma, 2009). Literacy outcomes at school are monitored using standardized tests twice every year from the start of Grade 1. Low literacy outcomes based on classroom instruction (Tier 1) should lead to intensified instruction (Tier 2), and further targeted intervention (Tier 3) if outcomes do not progress. The stepped care entails an increase in time spent on instruction and practice as well as differentiation in instruction. If performance remains ≤10th percentile on three consecutive measurements, despite increasingly intensive instruction, a child can be referred to specialist care. The 10th percentile is a nation-wide benchmark (Gijsel, Scheltinga, Druenen, & Verhoeven, 2011; Scheltinga et al., 2009). Importantly, given the time span of three subsequent measurements, the percentage of children referred to specialized centres is generally lower than 10% (Vloedgraven et al., 2010). Also, this procedure of three subsequent assessments means that a referral for dyslexia diagnosis cannot take place before Grade 2.

The second step in the diagnostic process is that children whose word-level literacy remains below the 10th percentile despite increasingly intensified instruction and intervention at school are referred by the school to specialized diagnostic and treatment centres. At the specialized centre, the further procedure for the diagnosis of dyslexia is based on national health care guidelines and protocols, which, at the time of the diagnostic reports under investigation in the present study, were those of the National Reference Centre Dyslexia (NRD, 2013) and the Dutch Dyslexia Foundation (SDN et al., 2016). At the behavioural level, assessment of word-level literacy takes place. Regarding word-level reading assessment, measures typically include word reading accuracy and fluency, and decoding. These findings provide solid and coherent insight into word-level reading abilities. Similarly, spelling is typically assessed by a spelling-to-dictation measure and complemented with a combined accuracy and fluency measure. At the cognitive level, phonological abilities are assessed, given the established relations between phonology and literacy (Hulme, Nash, Gooch, Lervåg, & Snowling, 2015; van Viersen et al., 2018). Measures are tasks tapping letter-sound knowledge, PA and RAN. Outcomes can be used to understand and confirm the literacy deficit. Furthermore, an assessment of (other) strengths and weaknesses of the child also takes place. This is done to exclude children for whom general learning difficulties or sensory deficits are the cause of persistent and severe literacy difficulties (e.g., DSM-5; Peterson & Pennington, 2015). A second purpose is to assess skills that relate to broader educational needs and could be relevant for dyslexia treatment and support. It usually entails the assessment of tasks spanning memory and (non-)verbal intelligence. The environment of the child is also taken into account in the diagnostic process: it is established that unfavourable educational and environmental issues cannot be the reason for the consistently poor literacy outcomes (in line with response-to-instruction) and it is ascertained whether there is a family history of dyslexia as a potential risk factor for dyslexia.

5.2 | Case file selection and group assignment

For the current study, we analyzed case files of children who were referred to a specialized centre for specific learning disabilities in the Netherlands for a diagnostic examination of dyslexia. This is the second step in the dyslexia diagnosis procedure described above. As part of the protocol of the institute, the parents of the children agreed to the use of the scores of this examination for scientific purposes. From all the children who visited the institute between January 2013 and June 2016 (over 2000 files) we randomly selected 242 files included in the current study. These files had to concern children who had actually been diagnosed with dyslexia somewhere between Grades 1 and 6. We excluded children with an IQ below 75 ($n = 5$). We then divided the case files into an early-diagnosed group ($n = 133$), with a diagnosis made in Grades 1–3 and a late-diagnosed group ($n = 104$), diagnosis between Grades 4–6.

Next, we excluded cases in the early-diagnosed group: for three cases, school literacy scores had not been entered in the diagnostic report. For 14 other children, a formal dyslexia diagnosis was made in Grade 2 but their

curriculum literacy outcomes were not in the bottom range and therefore did not warrant referral. All 14 children had been retained but the norm scores were based on the grade the children were attending. Their higher literacy scores, therefore, present an inflated picture of their literacy abilities. Nevertheless, to err on the side of caution, data of these children were excluded from the early-diagnosed sample. The early-diagnosed group consisted of case files of 116 early-diagnosed children.

The late-diagnosed group was divided into a LE and a LI group. The LI children were those who obtained three bottom scores (≤10th percentile) in a row on the standardized national reading and/or spelling tests in Grades 1 and 2 (see materials, below) based on the existing national school protocol (Gijsel et al., 2011). LE children were those who did not show these persistently low scores in the early school years. School literacy scores had not been entered in the diagnostic report for 5 children. Therefore, classification was possible for 99 of the 104 late-diagnosed children. The ensuing groups were 54 children with LE dyslexia and 45 with LI dyslexia.

5.3 | Participants

Participant information per group is presented in Table 1. All participants were fluent speakers of Dutch. None of the children had another diagnosed disorder (e.g., ADHD, DLD) next to dyslexia. The three groups did not differ on the distribution of boys/girls $\chi^2(2) = .869$, $p = .647$ and on mono/bilingual children $\chi^2(2) = 3.433$, $p = .180$. They also did not differ on the measure of home socio-economic status (SES), $F(2, 212) = .157$, $p = .855$ and school SES $F(2,212) = 1.107$, $p = .333$. These SES measures consisted of ratings of the postal code of the children's homes/ schools, reflecting average income and educational attainment, as well as the percentage of unemployment within the neighbourhood (Netherlands Institute for Social Research, 2017). Mean home and school SES for all three groups were higher than the national average of 0.28 ($SD = 1.09$).

The groups differed in age $F(2, 212) = 265.27$, $p < .001$ and in distribution of Grade $\chi^2(10) = 221.20$, $p < .001$, with the early-diagnosed children being younger and in a lower Grade at the time of diagnosis than both the LI and LE groups ($p < .001$). The LI and LE groups did not differ from each other in age ($p = .96$) and distribution of Grade at which the diagnosis was made $\chi^2(2) = 2.86$, $p = .240$.

5.4 | Materials

5.4.1 | Early language and literacy skills

The files of the children included both literacy and vocabulary outcomes from when children were in Grade 1 and 2. These scores were determined at school with the national curriculum-based tests that are administered in all regular Dutch schools both halfway through and at the end of each school year. The scores consisted of an indication of the student's level, ranging from 1 through 5 ((1) percentiles 75–100; (2) percentiles 50–75; (3) percentiles 25–50;

TABLE 1 Participant information for the three groups

		Late (n = 99)	
	Early (n = 118)	Late identified (n = 45)	Late emerging (n = 54)
Mean age of diagnosis	8;5 (8.1 months)	10;9 (10.1 months)	10;9 (10.2 months)
Nr boys	71 (61%)	29 (64%)	30 (55%)
Nr bilingual	3	4	4
SES home	0.94 (0.74)	0.94 (0.68)	1.0 (0.70)
SES school	0.95 (.74)	0.79 (0.71)	1.1 (0.74)

(4) percentiles 10–25; (5) percentiles 0–10). All tests have at least sufficient reliability and validity (COTAN, 2010). For group classification into early, LI and LE (see Procedure), word-reading and spelling scores in middle Grade 1, end Grade 1, middle Grade 2 and end Grade 2 were used. For comparisons of the LI group with early and LE groups on text-level reading and vocabulary, results of end Grade 1 are reported.

Word reading

The Cito word reading test is a fluency test (DMT; Krom, Jongen, Verhelst, Kamphuis, & Kleintjes, 2010). Children were presented with three lists of 150 words each; the first list contained monosyllabic words without consonant clusters (e.g., oom, *uncle*), the second monosyllabic words with consonant clusters (e.g., bloem, *flower*), and the third multisyllabic words (e.g., moeilijk, *difficult*). Children were asked to read each list of words as quickly and accurately as possible for 1 minute.

Text reading

The Cito text reading test is a combined accuracy and fluency test (AVI; Krom et al., 2010). The test consists of a total of 11 texts of increasing difficulty that are adapted to specific Grade levels. Children were presented with the text that fits their current Grade level and were asked to read the text as quickly and accurately as possible. For each text, a maximum number of errors and maximum reading speed is provided in the manual. If children met these criteria, they were presented with a more difficult text. If they did not, they were administered an easier test. This procedure was continued until the most difficult text they passed in terms of accuracy and speed was determined. The level of that text is the score (range 1–11). If children failed to meet the criteria for the easiest text, they received a score of 0.

Spelling

The Cito spelling test consisted of two parts of 25 items each (de Wijs, Kamphuis, Kleintjes, & Tomesen, 2010). The first part was a spelling-to-dictation task that was administered to the entire class. Based on the results of this first part, children received either the easy or the difficult follow-up test. The easy follow-up test consisted of another spelling-to-dictation task. The difficult follow-up test consisted of multiple-choice questions. Children were presented with four sentences that each included a target word and needed to identify which of the target words was spelled incorrectly.

Vocabulary

The Cito vocabulary test consisted of multiple-choice questions (van Berkel et al., 2010). The teacher read aloud a sentence with a target word. Children saw three pictures and had to choose the picture that best matched the target word in the sentence (e.g., "Inges moeder maakt een maaltijd. Welk plaatje past het best bij een maaltijd?" *Inge's mother makes a meal. What picture best matches a meal?* There were two parts of 25 items each.

Reading comprehension

The Cito reading comprehension test (Feenstra, Kleintjes, Kamphuis, & Krom, 2010) consisted of short text fragments, mainly fictional texts, and 50 related questions, divided into two parts. All questions were multiple choice and concerned a question about the actual text, a prediction about the text or a missing word or fragment of text to fill in. Similar to the spelling test, the first part was the same for all children, and children received either an easy or difficult follow-up test. The test does not assess global coherence as the fragments are short.

5.5 | Literacy and cognitive skills at diagnostic examination

Diagnostic testing for dyslexia included administering the 3DM (Blomert & Vaessen, 2009), a test battery for children in Grades 1 through 6, that consists of tasks for literacy, underlying cognitive and memory skills. The 3DM is a

computer-based test that allows registration of both accuracy and response times. Raw scores are automatically transformed into percentile scores. Additional diagnostic tasks were administered for sight word reading, decoding skills, spelling to dictation, and intelligence.

Word reading accuracy and speed

The word reading task of the 3DM (Blomert & Vaessen, 2009) consisted of three subtests containing respectively, high-frequent words, low-frequent words, and pseudowords. Each subtest consisted of five screens of 15 words each. Across screens, the items increased in difficulty, from monosyllabic words without consonant clusters to three-syllabic words with consonant clusters. Children were asked to read the words on each screen as quickly and accurately as possible for 30 seconds. Two scores were calculated; a score for (pseudo-)word reading speed, based on the total number of items read correctly within the time limit, and a score for word reading accuracy, reflecting the number of words read correctly out of all the words read. Test–retest reliability is .73 for accuracy and .95 for speed (Blomert & Vaessen, 2009).

Spelling recognition

The spelling recognition task of the 3DM consisted of three subtests containing respectively, phonologically consistent words, phonologically inconsistent words with spelling conventions based on phonemes, and phonologically inconsistent words with spelling conventions based on syllabic structure, each consisting of 18 items. Each item was presented orally and was presented on the computer screen simultaneously, without the grapheme(s) for one corresponding phoneme. Children were asked to select the omitted grapheme(s) from four alternatives (e.g., zwemmen [*to swim*] was presented as "zwe.en" with the alternatives vv, v, mm, m) as quickly and as accurately as possible with a maximum reaction time of 15 seconds. The task was discontinued after subtest 2 for children in Grade 1, and for children with less than 6 items correct on subtest 2. Two scores were calculated: spelling recognition accuracy and spelling recognition speed. Items with reaction times lower than 300 ms were not included. Internal consistency is .80 for accuracy and .94 for speed (Blomert & Vaessen, 2009).

Word reading fluency

Sight word reading was assessed with the Eén Minuut Test (*One Minute Test*; Brus & Voeten, 1999; $r = .89–.92$), which is a standardized test for word reading fluency. Children read aloud as quickly and accurately as possible a list of 116 words of increasing difficulty. The number of words read correctly in 1 minute was transformed to a standardized score with a mean of 10 and a standard deviation of 3.

Decoding skills

Decoding skills were assessed with the Klepel (van den Bos, Lutje Spelberg, Scheepstra, & de Vries, 1994; $r = .91$), which is a standardized test for pseudoword reading fluency. Children read aloud as quickly and accurately as possible a list of 116 pseudowords of increasing difficulty. The number of words read correctly in 2 minutes was transformed to a standardized score with a mean of 10 and a standard deviation of 3.

Spelling-to-dictation

The spelling of words was assessed with PI-dictee (Geelhoed & Reitsma, 1999; $r = .87–.91$), a standardized spelling-to-dictation task. The task included a total of 9 blocks of 15 items each. Each block consisted of items with the phoneme-grapheme connections, spelling rules, and exception words taught in a particular Grade. Testing started with the first block and was discontinued when less than 8 words of a block were spelled correctly. Each word was read aloud to the child, as well as a sentence including the word. Children were then asked to write down the target word (e.g., "hout. De tafel is van *hout* gemaakt", *wood. The table is made of* wood). The score consisted of the total number of items spelled correctly and was transformed into a percentile score.

Phoneme-grapheme connections

Phoneme-grapheme connections were assessed with a letter-sound identification task. The task consisted of 45 items. Children heard a phoneme and were asked to select the matching grapheme(s) from 4 alternatives with a maximum reaction time of 10 seconds (e.g., /d/ presented with the alternatives g, d, b, p). Accuracy and mean reaction time across items were established. Items with reaction times lower than 300 ms were not included. Internal consistency for accuracy is .72 and .90 for speed (Blomert & Vaessen, 2009).

Phoneme awareness

PA was assessed with a phoneme deletion task, containing 23 orally presented pseudowords with a CVC or CCVCC structure. From these pseudowords, children were asked to delete one of the consonants (e.g., /FOT/ without /F/ makes /OT/) with a maximum reaction time of 15 seconds. Both accuracy and speed (mean reaction time from stimulus offset) were scored. Items with reaction times lower than 300 ms were not included. A score for speed was not calculated if fewer than 5 items were answered correctly. Internal consistency is .76 for accuracy and .90 for speed (Blomert & Vaessen, 2009).

Naming speed

Naming speed was assessed as the rapid naming of letters (f, k, r, s, t) and digits (1, 4, 5, 6, 8). Each subtest consisted of two sheets of 15 items each. Children were asked to name all items as quickly and accurately as possible, with a maximum reaction time of 35 seconds. As errors are very uncommon, only the response time (seconds) was registered. A separate score was calculated for letters and digits, consisting of the mean reaction time for the two sheets. Reaction times lower than 1 second were excluded. Test–retest reliability is .80 for letters and .83 for digits (Blomert & Vaessen, 2009).

Memory span

The 3DM contains three tasks to assess memory span, two are verbal and one is nonverbal. To assess verbal memory span children were presented with strings of phonemes (all consonants) and syllables (CVC or CCV) that they were asked to repeat in the same order. For the nonverbal memory span task children were asked to reproduce the order in which four white squares lit up. In each task, 13 strings were presented; two strings of two items, three strings of three, four and five items, and two strings of six items. Accuracy was scored per item in the string. Internal consistency was .61 for phonemes, .73 for syllables, and .63 for the nonverbal task (Blomert & Vaessen, 2009).

Intelligence

Intelligence was assessed with the Dutch version of the Wechsler Intelligence Scale for Children, third edition (Kort et al., 2005; $r = .85-.93$), which is an individually administered intelligence test for children aged six to 16 years. The test included ten subtests that measured different intellectual abilities. Four composite scores, total IQ, verbal IQ, performance IQ and processing speed ($M = 100$; $SD = 15$), were included.

5.6 | Analyses

Given the categorical outcomes of Grade 1 measures (text-level reading, vocabulary), group comparisons were made through chi-squared analyses on the distributions of poor performers (≤10th percentile). Given the assumed skewness of the literacy and phonological measures at the time of diagnosis, group comparisons were made through chi-squared analyses. We distinguished whether the child did or did not belong to the lowest 10% (or had a standard score ≤ 6). For the IQ measures, we conducted ANOVAs.

Alpha levels were set at .05. For skills that were assessed with several measures, alpha levels were adjusted by dividing the alpha value of .05 through the number of outcome measures. As multiple measures were only available

for the time of diagnosis, adjusted alpha levels were only applied there. The adjusted values are $p < .0125$ for the four-word reading measures; $p < .0167$ for the three-word spelling measures, $p < .008$ for the six phonological skills measures, and $p < .0167$ for the thee memory measures.

6 | RESULTS

6.1 | Early vocabulary and text-level literacy skills (Grade 1)

6.1.1 | Vocabulary and text-level literacy skills end Grade 1

The percentage of poor performers (≤10th percentile) on the tasks at the end of Grade 1 is presented in Table 1. The distribution of performance across the five levels of the text reading, reading comprehension and vocabulary skills end Grade 1 for the three groups is presented in Figure S1. The percentage of poor performers in Table 2 corresponds to those obtaining bottom scores in Figure S1.

In terms of word-level literacy measures, chi-squared analyses show that the LE group differs from the early-diagnosed and LI groups in the percentage of poor performers (Table 2), confirming both the status of the LE and the LI groups. Furthermore, the analyses show overall significant differences between the three groups for text reading and reading comprehension, but not for vocabulary. Follow-up comparisons (Table 2, right-hand columns) show that the percentages of poor performers in the LI and the early identified group did not differ from text reading (but marginally so, $p = .052$) and reading comprehension. The LI group did not differ from the LE group on text reading, but the LI group did contain more children performing poorly on reading comprehension. The percentages of children at or below the 10th percentile in the LE and the early-identified group differed for text reading and reading comprehension, with fewer children in the LE group obtaining such low outcomes on these tasks.

6.2 | Literacy and cognitive skills at the diagnostic examination

The percentage of children performing poorly within each group (≤10th percentile or a standard score ≤ 6 on the literacy, literacy-related and memory tasks) at the diagnostic examination is presented in Table 3. Chi-squared statistics were conducted to examine the differences across groups in the percentages of weak performers (Table 3, right-

TABLE 2 Percentages of children performing at or below the 10th percentile for the early, late-identified (LI) and late-emerging (LE) groups at the end of first grade

	Means			Chi-square			
	Early	LI	LE	Overall	Early versus LI	LI versus LE	Early vs LE
Word reading	87.1	88.9	21.1	84.91***	ns	44.39***	70.20***
Word spelling	33.6	48.9	11.3	16.60***	ns	16.83***	9.26**
Text reading	57.1	40.0	27.5	13.36**	ns	ns	12.38**
Reading comprehension	30.5	36.9	10.0	9.66**	ns	9.17**	7.39**
Vocabulary	10.0	26.3	12.5	3.30	-	-	-

Note: *$p < .05$. **$p < .01$; ***$p < .001$. For text reading, reading comprehension and vocabulary data was available for a subset of the children (see Figure S1).
Abbreviation: ns, not significant.

hand columns). For completeness' sake, the mean scores of the literacy skills (word reading, spelling) and the risk (phonology) and protective factors (memory) at the time of diagnosis are presented in Table S1.

The percentage of children performing poorly on word reading and spelling-to-dictation is high in all three groups. This is expected, as all children had been formally diagnosed with dyslexia and their literacy performance had to meet the nationally set criteria (see Method). There are no overall differences in this percentage of poor performance on one-minute word reading fluency and pseudoword reading. For word reading accuracy and speed, there are overall differences. Follow-up analyses show that the LE group has a lower percentage of children performing poorly than the early-diagnosed group on both accuracy and speed. The group also contains fewer children with a poor accuracy score than the LI group. There are no differences between the early and LI groups. With respect to spelling, there are no overall differences between the percentages of poor performers on any of the three measures.

In terms of risk and protective factors, there are no overall differences between the groups in the percentages of poor performers on phonological and memory tasks. ANOVAs on the mean IQ outcomes (Table 4) only showed a significant effect on processing speed. Tukey HSD posthoc tests indicated that the LI group obtained significantly lower processing speed outcomes ($p = .016$) than the early group, but there were no differences between the LI and LE groups ($p = .323$) and between the early-diagnosed and LE groups ($p = .448$). Although the early-diagnosed group also showed numerically higher scores on total IQ, verbal IQ and performance IQ, these differences were not significant.

TABLE 3 Percentages of poorly performing children for the early, late-identified (LI) and late-emerging (LE) groups at the time of diagnosis

	Means			Ch-square			
	Early	LI	LE	Overall	Early versus LI	LI versus LE	Early versus LE
Word reading							
Word reading accuracy	77.6	75.6	50.0	14.18*	ns	6.78*	13.07*
Word reading speed	96.6	93.3	75.9	19.05*	ns	ns	17.42*
Sight word reading	96.6	93.3	92.5	ns	-	-	-
Decoding	94.4	91.1	83.3	ns	-	-	-
Spelling							
Spelling recognition acc	68.1	66.7	51.9	ns	-	-	-
Spelling recognition speed	55.2	66.7	63.0	ns	-	-	-
Spelling-to-dictation	97.4	88.9	88.9	ns	-	-	-
Phonology							
Letter-sound identific acc	47.4	33.3	46.3	ns	-	-	-
Letter-sound identific speed	47.4	46.7	42.6	ns	-	-	-
PA accuracy	62.1	62.2	63.0	ns	-	-	-
PA speed	48.0	73.8	57.1	ns	-	-	-
Naming speed letters	52.6	46.7	46.3	ns	-	-	-
Naming speed digits	57.8	46.7	37.0	ns	-	-	-
Memory							
Phonemes	45.7	44.4	57.4	ns	-	-	-
Syllables	41.4	40.0	27.8	ns	-	-	-
Nonverbal	14.7	28.9	31.5	ns	-	-	-

Note: Alpha-values were adjusted to account for multiple testing: word reading measures $p < .0125$; word spelling measures $p < .0167$ phonological skills measures, $p < .008$ and memory measures $p < .0167$. Scores are percentile scores, except for sight word reading and decoding, which refer to standard scores.

Abbreviations: *, significant; −, not tested; ns, not significant.

TABLE 4 Mean outcomes on intelligence for the early, late-identified (LI) and late-emerging (LE) groups at the time of diagnosis

	Early	LI	LE	F
Total IQ	102.48 (11.9)	98.20 (9.75)	100.78 (11.7)	2.305
Verbal IQ	103.46 (12.1)	99.69 (9.8)	101.43 (11.4)	1.874
Performance IQ	100.83 (12.4)	97.07 (10.8)	99.81 (12.1)	1.598
Processing speed[+]	100.17 (11.9)$_a$	93.95 (14.1)$_a$	97.65 (11.5)	3.964*

Note: *p < .05; Subscripts refer to posthoc testing for significant main effects: they indicate that group outcomes differ significantly from each other in these instances.
[+]Data is available for 111 early, 42 late-identified and 51 late-emerging children.

7 | DISCUSSION

In this study, we evaluated whether LI children with dyslexia differed from early-diagnosed children and children with LE dyslexia on measures of literacy, correlates of literacy and possible protective factors. Randomly selected case files of formal diagnoses of dyslexia were categorized as early diagnosed (Grades 1–3) and late diagnosed (Grade 4–6). The late-diagnosed group was divided into LI dyslexia (45% of the late-diagnosed sample), referring to children whose school literacy performance in Grades 1 and 2 already warranted referral to an institute for special learning disabilities, and LE dyslexia (55% of the late-diagnosed sample).

7.1 | Similarities and differences between the early-identified, late-identified and late-emerging groups

At the end of Grade 1, the LI group resembled the early-diagnosed group in word-level literacy, the selection criterion, but also in text reading and reading comprehension. As reading comprehension can be argued to largely reflect word-reading ability at this early phase of literacy instruction, the reading comprehension findings match those of the poor word reading skills of the groups. At the time of the diagnosis, word reading and spelling deficits were also similar in both groups. In contrast, the LE group contained fewer children with word reading accuracy and speed deficits at the time of diagnosis. These findings corroborate previous studies reporting more severe reading deficits for early than LE dyslexia (Bazen et al., 2020; Catts et al., 2012; Compton et al., 2008; Leach et al., 2003; Lipka et al., 2006; Torppa et al., 2015). This pattern seems robust: Our poor performance criterion was ≤10th percentile. Other studies used <25th percentile (Etmanskie et al., 2016; Lipka et al., 2006), <16th percentile (Catts et al., 2012; Compton et al., 2008; Leach et al., 2003) or < 10th percentile (Bazen et al., 2020; Torppa et al., 2015). The pattern is the same regardless of the criterion applied.

At the time of diagnosis, all three groups showed relatively low performance on tasks tapping the phonological deficit and there were no overall group differences These findings match previous studies that did not find differences between the early and LE groups on risk factors once literacy difficulties had been reported (Bazen et al., 2020; Leach et al., 2003).

7.2 | Compensation in late-emerging dyslexia?

With respect to the LE group, our findings do not agree easily with the assumption that dyslexia is late-emerging due to the presence of protective factors in the late-dyslexia group. Although the LE group showed less severe word reading difficulties at the time of diagnosis on some measures, they did not show outspoken strengths in phonology

or other cognitive abilities. This finding agrees with previous studies on LE dyslexia (Bazen et al., 2020; Leach et al., 2003; Lipka et al., 2006; Torppa et al., 2015). Similar to findings by Torppa et al. (2015), our LE group does not have a distinct poor early literacy profile, making the eventual dyslexia of these children difficult to predict. Possibly, the transition from accuracy to fluency is most difficult for this group of children.

7.3 | Why were children with late-identified dyslexia identified late?

An important finding of the current study is the existence of LI children in the random sample (45%). In terms of word reading and spelling outcomes at the end of Grade 1, there are no clear reasons why these children were missed, as they met the criteria for referral in the early grades. The percentages of poor performers in the LI and early-diagnosed groups also did not differ on text reading, but note that this finding points to marginal significance ($p = .052$). We therefore cannot establish whether (all) educators took text reading outcomes into account for deciding on children's referrals, a finding that was reported for German teachers (Schmitterer & Brod, 2021). Reliance on text reading would run counter to the school-literacy protocols and psychologist dyslexia protocols, which point to the importance of word-level outcomes (Gijsel et al., 2011; SDN et al., 2016).

The late referral of the LI group could also hypothetically be due to these children being better able to compensate for their literacy difficulties. However, this does not seem to be the case: literacy outcomes were poor, so the core deficit was not compensated. Furthermore, at the time of diagnosis, the LI group had numerically lower and not higher mean outcomes than the early-identified group on IQ measures, with significantly lower processing speed. Additionally, the LI group had a higher percentage of children showing poor performance on non-verbal memory than the early-identified group. Overall, these findings do not indicate that the late referral of the LI groups is due to compensation.

There is also no strong evidence for the reverse option, that lower outcomes on protective factors hampered the recognition of the literacy abilities in the LI group. The finding that the LI group contained a substantial proportion of children showing both low and high performance on vocabulary indicates that reasons for non-referral in the early grades are not straightforward. The slightly numerically lower IQ outcomes and the significantly lower processing speed outcomes attested at the time of diagnosis could also have surfaced in general learning ability in earlier grades, given the stability of IQ (Beaver et al., 2013). Clearly, this line of reasoning is highly speculative, as data on vocabulary, IQ and memory is not present for both time points, as we have no information on broader learning outcomes at school, and we have no information on the teachers' reasons for their late referral of the children.

7.4 | Implications

Whatever the background of LI dyslexia, it seems that attention needs to be devoted to the correct application of the response-to-intervention framework and the availability of the means to implement this approach to eliminate the existence of LI children, who can be taken to be instructional casualties. These children have only been provided with the appropriate dyslexia treatment late. This might have had a detrimental impact on their literacy development as well as their wellbeing, an area that is vulnerable for people with dyslexia (e.g., Francis, Caruana, Hudson, & McArthur, 2019; Gibby-Leversuch, Hartwell, & Wright, 2019).

Our findings not only relate to the existence of LI dyslexia (Barbiero et al., 2012; Graham, White, Tancredi, Snow, & Cologon, 2020), but also to research findings that point to the challenges that educators face in identifying and shaping literacy interventions: Studies have reported that schools, school psychologists, and teachers were not always sufficiently familiar with the adequate application of the response-to-intervention framework (e.g., Kratochwill, Volpiansky, Clements, & Ball, 2007; Vujnovic et al., 2014). It has also been reported that not all children requiring more intensive literacy instruction actually receive this help (Graham et al., 2020) and that teachers'

identification of reading difficulties in children does not always take place on the right information (Schmitterer & Brod, 2021). Consequently, the findings relate to the recent call of devoting more time and effort to translating research to education (Seidenberg, Cooper Borkenhagen, & Kearns, 2020) and supporting both educators' knowledge of literacy development and disorders and classroom implementation (e.g., Piasta, Connor, Fishman, & Morrison, 2009).

This attention to the application of the response-to-intervention framework should, on the one hand, limit the late referral of children for dyslexia diagnosis and specialized treatment and, on the other hand, ensure that children are not referred too early. Torppa et al. (2015) found that there was an instability in dyslexia between Grades 2 and 8. Specifically, in a larger sample of children ($n = 182$), 55 qualified as poor readers in one of the grades. Of these 55, there were 15 children (15/55 = 27%) whose literacy problems had disappeared at Grade 8. This means that a general sample of children in Grades 1 and 2 will consist of children who show a dip in literacy outcomes, but whose delay is not persistent (see also Vloedgraven et al., 2010). These children do not require specialized treatment outside school. There should thus be a balance between referring too early and too late, by providing the required levels of support (Tiers 1, 2 or 3) or help at school.

7.5 | Limitations and suggestions for further research

Our study is qualified by the limitation of missing information on the decisions for (late) referral. We do not know whether the LI children were receiving more specialized help at school prior to their late referral, for instance, or whether the schools the children attended differed in the speed or ease with which referrals were made. Second, we have not included all risk factors associated with dyslexia. For instance, we have not looked at visuospatial skills or visual attention span, which have also been associated with dyslexia (e.g., Bosse, Tainturier, & Valdois, 2007; Franceschini, Gori, Ruffino, Pedroll, & Facoetti, 2012; Peng & Fuchs, 2014). We cannot rule out that children's abilities on such measures might lead to teachers' interpretations of children's development and might possibly also influence the identification of literacy difficulties. Third, we cannot draw any conclusions about the prevalence of LI dyslexia, as we evaluated a random sample of case files of one institute for specific learning difficulties in the Netherlands. Although this centre has several locations across the Netherlands, we cannot generalize this to a general LI percentage. The percentage of LI children in this study was high (45%); there is no a priori reason why this percentage would be substantially lower in other centres. Nevertheless, a more comprehensive collection of case files is needed to provide more precise information on this matter. One avenue for further research is to evaluate the existence of LI dyslexia in other countries, specifically those also applying a response-to-intervention approach.

In sum, our diagnosis-based study confirms and extends the findings of previous research-based studies that dyslexia can be late-emerging. This calls for more attention to the way dyslexia surfaces across the lifespan and the required support. Our findings also indicate that instructional casualties occur, as a substantial number of children in this study were identified late with dyslexia. The literacy outcomes of these children resemble those of early-diagnosed children and are poorer than those of children with LE dyslexia. The findings thus call for further attention to shaping and supporting the literacy curriculum at schools. This proverbial stitch in time should prevent late identification from occurring as well as optimize general literacy instruction.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Elise H. de Bree https://orcid.org/0000-0001-5258-7518
Madelon van den Boer https://orcid.org/0000-0002-3330-2916

REFERENCES

Barbiero, C., Lonciari, I., Montico, M., Monasta, L., Penge, R., Vio, C., ... and for the Epidemiology of Dyslexia of Friuli Venezia Giulia working group (FVGwg). (2012). The submerged dyslexia iceberg: How many school children are not diagnosed? Results from an Italian study. *PLoS ONE*, *7*(10), e48082. https://doi.org/10.1371/journal.pone.0048082

Bazen, L., van den Boer, M., de Jong, P. F., & de Bree, E. H. (2020). Early and late diagnosed dyslexia in secondary school: Performance on literacy skills and cognitive correlates. *Dyslexia*, *26*, 359–376. https://doi.org/10.1002/dys.1652

Beaver, K. M., Schwartz, J. A., Connolly, E. J., Nedelec, J. L., Al-Ghamdi, M. S., & Kobeisy, A. N. (2013). The genetic and environmental architecture to the stability of IQ: Results from two independent samples of kinship pairs. *Intelligence*, *41*(5), 428–438. https://doi.org/10.1016/j.intell.2013.06.012

Blomert, L., & Vaessen, A. (2009). *3DM: Cognitieve analyse van lezen en spellen*. [3DM: Cognitive analysis of reading and spelling]. Amsterdam: Boom test uitgevers BV.

Bosse, M. L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition*, *104*(2), 198–230.

Brus, B. T., & Voeten, M. J. M. (1999). *Eén minuut test*. [One minute test]. Amsterdam, The Netherlands: Hartcourt Test Publishers.

Catts, H. W., Compton, D., Tomblin, J. B., & Bridges, M. S. (2012). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology*, *104*, 166–181. https://doi.org/10.1037/a0025323

Compton, D. L., Fuchs, D., Fuchs, L. S., Elleman, A. M., & Gilbert, J. K. (2008). Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. *Learning and Individual Differences*, *18*, 329–337. https://doi.org/10.1016/j.lindif.2008.04.003

COTAN. (2010). https://www.cotandocumentatie.nl/

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

de Wijs, A., Kamphuis, F., Kleintjes, F., & Tomesen, M. (2010). *Leerling- en onderwijsvolg-systeem. Spelling Groep 3 t/m 6*. [Student and education tracking system. Spelling Grade 1 through 4]. Arnhem: Cito.

Etmanskie, J. M., Partanen, M., & Siegel, L. S. (2016). A longitudinal examination of the persistence of late-emerging reading disabilities. *Journal of Learning Disabilities*, *49*, 21–35. https://doi.org/10.1177/0022219414522706

Feenstra, H., Kleintjes, F., Kamphuis, F., & Krom, R. (2010). *Leerling- en onderwijsvolg-systeem. Begrijpend lezen. Groep 3 t/m 6*. [Student- and educational tracking system: Reading comprehension. Grade 1 through 4]. Arnhem: Cito.

Franceschini, S., Gori, S., Ruffino, M., Pedroll, K., & Facoetti, A. (2012). A causal link between spatial attention and reading acquisition. *Current Biology*, *22*, 814–819.

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

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, *41*, 93–99. https://doi.org/10.1598/RRQ.41.1.4

Geelhoed, J., & Reitsma, P. (1999). *PI-dictee* [PI-dictation]. Lisse, The Netherlands: Swets Test Publishers.

Gibby-Leversuch, R., Hartwell, B. K., & Wright, S. (2019). Dyslexia, literacy difficulties and the self-perceptions of children and young people: A systematic review. *Current Psychology*, *1-18*, 5595–5612. https://doi.org/10.1007/s12144-019-00444-1

Gijsel, M., Scheltinga, F., Druenen, M. V., & Verhoeven, L. (2011). *Protocol Leesproblemen en Dyslexie voor groep 3* [Protocol for reading problems and dyslexia in Grade 1]. Nijmegen, The Netherlands: Expertisecentrum Nederlands.

Graham, L. J., White, S. L. J., Tancredi, H. A., Snow, P. C., & Cologon, K. (2020). A longitudinal analysis of the alignment between children's early word-level reading trajectories, teachers' reported concerns and supports provided. *Reading and Writing*, *33*, 1895–1923. https://doi.org/10.1007/s11145-020-10023-7

Haft, S. L., Myers, C. A., & Hoeft, F. (2016). Socio-emotional and cognitive resilience in children with reading disabilities. *Current Opinion in Behavioral Sciences*, *10*, 133–141. https://doi.org/10.1016/j.cobeha.2016.06.005

Hulme, C., Nash, H. M., Gooch, D., Lervåg, A., & Snowling, M. J. (2015). The foundations of literacy development in children at familial risk of dyslexia. *Psychological Science*, *26*, 1877–1886. https://doi.org/10.1177/0956797615603702

Kort, W., Schittekatte, M., Bosmans, M., Compaan, E. L., Dekker, P. H., Vermeir, G., & Verhaege, P. (2005). *Wechsler intelligence scale for children III-NL (WISC III-NL)*. Amsterdam, The Netherlands: Pearson.

Kratochwill, T. R., Volpiansky, P., Clements, M., & Ball, C. (2007). Professional development in implementing and sustaining multitier prevention models: Implications for response to intervention. *School Psychology Review*, *36*(4), 618–631.

Krom, R., Jongen, I., Verhelst, N., Kamphuis, F., & Kleintjes, F. (2010). *DMT en AVI. Groep 3 tot en met 8*. [DMT and AVI. Grade 1 through 6]. Arnhem: Cito.

Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology*, *95*, 211–224. https://doi.org/10.1037/0022-0663.95.2.211

Lipka, O., Lesaux, N. K., & Siegel, L. S. (2006). Retrospective analyses of the reading development of Grade 4 students with reading disabilities: Risk status and profiles over 5 years. *Journal of Learning Disabilities*, *39*, 364–378. https://doi.org/10.1177/00222194060390040901

Lyon, G.R. Learning disabilities and early intervention strategies. (2002). Retrieved from http://reidlyon.com/edpolicy/11-LEARNING-DISABILITIES-AND-EARLY-INTERVENTION-STRATEGIES.pdf.

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

Moll, K., Ramus, F., Bartling, J., Bruder, S., Neuhoff, N., & Valdois, S. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction*, *29*, 65–77.

Moojen, S. M. P., Gonçalves, H. A., Bassôa, A., Navas, A. L., de Jou, G., & Miguel, E. S. (2020). Adults with dyslexia: How can they achieve academic success despite impairments in basic reading and writing abilities? The role of text structure sensitivity as a compen-satory skill. *Annals of Dyslexia*, *70*(1), 115–140. https://doi.org/10.1007/s11881-020-00195-w

Netherlands Institute for Social Research (NISR) (2017). Status scores. https://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores.

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

NRD. Nationaal Referentiecentrum Dyslexie. *Protocol Dyslexie Diagnostiek en Behandeling 2.0*. [Protocol Dyslexia Treatment and Diagnosis 2.0]. (2013). Retrieved from https://www.nkd.nl/app/uploads/2017/10/protocol_dyslexie__diagnostiek_en_behandeling_2.0def.pdf

Peng, P., & Fuchs, D. (2014). A meta-analysis of working memory deficits in children with learning difficulties: Is there a difference between verbal domain and numerical domain? *Journal of Learning Disabilities*, *9*(1), 3–20. https://doi.org/10.1177/0022219414521667

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

Peterson, R. L., Pennington, B. F., Olson, R. K., & Wadsworth, S. (2014). Longitudinal stability of phonological and surface subtypes of developmental dyslexia. *Scientific Studies of Reading*, *18*(5), 347–362. https://doi.org/10.1080/10888438.2014.904870

Piasta, S. B., Connor, C. M., Fishman, B. J., & Morrison, F. J. (2009). Teachers' knowledge of literacy concepts, classroom practices, and student reading growth. *Scientific Studies of Reading*, *13*(3), 224–248. https://doi.org/10.1080/10888430902851364

Scheltinga, F., van der Leij, A., & Struiksma, C. (2009). Predictors of response to intervention of word reading fluency in Dutch. *Journal of Learning Disabilities*, *43*(3), 212–228. https://doi.org/10.1177/0022219409345015

Schmitterer, A. M. A., & Brod, G. (2021). Which data do elementary school teachers use to determine reading difficulties in their students? *Journal of Learning Disabilities*, *54*, 349–364. https://doi.org/10.1177/0022219420981990

SDN, de Bree, E. H., Henneman, K., Kleijnen, R., Loykens, E. H. M., Rolak, M., ... Wijnen, F. N. K. (2016). Dyslexie: Diagnostiek en behandeling: Brochure van de Stichting Dyslexie Nederland (SDN) [Dyslexia: diagnostics and treatment]. Retrieved from https://www.stichtingdyslexienederland.nl/images/publicaties/dyslexie_diagnostiek_en_behandeling_2016.pdf

Seidenberg, M. S., Cooper Borkenhagen, M., & Kearns, D. M. (2020). Lost in translation? Challenges in connecting reading science and educational practice. *Reading Research Quarterly*, *55*(S1), S119–S130. https://doi.org/10.1002/rrq.341

Snow, P. C. (2016). Elizabeth usher memorial lecture: Language is literacy is language - positioning speech-language pathology in education policy, practice, paradigms and polemics. *International Journal of Speech-Language Pathology*, *18*(3), 216–228.

Torppa, M., Eklund, K., van Bergen, E., & Lyytinen, H. (2015). Late-emerging and resolving dyslexia: A follow-up study from age 3 to 14. *Journal of Abnormal Child Psychology*, *43*, 1389–1401. https://doi.org/10.1007/s10802-015-0003-1

van Berkel, S., Hilte, M., Engelen, R., Kamphuis, F., Kleintjes, F., & Krom, R. (2010). *Leerling- onderwijsvolgsysteem. Woordenschat Groep 3 t/m 5*. [Student educational tracking system: Vocabulary Grade 1 through 3]. Arnhem: Cito.

van den Bos, K. P., Lutje Spelberg, H. C., Scheepstra, A. J. M., & de Vries, J. R. (1994). *De Klepel. Vorm A en B*. [Klepel non-word reading]. Amsterdam, The Netherlands: Pearson.

van Viersen, S., de Bree, E. H., & de Jong, P. F. (2019). Protective factors and compensation in resolving dyslexia. *Scientific Studies of Reading*, *23*, 1–17. https://doi.org/10.1080/10888438.2019.1603543

van Viersen, S., de Bree, E. H., Zee, M., Maassen, B., van der Leij, A., & de Jong, P. F. (2018). Pathways into literacy: The role of early oral language abilities and family risk for dyslexia. *Psychological Science*, *29*(3), 418–428. https://doi.org/10.1177/0956797617736886

van Viersen, S., Kroesbergen, E. H., Slot, E. M., & de Bree, E. H. (2016). High reading skills mask dyslexia in gifted children. *Journal of Learning Disabilities*, *49*, 189–199. https://doi.org/10.1177/0022219414538517

Vloedgraven, J. M. T., Keuning, J., & Verhoeven, L. T. W. (2010). Signalering van leesproblemen en dyslexie [Detection of reading problems and dyslexia]. In L. T. W. Verhoeven, F. N. K. Wijnen, & K. P. Bos (Eds.), *Zorg om Dyslexie (Care for Dyslexia)*. Maklu: Apeldooorn.

Vujnovic, R. K., Fabiano, G. A., Morris, K. L., Norman, K., Hallmark, C., & Hartley, C. (2014). Examining school psychologists' and teachers' application of approaches within a response to intervention framework. *Exceptionality*, *22*, 129–140. https://doi.org/10.1080/09362835.2013.865530

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: de Bree, E. H., van den Boer, M., Toering, B. M., & de Jong, P. F. (2022). A stitch in time...: Comparing late-identified, late-emerging and early-identified dyslexia. *Dyslexia*, *28*(3), 276–292. https://doi.org/10.1002/dys.1712