Annals of Dyslexia (2022) 72:276–300
https://doi.org/10.1007/s11881-022-00254-4

The reading profiles of late elementary English learners with and without risk for dyslexia

Jeremy Miciak[1] · **Yusra Ahmed**[1] · **Phil Capin**[2] · **David J. Francis**[1]

Received: 18 May 2021 / Accepted: 11 March 2022 / Published online: 24 May 2022
© The Author(s), under exclusive licence to The International Dyslexia Association 2022

Abstract
Few studies have systematically investigated the reading skill profiles of English learners (ELs) in late elementary school, a critical developmental period for language and literacy and the most common grades for initial identification with specific learning disabilities (O'Connor et al., *Learning Disabilities Research & Practice, 28*(3), 98–112, 2013). We investigated the reading skill profiles of 331 ELs in 3rd and 4th grades, including ELs with and without risk for dyslexia due to significant deficits in word and pseudo-word reading accuracy and fluency. We utilized latent profile analysis and factor mixture modeling to investigate (1) the nature and distribution of reading skill profiles; (2) whether these profiles were associated with differences in reading comprehension growth across one academic year; and (3) the stability of reading profiles across an academic year. We selected a two-class solution (reading disabled and typically developing) based on model fit indices, theoretical considerations, pattern of results across profile-solutions and time-points, and parameterizations, making the approach stronger and more generalizable. These classes demonstrated clear, consistent differences in performance across reading component skills, with the RD class scoring consistently below the TD class across code-based and meaning-based domains of reading. Across the year, the TD class demonstrated significantly higher patterns of growth in reading comprehension (χ^2 (1) = 206.21, $p < 0.001$). Class membership was largely stable (97% of participants maintain class membership). These results suggest that ELs with risk for dyslexia demonstrate multiple component skill deficits that may require long-term, comprehensive, intensive interventions to remediate.

Keywords English learners · Dyslexia · Latent profile analysis · Reading profiles

✉ Jeremy Miciak
Jeremy.Miciak@times.uh.edu

[1] Department of Psychology, Univeristy of Houston, 3811 Calhoun, Ste. 373, Houston, TX 77004, USA

[2] The University of Texas at Austin, Austin, TX, USA

The reading profiles of late elementary English learners with and without dyslexia

English learners (ELs) represent the fastest growing subgroup of the U.S. student population (National Center for Education Statistics, 2016). Unfortunately, studies indicate that current practice is not adequately addressing the educational needs of many ELs, who are at greater risk for academic difficulties than their monolingual peers (August et al., 2009; Lesaux et al., 2010). National assessments indicate that nearly three quarters of all ELs score below the basic level on reading assessments in late elementary and middle school (NCES, 2019). Further, evidence suggests that ELs and students from culturally and linguistically diverse backgrounds are at risk for disproportionate representation in special education (Donovan & Cross, 2002; Morgan et al., 2015) and there is considerable confusion related to parsing issues of second language acquisition and disability status, including dyslexia (Hall et al., 2019; Hoover et al., 2019). One reason for this confusion may be that comparatively fewer studies have investigated reading development among ELs with and without risk for dyslexia compared to their monolingual peers. These studies are valuable for comparing and contrasting typical and atypical processes for the development of reading component skill development in the context of second language acquisition. Thus, there remain questions about the reading profiles and instructional needs of ELs, including ELs with risk due to difficulties with accurate and fluent reading of words and text—primary symptomologies of dyslexia (Burr et al. 2015; Klingner et al., 2006). The present study investigates whether component reading skill profiles (based on word and pseudo-word reading accuracy, word and pseudo-word reading fluency, and language proficiency) emerge within a sample of late elementary ELs with and without risk for dyslexia. The overarching goal is to determine the nature and distribution of these component skill profiles. We further investigate whether students with specific reading profiles demonstrate differential patterns of reading comprehension growth across a single academic year and the stability of emergent reading profiles across one instructional year.

English learners in late elementary

First identified by Chall and Jacobs (1983), the so-called "fourth grade slump" describes a phenomenon in which students who had been considered successful readers in early elementary grades begin to experience difficulty in mid to late elementary school (Grades 3–5). The primary hypothesis for this delayed onset of reading difficulties relates to the changing reading task in mid to late elementary, as students shift from "learning to read" to "reading to learn." This shift taxes the entire reading system, as comprehending increasingly complex texts requires well-developed foundational word reading and reading fluency skills (Cho et al., 2019; Cirino et al., 2013); at the same time more complex factors such as language and background knowledge take greater predictive importance for reading comprehension (Catts et al. 2006; Cho et al., 2015; Miciak et al., 2014). For many students, difficulties in one or more of these component skills manifest as persistent difficulties reading and understanding grade-level text.

For ELs, this shift in the reading task and the essential contribution of different component skills may be particularly precarious, as students are increasingly required to apply English reading and language skills that may not be fully developed due to limited opportunity for second language acquisition. Studies indicate that many ELs continue to face academic difficulties and demonstrate academic language deficits in English, even when considered English proficient by schools (Francis et al., 2006; Kieffer & Lesaux, 2008; Lesaux & Kieffer 2010). Additionally, evidence suggests that many ELs experience academic failure during this critical period of transition: third and fourth grades are the most common grades for initial identification with LD for ELs (O'Connor et al. 2013). In a longitudinal study using data from the Early Childhood Longitudinal Study–Kindergarten Cohort, Samson and Lesaux (2009) investigated special education identification rates and potential predictors of special education placement with a nationally representative sample of kindergarten, first-, and third-graders. ELs were underrepresented in special education in kindergarten and first grade, but were overrepresented in third grade across all disability categories. These inconsistent rates of identification across different grades may reflect confusion among teachers and special educators in differentiating LDs (including dyslexia) from normal language acquisition and highlight the importance of studies that investigate late elementary reading outcomes in samples of ELs demonstrating varying risk.

The simple view of reading

The process of reading development involves a number of components or related skills. These components can be classified as either code-based or meaning-related, as highlighted by theories such as the Simple View of Reading (Gough & Tunmer, 1986) that have been extensively studied with monolingual English speakers (Hoover & Gough, 1990) but that are also applicable to the development of reading for ELs (Duke & Carlisle, 2011; Samson & Lesaux, 2015). Current evidence indicates that both code-based competencies (e.g., decoding, fluency) and meaning-related skills (e.g., vocabulary, linguistic comprehension) contribute to reading comprehension development for ELs (e.g., Geva & Yaghoub Zadeh, 2006; Gottardo & Mueller, 2009; Mancilla-Martinez & Lesaux, 2010). Much of this research has been conducted with early elementary-age students for whom code-based skills may play an outsized role in prediction. Among older ELs and monolingual English speakers, language skills take on greater importance in predicting overall reading proficiency (e.g., Bialystok, 2007; Jeon & Yamashita, 2014; Lesaux & Kieffer, 2010). Even when ELs develop adequate code-based skills, many may continue to demonstrate marked deficits in making meaning from text (Betts et al., 2009; Jean & Geva, 2009; Lesaux et al., 2006, 2010).

Despite the recognized importance of language in predicting reading comprehension among ELs, particularly ELs enrolled in grades beyond early elementary, there continue to be questions about ELs who demonstrate risk for dyslexia due to difficulties in accurate and fluent reading of words. Similar to monolingual English speakers with significant reading difficulties, a subset of ELs will demonstrate persistent difficulties related to code-based processes (Cho et al., 2019; Lesaux et al., 2010). However, and perhaps due to the well-placed emphasis on language development among ELs, fewer studies have directly investigated the characteristics and development of this group of ELs, who may demonstrate unique instructional needs.

In the present study, we consider reading development for a sample of Spanish speaking ELs in 3rd and 4th grade with and without risk for dyslexia. We consider both code-based (word and pseudo-reading accuracy and word and pseudo-word reading fluency) and meaning-based (linguistic comprehension) skills in English. These data permit us to investigate whether distinct reading profiles emerge, including whether a subsample of students with specific risk for dyslexia can be identified and differentiated from a subsample of students with relatively stronger code-based skills but deficits in reading comprehension due to deficits in linguistic comprehension. We then analyze patterns of reading comprehension development and the stability of these profiles over a single instructional year.

Reading profiles of students with reading difficulties

This study builds on previous studies that have attempted to identify reading profiles of students with reading difficulties using rigorous analytic methods like latent profile analysis. These studies differ in important ways, including differences in age, EL status, and sociodemographic characteristics of participants, as well as the specific way in which reading difficulties are defined. In a recent study identifying the reading and cognitive profiles of late elementary students with significant reading comprehension difficulties, Capin et al. (2021) argued that the central question that motivates studies of empirical reading profiles is to understand better the extent to which different reading profiles are marked by the *specificity* or *severity* of component skill deficits. This question has important implications for intervention design and delivery and informs longstanding debates about the nature of dyslexia. That dyslexia is a disorder primarily marked by deficits in word reading, reading fluency, and spelling words is uncontroversial (Miciak & Fletcher, 2020). However, there is considerable controversy related to the specificity of these deficits. For example, Shaywitz (2003) argued that specificity is a necessary marker of dyslexia—a specific weakness in a sea of strengths. Others contend that such notions are unhelpful and unsupported in empirical research (Elliott & Grigorenko, 2014; Miciak & Fletcher, 2020; Vellutino et al., 2000), as dyslexia is primarily marked by the severity of code-based reading deficits but often co-occurs with deficits in other correlated domains, such as language, comprehension, and other factors of cognition (Capin et al., 2021; Cho et al., 2019; Miciak et al., 2014).

Questions about the severity and specificity of component skill deficits for ELs with reading difficulties, including those with risk for dyslexia, are more complicated because the role of second language acquisition introduces additional questions. Latent profile analysis represents one rigorous method to investigate differences in component skill profiles for populations of interest, including struggling readers, ELs, and ELs with risk for dyslexia. Several recent studies have used latent profile analysis to investigate whether distinct reading skill profiles can be identified among students with reading difficulties enrolled in late elementary to ninth grade based on performance on reading component skills tasks (Brasseur-Hock et al., 2011; Capin et al., 2021; Clemens et al., 2017). Each of these studies identified latent profiles of students with reading difficulties based on performance on code-based tasks (i.e., decoding, reading fluency) and on language-based tasks (vocabulary, listening comprehension). All three studies successfully identified distinct profiles marked by either the severity or specificity of their component skill profiles, although the number of profiles and proportion of students in each profile group type varied. For example, Brasseur-Hock et al. (2011) and Capin et al. (2021) found that a large majority (85% and 91% respectively) of students belonged to profile groups marked only by the severity of component skill deficits (e.g., severe deficits in decoding, fluency, and vocabulary; moderate deficits in decoding, fluency, and vocabulary) and very few students exhibited a specific component skill deficit in isolation. Clemens et al. also found that most students (65%) belonged to profile groups identified by the severity of their deficits and few featured specific deficits, although a larger proportion of their sample belonged to profile groups marked by specificity than that observed in Brasseur-Hock et al. (2011) and Capin et al. (2021)

Specific to the reading profiles of ELs with and without reading comprehension difficulties, O'Connor et al. (2019) investigated the reading comprehension profiles of Grade 5 monolinguals and ELs to compare relationships among phonological skills, semantic knowledge, listening comprehension, and reading comprehension. In both the monolingual and EL group, latent profiles emerged for two groups: poor comprehenders and good comprehenders that were consistently differentiated across all component skills. This finding suggests that, within their sample of late elementary ELs and monolinguals, differences between poor comprehenders and good comprehenders were marked primarily by severity rather than specific deficits in reading and reading-related skills. In contrast, Li et al. (2021) examined the reading skill profile of eighth graders with and without reading difficulties in China who were learning English as a second language and found latent subgroups reflected specific difficulties. Using English word reading and reading comprehension performance to classify students in latent profiles, Li and colleagues (2021) found three subgroups: (1) typically developing readers with average word reading and reading comprehension, (2) poor comprehenders with difficulties in word reading, and (3) poor comprehenders without difficulties in word reading. These conflicting findings underscore the value of further research in this area.

Lesaux and Kieffer (2010) investigated whether distinct profiles emerged in a sample of sixth grade students with comprehension difficulties that included a large percentage of ELs (77%). Similar to the studies described above that utilized a simple view of reading framework to investigate the profiles of struggling readers, latent profile groups were formed based on students' performance on decoding, reading fluency, and vocabulary and multiple profiles were identified based on both the severity and specificity of deficits. In contrast to the studies above conducted with monolingual struggling readers, the majority (80%) of struggling readers in this study belonged to profile groups marked by specific deficits (e.g., average decoding and fluency, low vocabulary). The marked differences between the samples of the studies reviewed above offers a potential explanation difference in findings. For example, the studies included participants of different ages, with different EL status, and defined risk differently for inclusion in the study. Thus, additional investigation of the reading profiles of ELs in late elementary (3rd and 4th grades) is needed, particularly ELs with and without risk for dyslexia due to deficits in word and pseudo-word reading accuracy and fluency.

One potential limitation of the reviewed literature is that few previous latent profile analyses have utilized data collected at multiple time points to investigate questions of reading comprehension development and the stability of latent profile groups over time. In one notable study that investigated the prevalence and stability of bilingual profiles of younger students at multiple times across Grades 1–3, Swanson and colleagues (2019) utilized latent profile analyses based on performance on measures of English and Spanish language, achievement, and cognition. Four distinct groups emerged: (1) balanced bilingual, average achievers; (2) unbalanced bilingual, average achievers; (3) at risk for learning disabilities; and (4) English dominant. These latent profile groups generally demonstrated high stability (i.e., students remained in the same group over time); however one finding related to the group at risk for LDs is particularly noteworthy for the present study, which investigates latent profiles at two distinct time points in a single academic year. Among students in average achieving groups in Grade 1, almost one quarter (24%) of the balanced bilingual-achievers transitioned to the at-risk group in Grade 3. These data suggest that there is a critical period in third grade that predicts risk for learning disabilities. Additional work exploring latent profile groups at this critical age may help in disentangling these relations.

The present study

Recent research suggests that ELs and monolingual students vary in their sources of reading difficulties (Cho et al., 2019). Previous analyses examining the reading profiles of ELs have yielded conflicting information. For example, Li et al. (2021) found that two subgroups of ELs with reading difficulties emerged: one with difficulties in word reading and one without word reading difficulties. This finding suggests that subgroups may reflect specific difficulties in the component skills of reading. Conversely, O'Connor et al. (2019) found that ELs with reading difficulties scored consistently lower on component skills of reading than ELs with typical reading performance. In the present study, our goal is to determine the nature and distribution of the reading profiles that emerge, particularly the extent to which subgroups emerge with specific component skill deficits, as opposed to groups based strictly on the level of performance on their component skill profiles. Our study contributes to the understanding of reading profiles in several ways. First, we focus on all third and fourth grade ELs enrolled in participating schools and therefore include ELs with and without risk for dyslexia, which is marked by difficulties in word and pseudo-word reading accuracy and fluency. The focus on these grade levels is salient because these grades represent a critical period of reading development (e.g., Chall & Jacobs, 1983) and other studies examined the reading profiles of older ELs (Lesaux & Kieffer, 2010; Li et al., 2021; O'Connor et al., 2019). Second, participants are tested twice in a single academic year (Fall Grade 3 or 4; Spring Grade 3 or 4). This enables us to understand the extent to which group membership predicts gains in reading comprehension over the course of a school year and to examine the stability of group membership during this time period. Third, our study includes a large sample, uses psychometrically sound measures that represent important instructional targets, and uses robust analytic approaches including latent profile analysis and factor mixture modeling. Three research questions guide this study:

1. What latent profiles emerge when considering the English word and pseudo-word reading accuracy, word and pseudo-word reading fluency, and linguistic comprehension of the sample?
2. Does latent profile group membership predict differential gains in reading comprehension across a single academic year?
3. How stable is latent profile group membership over the course of a single academic year?

Methods

Participants

School participants

The setting for this study is two elementary schools in a large urban district in the southwest USA. These schools were recommended by district officials based on the demographic match for a study investigating outcomes with ELs with risk for dyslexia and a willingness to participate in external research. Both schools were majority Hispanic (82.7% and 74.9% of total enrollment), many of whom were considered ELs by state of Texas criteria (72.7% and 50.7%). In Texas, students whose parents report that a language other than English is spoken at home and who do not pass an initial English proficiency test upon enrollment are considered limited English proficient until they achieve proficiency on the Texas English Language Proficiency Assessment System (TELPAS) and meet academic benchmarks. Both schools enrolled a high percentage of students considered Economically Disadvantaged by the state of Texas (97.4% and 88.3%) and included many students who were below grade level on the state accountability exam (71% and 68% of total students).

The two schools differed in their instructional model for English learners, although student mobility and limited data prohibits fully characterizing the instructional history of each participant. School A featured a transitional bilingual program, in which students received 50% of their instruction in Spanish and 50% in English (although these percentages likely vary in practice due to teacher preferences). School B featured an all-English instructional model.

Student participants

Participants for the present study include all current and former ELs enrolled in Grades 3 and 4 at the participating schools who had a primary language of Spanish. Current ELs include students who are classified as limited English proficient based on Texas criteria. We also included students who were previously limited English proficient and who had been reclassified as English proficient within the last 2 years and were therefore participating in a 2-year monitoring phase following reclassification. The sample consisted of $n=331$ (56% female) Hispanic/Latinx students with Limited English Proficiency, all of whom qualified for free meals under the National School Lunch and Child Nutrition Program (a proxy for economic disadvantage). Twelve students did not have data at time 2 because they moved ($n=10$), did not have TELPAS data on file ($n=1$), or because the parent declined participation in the study at time 2 ($n=1$). The attritted sample did not differ from the non-attritted sample on measures of reading comprehension (χ^2 $(40)=17.66$, $p=0.999$), decoding fluency (χ^2 $(19)=15.2697$, $p=0.7053$), word reading fluency (χ^2 $(22)=23.079$, $p=0.401$), pseudo-word decoding (χ^2 $(57)=34.808$, $p=0.9911$), and English memory for sentences (χ^2 $(48)=35.83$, $p=0.902$), but the attritted sample differed on the measure of letter word recognition (χ^2 $(64)=91.49$, $p=0.0137$).

Measures

The assessment battery was chosen to comprehensively assess reading and reading component skills in both English and Spanish. All examiners were trained across two days, with a single day dedicated to explicit instruction on assessment principles and procedures and another day dedicated to guided practice. Prior to field data collection, all examiners passed a "check-out" with the assessment coordinator or an experienced member of the assessment team. During data collection, examiners were supervised by the assessment coordinator, who audited all data for inconsistencies during and after data collection.

Word reading

We administered two tests of untimed decoding: word and pseudo-word reading accuracy. Real word reading was assessed using The Kaufmann Test of Educational Achievement Third Edition Letter & Word Recognition (KTEA-3: Kaufman & Kaufman, 2014). The Letter and Word Recognition subtest is an individually administered assessment for use with students between ages 4 and 25. This subtest assesses the examinee's ability to read real words and letters accurately. For the first 18 items, students are asked to point to letters that correspond with what the examiner names, point to letters corresponding to sounds, and to give sounds of different letters. Beginning with item 19, examinees are asked to read words. The psychometric characteristics of this subtest are strong. Split-half reliabilities for ages 7–11 are 0.96–0.97, and standard errors of measurement in Grades 3 and 4 range from 2.56 to 2.78. Internal consistency estimates were acceptable for the current sample ($\alpha = 0.76$-0.79). The KTEA-3 Nonsense Word Decoding (Kaufman & Kaufman, 2014) subtest is an untimed assessment of decoding phonetically regular, nonsense words. The 50-item measure assesses the ability to pronounce nonsense words, asking students to apply phonics and structural analysis skills to decode nonsense words or increasing difficulty. Responses are recorded as correct or incorrect using a "phoneme key" for each word. The psychometric characteristics of this subtest are strong. Split-half reliability coefficients for ages 7–11 are 0.96–0.97, and standard errors of measurement in Grades 3 and 4 range from 2.79 to 2.80, but internal consistency was lower for the current sample ($\alpha = 0.68$).

Word reading fluency

We administered two tests of reading fluency: word and pseudo-word reading fluency. The KTEA-3 Word Recognition Fluency (Kaufman & Kaufman, 2014) is a timed assessment of single word reading fluency. The subtest asks students to read isolated words aloud as quickly as possible during two 15-s trials. The psychometric characteristics of this subtest are strong for fluency-type measures. Split-half reliability coefficients for ages 7–11 are 0.80–0.89, and the standard error of measurement in Grades 3 and 4 is 6.36. Internal consistency ranged from 0.79 to 0.80 in the current sample. The KTEA-3 Decoding Fluency (Kaufman & Kaufman, 2014) is a timed assessment of decoding fluency. This subtest requires examinees to read as many nonsense words aloud, as quickly as they can during two 15-s trials. This 51-item subtest is used for students in grades 3 through 12+. The psychometric characteristics of this subtest are strong for fluency-type measures. The split-half reliability coefficient for ages 8–11 is 0.82, and the standard error of measurement for Grades 3 and 4 is 6.54. Internal consistency was $\alpha = 0.81$ in the current sample.

Reading comprehension

The Gates-MacGinitie Reading Test Fourth Edition (GMRT-4; MacGinitie, 2000) reading comprehension subtest is a timed (35 min), group-administered assessment consisting of expository and narrative passages ranging in length from 3 to 15 sentences. Similar to many statewide reading achievement tests, students are asked to read each passage silently and answer multiple-choice questions. It consists of 11 passages with 48 multiple-choice questions that target the following areas: inference making, summarization, main idea, literal questions about text, and vocabulary. The GMRT-4 is commonly used in educational research and demonstrates high internal consistency, with reliability above 0.90 (the K-R 20 coefficient in Grade 4 = 0.93). Internal consistency ranged from 0.89 at time 1 to 0.92 at time 2 in the current study.

Linguistic comprehension

We assessed linguistic comprehension with two tests of students' language proficiency. We administered the WJ-III Sentence Recall (WJ-III: UC; Woodcock et al., 2007), an individually administered measure of students' ability to listen and recall sentences of increasing length and complexity. The subtest evaluates expressive syntax and requires the student to remember and repeat single words, phrases, and sentences presented orally, with increasing grammatical complexity. Memory for sentences is a robust indicator of overall language proficiency that requires knowledge of both vocabulary and syntax (Klem et al., 2015). The psychometric characteristics of the WJ-III subtests are adequate to strong. The median reliability coefficient at this age is 0.89, and the internal consistency for the study sample was 0.78–0.79. We also administered the WJ-III: Picture Vocabulary (WJ-III: PV; Woodcock et al., 2007), an individually administered measure of one-word vocabulary. Reliability coefficients range from 0.74 to 0.85 with this age and was 0.84 in the current study.

Analytic plan

Latent profile analysis (LPA) and factor mixture modeling (FMM) were used to empirically identify academic and language skill profiles of EL students, and latent transition analysis (LTA) was used to characterize the stability of profiles over two time-points estimated under the LPA and FMM approaches. Each model parameterization represents a special case of finite mixture modeling in which the categorical latent variable with *k* profiles (c_k in Fig. 1) represents the unobserved heterogeneity on the six continuous observed variables by two or more distinct subgroups of ELs. Thus, these models are person-centered approaches that do not require a grouping variable to be explicitly specified because it is empirically estimated (Henson et al., 2007). Models with two to six latent profiles were estimated for each parameterization in M-plus 8.6 (Muthen & Muthen, 2017) using full information maximum likelihood to handle missing data. Analyses were conducted with up to 5000 random sets of start values with 500 iterations for these random starts to avoid local solutions, and the 300 best solutions were retained for final stage optimization.

There is not a definitive statistical test for the comparison of models and profile solutions in LPA and MFF; selection of the optimal number of profiles requires inspection of fit indices and theory. As the models are not nested, relative goodness of fit for each model was evaluated using the following criteria, which take sample size, model fit, and number of parameters into account: Akaike's information criteria (AIC), the Bayesian information criteria (BIC), and sample-adjusted BIC (SABIC), with lower values indicating a better fit of the model. In addition, the optimum number of profiles was determined using the following statistical tests: Vuong-Lo-Mendell-Rubin likelihood ratio test (VLMR) and Lo-Mendell-Rubin adjusted likelihood ratio test (LMR) which utilize chi-square difference testing to compare the estimated model with *k* profiles versus one profile less (*k*-1), with lower *p* values indicating that the model with *k*-1 profiles

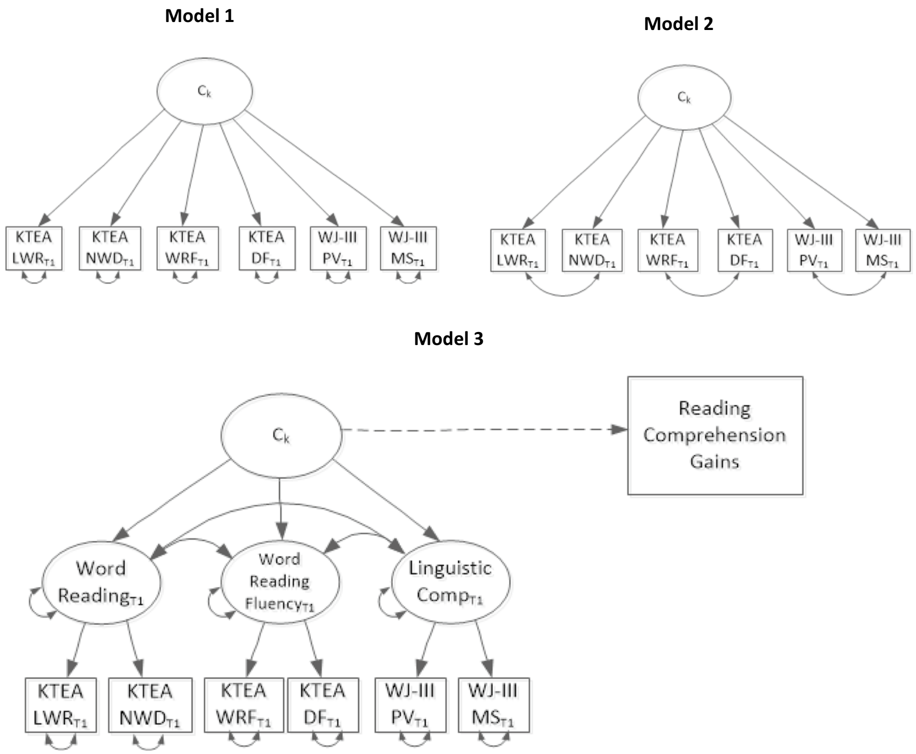

Fig. 1 Measurement models for the latent profile analysis (models 1 and 2) and factor mixture model (model 3) with auxiliary model shown only for model 3. Solid lines correspond to the measurement model and the dashed line corresponds to the auxiliary model. LWR = KTEA3 Letter and Word Recognition; NWR = KTEA3 Nonsense Word Reading; WRF = KTEA3 Word Reading Fluency; DF = KTEA3 Decoding Fluency; VOC = WJ-III Picture Vocabulary; MS = WJ-III Memory for Sentences

is rejected in favor of the model with *k* profiles. Model entropy was evaluated for quality of profile separation or effectiveness of classification, with values above 0.80 indicating adequate profile separation. In addition to model fit indices, theoretical justification was considered in determining the optimal number of profiles (Asparouhov and Muthén 2014; Nylund et al., 2007).

Measurement models

Model 1 consisted of an LPA model with six index variables as shown in Fig. 1. For this model, means and variances of the index variables were freely estimated in all profiles. A fundamental assumption in LPA is that indicators are conditionally independent of each other, given the latent profile membership. However, a failure to control for shared variance among indicators may preclude the extraction of clearly defined profiles when there is a strong theoretical reason to expect that relations among indicators exist over and above the expected latent profiles (Morin et al., 2016) or there is reason to expect an underlying factor structure to the observed indicators (Morin & Marsh, 2015). Because of the shared variance among indicators of the same theoretical constructs, a second measurement model (Model 2) was estimated in which residuals were correlated among measures of (a) word reading, (b) word reading fluency, and (c) linguistic comprehension (see Fig. 1).

A final model (Model 3) used factor mixture modeling to estimate profile membership with latent variables. As Fig. 1 shows, in this model three continuous latent variables captured shared variance that is free of measurement error associated with observed indicators (Lubke & Muthen, 2005). The intercepts and factor loadings of KTEA Letter Word Reading (word reading), KTEA Word Reading Fluency (word reading fluency) and WJ-III Picture Vocabulary (linguistic comprehension) were arbitrarily fixed at 0 and 1, respectively, to identify the three continuous latent variables. We fit models in which the latent variables were correlated and models in which the correlations were fixed at zero because the focus of the present study was on factor means rather than factor covariances. To estimate means of latent variables, one of the classes serves as a reference with its factor means set to zero. The factor means of the additional profiles can be interpreted as the difference in factor means across profiles (Clark et al., 2013; Leite & Cooper, 2010; Masyn et al., 2010). Factorial invariance across profiles was established by comparing models with different constraints. The following models were evaluated: weak factorial invariance (equal factor loadings), strong factorial invariance (equal factor loadings and thresholds), and strict factorial invariance (thresholds, loadings, and residual variances), with the strong invariance model retained for final analyses because this model is considered sufficient for comparing means across latent profiles (Leite & Cooper, 2010; Lubke and Muthén 2005).

Auxiliary models

Figure 1 displays both the measurement model (solid lines) and auxiliary models (dashed line) for the FMM (Model 3), although the auxiliary models were estimated for the LPA (Models 1–2) as well. After choosing the best class-solution for Models 1–3, a three-step method for modeling covariates was employed using the Bolck, Croon, and Hagenaars (BCH) approach (Bakk & Vermunt, 2016; Bolck et al., 2004; Croon, 2002; Vermunt, 2010) in M-plus to investigate whether growth in reading comprehension differed by student's probabilities of assignment to latent profiles (RQ 2), and this model was also estimated with reading comprehension as the distal outcome at each time point. The auxiliary models were estimated independently from the measurement models so that the auxiliary variables did not influence profile membership. The first step simply estimates the latent profile models using the six index variables. The second step of the BCH approach employs a weighted multiple group analysis in which the measurement errors are fixed using each individual's profile probabilities (Ferguson et al., 2020). The final step of the BCH approach connects the models from steps 1 and 2 with step 3 in a pairwise comparison of the mean of reading comprehension in one profile versus another. The same procedures were applied for the estimation of the auxiliary models for LPA and FMM.

Latent transition analysis (LTA)

Latent transition analysis (LTA) is the longitudinal extension of LPA. In the LTA models, two latent profile variables were measured at time 1 and time 2, respectively, and the autoregressive relation between the two latent profiles was estimated through logistic regression to predict changes in profile membership overtime (RQ 4; Asparouhov and Muthén

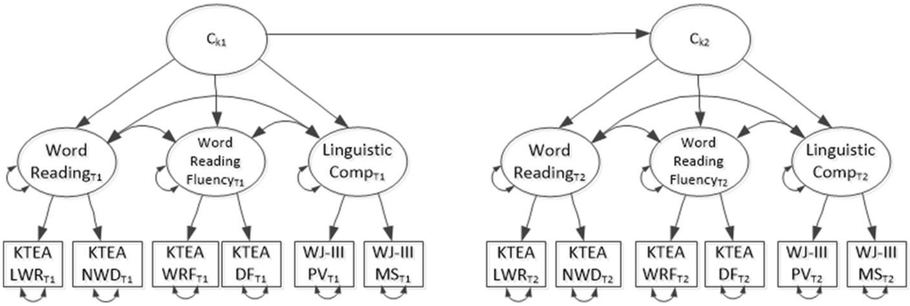

Fig. 2 Latent Transition Analysis for the Factor Mixture Model (Model 3). LWR = KTEA3 Letter and Word Recognition; NWR = KTEA3 Nonsense Word Reading; WRF = KTEA3 Word Reading Fluency; DF = KTEA3 Decoding Fluency; VOC = WJ-III Picture Vocabulary; MS = WJ-III Memory for Sentences

2014). This longitudinal extension was also applied to FMM (as shown in Fig. 2). Like the BCH approach (described above), the first steps consist of separately estimating the profile variables at time 1 and time 2, so that the measurement model at one time point is not affected by the other. The final steps fit the full model with prefixed error rates using the BCH weights. Longitudinal invariance was also evaluated by performing a sequence of invariance tests in which the means and residual variances were constrained to be equal across time. For example, a model with residuals correlated across time (e.g., KTEA Letter & Word Decoding at time 1 was correlated with KTEA Letter & Word Decoding at time 2) for the two-profile solution provided better class separation and lower information criteria (entropy = 0.96; AIC = 7895.43; BIC = 8043; SABIC = 7920.23) than a model without correlated residuals across time (entropy = 0.94; AIC = 8741.74; BIC = 8844.56; SABIC = 8758.91). Therefore, this model with correlated residuals was retained. For the FMM, longitudinal invariance was further established by constraining the factor loadings to be equal across time.

Results

In preliminary inspections of the data, outliers were defined as observations above or below three standard deviations and with high leverage. Two outliers were detected for WJ-III Picture Vocabulary at time 1, one outlier was detected for KTEA Letter Word Recognition at time 1, and one at time 2. However, these data points did not represent any major or minor reliability concerns (e.g., equipment failure) and the exclusion of these data points did not substantively impact the results. Therefore, we present the analyses with the outlier data included.

Table 1 presents the descriptive statistics and correlations for time 1 and time 2. There was a small standardized mean difference between time 1 and time 2 for English memory for sentences ($d = 0.19$), but standardized mean differences on all other measures were negligible. The highest correlations were among the KTEA measures, followed by correlations among the WJ-III measures. Spanish language proficiency (memory for sentences) was not significantly correlated with any of the English measures.

Table 1 Descriptive Statistics and Correlations at Time 1 (below diagonal) and Time 2 (above diagonal)

		1	2	3	4	5	6	7
Time 1								
	Mean	92.87	92.86	88.36	97.69	71.73	66.76	87.08
	SD	16.31	14.60	13.62	14.74	16.77	18.50	11.72
	n	326	331	328	326	302	303	329
Time 2								
	Mean	91.75	91.90	89.29	97.77	72.90	70.28	87.28
	SD	15.63	14.61	13.92	14.46	14.46	17.87	14.20
	n	315	321	321	320	281	278	319
Cohen's d Hedges' g	0.07	0.07	0.07	0.05	0.07	0.19	0.02	
English measures								
1	KTEA3 Letter and Word Recognition	1	0.76	0.78	0.71	0.58	0.57	0.63
2	KTEA3 Nonsense Word Decoding	0.76	1	0.75	0.79	0.45	0.49	0.54
3	KTEA3 Word Recognition Fluency	0.79	0.66	1	0.83	0.56	0.57	0.66
4	KTEA3 Decoding Fluency	0.67	0.70	0.79	1	0.40	0.44	0.55
5	WJ-III Picture Vocabulary	0.62	0.49	0.61	0.42	1	0.68	0.65
6	WJ-III Memory for Sentences	0.58	0.50	0.57	0.43	0.64	1	0.64
7	GM Reading Comprehension	0.52	0.40	0.57	0.42	0.48	0.48	1

$r > |.16|$, $p < 0.001$; $|.16| > r > |.10|$, $p < 0.05$. KTEA3, Kaufman Test of Educational Achievement, Third Edition; WJ-III, Woodcock Johnson, Third Edition; GM, Gates MacGinitie.

RQ 1: Latent profiles

Model selection

A comprehensive methodological approach was adopted in which latent profiles were estimated using observed and latent variables because of the well-known factor structures that underlie the variables included in the present study. Fit indices for the models at time 1 (see Table 2) and time 2 (see Table 3) showed that Model 1 (LPA without correlated residuals) and Model 3 (factor mixture model) provided better profile separation (i.e., entropy) than Model 2 (LPA with correlated residuals) for all solutions. However, Model 2 (LPA with correlated residuals) provided better relative fit (i.e., lower information criteria) compared to Models 1 and 3. In considering the substantive interpretations, we closely examined the results across all models and solutions and favored classification-based statistics (e.g., entropy) over likelihood-based statistics (e.g., AIC and BIC) because classification-based statistics are most frequently employed and are not affected by small sample sizes (Henson et al., 2007). Because Model 2 produced substantive results that were highly similar to those of Model 1, in what follows, we discuss the results of the best fitting LPA model (Model 1) and the FMM (Model 3).

Table 2 Model Fit Indices for the Latent Profile Models at Time 1

Classes	LL	AIC	BIC	SABIC	Entropy	Largest class %	Smallest class %	VLMR	LMR
Model 1: LPA									
2 classes	-2280.12	4610.24	4705.30	4626.00	0.91	53	47	896.99**	885.26**
3 classes	-2152.04	4380.09	4524.57	4404.03	0.86	41	27	256.16	252.81**
4 classes	-2086.98	4275.95	4469.86	4308.09	0.86	34	11	130.13	128.43
5 classes	-2042.15	4212.29	4455.63	4252.62	0.88	35	5	89.66	88.49
6 classes	-1998.82	4151.63	4444.39	4200.15	0.90	31	2	86.66	85.53
Model 2: LPA with correlated residuals									
2 classes	-2123.86	4303.72	4410.18	4321.36	0.85	57	43	424.67**	419.12**
3 classes	-2066.31	4214.62	4370.51	4240.46	0.79	40	25	115.1	113.59
4 classes	-2020.52	4149.05	4354.36	4183.07	0.81	37	9	91.57	90.38
5 classes	-1987.57	4109.13	4363.87	4151.35	0.86	35	3	65.92	65.06
6 classes	-1956.12	4072.25	4376.42	4122.66	0.87	31	2	62.88	62.06
Model 3: Factor mixture model									
2 classes	-2302.57	4643.14	4715.38	4655.11	0.91	55	45	852.1	831.63**
3 classes	-2198.64	4443.28	4530.73	4457.77	0.85	35	24	207.85	199.27
4 classes	-2136.50	4326.99	4429.65	4344.00	0.85	37	11	124.29	119.16
5 classes	-2099.51	4261.02	4378.88	4280.55	0.86	33	1	73.97	70.92
6 classes	-2066.98	4203.96	4337.03	4226.01	0.86	31	2	65.06	62.37

LL = Loglikelihood; AIC = Akaike Information Criteria; BIC = Bayesian Information Criteria; SABIC = Sample-adjusted BIC; VLMR = Vuong-Lo-Mendell-Rubin likelihood ratio test; LMR = Lo-Mendell-Rubin adjusted likelihood ratio test.

Number of profiles

The VLMR and LMR ratio tests suggested that the two-profile solution best represented the data for the latent profiles and factor mixture models, rather than solutions with more profile groups. The 2-profile solution also had the highest entropy values compared to the 3–6-profile solutions for Models 1–3. Results from these models suggested more similarities than differences, with most solutions showing ordered profiles, suggesting quantitative differences marked by severity (i.e., profiles presented higher or lower levels on all variables) rather than qualitative differences marked by specificity (Morin et al., 2016; Nylund, 2007). Figure 3 shows the ordered profiles for Model 1 for the 2-profile solution at time 1 and 2 in which all scores were standardized (z-scored) relative to the sample mean. Thus, the z-scores in Fig. 3 refer to relative performance rather than absolute normative performance.

For the 2-profile solution, the first latent profile described 47% and 43% of the sample of students as RD at time 1 and time 2, respectively, (i.e., on average 0.80 standard deviations below the sample mean). The second group, typically developing Els (i.e., standardized scores approximately 0.60 standard deviations above the sample mean), represented 53% (time 1) and 57% (time 2) of the total sample. A similar pattern of results emerged for the 2-profile solution for the FMM (Model 3), in which the at-risk group (45%) served as the referent group (see Fig. 4). Thus, in Model 3, the factor means for the TD profile were

Table 3 Model Fit Indices for the Latent Profile Models at Time 2

Classes	LL	AIC	BIC	SABIC	Entropy	Largest class %	Smallest class %	VLMR	LMR LRT
Model 1: LPA									
2 classes	−2185.09	4420.18	4514.46	4435.17	0.90	57	43	843.01**	831.92**
3 classes	−2055.85	4187.69	4331.01	4210.48	0.86	39	22	258.48	255.08
4 classes	−1978.72	4059.44	4251.78	4090.02	0.86	37	14	154.25	152.22
5 classes	−1942.52	4031.04	4254.41	4051.41	0.86	34	7	72.40	71.45
6 classes	−1907.21	3968.42	4258.82	4041.59	0.85	30	7	70.62	69.69
Model 2: LPA with correlated residuals									
2 classes	−2123.86	4303.72	4410.18	4321.36	0.85	57	43	424.68	419.12**
3 classes	−1955.23	3992.46	4147.09	4017.04	0.82	41	18	136.23	134.44
4 classes	−1905.67	3919.34	4123.00	3951.72	0.81	35	12	99.17	97.81
5 classes	−1871.96	3877.92	4130.61	3918.09	0.81	32	12	67.43	66.54
6 classes	−1846.34	3852.68	4154.40	3900.65	0.84	31	2	68.95	68.05
Model 3: Factor mixture model									
2 classes	−2197.11	4432.22	4503.87	4443.61	0.90	58	42	818.96**	799.18**
3 classes	−2080.36	4206.72	4293.47	4220.51	0.85	41	23	233.49	223.8
4 classes	−2017.92	4089.84	4191.67	4106.03	0.86	40	12	124.89*	119.70*
5 classes	−1999.24	4060.48	4177.39	4079.06	0.81	30	10	37.36	35.81
6 classes	−1982.10	4034.19	4166.19	4055.18	0.84	34	3	34.29	32.86

LL = Loglikelihood; AIC = Akaike Information Criteria; BIC = Bayesian Information Criteria; SABIC = Sample-adjusted BIC; VLMR = Vuong-Lo-Mendell-Rubin likelihood ratio test; LMR = Lo-Mendell-Rubin adjusted likelihood ratio test.

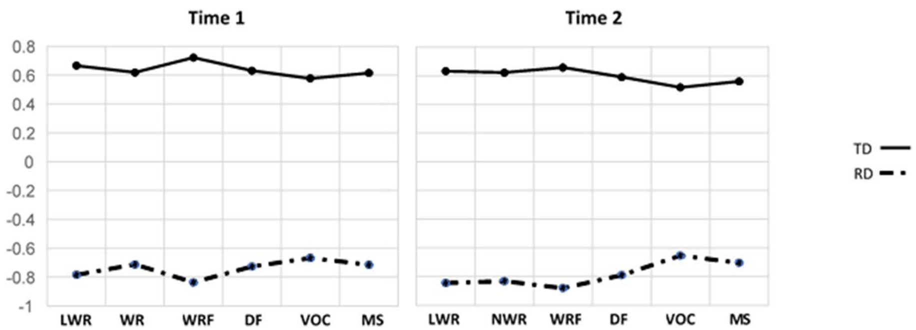

Fig. 3 Pattern of results for Model 1 (Latent Profile Analysis). Class counts for the 2-profiles solution are: T1 (RD = 154; TD = 177) and T2 (RD = 135; TD = 186). LWR = KTEA3 Letter Word Recognition; NWR = KTEA3 Non-Word Reading; WRF = KTEA3 Word Reading Fluency; DF = KTEA3 Decoding Fluency; VOC = WJ-III Picture Vocabulary; MS = WJ-III Memory for Sentences

estimated relative to the estimated factor means for the RD. These factor scores are also interpreted relative to the RD group mean rather than the normative population. The results for Model 3 showed ordered profiles, with slightly lower factor means in linguistic comprehension (T1 = 1.23; T2 = 1.18) than in word reading (T1 = 1.46; T2 = 1.48) or word reading

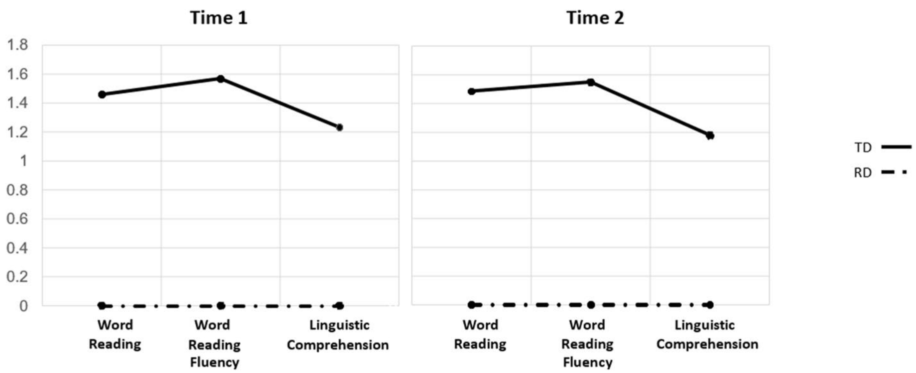

Fig. 4 Pattern of results for Model 3 (Factor Mixture Model). Class counts for the 2-profiles solution are: T1 (LD = 149; TD = 182), and T2 (LD = 133; TD = 188). The RD class served as the reference class with its factor means set to zero. Scores are Extended Scale Scores on Gates MacGinitie Reading Test. T1 = Time 1; T2 = Time 2

fluency (T1 = 1.57; T2 = 1.55). The descriptive statistics for the 2-profile solution from Model 1 are presented in Table 4. With the exception of memory for sentences in Spanish (d = 0.03–0.20), the at-risk and typical groups differed on all measures (d = 1.37–2.20). The standard scores in Table 4 help to explicate the sample's performance relative to normative expectations. For example, Table 4 shows that ELs with RD show standard scores between 77 and 87 on word reading measures at time 1, but they show standard scores across language measures of 55–63. Thus, it is important to note that normative differences are not immediately evident from the sample z-scores reported in Figs. 3–4.

RQ 2: Reading comprehension development

In answering research questions 2–3, we again present the results of the best fitting and most parsimonious model (2-profile solutions) because the solutions with more than two profiles did not provide empirical or theoretical advantages. Mean level differences in reading comprehension growth across the academic year emerged for at-risk and typically developing Els in the context of both observed index variables (χ^2 (1) = 11.74, $p < 0.001$), and factor mixture models (χ^2 (1) = 11.05, $p < 0.001$), with the typical readers increasing an average of 18 extended scale score points on the Gates MacGinitie (SD = 27) from time 1 to time 2 and the at-risk group increasing an average of 7 extended scale score points (SD = 30) in both the LPA (Model 1) and FMM (Model 3). Mean scores for the two-profile groups from LPA and FMM are presented in Fig. 5.

RQ 3: profile stability

Finally, the LTA models estimated the probability of transitioning from the at-risk profile to the typically developing profile (and vice versa). Both LTA analyses with latent profiles and factor mixtures provided evidence for a high degree of stability (e.g., with 97% of individuals staying in the same profile group and only 3% changing in Model 1).

Table 4 Descriptive statistics for latent profile groups at Time 1 and Time 2 (2-Class Solution Only)

	RD Profile Group					TD Profile Group					Cohen's d
	n	*Mean*	*SD*	*Min*	*Max*	*n*	*Mean*	*SD*	*Min*	*Max*	*TD vs. RD*
KTEA3 Letter and Word Recognition											
Time 1	150	78.0	11.8	40	114	176	102.6	11.5	84	158	2.11
Time 2	133	78.3	10.5	47	99	182	101.6	10.7	79	160	2.20
KTEA3 Nonsense Word Decoding											
Time 1	154	81.4	9.5	58	111	177	100.8	12.3	77	139	1.77
Time 2	135	79.5	8.5	58	98	186	100.9	11.1	75	134	2.16
KTEA3 Word Reading Fluency											
Time 1	151	77.64	9.67	40	97	177	97.51	9.06	76	131	2.13
Time 2	135	77.57	9.44	40	94	186	97.78	9.88	78	138	2.09
KTEA3 Decoding Fluency											
Time 1	149	87.03	11.59	40	108	177	106.67	10.53	82	135	1.83
Time 2	134	85.52	10.16	47	108	186	106.59	9.97	82	150	1.77
WJ-III Picture Vocabulary											
Time 1	140	60.2	15.5	11	93	162	80.8	10.2	55	102	1.57
Time 2	122	63.5	13.4	26	100	159	80.1	10.6	45	102	1.37
WJ-III Memory for Sentences											
Time 1	140	55.7	13.4	28	105	163	78.9	12.0	49	116	1.82
Time 2	121	57.5	13.4	25	90	157	80.1	14.3	47	115	1.63
GM Reading Comprehension											
Time 1	140	80.5	8.9	65	116	173	92.9	10.9	65	131	1.25
Time 2	135	77.4	9.1	65	110	184	94.5	12.9	65	135	1.53

KTEA3 = Kaufman Test of Educational Achievement, Third Edition; WJ-III = Woodcock Johnson, Third Edition; GM = Gates MacGinitie.

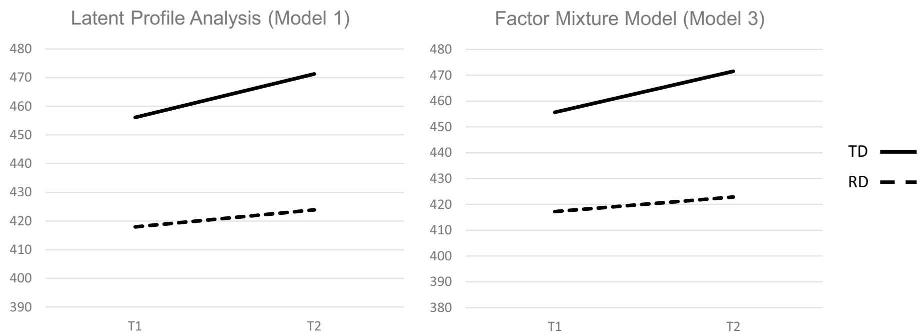

Fig. 5 Pattern of results for the covariate (reading comprehension) at times 1 and 2. Scores are Extended Scale Scores on Gates MacGinitie Reading Test. T1 = Time 1; T2 = Time 2.

Discussion

This study investigated the English reading profiles of a sample of ELs in Grades 3 and 4 with and without risk for dyslexia and other reading difficulties. Our first goal was to determine whether we could identify distinct latent profiles based on individual performance in word and pseudo-word reading accuracy and fluency and language comprehension in English. We further investigated if latent profile group membership was predictive of growth in reading comprehension from the beginning to end of the academic year. Finally, we investigated the stability of latent profile group membership over a single academic year. This study featured several attractive characteristics, which position it to inform our understanding of the reading development of ELs with and without risk for dyslexia, including a large sample, psychometrically strong measures collected at multiple times, and robust analytic methods including the use of factor mixture modeling to increase the generalizability of results beyond a single analytic method. In the sections that follow, we highlight key findings related to our three research questions and attempt to situate these findings within the context of previous research with ELs, particularly ELs at significant risk for dyslexia.

Latent profiles were marked by severity, not specificity

In recent years, several studies have investigated whether distinct profiles of readers (with and without reading difficulties) can be identified based on their performance on component reading skills. These studies are important because a better understanding of the unique and shared needs of individual readers (or profiles of readers) may point toward more effective screening and intervention strategies (Cho et al., 2019; Clemens et al., 2017). A central question in studies of this type concerns whether the latent profiles that emerge are marked by the severity (i.e., level) or specificity (i.e., specific intra-individual deficits) of students' reading skills (Capin et al., 2021). On this question, the findings of the present study are unambiguous: across analytic methods and class solutions, we consistently identified profiles differentiated by large, educationally meaningful differences in reading performance across all measured component skills. For example, within the Model 1 2-profile solution, mean standard scores in decoding and fluency for the TD group were within the average range (100–107), while mean scores for the RD group were generally more than one standard deviation below the population mean (ranged from 78 to 87). Although solutions with a larger number of profiles did identify small profile groups (e.g., less than 5% of the sample) with specific deficits (i.e., specific deficits in decoding), these solutions proved a poor fit to the data. Figures 3 and 4 provide a compelling visualization of these relative differences in reading level. Although we interpret the 2-profile solutions as the best fitting model, we also provide 3-profile solutions to demonstrate the consistency of our results. Across both LPA (Model 1) and FMM (Model 3) with both 2- and 3-profile solutions, relative performance across profile groups are strikingly parallel with differences emerging in the severity of reading difficulties, not specificity.

Our findings with Spanish-speaking ELs in high-poverty schools are generally consistent with previous latent profile analyses conducted with monolingual struggling readers, which find that the majority of struggling readers demonstrate difficulties across component reading skills and that profiles largely emerge on the basis of the severity rather than the specificity of reading deficits (Brasseur-Hock et al., 2011; Capin et al., 2021; Clemens et al., 2017; Lesaux & Kieffer, 2010). However, one key difference between the present study and previous studies is that the best fitting solutions in our study did not identify any latent profiles marked by specific component deficits. This finding is consistent with the findings of O'Connor and colleagues (2019), who also found a two-profile solution highlighting differences in severity between poor comprehenders and good comprehenders in both the EL group, as well as their monolingual English-only group.

To better understand why these different findings might emerge, it is important to consider at least two dimensions on which this study differs from latent profile analyses with struggling readers. Similar to O'Connor and colleagues (2019), we selected participants based on EL status, not their status as struggling readers, as was done in several recent latent profile analyses of reading skills in English (e.g., Brasseur-Hock et al., 2011; Capin et al., 2021; Clemens et al., 2017; Lesaux & Kieffer, 2010). The samples for O'Connor et al. and the present study featured ELs across a range of reading proficiency, meaning the magnitude of within-sample differences may differ from estimates in previous studies conducted solely with struggling readers. An analogy may serve to highlight the importance of this difference. Consider a study that aimed to compare the physical appearance of school-age children. The make-up of the study sample would have important implications for which differences emerge as most prominent. A sample with students aged 4–18 would be marked by significant, meaningful differences in size (height and weight). The size differences between four year olds and eighteen year olds are clear and important. Indeed, size differences might overwhelm and obscure other important physical differences. However, if the sample were limited to students in first grade, additional physical differences might become more prominent. While size differences will likely be present, smaller differences in appearance (e.g., hair length and type, complexion) may better differentiate within the sample. A similar phenomenon can occur when comparing latent profiles: if inclusion is limited to participants who perform similarly on some dimension (e.g., reading comprehension level), subgroup differences in other correlated dimensions (.e.g., word reading, fluency) may present a different picture than when participants with above average, average, and poor reading proficiency are included.

This study also differed from several previous latent profile analyses in the age of its sample (e.g., Brasseur-Hock et al., 2011; Clemens et al., 2017; Lesaux & Kieffer, 2010). To our knowledge, Capin et al. (2021) and O'Connor et al. (2019) are the only previous studies to employ latent profile analyses based on reading component skills with students in mid-to-late elementary. Notably, both studies that reported a very large proportion of the total sample were placed in groups marked by severity (91% and 100%, respectively) when compared to middle school samples. Capin and colleagues (2021) speculate that this finding may reflect the fact that late elementary students may experience more difficulties in foundational reading skills than middle school students who have had more time to solidify these skills. Our findings appear consistent with that hypothesis and are robust across analytic methods.

Growth in reading comprehension *differs by profile*

Our second research question investigated whether growth in reading comprehension across a single academic year differed for the RD and TD profile groups. Within our sample, we observed significant differences in growth in reading comprehension across the academic year, with the TD profile group outperforming the RD group in growth in reading comprehension across the year. Figure 4 visually presents observed changes in mean reading comprehension at times 1 and 2 for the 2-profile solutions of the latent profile analysis and factor mixture modeling. Both models show the gap in reading comprehension between the TD and RD group widening across the academic year, indicating that the ELs in this sample demonstrating significant risk for dyslexia and other reading disabilities are falling behind the benchmark sample. This finding is consistent with previous latent transition analyses (Swanson et al. 2020) and studies of disability identification (O'Connor et al. 2013; Samson & Lesaux, 2009) that highlight the heightened reading risk for ELs in these grades.

Latent profile group membership is largely stable

Our third research question concerned the stability of the latent profiles across a single academic year. Changes in profile membership over time have important implications, as findings may guide dynamic protocols for intervention and assessment. Our findings suggest that within our sample, the latent profile groups were quite stable, with 97% of the students remaining in the same profile at the end of the academic year. This stability in latent profile groups is higher than that observed by Swanson and colleagues (2019) and likely reflects three key differences between the studies. First, our optimal solution features only two latent profile groups, while Swanson and colleagues identified four profile groups. Importantly, Swanson and colleagues included a number of cognitive and linguistic variables not included in this study and featured extensive bilingual assessment. Given the larger number of profile groups and input variables, it is perhaps not surprising that we observed fewer transitions than Swanson and colleagues. The pattern of results for our 3–6 profiles solutions confirms this, as do the results of the factor mixture models, but across all models and solutions strong longitudinal stability was observed. Additionally, Swanson and colleagues evaluated latent profiles across three years in early elementary. The addition of multiple years, particularly in the dynamic years of early elementary, likely increases the odds of any one student transitioning from one latent profile group to another.

Implications for practice

The findings of this study should be interpreted in the context of the broader corpus of research investigating the reading profiles of students in elementary grades with and without risk for dyslexia. Within this context, we highlight the central finding in this study that the reading profiles that emerged were marked primarily by the severity of reading difficulty, not the specificity of those deficits. This finding has important implications for assessment, particularly screening processes to identify risk for dyslexia and reading difficulties more generally. The relatively consistent reading profiles we identified corroborate previous studies that find the vast majority of struggling readers demonstrate normative deficits across reading components (Capin et al., 2021; Cirino et al., 2013; O'Connor et al., 2019). This finding is also in line with the broader literature on screening for dyslexia and reading difficulties, which generally finds adequate predictive accuracy for general outcome measures or other broad indicators of reading achievement and therefore focuses on the efficiency of the measurement paradigm. Our findings suggest that similar paradigms hold promise for ELs, as well (Cummings et al., 2019; Johnson et al., 2009), although different benchmarks may be necessary. For instance, our results provide further evidence that ELs, with and without RD, demonstrate lower performance on linguistic comprehension measures relative to the normative population. As evidenced by the standard scores presented in Table 4, even typically developing ELs may show linguistic comprehension performance that is more than 1 SD below the normative mean. Our results revealed that the differences between ELs with typical reading skills and reading difficulties were most apparent, based on effect size differences, on measures of word reading. Given this and the critical role of word reading for students with dyslexia, it is worth wondering whether ELs may be best screened for reading difficulties using measures of word reading, like the ones used in this study.

Instructionally, the multiple component deficits observed in the RD group and the magnitude of normative deficits highlight the critical need for effective interventions for ELs in late elementary school and beyond. Past studies have rightly identified developing linguistic comprehension as a key area of difficulty for ELs (e.g., Lesaux & Kieffer, 2010; Buly & Valencia, 2002). Some may have interpreted this to suggest that English learners with difficulties only need English language support, such as academic language instruction. Our findings suggest that ELs with RD do have very low scores on English language measures; however, they also demonstrate underperformance in word reading. It is their performance in both of these areas that differentiates students with reading difficulties from those who are typically developing. These findings underscore the importance of providing ELs with evidence-based word reading instruction in the primary grades to prevent word reading difficulties and risk for dyslexia. They also highlight the need for long-term, multi-component reading interventions that simultaneously address word reading, fluency, and linguistic processes among students with reading difficulties. As opposed to working on these skills in isolation, optimal interventions may integrate word reading and fluency instruction within reading interventions that target building vocabulary, comprehension, and content knowledge (Authors, in development). For instance, within a single lesson, instructors may teach students to read difficult words and build their reading fluency using words and texts that are essential to the content that they are going to read when focused on developing vocabulary and comprehension. In this way, all of the instruction is integrated and focused on developing language, which may be particularly beneficial for ELs.

Limitations

The findings of this study are specific to the included sample and measures. The findings from this study should be interpreted in the context of ELs who are attending high-poverty schools who have a primary language of Spanish. A participant sample including Spanish-speaking ELs from additional schools and/or districts might increase the generalizability of these results. Specific to the measures included, it is important to note that none of the measures employed perfectly or completely measures the construct of interest. For example, we included two word reading fluency measures but did not include a measure of fluency with connected text. Similarly, we included a vocabulary and sentence recall task as indictors of linguistic comprehension and did not include a standardized listening comprehension. It is possible that the sentence recall task also taps cognitive skills such as short-term memory, which may have influenced results. We also observed within-sample internal consistency estimates that were lower than published reliabilities. Thus, our results should be interpreted specific to the measures we employed and with caution. Additionally, the analytic methods employed in this and other studies employing person-centered approaches involve some researcher discretion that influence results. Thus, generalizability in the context of latent profile analysis and mixture modeling is typically built on an accumulation of studies that differentiate the core set of academic/language skills profiles that emerge with regularity across samples of various characteristics (e.g., SES status, educational background, language proficiency, etc.) in order to differentiate them from profiles that emerge more spuriously (Morin & Litalien, 2019; Meyer & Morin, 2016). Further research may investigate the reading profiles of ELs in different educational contexts, including different grades, schools with greater demographic diversity, and schools with different instructional models for ELs. Additionally, there is a critical need for longitudinal studies that track ELs across late elementary grades to better parse issues of disability and second language acquisition.

Conclusion

We investigated the reading and language skill profiles of ELs with and without risk for dyslexia in 3rd and 4th grade—a critical period of transition in reading. Across analytic methods and class solutions, a consistent pattern of results emerged: the latent profiles that emerged demonstrated relatively consistent performance across reading tasks within groups, but latent profile groups differed in significant, educationally meaningful ways in their level of performance. These data suggest that within our sample of ELs, ELs at significant risk for dyslexia and other reading disabilities are defined primarily by the severity of their reading deficits rather than demonstrating specific deficits in single component skills. The latent profile groups were highly stable across the academic year and the RD group demonstrated less growth in reading comprehension across the year. Together with an observed drop in normative performance for the RD group across the academic year, this finding suggests that these grades represent a period of transition for ELs at risk for dyslexia and other reading difficulties in which they are particularly vulnerable to falling further behind. Long-term, intensive, multi-component reading interventions may be necessary for ELs at risk for dyslexia to prevent persistent reading difficulties.

Funding This research was supported by grant R324B170012 from the National Center for Special Education Research in the Institute of Education sciences and grants P50 HD052117 and R01HD096262 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Center for Special Education Research, the Institute of Education Sciences, the Eunice Kennedy Shriver National Institute of Child Health and Human Development, or the National Institutes of Health,

References

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Using the BCH method in Mplus to estimate a distal outcome model and an arbitrary secondary model. *Mplus Web Notes, 21*(2), 1–22.

August, D., Shanahan, T., & Escamilla, K. (2009). English language learners: Developing literacy in second language learners—report of the National Literacy Panel on Language-Minority Children and Youth. *Journal of Literacy Research, 41*(4), 432–452.

Bakk, Z., & Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling: A Multidisciplinary Journal, 23*(1), 20–31.

Betts, J., Bolt, S., Decker, D., Muyskens, P., & Marston, D. (2009). Examining the role of time and language type in reading development for English language learners. *Journal of School Psychology, 47*(3), 143–166.

Bialystok, E. (2007). Acquisition of literacy in bilingual children: A framework for research. *Language Learning, 57*, 45–77.

Bolck, A., Croon, M., & Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political analysis*, 3–27.
Brasseur-Hock, I. F., Hock, M. F., Kieffer, M. J., Biancarosa, G., & Deshler, D. D. (2011). Adolescent struggling readers in urban schools: Results of a latent class analysis. *Learning and Individual Differences, 21*(4), 438–452.
Buly, M. R., & Valencia, S. W. (2002). Below the bar: Profiles of students who fail state reading assessments. *Educational Evaluation and Policy Analysis, 24*(3), 219–239.
Burr, E., Haas, E., & Ferriere, K. (2015). Identifying and supporting english learner students with learning disabilities: Key issues in the literature and state practice. REL 2015–086. *Regional Educational Laboratory West*.
Capin, P., Cho, E., Miciak, J., Roberts, G., & Vaughn, S. (2021). Examining the reading and cognitive profiles of students with significant reading comprehension difficulties. *Learning Disability Quarterly*, *44*(3), 183–196.
Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading.
Chall, J. S., & Jacobs, V. A. (1983). Writing and reading in the elementary grades: Developmental trends among low SES children. *Language Arts, 60*(5), 617–626.
Cirino, P. T., Romain, M. A., Barth, A. E., Tolar, T. D., Fletcher, J. M., & Vaughn, S. (2013). Reading skill components and impairments in middle school struggling readers. *Reading and Writing, 26*(7), 1059–1086.
Cho, E., Capin, P., Roberts, G., Roberts, G. J., & Vaughn, S. (2019). Examining sources and mechanisms of reading comprehension difficulties: Comparing English learners and non-English learners within the simple view of reading. *Journal of Educational Psychology, 111*(6), 982.
Cho, E., Roberts, G. J., Capin, P., Roberts, G., Miciak, J., & Vaughn, S. (2015). Cognitive attributes, attention, and self-efficacy of adequate and inadequate responders in a fourth grade reading intervention. *Learning Disabilities Research & Practice, 30*(4), 159–170.
Clark, S. L., Muthén, B., Kaprio, J., D'Onofrio, B. M., Viken, R., & Rose, R. J. (2013). Models and strategies for factor mixture analysis: An example concerning the structure underlying psychological disorders. *Structural Equation Modeling: A Multidisciplinary Journal, 20*(4), 681–703.
Clemens, N. H., Simmons, D., Simmons, L. E., Wang, H., & Kwok, O. M. (2017). The prevalence of reading fluency and vocabulary difficulties among adolescents struggling with reading comprehension. *Journal of Psychoeducational Assessment, 35*(8), 785–798.
Croon, M. (2002). Ordering the classes. *Applied latent class analysis*, 137–162.
Cummins, J. (2000). Immersion education for the millennium: What we have learned from 30 years of research on second language immersion. In M. R. Childs & R. M. Bostwick (Eds.) *Learning through two languages: Research and practice*. *SecondKatoh Gakuen International Symposium on Immersion and Bilingual Education.* (pp. 34–47). Katoh Gakuen, Japan.
Donovan, S., & Cross, C. (Eds.). (2002). *Minority students in special and gifted education*. Washington, DC: National Academy Press.
Duke, N., & Carlisle, J. F. (2011). Comprehension Development. *Handbook of Reading Research, 4*, 199–228.
Elliott, J. G., & Grigorenko, E. L. (2014). *The dyslexia debate* (No. 14). Cambridge University Press.
Ferguson, S. L., Moore, G., & E. W., & Hull, D. M. (2020). Finding latent groups in observed data: A primer on latent profile analysis in Mplus for applied researchers. *International Journal of Behavioral Development, 44*(5), 458–468.
Francis, D. J., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). Based Recommendations for Instruction and Academic Interventions. Practical Guidelines for the Education of English Language Learners. Book 1 of 3. *Center on Instruction*.
Geva, E., & YaghoubZadeh, Z. (2006). Reading efficiency in native English-speaking and English-as-a-second-language children: The role of oral proficiency and underlying cognitive-linguistic processes. *Scientific Studies of Reading, 10*(1), 31–57.
Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6–10.
Gottardo, A., & Mueller, J. (2009). Are first-and second-language factors related in predicting second-language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology, 101*(2), 330.
Hall, C., Steinle, P. K., & Vaughn, S. (2019). Reading instruction for English learners with learning disabilities: What do we already know, and what do we still need to learn? *New Directions for Child and Adolescent Development, 2019*(166), 145–189.

Henson, J. M., Reise, S. P., & Kim, K. H. (2007). Detecting mixtures from structural model differences using latent variable mixture modeling: A comparison of relative model fit statistics. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(2), 202–226.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2*(2), 127–160.

Hoover, J. J., Erickson, J. R., Patton, J. R., Sacco, D. M., & Tran, L. M. (2019). Examining IEPs of English learners with learning disabilities for cultural and linguistic responsiveness. *Learning Disabilities Research & Practice, 34*(1), 14–22.

Jean, M., & Geva, E. (2009). The development of vocabulary in English as a second language children and its role in predicting word recognition ability. *Applied Psycholinguistics, 30*(1), 153–185.

Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning, 64*(1), 160–212.

Johnson, E. S., Jenkins, J. R., Petscher, Y., & Catts, H. W. (2009). How can we improve the accuracy of screening instruments? *Learning Disabilities Research & Practice, 24*(4), 174–185.

Kaufman, A. S., & Kaufman, N. L. (2014). *Technical & interpretive manual: Kaufman Test of Educational Achievement* (3rd ed.). NCS Pearson.

Kieffer, M. J., & Lesaux, N. K. (2008). The role of derivational morphology in the reading comprehension of Spanish-speaking English language learners. *Reading and Writing, 21*(8), 783–804.

Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S. A. H., Gustafsson, J. E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science, 18*(1), 146–154.

Klingner, J. K., Artiles, A. J., & Barletta, L. M. (2006). English language learners who struggle with reading: Language acquisition or LD? *Journal of Learning Disabilities, 39*(2), 108–128.

Leite, W. L., & Cooper, L. A. (2010). Detecting social desirability bias using factor mixture models. *Multivariate Behavioral Research, 45*(2), 271–293.

Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal, 47*(3), 596–632.

Lesaux, N. K., Lipka, O., & Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Reading and Writing, 19*(1), 99–131.

Lesaux, N. K., Crosson, A. C., Kieffer, M. J., & Pierce, M. (2010). Uneven profiles: Language minority learners' word reading, vocabulary, and reading comprehension skills. *Journal of Applied Developmental Psychology, 31*(6), 475–483.

Li, M., Kirby, J. R., Geva, E., Koh, P. W., & Zhang, H. (2021). Profiles of Poor Decoders, Poor Comprehenders, and Typically Developing Readers in Adolescents Learning English as a Second Language. *Journal of Learning Disabilities*, 00222194211023200.

Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods, 10*(1), 21.

MacGinitie, W. H. (2000). *Gates-MacGinitie Reading Tests.* Itasca. IL: Riverside.

Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority learners. *Journal of Educational Psychology, 102*(3), 701.

Masyn, K. E., Henderson, C. E., & Greenbaum, P. E. (2010). Exploring the latent structures of psychological constructs in social development using the dimensional–categorical spectrum. *Social Development, 19*(3), 470–493.

Meyer, J. P., & Morin, A. J. (2016). A person-centered approach to commitment research: Theory, research, and methodology. *Journal of Organizational Behavior, 37*(4), 584–612.

Miciak, J., & Fletcher, J. M. (2020). The critical role of instructional response for identifying dyslexia and other learning disabilities. *Journal of Learning Disabilities, 53*(5), 343–353.

Miciak, J., Stuebing, K. K., Vaughn, S., Roberts, G., Barth, A. E., & Fletcher, J. M. (2014). Cognitive attributes of adequate and inadequate responders to reading intervention in middle school. *School Psychology Review, 43*(4), 407–427.

Morin, A. J., & Marsh, H. W. (2015). Disentangling shape from level effects in person-centered analyses: An illustration based on university teachers' multidimensional profiles of effectiveness. *Structural Equation Modeling: A Multidisciplinary Journal, 22*(1), 39–59.

Morin, A., & Litalien, D. (2019). Mixture modeling for lifespan developmental research. In *Oxford Research Encyclopedia of Psychology*. Oxford University Press.

Morin, A. J., Boudrias, J. S., Marsh, H. W., Madore, I., & Desrumaux, P. (2016). Further reflections on disentangling shape and level effects in person-centered analyses: An illustration exploring the dimensionality of psychological health. *Structural Equation Modeling: A Multidisciplinary Journal, 23*(3), 438–454.

Morgan, P. L., Farkas, G., Hillemeier, M. M., Mattison, R., Maczuga, S., Li, H., & Cook, M. (2015). Minorities are disproportionately underrepresented in special education: Longitudinal evidence across five disability conditions. *Educational Researcher, 44*(5), 278–292.

Muthén, L.K., & Muthén B.O. (2017). *Mplus user's guide*, 8th ed., Muthén & Muthén.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(4), 535–569.

O'Connor, R. E., Bocian, K. M., Beach, K. D., Sanchez, V., & Flynn, L. J. (2013). Special education in a 4-year response to intervention (RtI) environment: Characteristics of students with learning disability and grade of identification. *Learning Disabilities Research & Practice, 28*(3), 98–112.

O'Connor, M., Geva, E., & Koh, P. W. (2019). Examining reading comprehension profiles of grade 5 monolinguals and English language learners through the lexical quality hypothesis lens. *Journal of Learning Disabilities, 52*(3), 232–246.

Samson, J. F., & Lesaux, N. K. (2009). Language-minority learners in special education: Rates and predictors of identification for services. *Journal of Learning Disabilities, 42*(2), 148–162.

Samson, J. F., & Lesaux, N. (2015). Disadvantaged language minority students and their teachers: A national picture. *Teachers College Record, 117*(2), 1–26.

Shaywitz, S. E. (2003). *Overcoming dyslexia: A new and complete science-based program for reading problems at any level*. Vintage.

Swanson, H. L., Arizmendi, G. D., & Li, J. T. (2020). The stability of learning disabilities among emergent bilingual children: A latent transition analysis. *Journal of Educational Psychology, 113*(6), 1244–1268. https://doi.org/10.1037/edu0000645.

Vellutino, F. R., Scanlon, D. M., & Reid Lyon, G. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities, 33*(3), 223–238.

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political analysis*, 450–469.

Woodcock, R., McGrew, K., Mather, N., & Schrank, F. (2007). *Woodcock-Johnson III NU tests of achievement*.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.