Schmidt, Claudia; Brandenburg, Janin; Busch, Jenny; Büttner, Gerhard; Grube, Dietmar; Mähler, Claudia; Hasselhorn, Marcus

Developmental trajectories of phonological information processing in upper elementary students with reading or spelling disabilities

Reading research quarterly 56 (2021) 1, S. 143-171

INTERNATIONAL LITERACY ASSOCIATION READING RESEARCH QUARTERLY

Developmental Trajectories of Phonological Information Processing in Upper Elementary Students With Reading or Spelling Disabilities

Claudia Schmidt
Carl von Ossietzky University, Oldenburg, Germany

Janin Brandenburg
DIPF | Leibniz Institute for Research and Information in Education, Frankfurt, Germany

Jenny Busch
Carl von Ossietzky University, Oldenburg, Germany

Gerhard Büttner
Goethe University Frankfurt and DIPF | Leibniz Institute for Research and Information in Education, Frankfurt, Germany

Dietmar Grube
Carl von Ossietzky University, Oldenburg, Germany

Claudia Mähler
University of Hildesheim, Germany

Marcus Hasselhorn
Goethe University Frankfurt and DIPF | Leibniz Institute for Research and Information in Education, Frankfurt, Germany

INTERNATIONAL LITERACY ASSOCIATION

Reading Research Quarterly, 56(1)
pp. 143–171 | doi:10.1002/rrq.299

ABSTRACT

Deficits in phonological information processing in upper elementary students with specific learning disabilities in reading or spelling may increase, decrease, or remain stable over time. The authors examined the development of phonological processing longitudinally in 209 students (109 with learning disabilities and 100 typically achieving; n = 127 boys) in Germany, from grade 3 to grade 5 (ages 8–11; mean age at recruitment = 8 years 6.78 months, SD = 5.39 months). Latent change score models revealed that the development of rapid automatized naming was best described as a decreasing deficit, whereas a persistent deficit in phonological awareness was observed. Differences between students with and without learning disabilities regarding the phonological loop increased over time. Further, there were no developmental differences as a function of reading versus spelling deficits. Theoretical and practical implications are discussed.

Although there is a substantial body of literature and research on phonological processing abilities in upper elementary students with reading or spelling difficulties (i.e., learning disabilities; LD), less is known about the developmental trajectories of these processes in these students. Previous longitudinal studies on LD focused mainly on developmental changes in reading and spelling (for an overview, see Pfost, Hattie, Dörfler, & Artelt, 2014) rather than the developmental changes in phonological processing. This lack of research is surprising given that problems in phonological processing are considered the core deficit of LD, so knowledge about how these phonological skills further develop in these students is crucial for a better theoretical understanding of LD. Reading difficulties are often found to be persistent, and several cross-sectional studies indicated that the same might be true for the underlying phonological processing deficit. However, cross-sectional studies do not allow conclusions about the developmental trajectories of phonological processing, that is, whether there may be changing relations in the manifestation of deficits in phonological processing. The purpose of the current longitudinal study was therefore to examine the developmental trajectories of rapid naming, phonological awareness, and the phonological loop (sometimes also referred to as verbal short-term memory) in a large sample of upper elementary German students, with and without LD, to determine how these skills develop over a period of three years.

The Importance of Phonological Information Processing for Reading and Spelling

The way young learners acquire reading and spelling skills is the topic of a number of developmental theories. Most of these theories describe the development of reading and spelling in young learners as proceeding in serial phases or stages (e.g., Ehri, 2005; Frith, 1986). According to Frith's (1986) theoretical framework, the development of literacy skills takes place in three phases. In the initial logographic phase, young learners recognize words only by their graphic characteristics. The alphabetical phase is then characterized by the use of phonological strategies. Young learners understand the arbitrary relation between the spoken and written language, that is, the relation between sequences of letters in the written language and the sounds that these letters represent. To read words, young learners typically sound out words and blend the sounds. Spelling, according to the alphabetic phase, is associated with young readers analyzing words with respect to their phonemes and the mapping of these phonemes to the corresponding graphemes. Repeated exposure to the same words consolidates an orthographic lexicon in which whole-word grapheme sequences are stored (i.e., orthographic representations).

The orthographic phase is reached when young readers no longer need to sound out known words but instead recognize words automatically by fast, direct, and instant access to their meanings without recourse to phonology. The orthographic lexicon is of special importance for spelling processes, as a systematic phoneme–grapheme mapping is not sufficient to write words orthographically correctly. As compared with English, the German orthography is considered relatively transparent, but orthographic regularity relates primarily to reading, not necessarily to spelling. The phoneme–grapheme relation (relevant in spelling processes) is most often less consistent than the grapheme–phoneme relation (relevant in reading processes). Most times, there are several possibilities for mapping a certain phoneme to a grapheme (combination), with only one possibility being orthographically correct; for example, [i:] could be spelled as *ie* as in *dieb* (thief), as *i* as in *igel* (hedgehog), as *ih* as in *ihm* (him), or as *ieh* as in *siehst* (see; see also Landerl & Wimmer, 2008). Consequently, orthographically correct spelling is the more difficult task in Germany.

Theories on the development of reading and spelling provide only descriptions of the way in which young learners' reading and spelling change with experience and do not explain the mechanism accounting for this change (Hulme & Snowling, 2009). According to the reading and spelling strategies described, phonological information processing—the use of phonological information when dealing with spoken or written language (Wagner & Torgesen, 1987)—provides a commonly adapted theoretical framework for explaining individual differences in literacy acquisition. Wagner and Torgesen (1987) subsumed phonological awareness, the phonological loop, and the retrieval of phonological codes from long-term memory (most often assessed by rapid automatized naming (RAN) tasks and, thus, hereafter referred to with this term) under phonological processing. The theoretical role of these three phonological components for the acquisition of reading and spelling are outlined next.

The Relation of RAN to Reading and Spelling

Naming speed is the speed of the pronunciation of the names of familiar stimuli (e.g., objects, colors, letters, digits). RAN tasks are used to measure phonological recoding in lexical access and can thus be defined as the speed that is required to map written symbols to their corresponding phonological equivalents (e.g., phonological codes associated with single letters, morphemes, or words) stored in long-term memory (Denckla & Rudel, 1974; Wagner & Torgesen, 1987).

RAN predicts reading ability in both opaque and transparent orthographies (Georgiou, Parrila, & Liao, 2008), even after controlling for other prominent reading-related variables such as phonological awareness, the phonological loop, or prior reading ability (Kirby, Georgiou, Martinussen, & Parrila, 2010). However, there are differences in the way RAN relates to each of the various reading processes, being generally more related to reading fluency than to accuracy or reading comprehension (Araújo, Reis, Petersson, & Faísca, 2015; Landerl et al., 2019). A number of theoretical explanations exist as to why reading and RAN are related (for a comprehensive overview, see Kirby et al., 2010), with most of them claiming that reading and RAN rely partly on similar domain-general processes. For example, in a study by Papadopoulos, Spanoudis, and Georgiou (2016), different theoretical accounts of the RAN–reading relation were contrasted simultaneously. The authors revealed that the relation is both direct (i.e., RAN was a unique predictor of oral reading fluency) and indirect (through the mediation of phonological awareness and orthographic processing). This complex pattern of effects indicates that RAN is crucial for both the actual reading performance and some of the basal processes underlying reading.

The relation of RAN and spelling is less well understood and has attracted far less attention than has the RAN–reading relation. According to Wolf and Bowers (1999), RAN is strongly involved in the process of building up word-specific orthographic representations in memory (i.e., they assume nonphonological processes indexed by RAN, a position contrary to the phonological processing argument of Wagner and Torgesen, 1987). Following this, and because orthographically correct spelling (especially the spelling of irregular words) is dependent on accessing those orthographic representations (Martin & Barry, 2012), RAN is supposed to have an impact on spelling, too. Indeed, research has found associations between RAN and spelling, with RAN being an independent predictor (Furnes & Samuelsson, 2011; Georgiou, Torppa, Manolitsis, Lyytinen, & Parrila, 2012; Savage, Pillay, & Melidona, 2008).

The Relation of Phonological Awareness to Reading and Spelling

Phonological awareness is the conscious focusing on and handling of phonological units such as syllables, onsets, and phonemes (Castles & Coltheart, 2004). The awareness of phonemes helps the novice reader and writer understand the correspondence between sounds and letters, which is especially crucial in the alphabetical phases of reading and spelling acquisition. Learning to read and the reading of unknown words involve the segmentation of letter strings, an alignment with corresponding phonemes, and blending of these phonemes together to read a whole word. Further, it is assumed that phonological awareness might be relevant for the buildup of stable orthographical representations that enable young learners to read fluently and write orthographically correctly (Ehri, 1995; Perfetti, 1992; Share, 2008), as it is assumed, especially for transparent orthographies such as German, that the orthographic representations are organized at the phonemic level (Goswami, 1997).

Whereas phonological awareness typically remains an important predictor for reading and spelling in opaque languages (e.g., English) far beyond elementary school, its importance seems to diminish in transparent orthographies after the first years of schooling (Furnes & Samuelsson, 2011; Georgiou et al., 2012). It has been argued that literacy acquisition demands less phonological awareness within consistent orthographies, which might be due to the combination of consistent grapheme–phoneme correspondence and systematic phonics teaching that enables young learners to become good and accurate readers (cf. Landerl & Thaler, 2006; Wimmer, 1993). The result of phonological awareness not contributing to reading or spelling after the first years of schooling in transparent orthographies may also be a consequence of the nature of the tasks used (Caravolas, Volin, & Hulme, 2005; Nikolopoulos, Goulandris, Hulme, & Snowling, 2006). For example, Caravolas and colleagues (2005) and Babayiğit and Stainthorp (2011) found phonological awareness to be predictive for reading and spelling in both the transparent Czech and Turkish orthographies and the opaque English orthography in older students when more demanding tasks (e.g., phoneme elision or spoonerism tasks) are used. Finally, there is further evidence from transparent orthographies that phonological awareness still contributes to higher order reading skills such as reading comprehension (Engen & Høien, 2002).

The Relation of the Phonological Loop to Reading and Spelling

The phonological loop is a subsystem of Baddeley's (1986) working memory model and is responsible for the short-term storage and processing of phonological information. The phonological loop is a modality-specific and capacity-limited storage system that consists of two components. First, within the phonological store, phonological information can be maintained for approximately 1.5–2 seconds before information decays (Baddeley, Thomson, & Buchanan, 1975). Second, there is the inner rehearsal process, an active mental repetition of phonological information to prevent information decay.

With respect to reading, the phonological loop is involved in the acquisition of letter–sound mapping rules and in the verbal recoding of visually presented letters by activating phonological codes through subvocal speech (i.e., articulatory coding; Baddeley, 1986). Efficient articulatory coding enables beginning readers to maintain accurate phonological representations while the blending process is executed and a particular word is read (Alloway, Gathercole, Willis, & Adams, 2004; Torgesen, Wagner, Rashotte, Burgess, & Hecht, 1997). Moreover, spelling requires an accurate perception of the heard words and the ability to segment and maintain phonemes and to map them to their corresponding graphemes. Based on these processes, not only is phonological awareness assumed to be critical for successful spelling, but so too is the phonological loop (cf. Moll et al., 2014).

A meta-analysis by Melby-Lervåg, Lyster, and Hulme (2012) revealed that the mean correlation between the phonological loop and reading was medium high ($r = .34$). Likewise, in another meta-analysis, O'Shaughnessy and Swanson (1998) reported that young students with LD showed lower verbal memory spans than did normal readers, with an effect size of 0.61 (Hedges's g). In a comprehensive study, Babayiğit and Stainthorp (2011) examined the predictive role of verbal short-term memory (and other cognitive skills) in emergent literacy in elementary school students acquiring a transparent orthography. Besides RAN and phonological awareness, the phonological loop emerged as a strong predictor of word decoding and text-reading speed, whereas its contribution to nonword reading was lower yet still stastitically significant. Likewise, the phonological loop explained reliable variance in both word and sentence spelling.

The Phonological Deficit Theory of LD

Due to its crucial involvement in the development of literacy skills, poor phonological processing is considered the core deficit of LD, as expressed in the phonological deficit theory (for a review, see Vellutino, Fletcher, Snowling, & Scanlon, 2004). It has been well established that individuals with LD perform poorly on tasks measuring RAN, phonological awareness, and the phonological loop (Melby-Lervåg et al., 2012; Swanson, Zheng, & Jerman, 2009; Wolf, Bowers, & Biddle, 2000). It is assumed that these phonological impairments stem from underspecified phonological representations (the phonological representation hypothesis; for an overview, see Snowling, 2000); that is, learners with LD may take longer to retrieve phonological representations (RAN deficit), maintain them (phonological loop deficit), and struggle with analyzing and operating on them (phonological awareness deficit). In contrast and based on other studies, Ramus and colleagues (Ramus, Marshall, Rosen, & van der Lely, 2013; Ramus & Szenkovits, 2008) argued that the phonological representations of individuals with LD may be intact but not the access to these representations: "The phonological deficit surfaces only as a function of certain task requirements, notably short-term memory, conscious awareness, and time constraints" (Ramus & Szenkovits, 2008, p. 139).

Besides the ambiguity of the nature of the phonological deficit in young learners with LD (i.e., whether it is a deficit in the phonological representation or a deficit in the processing of phonological representation), the phonological deficit theory is widely accepted. Yet, the theory raises no clear assumptions about the developmental trajectories that the phonological processing deficits take in learners with LD. Despite its importance for theoretical explanations of LD, this question has received relatively little attention in research because previous longitudinal studies on LD have mainly been concerned with developmental changes in academic achievement, such as reading and spelling (for an overview, see Pfost et al., 2014), rather than the development of potential underlying deficits in phonological processing.

Developmental Trajectories: Deficit Model Versus Developmental Lag Model

From a longitudinal perspective, (at least) two different developmental models have been considered in research focused on the academic development in young learners with LD (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Pfost et al., 2014). Both models rely on the assumption that the relation between phonological processing and reading is bidirectional and causally related, as several studies revealed that phonological processing affects subsequent reading and spelling, and vice versa (e.g., Bast & Reitsma, 1998; Castles & Coltheart, 2004; Torgesen et al., 1997).

One theoretical perspective is the developmental lag model (Francis et al., 1996; Stanovich, Nathan, & Zolman, 1988). At the heart of this model is the assumption that learners differing in literacy skills vary in their developmental rate with respect to their underlying cognitive and phonological skills: It is assumed that these skills will develop over time. That is, as learners get older, they may overcome their initial cognitive deficits and catch up in their proficiency levels. A strong version of the developmental lag model would even claim that learners with LD might catch up with their average-achieving peers. Consequently, performance differences between learners with and without LD should be weaker in older learners than in younger learners, or even absent.

In contrast, the (developmental) deficit model (Francis et al., 1996; Stanovich & Siegel, 1994) assumes that young students with LD perform poorly on reading and phonological tasks because skills have not developed adequately. According to Denckla (1979), the deficit model implies an atypical underlying cognitive or neurological structure in learners with LD. Within this approach, two assumptions are prevalent. First, there is the assumption of high interindividual rank-order stability. Importantly, this does not imply a stagnation of reading or spelling development and related cognitive skills in students with LD. Instead, those students further develop their skills but usually continue to lag behind their typically achieving peers, so there remains a constant gap between the two groups.

The second assumption is that the differences between learners with LD and without LD might increase over time. Those increasing differences might be attributed either to a cascading deficit model or to a "rich get richer and poor get poorer" effect (i.e., Matthew effect; Stanovich, 1986), respectively. According to the former model, a lack of phonological processing (or likewise, a lack of reading and spelling skills) in early development leads to a lack of effectively using the learning opportunities provided in school. This, in turn, leads to reduced (i.e., flatter) developmental trajectories in these young learners as compared with their same-age peers. In other words, the performance of learners with LD still improves throughout development, but their developmental rates get smaller as compared with the developmental rates of their peers. According to the "rich get richer" effect (Stanovich, 1986), differences in reading and other reading-related skills between young learners with and without LD increase as students without LD show cumulative performance gains over time. Likewise, the "poor get poorer" effect would suggest that students with LD show progressively worsening of reading and reading-related skills. In other words, throughout development, the absolute skill level of these students would decrease because their developmental rates are negative. Different and partly interrelated mechanisms contribute to the described increasing group differences: First, high-performing readers are more intrinsically motivated to read and write, and hence, they receive more practice in these skills than do low-performing readers. Second, because of their larger knowledge base, skilled readers take advantage of the learning opportunities at school faster and more efficiently than do less skilled readers. Third, some cognitive skills (e.g., phonological processing) share a reciprocal relation with reading, and poor reading skills may therefore inhibit growth in these cognitive skills. For example, young readers with early deficits in phonological processing develop LD (i.e., they read less fluently, comprehend less, and misspell words), leading to less motivation to read, possibly meaning that they learn and comprehend less, in turn impeding their reading and cognitive development further (i.e., the gap widens; Bast & Reitsma, 1998; Stanovich, 1986; Wagner, Torgesen, & Rashotte, 1994).

The Development of Phonological Processing in Students With LD

Against this backdrop, the question arises as to whether the cognitive mechanisms that underlie LD can also be described based on the developmental pattern observed regarding academic development in students with LD. This information would be relevant for intervention and diagnostics: If, for example, differences in phonological processing abilities between students with and without LD decrease over time, these phonological processes could only be a relevant diagnostic feature at a very specific age, when the deficit is still present (i.e., developmentally limited diagnostic features as supposed by Stanovich, 1986). Further, it could be argued that training in phonological skills would not be necessary or would only accelerate a development that would happen anyway (Francis et al., 1996). In contrast, if a persistent deficit in students with LD was to be expected, these students could be identified with the help of phonological processing irrespective of age and might benefit from intervention of phonological processing at any time.

However, extant empirical evidence regarding the development of phonological processing in young students with LD is still scarce in comparison with empirical evidence for the development of literacy skills. Nevertheless, one major finding from cross-sectional studies is that deficits in phonological processing continue to characterize older students or adults with LD as compared with typically reading peers (for an overview, see Melby-Lervåg et al., 2012; Swanson et al., 2009; Swanson & Hsieh, 2009), thus pointing to the deficit model. However, cross-sectional studies do not allow conclusions about whether differences between learners with and without LD decrease, increase, or remain stable over time.

Among the few existing longitudinal studies focused on the development of phonological processing in young students with LD, some have provided empirical evidence for the deficit model, and others have provided empirical evidence for the developmental lag model. For example, de Jong and van der Leij (2003) analyzed the development of Dutch students with and without LD from kindergarten through sixth grade. The authors reported increasing differences regarding RAN and phonological awareness at the phoneme level from kindergarten through grade 1. By the end of elementary school, however, phonological awareness deficits were only evident in high-demanding tasks that required the complex and dynamic manipulation and processing of phonemes (e.g., nonword phoneme deletion), whereas no difference between the LD and non-LD groups was found in more simple phonological awareness tasks (e.g., rhyme categorization). Moreover, differences regarding RAN increased from grade 1 to grade 6, indicating that the developmental trajectory of this cognitive skill in students with LD fits a deficit model. Similar results were found by Korhonen (1995): Initial deficits in nonalphanumeric RAN, reading, and spelling at the age of 9 persisted until the age of 18 in Finnish students with reading and naming deficits.

These latter results are, however, in contrast to those from a study conducted by Georgiou and Stewart (2013), who found decreasing yet still statistically significant performance differences in alphanumeric and nonalphanumeric RAN between good and poor readers from first to third grade, which would support a developmental lag model. Likewise, in their cohort study, Kuppen and Goswami (2016) assessed students ages 6–10 and found further support for a developmental lag model of RAN, as initial naming deficits between poor and good readers decreased with age.

In a longitudinal study by Dandache, Wouters, and Ghesquière (2014), the development of literacy and phonological skills of Dutch students with LD, typically achieving students at risk for LD, and typically achieving students with a low risk for LD were followed from kindergarten through grade 6. The authors reported persistent deficits in students with dyslexia in phonological awareness and RAN in comparison with both groups of students without LD between first and sixth grades. Differences among the three groups regarding the phonological store component of the phonological loop (assessed with nonword repetition tasks) were no longer present at grade 6; thus, students with LD had caught up with their peers, supporting the developmental lag model. Challenging this result, Fischbach, Könen, Rietz, and Hasselhorn (2014) followed German students with and without LD from grade 4 to grade 6 and thoroughly examined the developmental trajectories in the various subprocesses of the phonological loop. The results revealed that different developmental patterns are in place: Whereas the structural capacity of the phonological store (assessed with nonword repetition) seemed to be intact in students with LD (in line with the results reported by Dandache et al., 2014), deficits in the functional capacity of the phonological loop (assessed with digit span) persisted from grade 4 to grade 6. More importantly, during the course of grade 5, students with LD started to develop an additional deficit in the rehearsal component of the phonological loop (as suggested by lower articulation rates), which was yet not apparent in grade 4. Likewise, results of the aforementioned study by Korhonen (1995) also pointed to a persisting deficit in phonological short-term storage in LD between age 9 and age 18, thus supporting the deficit model.

The Present Study

To sum up, phonological processing deficits in students with LD are frequently observed. However, the pattern of developmental trajectories of phonological processing in students with LD is not clear, as the empirical evidence is not only scarce but also relatively inconsistent. Not all of the previous studies outlined the growth of phonological processing skills over the elementary school years, as only a few studies focused longitudinally on the developmental trajectories of phonological processing.

In addition, there were some weaknesses in the longitudinal studies on LD (e.g., small sample sizes, not considering all aspects of phonological processing). Consequently, our goal in the present study was to examine the developmental trajectories of phonological processing in upper elementary students with LD and typically achieving peers from grade 3 to grade 5. We aimed to investigate in which aspects of phonological processing the two groups of students differ at grade 3 (research question 1) and how these phonological skills develop over the considered period (research question 2). Specifically, we asked which theoretical model of development (i.e., deficit vs. developmental lag model) best fits the longitudinal changes in phonological information processing in students with and without LD. Following existing empirical evidence, we expected performance differences in phonological processing in favor of typically achieving students at grade 3. We further expected a persistent deficit in RAN, phonological awareness, and the phonological loop.

Last but not least, previous heterogeneous results concerning phonological processing in young learners with LD may be the consequence of not considering the comorbidity of spelling difficulties in learners with reading difficulties, and vice versa. This seems to be particularly important when examining transparent orthographies such as German, as dissociations between reading and spelling skills are much more common (Moll & Landerl, 2009) than in English and have been found to be associated with slightly different cognitive profiles. For example, in a study by Wimmer and Mayringer (2002), deficits in phonological awareness prior to school entry were only found in German poor spellers, not in dysfluent readers. In contrast, a nonalphanumeric RAN deficit prior to school entry was only reported for children who later went on to develop a deficit in reading fluency, not in spelling. Deficits in the phonological loop, however, seem to be evident in both German reading and spelling deficits (Brandenburg et al., 2015). Consequently, examination of potential differences in the respective developmental trajectories is warranted.

In fact, spelling difficulties were obvious in some of the previous studies on LD, and some even reported statistically significant differences in spelling (Dandache et al., 2014; de Jong & van der Leij, 2003; Korhonen, 1995; see also the studies analyzed by Melby-Lervåg et al., 2012). Hence, it remains unclear whether deficits in phonological processing or different developmental trajectories for young learners with and without LD are indeed related to reading difficulties or might (also) be related to comorbid spelling difficulties. To consider this possibility, we assessed both reading and spelling in our LD sample, as this allowed us to examine potential differences that may exist between various subtypes of LD.

Method

Recruitment of the Participants

To recruit students with and without LD, 2,195 students (mean age = 8 years 8 months, standard deviation [*SD*] = 5 months; 49% girls) attending regular elementary schools in Germany were screened in 2011 at the end of grade 2 or at the beginning of grade 3. By this means, LD was diagnosed at the earliest point at which the diagnosis is considered reliable. Students completed standardized and norm-referenced tests for nonverbal intelligence, reading, spelling, and mathematics.

Inclusion Criteria

For defining students with LD, we used cutoff criteria recommended in German diagnostic guidelines (Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie, Psychosomatik und Psychotherapie (German Society for Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy; 2015) and most frequently used in German educational or clinical settings (Moll, Wallner, & Landerl, 2012). Students were diagnosed as having LD if they scored at least one standard deviation below average ($T < 40$; the mean of the T-scale is 50, with a SD of 10) in the standardized reading or spelling test while having at least average nonverbal intelligence (IQ ≥ 85) and mathematics scores of at least $T \geq 45$. We set the cutoff score for mathematical abilities at $T \geq 45$ to ensure that students with LD did not have additional math difficulties, which might have resulted in rather heterogeneous cognitive deficits. According to the latest International Classification of Diseases (World Health Organization, 2018), we additionally considered the IQ-discrepancy criterion for defining LD: Students showed a discrepancy of at least 1.2 standard deviations between their literacy achievement and their nonverbal IQ. By using these criteria for defining LD, the sample best represented the subpopulation of schoolchildren in Germany usually referred to as having LD. In contrast, students were assigned to the control group if they showed average reading, spelling, and mathematics skills (i.e., each $T \geq 45$) and had an average nonverbal intelligence (IQ ≥ 85).

Screening Measures

To assess reading skills, a test covering reading fluency and reading comprehension was used (Ein Leseverständnistest für Erst- bis Sechstklässler [ELFE 1–6; A reading comprehension test for grades 1 to 6]; Lenhard & Schneider, 2006). Decoding speed was assessed with the first subtest: The students' task was to identify the correct word out of four words that corresponded to a given picture (72 items). In the second subtest, students needed to complete 28 sentences by choosing one out of five words. In the last subtest, 20 short texts were presented, and students had to answer multiple-choice questions about them. As all subtests are limited in time, the ELFE 1–6 is designed as a speed test rather than a power test. Whereas the first subtest measures primarily decoding speed, the second and third subtests also measure reading comprehension. Internal consistency is high, with Cronbach's alphas between .92 and .94.

Spelling skills were assessed using a dictation for second and third graders (Weingartener Grundwortschatz Rechtschreib-Test für zweite und dritte Klassen [Weingarten's spelling test of basic vocabulary for second and third grades]; Birkel, 2007). In this test, 43 orally dictated familiar words have to be written. Internal consistency is high (Cronbach's $\alpha = .94$). To control for potential comorbid difficulties in mathematics, a standardized math test was administered (Cronbach's $\alpha = .91$; Deutscher Mathematiktest für zweite Klassen [German mathematical test for second grade]; Krajewski, Liehm, & Schneider, 2004). Fluid intelligence was measured with the nonverbal Culture Fair Intelligence Test (Cattell, Weiß, & Osterland, 1997), consisting of figural material (e.g., identification of similarities, figure classifications). Split-half reliability is .92 for the age groups studied.

Main Study

Participants

Subsequent analyses are based on a sample of 209 students, with an oversampling of students with LD ($n = 109$). Some of the students with LD in our study showed deficits in reading only ($n = 35$), in spelling only ($n = 36$), or in both domains ($n = 38$). Students with LD did not differ from students in the control group in terms of age, $F(1, 207) = 2.59$, $p = .11$, $d = 0.22$; nonverbal intelligence, $F(1, 207) = 0.71$, $p = .40$, $d = 0.12$; and mathematics, $F(1, 207) = 0.47$, $p = .49$, $d = -0.10$. As intended, groups significantly differed with regard to reading, $F(1, 207) = 186.73$, $p < .001$, $d = -1.93$; and spelling performance, $F(1, 207) = 263.07$, $p < .001$, $d = -2.25$; as typically achieving students outperformed students with LD. Sample details are provided in Table 1.

Slightly more boys (61%) than girls were in the sample, $\chi^2(1, N = 209) = 9.69$, $p = .002$. Approximately 7% of the students claimed to speak with at least one parent a language other than German, but there was no systematic relation between the language spoken at home and group membership, $\chi^2(1, N = 182) = 3.97$, $p = .14$. To obtain a

TABLE 1
Descriptive Statistics as a Function of Group

	Control group (n = 100)		LD group (n = 109)	
Measure	*M*	*SD*	*M*	*SD*
Age at screening	8 years 6.16 months	4.86 months	8 years 7.36 months	5.80 months
Socioeconomic status	8.79	1.08	8.60	1.19
Intelligence (IQ score)	106.86	11.19	108.07	9.68
Mathematics (*T*-score)	53.98	5.49	53.40	6.53
Reading (*T*-score)	53.53	5.91	39.86	8.03
Spelling (*T*-score)	51.36	5.75	37.77	6.31

Note. LD = students with specific learning disabilities in reading or spelling.

measure of socioeconomic status, we computed a composite in which the graduation level (1 = *no graduation*; 2 = *graduation after grade 9 or 10*; 3 = *graduation after grade 11 or higher*) and the employment status (1 = *no employment*; 2 = *full- or part-time employment*) of the father was considered. Accordingly, higher values represent higher socioeconomic status, with a maximum of 5. The mean socioeconomic status of the sample was 4.42 (SD = 0.63), with no differences between the groups, $F(1, 167) = 2.08$, $p = .15$, $d = 0.22$. Written informed parental consent was obtained for all students prior to testing. Participation was voluntary, and consent could be withdrawn at any time without giving a reason.

There were three measurement points of phonological processing in the main study, with individual measurements at grade 3 (wave 1; approximately half a year after the screening), grade 4 (wave 2), and grade 5 (wave 3). Students were tested once a year, approximately 11–12 months apart. During the course of the study, 38% of students dropped out of the study (34 at wave 2 and 48 at wave 3). On most variables, there were no statistically significant differences between students who dropped out prematurely from the study and those who completed the study. Further details of the sample (i.e., extent of attrition and strategies to deal with it) are provided in Appendix A.

Assessment of Phonological Processing

To measure lexical access from long-term memory, four different RAN tasks were used. Specifically, students had to name alphanumeric (letters and digits) and nonalphanumeric items (colors and objects). Items were presented in five rows, with 10 items per row, on a piece of paper. We used naming times (in seconds) for each of the four subtests as dependent variables. The Cronbach's alphas for RAN were .71 at wave 1, .80 at wave 2, and .79 at wave 3.

Phonological awareness was measured by three tasks (vowel substitution, vowel length, and phoneme reversal) from a German standardized test (Basiskompetenzen für Lese-Rechtschreibleistungen: ein Test zur Erfassung der phonologischen Bewusstheit vom ersten bis vierten Grundschuljahr [Basic skills for reading and spelling: A test for measuring phonological awareness from first to fourth grade]; Stock, Marx, & Schneider, 2003). All items were presented using a computer with speakers. Preceding each task, students received an explanation and two practice trials. If a student answered three consecutive items within a subtest incorrectly, testing was stopped. In the vowel substitution task, students were asked to substitute every occurrence of /a/ with /i/ within eight orally presented words and four nonwords. To complete this task properly, students needed to identify the vowel sounds to be replaced and insert the new vowel sound in the correct position(s). In addition to the occurrence of /a/ sounds, other vowel sounds were presented, so students needed to be careful not to change those other vowel sounds (e.g., *mathematik* [mathematics]). Sample-based internal consistencies (based on the Kuder–Richardson formula due to the dichotomy of the item scoring) were .81 at wave 1, .69 at wave 2, and .70 at wave 3. To solve the vowel length task, which consisted of 10 trials, students had to decide which nonword out of four orally presented nonwords with the same vowel sounded different due to differences in vowel length (e.g., /re:m/, /fe:r/, /nɛl/, /be:f/). Internal consistencies were .76, .74, and .67 for the waves studied, respectively. In the phoneme reversal task, students had to reverse the phoneme sequences of orally presented words and nonwords so (four) new words and (14) nonwords emerged. Internal consistencies were .91 at wave 1, .90 at wave 2, and .87 at wave 3. For each task, we used the number of correct trials as a dependent variable for subsequent analyses.

Three subtests (two word span tasks and a digit span task) of a computer-based, standardized German test (Arbeitsgedächtnistestbatterie für Kinder von 5 bis 12 Jahren [Working memory test battery for children ages 5 to 12 years]; Hasselhorn et al., 2012) were used to measure storage capacity of the phonological loop. Computer-presented instructions and two practical trials preceded each subtest. The word span tasks and digit span task each included 10 trials following an adaptive algorithm. In the word span tasks, sequences of two to nine words (one or three syllables long depending on the subtest) were presented acoustically, and students had to recall them immediately afterward in the same order in which they had been presented. Sample-based internal consistencies for the monosyllabic word span task were .91 at wave 1, .94 at wave 2, and .91 at wave 3. For the trisyllabic word span task, internal consistencies were .85, .88, and .89, respectively. The digit span task worked in the same way, with sequences of two to nine digits instead of words. Internal consistencies for the three waves were .92, .89, and .90, respectively. For each subtest, we used the mean of the last eight trials as a dependent variable.

Statistical Analyses

We used M*plus* version 7.11 (L.K. Muthén & Muthén, 2010) for all analyses. For all statistical analyses, we set the alpha level at $p < .05$.

Due to attrition (i.e., dropout, nonmonotone missing values), approximately 20% of the data were missing. Little's missing completely at random test resulted in a statistically nonsignificant result, $\chi^2(545, N = 209) = 88.50$, $p = .10$; thus, indicating that no identifiable pattern exists to missing data. We applied statistical methods that analyze the data without loss of information. Thus, we used the multiple linear regression (MLR) estimator, which treats missing values with the full information maximum likelihood (FIML) method.

Before we started our analyses, we had to ensure that it was justified to subsume the low-achieving students together in one LD category. This was necessary because our classification criteria selected students with deficits in reading and/or spelling. As outlined earlier, because some research has indicated that reading difficulties in transparent orthographies might be accompanied with other phonological processing deficits than spelling difficulties, subsuming these various subtypes of LD is not justified a priori and should be tested statistically instead.

We thus conducted a multiple-indicators, multiple-causes (MIMIC) model, in which reading and spelling difficulties were both entered as covariates and then tested for equality of their parameter estimates. This analysis (as shown in Appendix B) revealed no performance differences in phonological processing between reading and spelling difficulties, neither for the initial performance level at grade 3 nor for the successive development. These findings let us to combine the students into one LD group, whose performance was then compared with the control group via multigroup latent change score (LCS) modeling.

There were two principal steps to this main analysis. First, we established the measurement model for each phonological processing construct by testing separate three-factor oblique models in which the factors represented the three testing waves. For RAN, we modeled two measurement models, one representing the alphanumeric measures and the other representing the nonalphanumeric measures. We introduced this separation by stimulus type because previous studies have demonstrated performance differences between alphanumeric and nonalphanumeric stimuli in school-age learners with LD (e.g., van Daal & van der Leij, 1999; van den Bos, Zijlstra, & lutje Spelberg, 2002). We were therefore interested in whether this was also true for the developmental course of these skills. To ensure that the latent factors represented the same underlying construct across groups and time, we set both the factor loadings and the manifest intercepts as invariant across groups and testing waves (Byrne, 2013). In addition, we also included factors that accounted for indicator-specific variance across time (i.e., common variance between one and the same phonological processing task across the three testing waves), as suggested by Geiser (2012). Following Hu and Bentler (1999), a good fit to data was indicated by (a) a comparative fit index (CFI) with values of approximately 0.95, (b) a root mean square error of approximation (RMSEA) with values of approximately 0.06 or less, and (c) a statistically nonsignificant chi-square test. Additionally, we used the ratio χ^2/df, where a ratio of <2 indicates an acceptable model fit (Ullman, 2007).

Second, having established the measurement models, we next applied LCS models (McArdle, 2009) to examine and compare the longitudinal development of phonological processing across the two groups. LCS has several advantages over more traditional longitudinal models, such as a repeated-measures analysis of variance. For example, the developmental differences between measurement points can be directly estimated in the LCS model, whereas in analysis of variance, developmental change is tested in a more indirect way, namely, via the interaction between groups and measurement points. Thus, LCS modeling allows researchers to test their hypotheses on developmental trends in a more flexible manner. LCS models, as with other structural equation modeling techniques, have the additional advantages that the hypothesized models can be tested against the data and evaluated through goodness-of-fit indexes and that models of varying complexity can be compared (Heck & Thomas, 2015).

In LCS, developmental change is modeled by introducing latent variables (i.e., latent change factors), which represent the discrete performance changes between two timepoints. (See Appendix C as an example of the latent change model for the nonalphanumeric RAN measures.) For LCS modeling, we applied a three-step testing procedure in which different equality constraints were imposed on the LCS models. In the second and third testing steps, we used nested model comparison to assess whether the more restrictive model was preferable to the less restricted model. Because we used the MLR estimator, changes in model fit were examined with the Satorra–Bentler scaled chi-squared difference test (ΔSB-χ^2; Satorra & Bentler, 2001). A statistically nonsignificant value for this statistic implies that the restrictive model fits the data just as well as the less restrictive comparison model (Wang & Wang, 2012).

The three-step testing procedure was as follows: (1) We started by testing the unrestricted model, in which the baseline factor and the two latent change factors were estimated freely across groups and time (i.e., without imposing equality constraints on latent factor means). This unrestricted model served as a reference to which we compared the more restrictive models that followed. (2) Next, we examined whether the groups showed comparable performance levels at grade 3. To this end, we constrained the latent mean of the baseline factor to be equal across both groups of students. We then statistically compared this restricted baseline model with the unrestricted model. A statistically nonsignificant result would indicate that the performance levels at grade 3 were comparable between the groups, whereas a statistically significant result would suggest that students with LD showed lower performance at grade 3 than did students in the control group. (3) Finally, we investigated the longitudinal development across the groups. To this end, we constrained the means of the latent change factors to be equal across both time and group. This maximum restricted longitudinal model implies linear development across testing waves, as well as parallel development across the two groups. Again, we statistically compared this model with the unrestrictive model (as the latter model implies nonsystematic development across both group and time). A statistically nonsignificant result would indicate that the maximum restricted longitudinal model with linear and parallel development best described the longitudinal development across time and group. In contrast, a statistically significant result would indicate that the maximum restricted model did not hold. In this case, variations of this model would also be tested to trace the developmental trajectories that best fit the data.

Results

Data Screening

Table 2 shows descriptive statistics of the phonological processing measures as a function of group and testing wave. First, we evaluated whether the data met basic assumptions of structural equation modeling: There was no problem with multicollinearity because none of the zero-order correlations between the manifest variables within testing wave was above .80 (see Appendix D). Furthermore, we checked the data for univariate outliers, classified as cases more than 3.5 standard deviations from the sample's means: Of the 6,270 values in the data set, 19 were univariate outliers (eight students in the LD group and 11 students in the the control group). We deleted these values from further analyses. In addition, we identified two cases (one for each group) as multivariate outliers through Mahalanobis distance and also deleted them. There was no evidence that the assumption of univariate normality distribution was violated, because all measures showed skewness less than 3 and kurtosis less than 4. Nevertheless, Mardia's test of multivariate normality (Mardia, 1974) revealed a violation of the assumptions of multivariate skewness and kurtosis. Therefore, we tested models using the MLR estimator, which is robust to nonnormal data (Wang & Wang, 2012).

Measurement Models

Standardized factor loadings of the measurement models are provided in Table 2. For alphanumeric RAN, the measurement model provided a good fit to the data, $\chi^2(17) = 21.79$, $p = .19$; $\chi^2/df < 2$; RMSEA = 0.05, 90% confidence interval (CI) [0.00, 0.11]; CFI = 0.99. The same applied to the nonalphanumeric RAN model, $\chi^2(17) = 12.81$, $p = .74$; $\chi^2/df < 2$; RMSEA = 0.00, 90% CI [0.00, 0.08]; CFI = 1.00. For phonological awareness, however, the measurement model showed a slightly worse fit, $\chi^2(58) = 79.36$, $p = .03$; $\chi^2/df < 2$; RMSEA = 0.06, 90% CI [0.02, 0.09]; CFI = 0.93. Whereas the χ^2/df ratio, the CFI, and the RMSEA were in the acceptable range, the chi-squared statistic slightly missed conventional levels. However, the chi-squared statistic is relatively sensitive to sample size and thus no longer recommended as a sole basis for evaluating model fit (Schermelleh-Engel, Moosbrugger, & Müller, 2003). Modification indexes revealed that a correlated residual between vowel substitution and vowel length at grade 5 could be included in the model to improve fit. However, because we had no theoretical or empirical reason to do so, we did not engage in this post hoc model specification and instead decided to go with this slightly worse fit. For the phonological loop, the measurement model provided a good fit to the data, $\chi^2(58) = 67.21$, $p = .19$; $\chi^2/df < 2$; RMSEA = 0.04, 90% CI [0.00, 0.08]; CFI = 0.99.

LCS Models

We next modeled the unrestricted LCS models. As this type of model is statistically equivalent to the measurement models established before, the models show the same model fit and are thus not further described here.

Group Differences at Grade 3

With our first research question, we aimed to investigate in which aspects of phonological processing the two groups of students differed at grade 3. We therefore specified the restricted baseline models and compared this type of model to the unrestricted LCS models.

Concerning alphanumeric RAN, the restricted baseline model showed the following fit indexes: $\chi^2(18) = 37.18$, $p = .01$; $\chi^2/df > 2$; RMSEA = 0.10, 90% CI [0.05, 0.15]; CFI = 0.95. When we compared this model with the unrestricted model, the chi-squared difference test was statistically significant, $\Delta SB\text{-}\chi^2 = 59.46$, $\Delta df = 1$, $p < .001$; thus, we rejected the restricted baseline model. This result indicates that the LD group showed slower alphanumeric naming speed at grade 3 than did the control group. We therefore released the equality constraints on the baseline factors because the unrestricted model holds.

For the nonalphanumeric RAN, model fit for the restricted baseline model was as follows: $\chi^2(18) = 14.10$, $p = .07$; $\chi^2/df < 2$; RMSEA = 0.00, 90% CI [0.00, 0.07]; CFI = 1.00. When we compared this model with the unrestricted model, the chi-squared difference test showed a statistically nonsignificant result, $\Delta SB\text{-}\chi^2 = 1.34$, $\Delta df = 1$, $p = .25$. That is, the restricted baseline model could not be rejected. Hence, this result suggests that the control group and the LD group did not differ in their nonalphanumeric naming speed at grade 3; their initial performance level was comparable. The equality constraints on the baseline factors could thus not be released because the restricted baseline model holds. Accordingly, with respect to our second research question (see the third testing step below), we used this restricted baseline model rather than the unrestricted model as a reference model.

With respect to phonological awareness, the restricted baseline model showed the following fit indexes: $\chi^2(59) = 132.25$, $p < .001$; $\chi^2/df > 2$; RMSEA =

TABLE 2
Descriptive Statistics and Standardized Factor Loadings for Phonological Processing Measures as a Function of Group

	Control group					LD group				
Measure	*M*	*SD*	Actual range	Skewness	Estimate[a]	*M*	*SD*	Actual range	Skewness	Estimate[a]
Rapid automatized naming										
t1 letter naming	31.12	6.66	20–55	1.10	0.81	34.35	7.18	23–67	1.41	0.67
t2 letter naming	26.05	4.63	17–60	2.65	1.00	28.87	4.91	20–54	1.32	0.91
t3 letter naming	24.28	5.32	15–49	1.62	0.98	24.98	4.68	18–37	0.79	0.95
t1 digit naming	28.65	5.81	19–55	1.28	0.67	34.25	7.29	23–71	1.24	0.58
t2 digit naming	25.61	5.33	17–56	2.13	0.70	29.35	6.64	19–51	0.73	0.68
t3 digit naming	24.23	5.50	15–51	1.67	0.77	25.41	5.57	16–43	1.28	0.68
t1 object naming	47.42	7.53	33–78	0.87	0.88	48.18	7.37	35–89	1.67	0.84
t2 object naming	42.95	7.02	31–73	1.38	0.89	45.51	7.48	31–71	0.96	0.91
t3 object naming	39.60	7.58	21–71	0.80	0.88	41.48	6.92	32–62	1.08	0.93
t1 color naming	49.93	10.20	31–84	0.79	0.66	51.67	10.43	33–81	0.81	0.63
t2 color naming	45.87	9.68	28–73	0.91	0.65	48.29	9.77	33–83	0.95	0.75
t3 color naming	42.22	9.56	28–69	1.08	0.76	43.85	8.69	25–74	0.87	0.75
Phonological awareness										
t1 phoneme reversal	9.41	4.79	1–18	−0.04	0.72	5.21	4.16	0–17	0.84	0.70
t2 phoneme reversal	11.22	4.92	0–18	−0.52	0.80	7.40	4.43	0–17	0.37	0.76
t3 phoneme reversal	13.06	3.79	2–18	−1.29	0.57	9.11	4.74	1–18	0.01	0.76
t1 vowel length	4.63	2.76	0–10	0.05	0.49	3.45	2.25	0–9	0.31	0.49
t2 vowel length	5.77	2.73	0–10	−0.36	0.58	4.21	2.70	0–10	0.22	0.51
t3 vowel length	6.19	2.66	0–10	−0.51	0.37	5.46	2.42	0–10	−0.28	0.60
t1 vowel substitution	9.46	2.28	0–12	−1.29	0.45	7.06	2.94	0–12	−0.51	0.30
t2 vowel substitution	9.91	1.84	5–12	−0.70	0.58	8.40	2.77	1–12	−0.98	0.39
t3 vowel substitution	10.54	1.78	3–12	−1.65	0.38	8.82	2.52	0–12	−1.16	0.46

(continued)

TABLE 2
Descriptive Statistics and Standardized Factor Loadings for Phonological Processing Measures as a Function of Group (*continued*)

	Control group					LD group				
Measure	*M*	*SD*	Actual range	Skewness	Estimate[a]	*M*	*SD*	Actual range	Skewness	Estimate[a]
Phonological loop										
t1 word span for one-syllable words	3.95	0.65	2.3–5.6	0.08	0.84	3.74	0.58	2.4–5.0	−0.27	0.74
t2 word span for one-syllable words	4.21	0.73	2.9–6.3	0.55	0.83	3.94	0.73	2.3–5.6	0.03	0.75
t3 word span for one-syllable words	4.28	0.77	2.8–6.1	0.32	0.81	4.14	0.66	2.2–5.8	0.50	0.79
t1 word span for three-syllable words	3.12	0.44	2.1–4.3	0.37	0.74	3.00	0.37	2.0–4.0	0.17	0.68
t2 word span for three-syllable words	3.22	0.47	2.0–5.3	0.58	0.76	3.06	0.39	2.1–5.4	1.59	0.75
t3 word span for three-syllable words	3.39	0.51	2.1–4.5	0.11	0.73	3.16	0.42	2.1–4.3	0.50	0.68
t1 digit span	4.61	0.60	2.8–6.3	−0.41	0.81	4.17	0.58	2.7–5.6	−0.36	0.71
t2 digit span	4.91	0.69	3.1–6.5	0.36	0.85	4.49	0.62	2.0–5.9	−0.47	0.77
t3 digit span	5.11	0.68	3.0–7.0	0.46	0.84	4.64	0.69	2.6–6.7	−0.23	0.73

Note. LD = students with specific learning disabilities in reading or spelling; *M* = mean; *SD* = standard deviation; t1 = first testing wave; t2 = second testing wave; t3 = third testing wave.
[a]For example, factor loadings.

0.11, 90% CI [0.08, 0.13]; CFI = 0.77. When we compared this model with the unrestricted model, the chi-squared difference test was statistically significant, ΔSB-χ^2 = 28.18, Δdf = 1, p < .001. This result indicates that the LD group showed lower performance in phonological awareness at grade 3 than did the control group. We therefore released the equality constraints on the baseline factors because the unrestricted model holds.

Finally, with respect to the phonological loop, the restricted baseline model provided the following fit indexes: $\chi^2(59)$ = 71.75, p = .12; χ^2/df < 2; RMSEA = 0.05, 90% CI [0.00, 0.08]; CFI = 0.98. When we compared this model with the unrestricted model, the chi-squared difference test was statistically significant, ΔSB-χ^2 = 4.65, Δdf = 1, p = .03; thus, we rejected the restricted baseline model. This result suggests that the LD group showed lower phonological loop at grade 3 than did the control group. Hence, we released the equality constraints on the baseline factors because again the unrestricted model holds.

Developmental Trajectories of Phonological Processing

With our second research question, we wanted to investigate how the phonological skills develop in students with and without LD over the considered period (from grade 3 to grade 5). The latent means of the phonological processing variables for the two groups are displayed in Figure 1.

With respect to alphanumeric RAN, the maximum restricted longitudinal model showed the following fit indexes, $\chi^2(20)$ = 41.18, p < .001; χ^2/df > 2; RMSEA = 0.10, 90% CI [0.06, 0.15]; CFI = 0.94. When we compared this model with the unrestricted model, the chi-squared difference test was statistically significant, ΔSB-χ^2 = 25.82, Δdf = 3, p < .001. This result suggests that the maximum restricted longitudinal model does not hold because it poses too many equality constraints on the developmental trajectories. We therefore specified less restrictive variations of this model. However, the restrictive model of linear but nonparallel development showed only a poor fit to the data, $\chi^2(19)$ = 33.01, p = .02; χ^2/df < 2; RMSEA = 0.08, 90% CI [0.03, 0.13]; CFI = 0.96; as did the restrictive model of nonlinear development across time but parallel development between groups, $\chi^2(19)$ = 32.57, p = .03; χ^2/df < 2; RMSEA = 0.08, 90% CI [0.03, 0.13]; CFI = 0.96. Inspection of the latent means revealed that the control group showed only marginal performance gains from grade 4 to grade 5, which allowed the LD group to catch up, so they reached a performance level at grade 5 comparable to their typically achieving peers. This latent mean pattern is best described by decreasing group differences. To test this model statistically, we fixed all latent change factors to be equal except the second one, the latent difference in RAN for alphanumeric stimuli between the third testing wave (time 3; t3) and the second testing wave (time 2; t2; $t3RAN_a - t2RAN_a$), in the control group. This model provided a good fit to the data, $\chi^2(19)$ = 24.66, p = .17; χ^2/df < 2; RMSEA = 0.05, 90% [CI [0.00, 0.11]; CFI = 0.99. In addition, when we tested this model against the unrestricted model, the chi-squared difference test was statistically nonsignificant, ΔSB-χ^2 = 2.94, Δdf = 2, p = .23; providing further support in favor of decreasing differences. That is, initial group differences in alphanumeric naming speed at grade 3 had disappeared by grade 5: for time 1 (t1) performance (LD group M – control group M), 34.05 – 30.87 = Δ3.1; for t3 performance (LD group M – control group M), 24.49 – 24.53 = Δ–0.04.

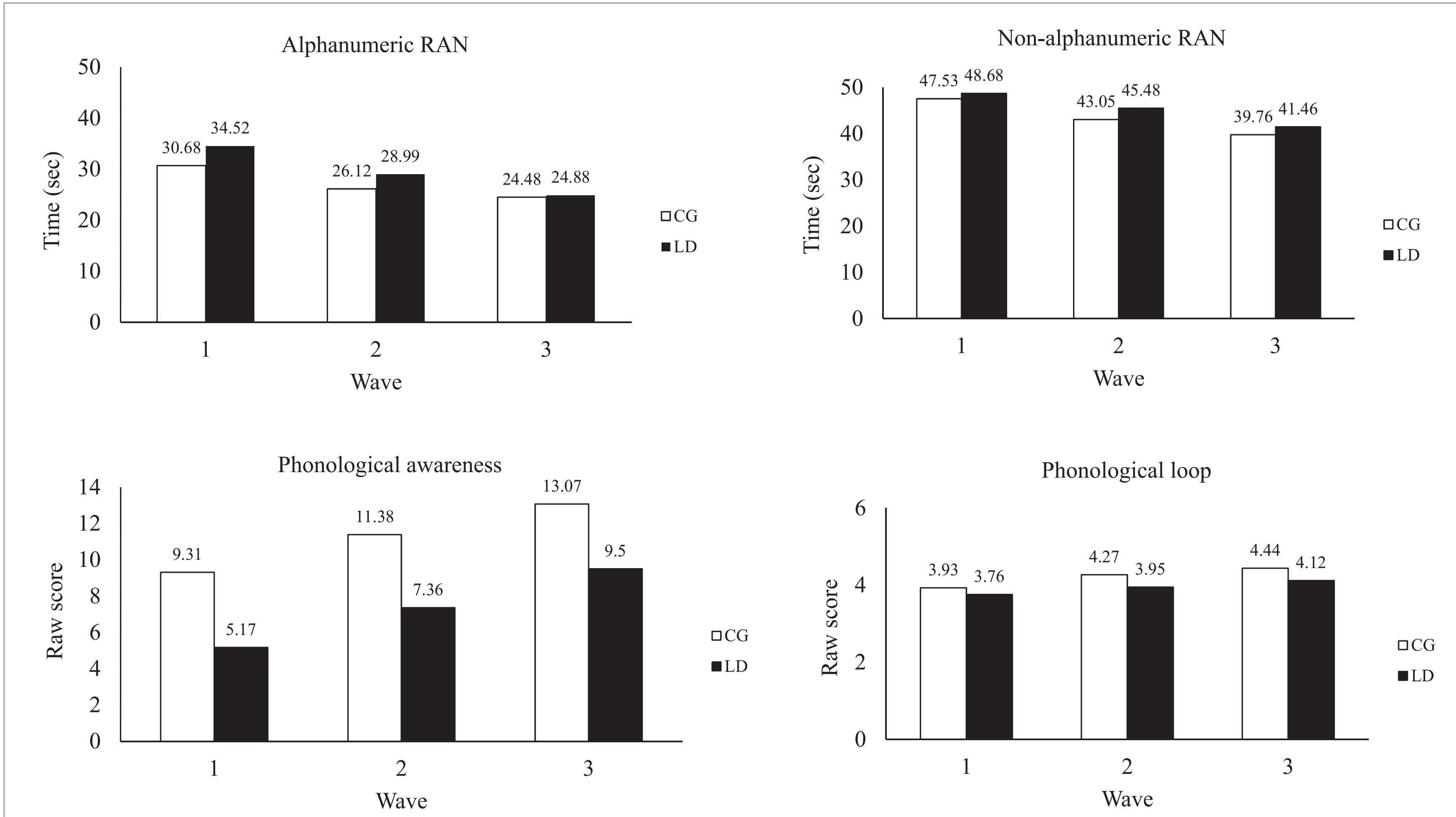
For nonalphanumeric RAN, model fit of the maximum restricted longitudinal model was as follows: $\chi^2(21)$ = 19.87, p = .53; χ^2/df < 2; RMSEA = 0.00, 90% CI [0.00, 0.08]; CFI = 1.00. When we compared this model with the restricted baseline model, the chi-squared difference test showed a statistically nonsignificant result, ΔSB-χ^2 = 5.52, Δdf = 3, p = .14. This result suggests that the control group and the LD group showed the same linear developmental trajectories for nonalphanumeric stimuli.

The maximum restricted model of the phonological awareness measures had the following fit indexes: $\chi^2(61)$ = 80.14, p = .05; χ^2/df < 2; RMSEA = 0.06, 90% CI [0.00, 0.09]; CFI = 0.94. It is worth mentioning that this model overcomes the initial slightly poorer fit of the unrestricted model and the measurement model, respectively. In line with this, when we compared this model with the unrestricted model, the chi-squared difference test was statistically nonsignificant, ΔSB-χ^2 < 1, Δdf = 3, p = .82. This result indicates that the developmental trajectories of phonological awareness are best described by linear development across testing waves and parallel development across groups.

Finally, concerning the phonological loop, the maximum restricted longitudinal model provided the following fit indexes: $\chi^2(61)$ = 77.71, p = .07; χ^2/df < 2; RMSEA = 0.05, 90% CI [0.00, 0.08]; CFI = 0.98. When we compared this model with the unrestricted model, the chi-squared difference test was statistically significant, ΔSB-χ^2 = 11.83, Δdf = 3, p = .01. As before with the alphanumeric RAN model, this result suggests that the maximum restricted longitudinal model does not hold because it poses too many equality constraints on the developmental trajectories. We therefore specified two less restrictive variations of this model.

On the one hand, we tested a restrictive model of parallel but nonlinear development, which implies parallel development across groups (i.e., latent change factors constrained to be equal across groups) but nonlinear development across time (i.e., latent change factors of different timepoints not fixed to be equal). However, this model provided an even poorer fit because the chi-squared statistc now reached statistical significance level, $\chi^2(60)$ = 79.35, p = .05; χ^2/df < 2; RMSEA = 0.06, 90% CI [0.01, 0.09]; CFI = 0.98. On the other hand, we tested a restrictive model

FIGURE 1
Latent Means of Phonological Processing by Grade Level



Note. CG = control group; LD = students with specific learning disabilities in reading or spelling; RAN = rapid automatized naming.

of linear but nonparallel development, which implies linear development across testing waves (i.e., latent change factors of different timepoints constrained to be equal) but nonparallel development across groups (i.e., latent change factors of different groups not fixed to be equal). This model provided an excellent fit to the data, $\chi^2(60) = 72.14$, $p = .14$; $\chi^2/df < 2$; RMSEA = 0.04, 90% CI [0.00, 0.08]; CFI = 0.99. In addition, when we tested this model against the unrestricted model, the chi-squared difference test was statistically nonsignificant, $\Delta\text{SB-}\chi^2 = 4.84$, $\Delta df = 2$, $p = .09$; providing further support that this model best described the underlying developmental trajectories. That is, developmental increase from grade 3 to grade 4 and from grade 4 to grade 5 was comparable within groups but not between groups. When looking at the latent means in Table 3, it becomes evident that this nonparallel development across groups is due to increasing group differences in the phonological loop from grade 3 to grade 5: for t1 performance (control group M – LD group M), 3.93 – 3.76 = Δ0.17; for t3 performance (control group M – LD group M), 4.46 – 4.12 = Δ0.34.

Discussion

In the present study, we addressed a critical issue by investigating the phonological processing developmental trajectories in students with LD and compared these students with their typically achieving peers by using multigroup LCS. In particular, German students with and without LD were followed from grade 3 to grade 5, and their phonological processing was assessed three times within this period.

After ensuring that there were no differences in phonological processing as a function of reading versus spelling difficulties, we addressed our first research question with the aim of replicating a growing body of empirical evidence supporting the idea of an underlying phonological deficit in students with LD. Indeed, our results are in line with those of previous research (Melby-Lervåg et al., 2012; Swanson, 2003). Accordingly, there were performance differences in phonological processing in favor of the typically achieving students at grade 3 in alphanumeric RAN, phonological awareness, and the phonological loop. There were, however, no group differences in nonalphanumeric RAN, which further supports the need for dividing RAN into alphanumeric and nonalphanumeric performance, as students with LD performed differently depending on stimulus type. This result of better performance on nonalphanumeric RAN than on alphanumeric measures has been reported in prior studies on LD (van Daal & van der Leij, 1999; van den Bos et al., 2002) and seems to be related, at least to some extent, to the age and literacy

TABLE 3
Latent Means for Baseline and Change Factors for the Final Latent Change Score Models as a Function of Group

Parameter	Control group	LD group
Rapid automatized naming for alphanumeric stimuli (RAN_a)		
Baseline factor $t1RAN_a$	30.87	34.05
Latent change factor $t2RAN_a$ – $t1RAN_a$	−4.78	−4.78
Latent change factor $t3RAN_a$ – $t2RAN_a$	−1.56	−4.78
Rapid automatized naming for nonalphanumeric stimuli (RAN_n)		
Baseline factor $t1RAN_n$	48.07	48.07
Latent change factor $t2RAN_n$ – $t1RAN_n$	−3.72	−3.72
Latent change factor $t3RAN_n$ – $t2RAN_n$	−3.72	−3.72
Phonological awareness (PA)		
Baseline factor t1PA	9.21	5.25
Latent change factor t2PA – t1PA	2.01	2.01
Latent change factor t3PA – t2PA	2.01	2.01
Phonological loop (PL)		
Baseline factor t1PL	3.93	3.76
Latent change factor t2PL – t1PL	0.26	0.18
Latent change factor t3PL – t2PL	0.26	0.18

Note. LD = students with specific learning disabilities in reading or spelling; t1 = first testing wave; t2 = second testing wave; t3 = third testing wave; t2 − t1 = latent difference between the second and first testing waves; t3 − t2 = latent difference between the third and second testing waves.

experience of the sample. Whereas nonalphanumeric RAN seems to highly differentiate between LD and non-LD in kindergarten and the early grades, its relation to literacy diminishes (as compared with alphanumeric RAN) during elementary school when students become more and more literate (van den Bos et al., 2002).

Our second and main focus in this study was to determine whether the developmental trajectories of phonological processing differ between students with and without LD. To determine the specific developmental pattern across the two groups, we considered two theoretical models: the deficit model and the developmental lag model.

A Developmental Lag for Alphanumeric RAN

For RAN, we expected persistent deficits in students with LD as compared with typically achieving students over the course of the three school years. In nonalphanumeric RAN (i.e., the component in which we did not find initial performance differences between groups at grade 3), the students in the control group and the students with LD continued to show parallel development in the subsequent years. Contrary to our expectations, we did not find evidence for a persistent deficit in alphanumeric RAN but instead found a developmental lag. That is, initial differences in alphanumeric RAN decreased as students with LD caught up with their typically achieving peers. This finding is in line with those of several studies with Anglo-American students (e.g., Kuppen & Goswami, 2016; Mazzocco & Grimm, 2013). If compared with results hailing from two studies in non-Anglo-American countries (de Jong & van der Leij, 2003; Korhonen, 1995), our result is divergent, as deficits in alphanumeric RAN continued to characterize older students with LD in these studies. In both of these studies, however, reading abilities were assessed using a pure measure of reading fluency, whereas we used a combined test of reading comprehension and reading fluency and a spelling test for sample selection. Because RAN is more closely related to reading fluency than reading comprehension or spelling (Araújo et al., 2015), this might explain the divergent results. In addition, the LD group in de Jong and van der Leij's (2003) study belonged to the bottom 3.5% of their grade level and was, hence, much more severely affected than the students with LD in the present study, whose academic performance in reading and spelling belonged to the bottom 16%. One might speculate that in very low performing students, initial RAN deficits may continue to persist a little bit longer.

According to several studies' findings, the development in RAN is not gradual, but rather there are developmental phases in which students improve and those in which students' development rests. For example, Siddaiah, Saldanha, Venkatesh, Ramachandra, and Padakannaya (2016) investigated the development of RAN in a cohort study with Indian students from grade 1 to grade 10. The authors reported that the rate of increment was not uniform across all successive grade levels. Whereas the students' letter-naming speed gradually increased up to grade 4 (and again from grade 5 to grade 6), there was a developmental pause between grades 4 and 5. A possible explanation for this nonuniform growth pattern might lie in the alternation of speedup processes versus restructuring processes. In Georgiou and Stewart's (2013) longitudinal study, Greek students with and without LD were followed from first to third grade. The authors reported that only speedup processes were present in this time period and that there was a two-year gap between the students with and without LD; that is, the improvement in the efficiency with which alphanumeric RAN tasks were performed took place earlier in non-LD students than in LD students. Taking the results of these two studies together, we could assume that between grades 4 and 5, no increase in RAN performance occurred in our control group because the speedup processes, which are responsible for the performance increase, possibly reached a temporary limit. The period between grades 4 and 5 could thus be seen as a developmental break for the control group, with possibly a further performance increase in RAN after grade 5. Likewise, students with LD were able to catch up to the control group because their speedup processes possibly took place later and had not yet reached a limit or developmental break.

It is worth mentioning that the result of a catch-up by students with LD is very unlikely to be attributable to either a RAN intervention effect in the LD group or a ceiling effect in alphanumeric item naming in the control group. There is some evidence that naming speed cannot easily be improved through intervention. For example, an intervention study by de Jong and Vrielink (2004) revealed no effects of a specific RAN training on either subsequent RAN performance or reading skills in Dutch students.

Further, there are at least three reasons why we assumed no ceiling effect in alphanumeric RAN in our control group. First, ceiling effects in timed data usually occur if pruned data are used for time recoding (i.e., participants' scores cannot exceed or fall below a specific value because of an a priori specified threshold). However, we did not use censored data to assess naming speed and thus did not restrict the possible range of the data a priori. Second, inspection of histograms revealed that our naming time data were equally well distributed across testing waves and that there was no bump at the lower end of distribution. Third, we applied the standard version of RAN that consists of 50 to-be-named items per subtest. This RAN version has been used in many studies with a wide age range and has demonstrated sensitivity up to adulthood. For example, Mazzocco and Grimm (2013) observed an average naming time of alphanumeric RAN of only 13 to14 seconds among eighth-grade students. Thus, although, at grade 5, students with LD named digits and letters as fast as control group students did, this does not necessarily imply that the development of RAN had reached an asymptote in our sample, given that much lower naming times have been reported for older students.

A Persistent Deficit in Phonological Awareness

Regarding phonological awareness, we found evidence of a persistent deficit in students with LD throughout elementary school. Although students with LD improved in phonological awareness over the considered years, they continued to lag behind their typically achieving peers. This finding is in line with those of cross-sectional and longitudinal studies (e.g., Dandache et al., 2014; Korhonen, 1995) reporting phonological awareness deficits in older students or adults with LD.

De Jong and van der Leij (2003) reported that deficits in phonological awareness in Dutch students with LD tended to disappear at the end of elementary school. Although the students in their study were of comparable age to ours and acquired a relatively transparent orthography, there are at least two significant differences. First, de Jong and van der Leij selected students on the basis of a pure measure of reading fluency only, whereas our test covered both reading fluency and reading comprehension, and we additionally considered spelling performance for sample selection. Second, the authors used a spoonerism task that requires segmentation at the onset-rime level only and could thus be considered a relatively low-demand task. In contrast, we used phonological awareness tasks that require processing at the level of phonemes (i.e., recognizing, manipulating, and reversing of phonemes), which are more demanding and complex. In a second cross-sectional study reported in the same article by de Jong and van der Leij, grade 4 students with LD exhibited severe phonological awareness deficits when task demands increased by the inclusion of a nonword phoneme deletion task, a task quite similar to our vowel substitution task. Taken together, a persistent deficit in phonological awareness in older students with LD in the learning of a relatively regular writing system might thus be limited to the level of phonemes and to tasks that are more complex (see also Landerl & Wimmer, 2000).

This notion raises the question of whether the persistent deficits in phonological awareness that we (and others) revealed in older students with LD acquiring a transparent orthography are genuine deficits in phonological awareness or whether they are about memory demands (i.e., phonological loop, working memory; de Jong & van der Leij, 2003; Landerl & Wimmer, 2000). According to Ramus and Szenkovits (2008), the primary deficit in students with LD is an impaired verbal short-term memory, leading to deficits in the processing of phonological information because "the auditory…representations of people with dyslexia are intact, but…they have difficulties accessing them under certain conditions involving storage in short-term memory, speeded or repeated retrievals,… and other task difficulty factors" (p. 139). Thus, our finding of persistent deficits in rather complex phonological awareness tasks in older students with LD acquiring the transparent orthography of German may be attributable to deficits in the phonological loop or potential deficits in working memory in general rather than in phonological awareness per se (de Jong & van der Leij, 2003; Landerl, Wimmer, & Frith, 1997).

From a conceptual view, there is one important distinctive feature in the differentiation between phonological awareness tasks and memory tasks, which is the distinction between implicit and explicit phonological processing (cf. Hulme & Snowling, 2009). Phonological loop tasks (e.g., digit or word span) are implicit, as these tasks require only the access to and the maintenance of phonological representations but no additional awareness of the sound structure of these representations because participants are instructed to simply repeat the presented item order rather than manipulate the stimuli. Further, in the case of digit and word span tasks, the phonological representations are likely to be processed at the lexical level, as the stimuli are quite familiar.

In contrast, phonological awareness tasks are explicit, as they require (a) an awareness of (i.e., conscious reflection on) language being composed of sounds and (b) the manipulation of these sounds in consciousness. Accordingly, they have much in common with tasks capturing the central executive of Baddeley's (1986) working memory model (i.e., the superior system in working memory that controls ongoing cognitive processing). Performing those executive tasks requires one to remember some task elements (i.e., whole words instead of phonemes, as in the case of phonological awareness tasks) and ignore or inhibit other elements while completing task-relevant operations (Swanson et al., 2009). In fact, a number of attempts have already been made to specify the nature of different phonological awareness tasks (e.g., Oakhill & Kyle, 2000; Snowling, Hulme, Smith, & Thomas, 1994; Yopp, 1988). Of interest here is the classification provided by Yopp (1988) as she separated a simple phonological awareness factor (with tasks such as segmentation, blending, or sound isolation) and a compound phonological awareness factor in tasks requiring the short-term maintenance of sounds during performance of additional operations in working memory (e.g., reversing or odd-one-out tasks; see also Baddeley, 1986; Tunmer & Hoover, 1992; Wagner & Torgesen, 1987). According to this classification, the rather complex phonological awareness tasks that we used in our study can be assigned to the compound factor and thus tap working memory to a greater extent.

Increasing Group Differences in the Phonological Loop

Concerning the development of the phonological loop, we also found evidence in favor of the deficit model, which pointed to a cascading deficit model: Group differences in the capacity of the phonological loop increased from grade 3 to grade 5. In other words, the gap between typically achieving students and students with LD regarding the ability to represent, store, and retrieve phonological information opened over time. This was due to the fact that LD's growth from grade 4 to grade 5 was smaller than the respective growth shown from grade 3 to grade 4. Although the phonological loop of the LD group increased over the three-year period, its positive growth was reduced (i.e., flatter) from grade 4 to grade 5 when compared with that demonstrated by the control group.

A deficit in the development of the phonological loop capacity is in accordance with the findings of several cross-sectional studies reporting differences among adults with and without LD (Siegel, 1994), the reported meta-analyses (Melby-Lervåg et al., 2012; Swanson, 2003; Swanson et al., 2009), and the longitudinal studies by Korhonen (1995) and Fischbach et al. (2014). The question arises, however, as to why differences between students with and without LD in the functionality of the phonological loop increased over the considered period of three years. At least two factors may contribute to this effect.

First, developmental maturation processes within the phonological loop that differ between LD and non-LD students may be responsible for this effect. Specifically, there is evidence that developmental increases in the phonological loop beyond the age of 7 are merely due to increased speed and efficiency of the rehearsal process (Gathercole, 1998). For example, when rehearsal rate increases in students, more items can be maintained within the phonological store, thus enhancing overall functionality of the phonological loop. Important with respect to the reported effect is the longitudinal study by Fischbach et al. (2014), which showed that students with LD, as compared with their typically achieving peers, developed a deficit in the automatization of subvocal rehearsal during the course of grade 5, which was not apparent in grade 4. This result suggests that students in the control group, but not the students with LD, underwent a critical maturation process in their articulation speed, which in turn resulted in an increased functional capacity of the phonological loop. Because the increasing differences that we found in our study developed between grade 4 and grade 5, differences in the automatization of subvocal rehearsal that, according to Fischbach et al. (2014), occur during this age might explain the results. Also of interest is a study conducted by Poloczek, Henry, Messer, and Büttner (2019), which showed that at approximately 10 years of age, cumulative rehearsal (as opposed to single naming) increasingly emerges as a memory strategy to boost memory performance further. More importantly, the authors revealed that high-performing 10-year-old students used cumulative rehearsal to improve immediate memory more often and more efficiently than their low-performing peers. Taken together, it is likely to suggest that the divergent development in our study might be partly due to those different rehearsal dynamics in the low-performing LD group and the high-performing control group that start around the age of 10 (grades 4 and 5).

The second factor contributing to the increasing differences in functionality of the phonological loop between the two groups is that cumulative experience in reading and spelling seems to have an impact on the development of the phonological loop, too. The limited phonological loop growth of the LD group may be a consequence of an individual's reading level rather than the cause of LD. For example, Demoulin and Kolinsky (2016) synthesized the growing body of evidence that developmental changes in the phonological loop may be, at least to some extent, a consequence of reading skill and practice. Specifically, the authors showed that increasing efficiency in decoding abilities boosts subvocal rehearsal mechanisms in the phonological loop, which results in increased performance in memory span tasks. Because good readers are more intrinsically motivated to read and, hence, receive more automaticity in reading than low-performing readers are, good readers' subvocal rehearsal processes are more likely to benefit from reading practice, and increasing differences between the LD and control groups in the phonological loop are the consequence. Likewise, learning to read improves phonemic representations, which in turn influence the performance in tasks assessing the phonological loop. The quality of the phonological representations is likely to be influenced by phonological awareness and literacy and, in turn, increases the quality of phonological coding in memory tasks (e.g., tasks such as digit or word span tasks involve phonological coding of heard information in a sound-based representation system; Baddeley, 1986). Supporting this hypothesis, Melby-Lervåg and Hulme (2010) and Park, Ritter, Lombardino, Wiseheart, and Sherman (2014) demonstrated that phonemic awareness training in young students improved performance in serial recall tasks (in both digit and word span tasks, that is, in tasks that are most dependent on phonological representations).

In simple terms, students with early poor phonological processing skills were impeded in their subsequent literacy acquisition and hence read less than students with adequate phonological processing. Reading less may have led to relatively limited development of the phonological loop in students with early poor phonological processing skills (i.e., less rehearsal or poorly specified phonological representations leading to less efficient reintegration or phonological coding), which in turn increased the differences between students with and without LD.

Limitations

Some limitations of the present study have to be acknowledged. The first limitation concerns our attrition rates. Longitudinal studies are complicated by the (almost inevitable) loss of individuals between testing waves due to successive dropout. In our study, between grades 3 and 5, 20% of the data were missing. We therefore dealt with the attrition rate by applying statistical methods that analyze data without loss of information. Second, for some of the phonological indicators, internal consistency was quite low (<0.80). We therefore used latent variable modeling and several indicators per factor to address this issue. Third, it has to be kept in mind that the relation between phonological processing and literacy seems to be reciprocal and that literacy itself can have differential developmental trajectories just as phonological processing does. A related issue is that we did not track students' academic performance longitudinally and hence do not know how their reading and spelling skills developed further with respect to norm-referenced tests. However, students' final German grades at the end of elementary school demonstrate that the students with LD were still struggling with reading and spelling: Approximately 80% of the students with LD obtained grade 3 (satisfactory) or grade 4 (sufficient) in their report card. In contrast, approximately 80% of the control group students received grade 2 (good) or grade 1 (excellent). This is in accordance with longitudinal studies pointing to a high interindividual rank-order stability of reading and spelling skills (Pfost et al., 2014).

A fourth limitation concerns the generalization of our results, as they are limited to alphabetic writing systems having a relatively transparent orthography, such as German, and are further limited to the specific group of students in our sample. That is, our results on LD have to be interpreted against the backdrop of the diagnostic criteria and measures used. Although our LD group was carefully selected as part of an extreme group design and although our cutoff scores followed the diagnostic guidelines commonly used in German educational practice, thresholds for defining LD are to some extent arbitrary and controversial. We are aware that only using standardized tests to find students with LD is not sufficient for a clinical diagnosis, as it underscores the complexity of diagnosing LD. Other studies may apply different classification criteria or diagnostic measures (e.g., other degrees of discrepancy, a Response to Intervention approach, family history), which may lead to slightly different results. Our diagnostic criteria eliminated many of the lowest performing readers and spellers because we focused on homogeneous groups of students with LD (with average intelligence and average math performance). Finally, the number of students with LD who could have been easily remediated by an early intervention cannot be known.

Theoretical Implications

With regard to theoretical implications, our study demonstrated that the various manifestations of a phonological deficit in students with LD followed distinct developmental trajectories. Deficits in phonological awareness and the phonological loop persisted or increased, whereas the alphanumeric RAN deficit was developmentally limited. This finding of decreasing differences in alphanumeric RAN may question the nature of the phonological deficit in students with LD. However, one may argue against this conclusion that a RAN deficit is only one of the several possible manifestations of a phonological core deficit. Or, as supposed by other researchers, differences in RAN between students with and without LD may reflect a partly different, phonology-independent deficit (Kirby et al., 2010). Whatever the nature of the phonological deficit, we agree with de Jong and van der Leij (2003), who claimed that "the manifestations of a phonological deficit can change over time" (p. 36).

We observed two main impairments in students with LD (i.e., persistent deficits in phonological awareness, increasing deficits in the phonological loop). Accordingly, students with LD demonstrated intractable deficits in the short-term storing and manipulating of phonological representations. As outlined earlier, our phonological awareness tasks are rather complex in terms of concurrent memory demands. Even our students with LD would probably be able to solve more simple phonological awareness tasks, such as phoneme segmentation, because research has shown that the combination of the transparent German orthography with systematic phonics teaching enables even low-performing readers to acquire an adequate level of phonological awareness (cf. Landerl & Wimmer, 2000). This would be in line with the idea proposed by others (de Jong & van der Leij, 2003; Ramus & Szenkovits, 2008) that in older students with LD, phonological awareness at the level of simple tasks is intact but fragile, and a phonological awareness deficit in these students becomes only manifest if task demands (and concurrent memory load) increase. One could hypothesize as to whether this vulnerability is due to a possible underlying deficit of LD in the central executive in terms of Baddeley's (1986) working memory model, but this cannot be determined based on our study. Further research is needed to address this specific question. Either way, our results go well with the phonological access hypothesis proposed by Ramus and Szenkovits (2008), that students with LD struggle persistently in phonological tasks that involve either explicit and complex mental operations on phonological representations or a high load on the phonological loop.

Implications for Educational Practice

Finally, we discuss some practical implications of our results. A common belief among teachers and parents is that deficits in students with LD disappear over time. Based on our results, however, this is not applicable for at least some of the phonological deficits that are associated with LD (i.e., deficits in phonological awareness and the phonological loop). As we outlined in the introduction, an (additional) intervention and diagnosis of phonological processing subskills, as recommended by some clinical practice guidelines (e.g., Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie, Psychosomatik und Psychotherapie, 2015), in older students would only make sense if these phonological deficits were still present in these students. With regard to this issue, our results point to the conclusion that as far as students in grades 3–5 are concerned, interventions that help these phonological processing skills develop further are still to be commended (except for RAN).

Two different approaches might help in accomplishing this goal. First, given the reciprocal relation between phonological processing and literacy development, interventions that concentrate on reading and spelling activities might also transfer to phonological processing. Second, according to a meta-analysis by Suggate (2016), interventions focused on reading/spelling and phonological awareness seem to be promising to help students with LD. A theoretical account of why those combined interventions are effective is expressed in the phonological linkage hypothesis proposed by Hatcher, Hulme, and Ellis (1994). According to the authors, those combined interventions help students form explicit relations between phonology and written language by providing learning activities that directly link phonemes to words. Because of their reciprocal relation, phonological processing and literacy boost each other in this learning process and make the intervention maximally effective. In contrast, the capacity of the phonological loop is hardly influenced by interventions (Melby-Lervåg & Hulme, 2013). A more promising approach to help students with LD maintain phonological information is by applying teaching principles that reduce the verbal memory demands in worksheets and tasks (e.g., Alloway, 2006; Mähler, Jörns, Radtke, & Schuchardt, 2015).

Conclusion

In summary and following Pfost et al. (2014), we conclude that there is no support for the overall validity of a single developmental pattern in German students with LD regarding the development of phonological processing. Depending on the aspect of phonological information processing that is considered, we found evidence in favor of persistent deficits (with respect to phonological awareness), increasing differences (the phonological loop), and decreasing deficits (alphanumeric RAN).

NOTES

The present study was part of the Entwicklungsstörungen schulischer Fähigkeiten (Developmental Disorders of Scholastic Skills) research initiative and was funded by Germany's Federal Ministry of Education and Research (grant 01GJ1012 A-D). We thank Christina Balke-Melcher, Anne Fischbach, Julia Klesczewski, and Kirsten Schuchardt for their involvement in carrying out this study. We further thank Sina Simone Huschka and Jeremy Thorne for their involvement in the final stages of the manuscript process.

REFERENCES

Alloway, T.P. (2006). How does working memory work in the classroom? *Educational Research Review*, *1*(4), 134–139.

Alloway, T.P., Gathercole, S.E., Willis, C., & Adams, A.M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, *87*(2), 85–106. https://doi.org/10.1016/j.jecp.2003.10.002

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

Arbuckle, J.L. (1996). Full information estimation in the presence of incomplete data. In G.A. Marcoulides & R.E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 243–278). New York, NY: Erlbaum.

Babayiğit, S., & Stainthorp, R. (2011). Modeling the relationships between cognitive–linguistic skills and literacy skills: New insights from a transparent orthography. *Journal of Educational Psychology*, *103*(1), 169–189. https://doi.org/10.1037/a0021671

Baddeley, A.D. (1986). *Working memory*. Oxford, UK: Clarendon.

Baddeley, A.D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575–589. https://doi.org/10.1016/S0022-5371(75)80045-4

Baraldi, A.N., & Enders, C.K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology*, *48*(1), 5–37. https://doi.org/10.1016/j.jsp.2009.10.001

Bast, J., & Reitsma, P. (1998). Analyzing the development of individual differences in terms of Matthew effects in reading: Results from a Dutch longitudinal study. *Developmental Psychology*, *34*(6), 1373–1399. https://doi.org/10.1037/0012-1649.34.6.1373

Birkel, P. (2007). *WRT 2+: Weingartener Grundwortschatz Rechtschreib-Test für zweite und dritte Klassen: Manual* [Weingarten's spelling test of basic vocabulary for second and third grades: Manual]. Göttingen, Germany: Hogrefe.

Brandenburg, J., Klesczewski, J., Fischbach, A., Schuchardt, K., Büttner, G., & Hasselhorn, M. (2015). Working memory in children with learning disabilities in reading versus spelling: Searching for overlapping and specific cognitive factors. *Journal of Learning Disabilities*, *48*(6), 622–634. https://doi.org/10.1177/0022219414521665

Byrne, B.M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York, NY: Routledge.

Caravolas, M., Volin, J., & Hulme, C. (2005). Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: Evidence from Czech and English children. *Journal of Experimental Child Psychology*, *92*(2), 107–139. https://doi.org/10.1016/j.jecp.2005.04.003

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

Cattell, R.B., Weiß, R.H., & Osterland, J. (1997). *Grundintelligenztest Skala 1—CFT 1: Handanweisung für die Durchführung, Auswertung und Interpretation* [Basic intelligence test scale 1—CFT 1: Instructions for implementation, evaluation and interpretation] (15th ed.). Göttingen, Germany: Hogrefe.

Dandache, S., Wouters, J., & Ghesquière, P. (2014). Development of reading and phonological skills of children at family risk for dyslexia: A longitudinal analysis from kindergarten to sixth grade. *Dyslexia*, *20*(4), 305–329. https://doi.org/10.1002/dys.1482

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

de Jong, P.F., & Vrielink, L.O. (2004). Rapid automatic naming: Easy to measure, hard to improve (quickly). *Annals of Dyslexia*, *54*(1), 65–88. https://doi.org/10.1007/s11881-004-0004-1

Demoulin, C., & Kolinsky, R. (2016). Does learning to read shape verbal working memory? *Psychonomic Bulletin & Review*, *23*(3), 703–722. https://doi.org/10.3758/s13423-015-0956-7

Denckla, M.B. (1979). Childhood learning disabilities. In K.M. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology* (pp. 535–537). New York, NY: Oxford University Press.

Denckla, M.B., & Rudel, R. (1974). Rapid "automatized" naming of pictured objects, colors, letters and numbers by normal children. *Cortex*, *10*(2), 186–202. https://doi.org/10.1016/S0010-9452(74)80009-2

Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie, Psychosomatik und Psychotherapie [German Society for Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy]. (2015). *S3 Leitlinie: Diagnostik und Behandlung von Kindern und Jugendlichen mit Lese- und/oder Rechtschreibstörung* [S3 guideline: Diagnostics and treatment of children and adolescents with reading and/or spelling disorders]. Retrieved from https://www.awmf.org/uploads/tx_szleitlinien/028-044l_S3_Lese-Rechtschreibst%c3%b6rungen_Kinder_Jugendliche_2015-06.pdf

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

Enders, C.K., & Bandalos, D.L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, *8*(3), 430–457. https://doi.org/10.1207/S15328007SEM0803_5

Engen, L., & Høien, T. (2002). Phonological skills and reading comprehension. *Reading and Writing*, *15*(7/8), 613–631. https://doi.org/10.1023/A:1020958105218

Fischbach, A., Könen, T., Rietz, C.S., & Hasselhorn, M. (2014). What is not working in working memory of children with literacy disorders? Evidence from a three-year-longitudinal study. *Reading and Writing*, *27*(2), 267–286. https://doi.org/10.1007/s11145-013-9444-5

Francis, D.J., Shaywitz, S.E., Stuebing, K.K., Shaywitz, B.A., & Fletcher, J.M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology*, *88*(1), 3–17. https://doi.org/10.1037/0022-0663.88.1.3

Frith, U. (1986). A developmental framework for developmental dyslexia. *Annals of Dyslexia*, *36*(1), 67–81. https://doi.org/10.1007/BF02648022

Furnes, B., & Samuelsson, S. (2011). Phonological awareness and rapid automatized naming predicting early development in reading and spelling: Results from a cross-linguistic longitudinal study. *Learning and Individual Differences*, *21*(1), 85–95. https://doi.org/10.1016/j.lindif.2010.10.005

Gathercole, S.E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry*, *39*(1), 3–27. https://doi.org/10.1111/1469-7610.00301

Geiser, C. (2012). *Data analysis with Mplus*. New York, NY: Guilford.

Georgiou, G.K., Parrila, R., & Liao, C.H. (2008). Rapid naming speed and reading across languages that vary in orthographic consistency. *Reading and Writing*, *21*(9), 885–903. https://doi.org/10.1007/s11145-007-9096-4

Georgiou, G.K., & Stewart, B. (2013). Is rapid automatized naming automatic? *Preschool & Primary Education*, *1*(1), 67–81. https://doi.org/10.12681/ppej.46

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

Goswami, U. (1997). Learning to read in different orthographies: Phonological awareness, orthographic representations and dyslexia. In C. Hulme & M.J. Snowling (Eds.), *Dyslexia: Biology, identification and intervention* (pp. 131–152). London, UK: Singular.

Hasselhorn, M., Schuman-Hengsteler, R., Gronauer, J., Grube, D., Mähler, C., Schmid, I., … Zoelch, C. (2012). *AGTB 5–12: Arbeitsgedächtnistestbatterie für Kinder von 5 bis 12 Jahren: Manual* [AGTB 5–12: Working memory test battery for children ages 5 to 12 years: Manual]. Göttingen, Germany: Hogrefe.

Hatcher, P.J., Hulme, C., & Ellis, A.W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development*, *65*(1), 41–57. https://doi.org/10.2307/1131364

Heck, R.H., & Thomas, S.L. (2015). *An introduction to multilevel modeling techniques: MLM and SEM approaches using Mplus* (3rd ed.). New York, NY: Routledge.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hulme, C., & Snowling, M.J. (2009). *Developmental disorders of language learning and cognition*. Chichester, UK: Wiley-Blackwell.

Jeličić, H., Phelps, E., & Lerner, R.M. (2009). Use of missing data methods in longitudinal studies: The persistence of bad practices in developmental psychology. *Developmental Psychology*, *45*(4), 1195–1199. https://doi.org/10.1037/a0015665

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

Kline, R.B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: Guilford.

Korhonen, T.T. (1995). The persistence of rapid naming problems in children with reading disabilities: A nine-year follow-up. *Journal of Learning Disabilities*, *28*(4), 232–239. https://doi.org/10.1177/002221949502800405

Krajewski, K., Liehm, S., & Schneider, W. (2004). *DEMAT 2+: Deutscher Mathematiktest für zweite Klassen: Manual* [DEMAT 2+: German mathematical test for second grade: Manual]. Göttingen, Germany: Beltz Test.

Kuppen, S.E., & Goswami, U. (2016). Developmental trajectories for children with dyslexia and low IQ poor readers. *Developmental Psychology*, *52*(5), 717–734. https://doi.org/10.1037/a0040207

Landerl, K., Freudenthaler, H.H., Heene, M., de Jong, P.F., Desrochers, A., Manolitsis, G., … Georgiou, G.K. (2019). Phonological awareness and rapid automatized naming as longitudinal predictors of reading in five alphabetic orthographies with varying degrees of consistency. *Scientific Studies of Reading*, *23*(3), 220–234. https://doi.org/10.1080/10888438.2018.1510936

Landerl, K., & Thaler, V. (2006). Reading and spelling acquisition and dyslexia in German. In R.M. Joshi & P.G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 121–134). New York, NY: Erlbaum.

Landerl, K., & Wimmer, H. (2000). Deficits in phoneme segmentation are not the core problem of dyslexia: Evidence from German and English children. *Applied Psycholinguistics*, *21*(2), 243–262. https://doi.org/10.1017/S0142716400002058

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology*, *100*(1), 150–161. https://doi.org/10.1037/0022-0663.100.1.150

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

Lenhard, W., & Schneider, W. (2006). *ELFE 1–6: Ein Leseverständnistest für Erst- bis Sechstklässler: Manual* [ELFE 1–6: A reading comprehension test for grades 1 to 6: Manual]. Göttingen, Germany: Hogrefe.

Mähler, C., Jörns, C., Radtke, E., & Schuchardt, K. (2015). Chancen und Grenzen eines Trainings des Arbeitsgedächtnisses bei Kindern mit und ohne Lese-/Rechtschreibschwierigkeiten [Opportunities and limitations of training working memory in children with and without reading and writing disabilities]. *Zeitschrift für Erziehungswissenschaft*, *18*(3), 453–471. https://doi.org/10.1007/s11618-015-0643-5

Mardia, K.V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhyā B*, *36*(2), 115–128.

Martin, D.H., & Barry, C. (2012). Writing nonsense: The interaction between lexical and sublexical knowledge in the priming of nonword spelling. *Psychonomic Bulletin & Review*, *19*(4), 691–698. https://doi.org/10.3758/s13423-012-0261-7

Mazzocco, M.M., & Grimm, K.J. (2013). Growth in rapid automatized naming from grades K to 8 in children with math or reading disabilities. *Journal of Learning Disabilities*, *46*(6), 517–533. https://doi.org/10.1177/0022219413477475

McArdle, J.J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, *60*, 577–605. https://doi.org/10.1146/annurev.psych.60.110707.163612

Melby-Lervåg, M., & Hulme, C. (2010). Serial and free recall in children can be improved by training: Evidence for the importance of phonological and semantic representations in immediate memory tasks. *Psychological Science*, *21*(11), 1694–1700. https://doi.org/10.1177/0956797610385355

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? *A meta-analytic review. Developmental Psychology*, *49*(2), 270–291. https://doi.org/10.1037/a0028228

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

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

Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., … Landerl, K. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction*, *29*(1), 65–77. https://doi.org/10.1016/j.learninstruc.2013.09.003

Moll, K., Wallner, R., & Landerl, K. (2012). Kognitive Korrelate der Lese-, Leserechtschreib- und der Rechtschreibstörung [Cognitive correlates of deficits in reading and/or spelling]. *Lernen und Lernstörungen*, *1*(1), 7–19. https://doi.org/10.1024/2235-0977/a000002

Muthén, B.O. (2002). Beyond *SEM*: General latent variable modeling. *Behaviormetrika*, *29*(1), 81–117. https://doi.org/10.2333/bhmk.29.81

Muthén, L.K., & Muthén, B.O. (2010). *Mplus: Statistical analysis with latent variables: User's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

Nikolopoulos, D., Goulandris, N., Hulme, C., & Snowling, M.J. (2006). The cognitive bases of learning to read and spell in Greek: Evidence from a longitudinal study. *Journal of Experimental Child Psychology*, *94*(1), 1–17. https://doi.org/10.1016/j.jecp.2005.11.006

Oakhill, J., & Kyle, F. (2000). The relation between phonological awareness and working memory. *Journal of Experimental Child Psychology*, *75*(2), 152–164. https://doi.org/10.1006/jecp.1999.2529

O'Shaughnessy, T.E., & Swanson, H.L. (1998). Do immediate memory deficits in students with learning disabilities in reading reflect a developmental lag or deficit? A selective meta-analysis of the literature. *Learning Disability Quarterly*, *21*(2), 123–148. https://doi.org/10.2307/1511341

Papadopoulos, T.C., Spanoudis, G.C., & Georgiou, G.K. (2016). How is RAN related to reading fluency? A comprehensive examination of the prominent theoretical accounts. *Frontiers in Psychology*, *7*, article 1217. https://doi.org/10.3389/fpsyg.2016.01217

Park, J., Ritter, M., Lombardino, L., Wiseheart, R., & Sherman, S. (2014). Phonological awareness intervention for verbal working memory skills in school-age children with specific language impairment and concomitant word reading difficulties. *International Journal of Research Studies in Language Learning*, *3*(4), 3–22. https://doi.org/10.5861/ijrsll.2013.534

Perfetti, C.A. (1992). The representation problem in reading acquisition. In P.B. Gough, L.C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Hillsdale, NJ: Erlbaum.

Pfost, M., Hattie, J., Dörfler, T., & Artelt, C. (2014). Individual differences in reading development: A review of 25 years of empirical research on Matthew effects in reading. *Review of Educational Research*, *84*(2), 203–244. https://doi.org/10.3102/0034654313509492

Poloczek, S., Henry, L.A., Messer, D.J., & Büttner, G. (2019). Do children use different forms of verbal rehearsal in serial picture recall tasks? A multi-method study. *Memory*, *27*(6), 758–771. https://doi.org/10.1080/09658211.2018.1563615

Ramus, F., Marshall, C.R., Rosen, S., & van der Lely, H.K.J. (2013). Phonological deficits in specific language impairment and developmental dyslexia: Towards a multidimensional model. *Brain*, *136*(2), 630–645. https://doi.org/10.1093/brain/aws356

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

Satorra, A., & Bentler, P.M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507–514. https://doi.org/10.1007/BF02296192

Savage, R., Pillay, V., & Melidona, S. (2008). Rapid serial naming is a unique predictor of spelling in children. *Journal of Learning Disabilities*, *41*(3), 235–250. https://doi.org/10.1177/0022219408315814

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

Share, D.L. (2008). Orthographic learning, phonological recoding, and self-teaching. *Advances in Child Development and Behavior*, *36*, 31–82. https://doi.org/10.1016/S0065-2407(08)00002-5

Siddaiah, A., Saldanha, M., Venkatesh, S.K., Ramachandra, N.B., & Padakannaya, P. (2016). Development of rapid automatized naming (RAN) in simultaneous Kannada-English biliterate children. *Journal of Psycholinguistic Research*, *45*(1), 177–187. https://doi.org/10.1007/s10936-014-9338-y

Siegel, L.S. (1994). Working memory and reading: A life-span perspective. *International Journal of Behavioral Development*, *17*(1), 109–124. https://doi.org/10.1177/016502549401700107

Snowling, M.J. (2000). *Dyslexia*. Oxford, UK: Blackwell.

Snowling, M.J., Hulme, C., Smith, A., & Thomas, J. (1994). The effects of phonetic similarity and list length on children's sound categorization performance. *Journal of Experimental Child Psychology*, *58*(1), 160–180. https://doi.org/10.1006/jecp.1994.1030

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading*

Research Quarterly, *21*(4), 360–407. https://doi.org/10.1598/RRQ.21.4.1
Stanovich, K.E., Nathan, R.G., & Zolman, J.E. (1988). The developmental lag hypothesis in reading: Longitudinal and matched reading-level comparisons. *Child Development*, *59*(1), 71–86. https://doi.org/10.2307/1130390
Stanovich, K.E., & Siegel, L.S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology*, *86*(1), 24–53. https://doi.org/10.1037/0022-0663.86.1.24
Stock, C., Marx, P., & Schneider, W. (2003). *BAKO 1–4: Basiskompetenzen für Lese-Rechtschreibleistungen: ein Test zur Erfassung der phonologischen Bewusstheit vom ersten bis vierten Grundschuljahr* [BAKO 1–4: Basic skills for reading and spelling: A test for measuring phonological awareness from first to fourth grade]. Göttingen, Germany: Hogrefe.
Suggate, S.P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities*, *49*(1), 77–96. https://doi.org/10.1177/0022219414528540
Swanson, H.L. (2003). Age-related differences in learning disabled and skilled readers' working memory. *Journal of Experimental Child Psychology*, *85*(1), 1–31. https://doi.org/10.1016/S0022-0965(03)00043-2
Swanson, H.L., & Hsieh, C.J. (2009). Reading disabilities in adults: A selective meta-analysis of the literature. *Review of Educational Research*, *79*(4), 1362–1390. https://doi.org/10.3102/0034654309350931
Swanson, H.L., Zheng, X., & Jerman, O. (2009). Working memory, short-term memory, and reading disabilities: A selective meta-analysis of the literature. *Journal of Learning Disabilities*, *42*(3), 260–287. https://doi.org/10.1177/0022219409331958
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
Tunmer, W.E., & Hoover, W.A. (1992). Cognitive and linguistic factors in learning to read. In P.B. Gough, L.C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 175–214). Hillsdale, NJ: Erlbaum.
Ullman, J.B. (2007). Structural equation modeling. In B.G. Tabachnick & L.S. Fidell (Eds.), *Using multivariate statistics* (5th ed., pp. 676–780). Boston, MA: Pearson.
van Daal, V., & van der Leij, A. (1999). Developmental dyslexia: Related to specific or general deficits? *Annals of Dyslexia*, *49*(1), 71–104. https://doi.org/10.1007/s11881-999-0020-2
van den Bos, K.P., Zijlstra, B.J.H., & lutje Spelberg, H.C. (2002). Life-span data on continuous-naming speeds of numbers, letters, colors, and pictured objects, and word-reading speed. *Scientific Studies of Reading*, *6*(1), 25–49. https://doi.org/10.1207/S1532799XSSR0601_02
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. https://doi.org/10.1046/j.0021-9630.2003.00305.x
Wagner, R.K., & Torgesen, J.K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*(2), 192–212. https://doi.org/10.1037/0033-2909.101.2.192
Wagner, R.K., Torgesen, J.K., & Rashotte, C.A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*, *30*(1), 73–87. https://doi.org/10.1037/0012-1649.30.1.73
Wang, J., & Wang, X. (2012). *Structural equation modeling: Applications using Mplus*. Chichester, UK: John Wiley & Sons.
Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied Psycholinguistics*, *14*(1), 1–33. https://doi.org/10.1017/S0142716400010122
Wimmer, H., & Mayringer, H. (2002). Dysfluent reading in the absence of spelling difficulties: A specific disability in regular orthographies. *Journal of Educational Psychology*, *94*(2), 272–277. https://doi.org/10.1037/0022-0663.94.2.272
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
Wolf, M., Bowers, P.G., & Biddle, K. (2000). Naming-speed processes, timing, and reading: A conceptual review. *Journal of Learning Disabilities*, *33*(4), 387–407. https://doi.org/10.1177/002221940003300409
World Health Organization. (2018). *International classification of diseases* (11th rev. ed.). Retrieved from https://www.who.int/classifications/icd/en/
Yopp, H.K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly*, *23*(2), 159–177. https://doi.org/10.2307/747800

Submitted September 26, 2018
Final revision received December 1, 2019
Accepted December 12, 2019

CLAUDIA SCHMIDT (corresponding author) is a doctoral student at the Carl von Ossietzky University Oldenburg, Germany; email claudia.schmidt@uol.de. Her research interests include learning disabilities in reading and spelling and associated cognitive correlates, diagnostics, and individual support for students with learning disabilities.

JANIN BRANDENBURG is a postdoctoral researcher at the Center for Research on Individual Development and Adaptive Education of Children at Risk at the DIPF | Leibniz Institute for Research and Information, Frankfurt, Germany; email brandenburg@dipf.de. Her research focuses on learning disabilities, working memory, and development of language competencies in monolingual and bilingual students.

JENNY BUSCH is an educational psychologist in Oldenburg, Germany; email jenny.busch@uol.de. At the time of this study, she was a research associate at the Carl von Ossietzky University Oldenburg, Germany. She is interested in assessment of learning disabilities in math and basic arithmetic facts.

GERHARD BÜTTNER is a professor in the Department of Educational Psychology at Goethe University Frankfurt and a project leader at the Center for Research on Individual Development and Adaptive Education of Children at Risk at the DIPF | Leibniz Institute for Research and Information, Frankfurt, Germany; email buettner@paed.psych.uni-frankfurt.de. The main focus of his work is self-regulated learning, development of working memory, metacognition, learning disabilities, and intellectual disabilities.

DIETMAR GRUBE is a professor in the Institute for Education at the Carl von Ossietzky University Oldenburg, Germany; email dietmar.grube@uol.de. His research focuses on the development of arithmetic skills, basic arithmetic facts, learning disabilities in math, and working memory.

CLAUDIA MÄHLER is a professor in the Institute of Psychology at the University of Hildesheim, Germany; email maehler@uni-hildesheim.de. The current foci of her research are learning and intellectual disabilities, diagnosis and training of working memory functioning, and cognitive development in early childhood.

MARCUS HASSELHORN is a professor in the Department of Education and Development at Goethe University Frankfurt and the director of the Department of Education and Human Development at the DIPF | Leibniz Institute for Research and Information in Education, Frankfurt, Germany; email hasselhorn@dipf.de. His research focuses on the ontogenesis of individual conditions of successful learning, learning disabilities, early education, and educational diagnostics.

APPENDIX A

Detailed Description of the Sample: Attrition and Missing Data

The attrition rate in our sample was rather high (approximately 40%). Unfortunately, in many longitudinal studies, attrition rate is not mentioned, although missing data are (to some degree) almost inevitable (Jeličić, Phelps, & Lerner, 2009). To consider this problem, we performed various analyses. First, we tested for potential differences in several variables between students who dropped out of the study and those who remained. As can be seen in Table A1, students who dropped out of the study and those who remained were nearly identical. There were no differences regarding sociodemographic variables (i.e., age, socioeconomic status), academic achievement at grade 2 (i.e., intelligence, reading, spelling, mathematics), and phonological processing at grade 3 (i.e., RAN, phonological awareness, and the phonological loop for the LD group). In addition, the gender distribution was similar for students remaining in the study and those leaving, $\chi^2(1, N = 100) = 2.72$, $p = .10$ for the control group, and $\chi^2(1, N = 109) = 0.30$, $p = .87$ for the LD group. There was no systematic relation between the spoken language at home and attrition, $\chi^2(1, N = 86) = 1.98$, $p = .37$ for the control group, and $\chi^2(1, N = 96) = 0.59$, $p = .74$ for the LD group.

Second, we checked whether attrition rate was similar in both groups: During the course of the study, 40% of the control group students (20 at wave 2 and 20 at wave 3) and 31% of the students with LD (14 at wave 2 and 30 at wave 3) dropped out of the study. Chi-squared tests revealed that attrition rate was similar for both groups at wave 2, $\chi^2(1, N = 209) = 1.96$, $p = .16$, and at wave 3, $\chi^2(1, N = 209) = 1.02$, $p = .31$. Due to attrition (i.e., dropout, nonmonotone missing values), 20% of the data were missing, with no differences in frequency between the LD group (19%) and the control group (20%), $\chi^2(1, N = 209) = 0.12$, $p = .73$.

Third, we conducted Little's missing completely at random test (for the whole sample and separate for each group) to check whether the missing (completely) at random assumption was tenable. Little's test resulted in $\chi^2(545, N = 209) = 588.50$, $p = .10$ for the whole sample: $\chi^2(364, N = 109) = 372.11$, $p = .36$ for the LD group and $\chi^2(259, N = 100) = 284.33$, $p = .13$ for the control group. This indicated that no identifiable pattern existed in the missing data.

Given the results of these analyses, there is hardly any indication that the reason for missing data in our study is systematic. In addition, we also addressed the issue of attrition in our structural equation modeling analyses by using the MLR estimator. Specifically, MLR treats missing values with the FIML method. FIML is a state-of-the-art approach that is superior to traditional methods of handling missing values such as listwise deletion. Also, Enders and colleagues (Baraldi & Enders, 2010; Enders & Bandalos, 2001) and Arbuckle (1996) recommended using the FIML method, especially in the context of structural equation modeling. In addition, MLR is especially suitable for small and medium sample sizes (B.O. Muthén, 2002) and "allows missing completely at random (MCAR) and missing at random (MAR)" (Wang & Wang, 2012, p. 16). The FIML method is considered "a reasonable option for conducting the analysis with an incomplete data set" (Kline, 2016, p. 88).

TABLE A1
Descriptive Statistics and Group Differences for Students Who Did or Did Not Drop Out of the Study as a Function of Group

Measure	Control group						LD group					
	Dropout t1–t3		Non-dropout				Dropout t1–t3		Non-dropout			
	M	*SD*	*M*	*SD*	*F*	*p*	*M*	*SD*	*M*	*SD*	*F*	*p*
Age	8 years 6.26 months	3.81 months	8 years 6.10 months	5.43 months	0.03	.87	8 years 7.00 months	5.71 months	8 years 7.59 months	5.89 months	0.27	.61
Socioeconomic status	4.43	0.57	4.53	0.61	0.51	.48	4.20	0.72	4.45	0.60	3.28	.07
Intelligence (IQ score)	107.92	11.49	106.21	11.04	0.55	.46	107.63	9.11	108.36	10.09	0.15	.70
Mathematics (*T*-score)	53.68	4.76	54.16	5.93	0.18	.68	52.79	7.05	53.80	6.19	0.62	.43
Reading (*T*-score)	53.37	5.92	53.03	5.92	0.54	.46	38.89	7.63	40.67	8.14	1.27	.26
Spelling (*T*-score)	51.21	6.04	51.45	5.62	0.04	.84	37.40	5.90	38.02	6.60	0.25	.62
t1 letter naming	30.55	5.18	31.76	7.23	0.72	.40	35.12	7.98	34.38	7.77	0.23	.64
t1 digit naming	28.97	5.41	29.24	7.42	0.04	.85	34.88	9.14	34.43	7.44	0.07	.79
t1 object naming	47.08	6.94	48.13	8.77	0.39	.53	48.77	9.56	49.20	8.92	0.05	.82
t1 color naming	49.53	9.79	50.18	10.51	0.10	.76	51.95	11.59	51.80	9.99	0.01	.95
t1 phoneme reversal	9.74	4.42	9.21	5.04	0.28	.60	5.40	4.60	5.09	3.88	0.14	.71
t1 vowel length	4.82	2.52	4.51	2.91	0.29	.59	3.24	1.95	3.59	2.42	0.63	.43
t1 vowel substitution	9.29	2.58	9.41	2.40	0.06	.81	6.48	2.95	7.44	2.90	2.80	.10
t1 word span for one-syllable words	4.14	0.59	3.84	0.66	5.20	.03	3.77	0.50	3.71	0.64	0.27	.61
t1 word span for three-syllable words	3.22	0.43	3.06	0.44	3.12	.08	2.96	0.38	3.02	0.36	0.63	.43
t1 digit span	4.78	0.49	4.50	0.65	5.18	.03	4.18	0.59	4.17	0.57	0.01	.93

Note. LD = students with learning disabilities in reading and/or spelling; *M* = mean; *SD* = standard deviation; t1 = first testing wave; t3 = third testing wave.

APPENDIX B

Detailed Description of the MIMIC Model

TABLE B1
Testing the Equality of Baseline Factors for Specific Reading Disability and Specific Spelling Disability: Fit Indexes for Nested Model Comparison

Model	x^2	*df*	p_1	RMSEA	CFI	ΔCFI	ΔSB-x^2	Δ*df*	p_2
RAN_a freely estimated	43.72	22	.004	0.07	0.949	0.007	3.79	1	.052
RAN_a constrained	47.64	23	.002	0.07	0.942				
RAN_n freely estimated	48.33	23	.002	0.07	0.942	0.001	<1	1	.73
RAN_n constrained	48.89	24	.002	0.07	0.943				
PA freely estimated	74.46	50	.01	0.05	0.955	0.000	<1	1	.86
PA constrained	74.37	51	.02	0.05	0.955				
PL freely estimated	73.76	49	.01	0.05	0.975	0.000	<1	1	.59
PL constrained	74.32	50	.01	0.05	0.975				

Note. ΔSB-x^2 = Satorra–Bentler scaled chi-squared difference test; CFI = comparative fit index; p_1 = probability value of model fit; p_2 = probability value obtained in the SB chi-squared difference test; PA = phonological awareness; PL = phonological loop; RAN_a = rapid automatized naming for alphanumeric stimuli; RAN_n = rapid automatized naming for nonalphanumeric stimuli; RMSEA = root mean square error of approximation.

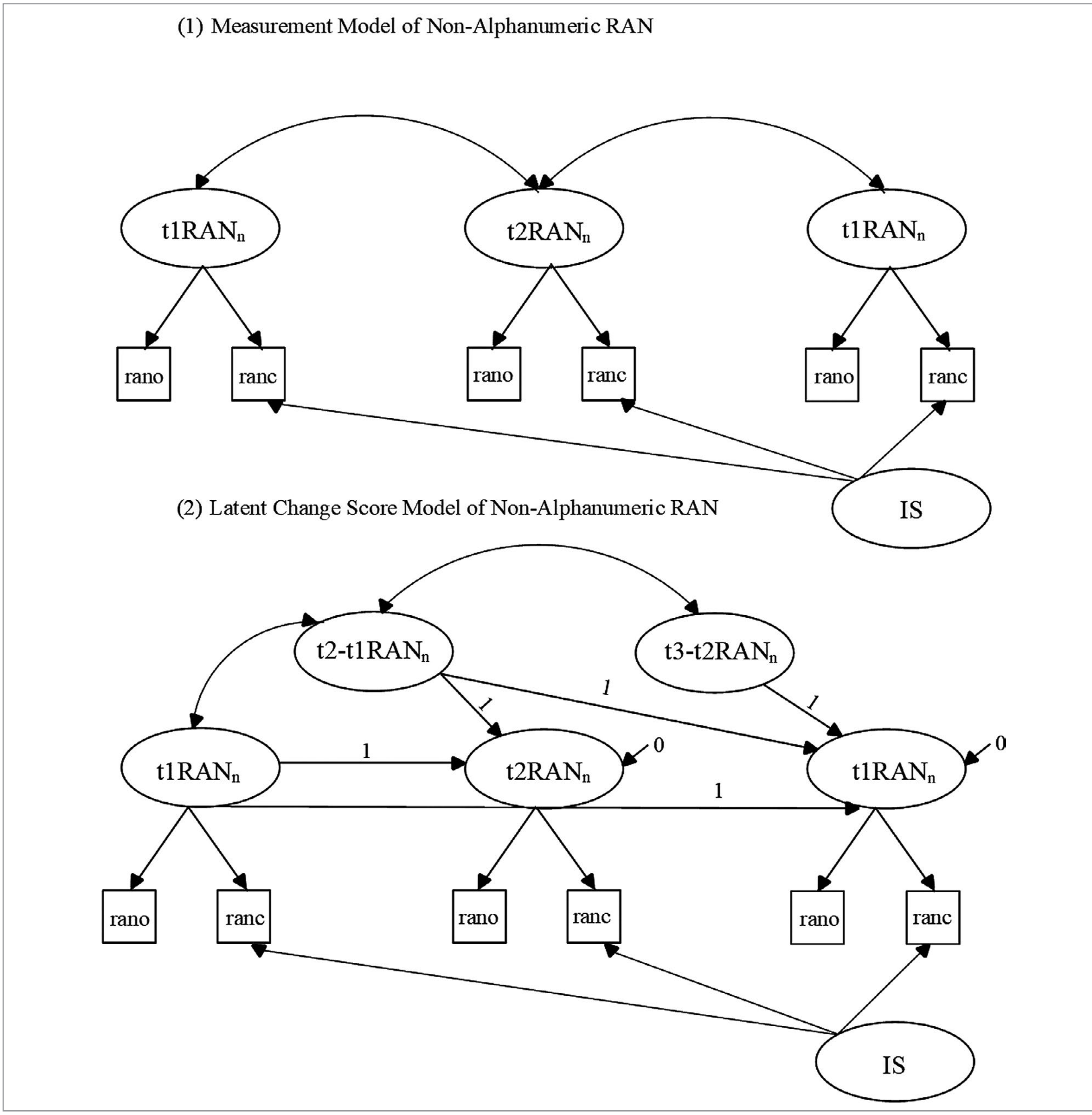
TABLE B2
Testing the Equality of Latent Change Factors for Specific Reading Disability and Specific Spelling Disability: Fit Indexes for Nested Model Comparison

Model	x^2	*df*	p_1	RMSEA	CFI	ΔCFI	ΔSB-x^2	Δ*df*	p_2
RAN_a freely estimated	45.18	20	.001	0.08	0.941	0.001	2.55	3	.47
RAN_a constrained	47.64	23	.002	0.07	0.942				
RAN_n freely estimated	46.69	21	.001	0.08	0.941	0.002	2.13	3	.55
RAN_n constrained	48.89	24	.002	0.07	0.943				
PA freely estimated	71.07	48	.02	0.05	0.955	0.000	3.27	3	.35
PA constrained	74.37	51	.02	0.05	0.955				
PL freely estimated	71.71	47	.01	0.05	0.975	0.000	2.58	3	.46
PL constrained	74.32	50	.01	0.05	0.975				

Note. ΔSB-x^2 = Satorra–Bentler scaled chi-squared difference test; CFI = comparative fit index; p_1 = probability value of model fit; p_2 = probability value obtained in the SB chi-squared difference test; PA = phonological awareness; PL = phonological loop; RAN_a = rapid automatized naming for alphanumeric stimuli; RAN_n = rapid automatized naming for nonalphanumeric stimuli; RMSEA = root mean square error of approximation.

APPENDIX C

Measurement Model (1) and Latent Change Score Model (2) of Rapid Automatized Naming (RAN) for Nonalphanumeric (RANn) Stimuli



Note. IS = indicator-specific factor; ranc = rapid automatized naming of colors; rano = rapid automatized naming of objects; t1 = first testing wave; t2 = second testing wave; t3 = third testing wave; t2 − t1 = latent change factor representing the difference between the first and second testing waves; t3 − t2 = latent change factor representing the difference between the second and third testing waves. As displayed in model 2, nonalphanumeric naming speed at grade 4 ($t2RAN_n$) is predicted by the baseline performance at grade 3 ($t1RAN_n$) and the discrete performance change between grades 3 and 4 ($t2RAN_n - t1RAN_n$). Accordingly, nonalphanumeric RAN at grade 5 ($t3RAN_n$) is predicted by the baseline performance at grade 3 ($t1RAN_n$) and by the discrete performance changes between grades 3 and 4 (represented by the latent change factor $t2RAN_n - t1RAN_n$) and the performance changes between grades 4 and 5 (represented by the latent change factor $t3RAN_n - t2RAN_n$). For better clarity, residual variances are not shown.

APPENDIX D

Correlations Between Phonological Processing Measures

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1. t1WS1	—	.47	.59	.59	.55	.56	.61	.49	.62	.21	.08	.10	.16	.13	.12	.33	.13	.13	−.19	−.02	−.02	−.12	−.13	.00	−.01	−.17	−.08	−.05	.08	.18
2. t1WS3	.63	—	.52	.52	.68	.45	.62	.59	.56	.20	.22	.28	.27	.15	.23	.22	.20	.35	−.09	−.04	.10	.07	−.04	.03	.04	−.02	.02	.08	.02	.11
3. t1DS	.66	.67	—	.58	.53	.63	.52	.38	.58	.20	.17	.17	.30	.30	.19	.34	.20	.36	−.05	−.05	−.04	−.08	−.02	.03	−.11	−.17	.01	.16	−.15	.03
4. t2WS1	.65	.60	.55	—	.59	.60	.55	.49	.61	.15	.19	.22	.25	.12	.16	.26	.15	.17	−.06	.04	−.03	−.11	.01	.04	−.04	−.16	.17	.09	−.07	.05
5. t2WS3	.62	.72	.58	.66	—	.48	.57	.64	.58	.20	.12	.18	.32	.15	.31	.37	.26	.26	−.05	.08	.03	−.05	.04	.04	−.05	−.09	.17	.12	.01	.11
6. t2DS	.76	.60	.69	.66	.68	—	.66	.46	.59	.14	.07	.06	.18	.02	.21	.22	.23	.27	−.02	−.09	.06	−.07	.09	.06	−.05	−.07	.04	.10	.01	.12
7. t3WS1	.68	.62	.65	.77	.64	.71	—	.57	.59	.17	.16	.06	.22	.14	.18	.31	.13	.22	−.08	.15	.05	−.01	−.01	.06	−.02	−.15	−.09	−.00	−.06	.05
8. t3WS3	.59	.59	.58	.70	.66	.67	.71	—	.48	.08	.31	−.02	.03	.07	.16	.13	.03	.09	.02	.29	.19	.02	.14	.26	.17	.05	.18	.20	.00	.00
9. t3DS	.67	.69	.70	.68	.67	.78	.65	.71	—	.33	.31	.09	.30	.24	.30	.35	.14	.15	−.21	−.10	−.02	−.11	−.19	−.12	−.12	−.16	.12	.13	.05	.03
10. t1VS	.21	.22	.14	.25	.19	.32	.20	.27	.40	—	.19	.40	.37	.38	.34	.29	.32	.28	−.15	−.11	−.07	.04	−.19	−.02	.02	.00	−.26	−.11	−.06	−.06
11. t1VL	.25	.26	.27	.13	.17	.38	.28	.32	.34	.14	—	.22	.13	.46	.09	.16	.37	.18	−.11	−.02	−.16	.01	−.01	.18	−.00	−.01	−.03	.01	−.13	−.19
12. t1PR	.06	.15	.09	.06	.16	.18	−.02	.11	.06	.30	.39	—	.37	.35	.52	.14	.31	.33	−.04	−.08	−.13	−.08	−.17	.05	−.11	−.14	−.02	−.08	−.09	.01
13. t3VS	.12	.19	.17	.19	.22	.16	.08	.26	.15	.36	.11	.28	—	.28	.30	.39	.45	.27	−.10	−.03	.02	.01	−.04	−.07	−.08	−.01	.04	.13	−.05	.04
14. t3VL	.28	.26	.37	.38	.32	.47	.46	.45	.36	.30	.38	.19	.37	—	.41	.25	.33	.33	−.08	−.01	−.14	−.02	−.18	.10	−.04	−.17	−.10	−.04	.00	−.05
15. t3PR	.34	.37	.36	.43	.39	.42	.54	.42	.33	.28	.38	.47	.41	.53	—	.41	.41	.60	.05	.03	.11	.10	−.08	.02	−.10	.01	.20	.08	.11	.09
16. t5VS	.36	.23	.37	.32	.38	.27	.27	.38	.29	.42	.26	.20	.25	.25	.14	—	.37	.32	.03	.126	−.06	.02	.01	.07	.04	−.04	.01	.00	.04	.08
17. t5VL	.28	.31	.29	.42	.38	.48	.32	.49	.45	.41	.47	.38	.37	.60	.37	.47	—	.44	.06	.09	.09	.18	.05	.17	.03	.17	.115	.13	.00	.14
18. t5PR	.25	.29	.11	.22	.13	.19	.30	.15	.10	.35	.18	.33	.07	.13	.35	.16	.14	—	.08	.13	.10	.01	.01	.14	−.05	−.01	.156	.16	−.03	.08
19. t1RAN$_{o}$	.05	.07	−.04	−.03	−.07	−.09	−.06	−.01	−.03	.11	−.16	−.06	.10	−.22	−.17	.08	.03	.10	—	.48	.25	.37	.70	.47	.19	.41	.35	.22	.02	.30
20. t1RANc	−.13	−.02	−.13	−.34	−.26	−.26	−.36	−.30	−.24	.16	−.10	.06	−.01	−.22	−.31	−.14	−.05	.05	.60	—	.29	.45	.47	.68	.27	.40	.44	.54	.10	.27

(continued)

Correlations Between Phonological Processing Measures (*continued*)

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
21. t1RAN$_l$	−.07	−.00	−.04	−.20	−.15	−.19	−.27	−.26	−.17	.17	.03	.08	.08	−.16	−.03	.04	−.02	.11	.35	.43	—	.44	.16	.14	.45	.40	.15	.26	.39	.45
22. t1RAN$_d$	−.08	.03	−.07	−.02	−.09	−.18	−.14	.00	−.05	.14	−.05	.09	.21	−.03	.04	.13	.07	.23	.29	.38	.52	—	.47	.50	.56	.76	.26	.27	.37	.56
23. t2RAN$_o$	−.06	−.16	−.17	−.17	−.18	−.18	−.16	−.17	−.20	.12	−.10	.17	−.04	−.17	−.11	.09	.01	.09	.72	.50	.45	.28	—	.61	.41	.56	.43	.26	.11	.30
24. t2RAN$_c$	−.23	−.13	−.31	−.20	−.15	−.19	−.23	−.13	−.11	.06	−.12	.16	.02	−.18	−.13	−.14	.04	.03	.53	.72	.41	.49	.56	—	.41	.53	.37	.46	.19	.36
25. t2RAN$_l$	.01	.07	.06	−.16	−.08	−.11	−.24	−.22	−.10	.14	.05	.23	.18	−.05	.11	.02	.02	.10	.33	.45	.72	.39	.39	.47	—	.62	−.00	−.06	.40	.33
26. t2RAN$_d$	−.07	.12	−.02	−.06	.02	−.14	−.22	−.10	−.11	.18	.05	.26	.17	−.10	.18	.08	.02	.11	.36	.42	.55	.68	.37	.50	.69	—	.28	.18	.37	.50
27. t3RAN$_o$	−.25	−.10	−.24	−.29	−.31	−.14	−.33	−.31	−.15	.08	.02	.02	−.03	−.19	−.14	−.16	−.05	−.23	.25	.37	.35	.19	.38	.32	.24	.29	—	.72	.26	.36
28. t3RAN$_c$	−.46	−.31	−.47	−.33	−.30	−.32	−.47	−.37	−.15	.07	−.11	.00	.00	−.22	−.24	−.18	−.05	−.22	.31	.57	.36	.33	.35	.58	.32	.44	.68	—	.26	.28
29. t3RAN$_l$	−.13	.03	.01	−.21	−.08	−.07	−.26	−.15	.05	.25	.31	.32	.13	−.03	.10	.11	.12	−.08	.25	.46	.47	.48	.30	.44	.54	.56	.52	.53	—	.68
30. t3RAN$_d$	−.13	.02	−.06	−.19	−.08	−.08	−.24	−.06	.03	.08	.07	.15	.01	−.18	.03	.11	.03	−.11	.36	.40	.28	.52	.25	.40	.32	.57	.54	.45	.76	—

Note. DS = digit span; PR = phoneme reversal; RAN$_c$ = rapid automatized naming of colors; RAN$_d$ = rapid automatized naming of digits; RAN$_l$ = rapid automatized naming of letters; RAN$_o$ = rapid automatized naming of objects; t1 = first testing wave; t2 = second testing wave; t3 = third testing wave; VL = vowel length; VS = vowel substitution; WS1 = word span for one-syllable words; WS3 = word span for three-syllable words. Intercorrelations for the control group are presented below the diagonal, and those for students with learning disabilities in reading and/or spelling are presented above the diagonal. All correlation coefficients of $r > .20$ are statistically significant at $p < .05$.