The Stability of Learning Disabilities Among Emergent Bilingual Children: A Latent Transition Analysis

H. Lee Swanson

Genesis D. Arizmendi

Jui-Teng Li

University of New Mexico

H. Lee Swanson, Ph.D., Educational Psychology, University of New Mexico (HLswanson@unm.edu) and University of California-Riverside (Lee.Swanson@ucr.edu). Genesis Arizmendi, Educational Psychology, Ph.D., post-doctoral Fellow University of New Mexico (garizmendi@unm.edu) and Jui-Teng Li, Ph.D., Educational Psychology, Ph.D., post-doctoral Fellow, University of New Mexico (Jui-tengli@unm.edu).

This research was support by an NSF grant, Division of Research on Learning (award number 1660828) Sciences awarded to the first author. This study does not necessarily reflect the views of NSF, or the participating school districts. Special appreciation is given to Drs. Stefania Petcu and Jennifer Kong as project directors. Also special appreciation is given to Monica Ascencio, Mario Del Angel Guevara, Lorena Luevano, Julianna Massa, Catherine Riskie, Tara Curtis, Christy Yan, Karen Flores, Sindy Zambrano, Karen Cruz, and Bernadette Hall-Cuaron and Michael Orosco, and School District Laisons and Consultant: Rosie Gonzales and Erin Bostick Mason for data collection and/or analysis.

Correspondence concerning this article should be addressed to H. Lee Swanson, Educational Psychology, College of Education and Human Sciences, University of New Mexico, and Albuquerque, NM 87131.
E-mail:HLswanson@unm.edu or Lee.Swanson@ucr.edu

Swanson, H. L., Arizmendi, G. D., & Li, J. (2021). The stability of learning disabilities among emergent bilingual children: A latent transition analysis. *Journal of Educational Psychology, 113*(6), 1244-1268. doi:http://dx.doi.org/10.1037/edu0000645

Abstract

This study investigated the prevalence and stability of latent classes among elementary-aged English learning (EL) children whose first language is Spanish. To this end, EL children (*N* = 267) in Grades 1, 2, and 3 at Wave 1 (Year 1) were administered a battery of vocabulary, reading, math, and cognitive measures (short-term memory, working memory, rapid naming, inhibition) in both Spanish and English. These same measures were also administered one year later (Wave 2). Four important findings occurred. First, four latent classes (balanced bilinguals-average achievers, unbalanced bilinguals-average achievers, children at risk for learning disabilities, English dominant) at both testing waves emerged. Second, probability estimates indicated that 20% of the total sample was at risk for learning disabilities at Wave 1, with late-emerging academic difficulties increasing the learning disabilities latent class by 5% at Wave 2. Third, the incidence of late-emerging children at risk for learning disabilities was higher among balanced bilingual average achievers, especially for those children transitioning to and/or from grade 3. Finally, the cognitive measures for predicting the odds of children being correctly classified in the final wave of testing included measures of naming speed and working memory. The results support the notion that statistically distinct latent classes emerge under the umbrella of language and academic performance and that children at risk for learning disabilities can be separated among a heterogeneous sample of children who are English language learners.

Keywords: reading and math disabilities, English learners, working memory, bilingual, latent class analysis

Educational Impact and Implications

Children who are learning in their second language with potential learning disabilities are at high risk for more advanced reading and math difficulties, school drop out, and future employment. Such children are also hard to identify amongst children who are struggling to become proficient in their second language. The present study identified a discrete and stable group of English learning (EL) children at risk for learning disabilities over two time periods. Consistent with findings on English monolingual children, EL children at risk for learning disabilities showed a combination of reading and math difficulties related to specific inefficiencies in processing speed and working memory. The results also indicated that there is a critical period during the third grade that plays an important role in determining late-emerging risk for learning disabilities, as well as variability in performance among average achieving groups after one year.

The Stability of Learning Disabilities among English Learners: A Latent Transition Analysis

Children with Spanish as a first language in the United States have been found to yield low achievement scores in reading and mathematics when compared to other English learners (EL) on national assessments across several years (e.g. August & Hakuta, 1997; Hemphill & Vanneman, 2011; National Assessment of Education Progress, 2011, 2017; 2019). Although closing achievement gaps has been a goal in national and state education policies, achievement scores in reading and math for non-EL students in grades 4 and 8 have been higher than the scores of then EL student whose first language is Spanish since 1996 (e.g., August & Hakuta, 1997; Bumgarner, Martin, Brooks-Gunn, 2013). Cross-sectional studies have shown that ELs whose first language is Spanish disproportionately experience academic difficulties across various age levels (U.S. Department of Education, National Center for Education Statistics, 2009). Factors that compound these challenges further include the findings that many of these EL children are at an increased risk for having reading and/or math disabilities (e.g., Farnia & Geva, 2019; Kieffer, 2011; Mancilla-Martinez, Hwang, Oh, & McClain, 2019; Swanson, Kudo, & Guzman-Orth, 2016). These findings are further vexing because several studies suggest that the gap in reading and/or math performance between monolingual children and Spanish speaking ELs increases across age groups (Kieffer, 2011; Mancilla-Martinez & Lesaux, 2017).

The reason behind the prevalence of low achievement in ELs whose first language is Spanish in the public schools is unclear because a consistent and accurate identification system across states, as well as school districts, does not exist (McCardle, Keller-Allen, & Shuy, 2008). Large scale studies have also revealed that ELs are underrepresented overall in special education, meaning that a smaller percentage of these students are not receiving educational services (e.g., Morgan & Farkas, 2016). However, there are also reports of overrepresentation of EL children in special education (e.g., Artiles, Rueda, Salazar, & Higareda, 2005). These challenges underscore the need for a better method for accurately identifying EL children at risk for academic difficulties.

This study had two purposes. The first purpose of this study was to determine if children at risk for academic difficulties within an EL sample reflect a discrete latent class and if the latent classes are stable over time. The second purpose was to determine the cognitive processes that can predict children at risk for learning disabilities within an EL sample. Traditionally, children have been defined as at risk by performing below a cut-off score point on standardized reading and/or math measures (e.g., Brandenburg et al, 2017; Branum-Martin, Fletcher, & Stuebing, 2013; Geary, Hoard, Nugent, & Bailey, 2012). For example, performance between the 16^{th} and 25^{th} percentile (85 to 90 standard score) on standardized norm-referenced reading and/or math measures is a common standard to identify children at risk for reading difficulties and/or math difficulties (e.g., Fuchs et al., 2006; Siegel & Ryan, 1989; Stanovich & Siegel, 1994), whereas scores below a standard deviation on norm-referenced measures (< 85 standard scores) across multiple years is a common designation for children at risk for learning disabilities in reading and/or math (e.g., Fuchs et al., 2006; Geary et al., 2012; Murphy, Mazzocco, Hanich, & Early, 2007; Swanson, Sáez, Gerber, Leftsted, 2004; Swanson, Olide, & Kong, 2018; Yeung, 2018). Multiple measures of reading and math are also included in the identification process because children with reading disabilities (RD) and math disabilities (MD) share similar processing difficulties and/or common dimensions that underlie their risk status (i.e., cognitive processes related to reading and math overlap [e.g., Swanson, 2020; Child et al., 2019; Mann, & Miller, 2013]).[1]

However, this selection process of determining EL children as at-risk has been criticized because of a reliance on arbitrary cut-off scores (e.g., Branum-Martin et al., 2013; Cirino, Fuchs, Elias, Powell, & Schumacher, 2015).[2] These cut-off standards have also been exacerbated when defining risk status among EL students because such children are not tested in their first language (e.g., Peña, Bedore, & Kester, 2016). This is because children who are learning two languages must attain a threshold in their first language before cognitive and academic skills can be assessed, as a child's mastery of their second language is a function of the child's linguistic competence in their first language (e.g., Cummins, 1979; however see Prevoo Malda, Mesman & van IJzendoorn, 2016, for a meta-analysis of this literature). An EL child's experience with the English language is far less rich than their monolingual English-speaking peers. Thus, achieving performance above cut-off scores on an English test comes with greater ease for a monolingual English speaking child.

Recent methodological advances may contribute to our understanding of children's reading and math proficiency, as it relates to children at risk who are also second language learners. These advances include modeling the development of discrete processes based on latent class analysis (e.g., Asparouhov & Muthén, 2014, 2015; Collins & Lanza, 2010; Muthén, 2006; Vermunt, 2007). Latent class analysis (LCA) is a statistical method used to identify subgroups of individuals characterized by similar multidimensional patterns of responses (e.g., Muthén & Muthén, 2012). A rationale for using latent class or mixture modeling is that although reading and/or math performance in children at risk for learning disabilities can be represented as a continuous outcome variable, the sample may be composed of different groups (or classes) of individuals (Swanson et al., 2016; 2018; Swanson, Kong, Petcu & Pimental, 2020). This group membership is not directly observed in continuous outcomes variables, even though the distribution of children's reading and/or math proficiency may reflect at least two different latent classes (e.g., children at risk and not at risk for academic difficulties). Thus, when focusing only on continuous variables such as reading and/or math achievement, the distributions of these latent classes are not observed, even though the continuous variable may reflect a mixture of distributions (see, Masyn, 2013; Morin, Bujacz, & Gagné, 2018 for a discussion).

The first purpose of this study is to test the notion that discrete latent classes or mixtures, representing different states of reading and math skills, exist in EL children who may be identified as at-risk or not at risk for RD and/or MD. An assessment of whether a mixture distribution exists can be judged via a measure of data-to-model fit (e.g., Asparouhov & Muthén, 2014; Nylund, Asparouhov, & Muthén, 2007). In addition, further analysis, via mixture modeling, allows for a determination as to whether the state of risk for RD and/or MD in EL children is extremely transient or stable. Latent Transition Analysis (LTA) can be used to test models involving sequential development and/or those individuals do not show typical change over time.

The second purpose of this study was to determine which cognitive processes can predict children at risk for RD and/or MD within an EL sample. Several studies suggest that children with potential learning difficulties in reading and/or math experience cognitive constraints, which impedes their ability to perform efficiently on achievement measures (e.g., Geary, Nicholas, Li, & Sun, 2017; Lesaux, Lipka, & Siegel, 2006; Swanson, Jerman & Zheng, 2008). Assuming a discrete subgroup of EL children at risk for learning disabilities in reading and/or math emerges in our study, it is important to know which of those cognitive processes are associated with these at-risk groups.

One of the most often referred to cognitive processes underlying both RD and MD is working memory (WM; Cowan, 2014; David, 2012; Peng et al., 2016a, 2016b, 2018; Swanson et al., 2008), which has also been related to achievement difficulties in ELs (e.g., de Abreu, 2011; de Abreu & Gathercole, 2012; Linck, Osthus, Koeth, & Bunting, 2013; Swanson, Sáez, & Gerber, 2006; Swanson, Orosco, & Lussier, 2015). These domain-general processes have been shown to contribute significant variance to reading and math achievement, even when domain-specific processes related to reading (e.g., phonological awareness; Swanson et al., 2015) and math (e.g., estimation, numeracy; Swanson, Kong, & Petcu, 2019a) are included in the regression modeling. Even though the association between WM and reading and/or math has been consistently established in the literature for children at risk for learning disabilities (e.g., Cowan, 2014; David, 2012; Peng et al., 2016a, 2016b, 2018; Swanson et al. 2008; 2015), the processes of WM that underlie predictions of reading and/or math performance are unclear (see Peng et al., 2016a, 2016b, 2018, for review). Some studies have suggested that the storage component of WM (referred to as verbal short-term memory, STM) plays a major role in academic performance (e.g., Peng et al., 2016b). Other studies have noted that academic difficulties are tied to the executive component of WM (e.g., Swanson et al., 2015; Swanson & Fung, 2016).

A model to capture individual differences among latent classes is Baddeley's multicomponent working memory model (2012; Baddeley & Logie, 1999). According to Baddeley's model (Baddeley & Logie, 1999), WM is comprised of a central executive controlling system that interacts with a set of two subsidiary storage systems: the visual-spatial sketchpad and the speech-based phonological loop. The central executive is involved in the control and regulation of the WM system. The visual-spatial sketchpad is responsible for the storage of visual-spatial information over brief periods of time and plays a key role in the generation and manipulation of mental images. The speech-based phonological loop is responsible for the temporary storage of verbal information; items are held within a phonological store of limited duration and are maintained within the store through a subvocal articulation process. The speech-based phonological loop is commonly associated with short-term memory (STM) because it involves two major components discussed in the STM literature: a speech-based phonological input store and a rehearsal process.

Given the above literature, we considered two models in accounting for the language and academic proficiency in EL children: one focuses on processing efficiency at a phonological level and the other focuses on executive processes (see Swanson et al.,2015; 2016 for further details on two models). The first model views children's reading and math skills as being related to processing difficulties at the phonological level (e.g., phonological storage of verbal information). A key mechanism that underlies phonological STM storage is naming speed. Naming speed has been considered a measure of how quickly items can be encoded and rehearsed within the STM system (referred to as the phonological loop, e.g., Bonifacci, Giombini, Bellocchi, & Contento; 2011; Georgio, Tziraki, Manolitis, & Fella, 2013; McDougall, Hulme, Ellis, & Monk, 1994). An alternative model suggests that cognitive processes separate from the phonological system, especially those related to the central executive controlling system of WM, play an equally important role. A key mechanism in this executive processing is the inhibition of the competing language system (e.g., Bialystok, 2011; Bonifacci et al., 2011). This is because inhibition has been associated with academic performance (e.g., Bonifacci et al., 2011; Lonigan et al., 2017) and WM (e.g., Friedman, Haberstick, Willcutt, Miyake, Young, & Hewitt, 2007), and therefore individual differences related to inhibition may play an important role in EL children's academic performance.

In summary, the present study attempted to broaden our understanding of risk status for learning disabilities in reading and/or math among EL children by considering the cognitive processes that may play a role in transitioning from one classification to another. To this end, three cohorts of children identified as English learners by their school districts were followed for two years. Children in grades 1, 2, and 3 were targeted in Wave 1 of the study (Year 1) and followed into Wave 2 (Year 2). To extend this literature, the study sought to answer three questions:

1. **Can a latent classification of children at risk for learning disabilities be identified among a heterogeneous sample of EL children?**

The present study determined the probability of identifying a latent class of participants at risk for learning disabilities using the 16th percentile (85 standard score) as a cut-off point within a sample that includes a test battery of reading, math, and cognitive abilities. This cut-off, 16th percentile, was considered a conservative cut-off point because it could capture performance below what is considered the average range in normative standard score distributions. As indicated by Barnes et al. (2020), "the 16th percentile falls 1*SD* below the mean, falls outside of the interquartile range, and represents a significant departure from the average range of achievement (p.689)". Also, LCA was used to assess probabilities, via modeling testing, to determine if some children who fell below the 16th percentile reflected a discrete latent. We assumed this latent class would follow patterns in achievement scores noted in the literature on English monolingual children with RD and/or MD.

Our predictions are based on studies finding that reading and math skills are comparable among EL and monolingual children at risk (e.g., Chiappe & Siegel, 2006; Lesaux et al., 2006). Thus, we predicted that a latent class of EL children who are at risk would emerge from the pool of ELs who are not at risk. A refinement in describing this sample includes establishing that such children's academic difficulties are not due to general intellectual difficulties and/or biased aptitude measures (e.g., Ferrer et al., 2010; Lohamn & Gambrell, 2012). Also, it is necessary to establish that risk status resides in the academic domain and not in one language (i.e., L1), per se. This refinement included in the sample of EL children at risk for learning disabilities that they perform above the cut-off scores (> 16th percentile) on vocabulary measures in L1 (i.e., Spanish).

Based on the assumption that that variations in language skills in English and Spanish would influence reading and math achievement (e.g., Cummins, 1979; however see Prevoo et al., 2016 for a critical review), we also predicted from our earlier work that at least two discrete latent classes would emerge that reflected balanced and unbalanced bilingualism (Swanon et al. 2004). Balanced bilingual children have comparable levels of competencies in both languages, whereas unbalanced bilinguals demonstrate notably higher proficiency in one language when compared to the other language. The phrase "unbalanced bilingual" refers to children with higher vocabulary in one language versus the other language (Cummins, 1976;1979).

2. Does membership in a latent class of EL children at risk change over time?

There have been few studies that have taken a discrete multidimensional approach to capture categorical differences in EL children at risk for serious academic problems across time (however, see Swanson et al., 2016). No studies we are aware of have used a multidimensional approach to address language acquisition, reading and mathematical difficulties in EL children in both English and Spanish. The majority of these studies focus on monolingual learners in the domain of reading and/or math or the testing of EL children in their second language (L2; e.g., Chiappe, Siegel, Wade-Woolley, 2002; Chiappe & Siegel, 2006; Farnia & Geva, 2019; Lesaux & Harris, 2017; Swanson et al., 2016). These studies have shown that some children who do not show signs of reading or language deficits in the first and second grade, do show them in the later grades (Catts, Compton, Tomblin, & Bridges, 2012; Compton, Fuchs, Fuchs, Elleman, & Gilbert, 2008; Farnia & Geva, 2019; Kieffer, 2011). Although these studies focus on reading exclusively, the incidence figure of late-emerging reading disabilities in monolingual children is estimated at approximately 13% (Catts et al., 2012), whereas this figure is substantially lower in EL children (< 1%, Swanson et al., 2016). These figures are difficult to compare because EL children in these studies where best classified as experiencing reading difficulties (cut-off scores for determining risk were in the normal range, i.e., between an 85 and 90 or 95 standard score) and not reading disabilities. More importantly, these studies did not consider difficulties in other academic domains, such as mathematics, and therefore the supposed late-emerging difficulties are isolated to reading may be incorrect. Thus, it is unclear in these studies as to whether late-emerging difficulties were related to reading measures specifically or reflect comorbid difficulties.

In the present study, we determine if the variables that predict latent classes at Wave 1 are stable or if the later classification of risk status primarily reflects the processes related to higher-order processing (comprehension and problem-solving). The literature has suggested that although EL children at risk for RD and/or MD make progress in acquiring basic reading and math skills (e.g., Kieffer & Thompson, 2018; Proctor, Carlo, August, & Snow, 2005), difficulties persist in higher-level processing such as reading comprehension and math word problem-solving (e.g., Lesaux & Harris, 2017; Swanson, Orosco, Lussier, Gerber & Guzman-Orth, 2011). Thus, the present study determined whether the classification of risk for learning disabilities was isolated to low scores in basic skills of reading and math (word recognition, computation) in the earlier testing wave and high order (reading comprehension and math problem-solving) difficulties in the later testing waves.

3. **Do specific cognitive measures predict latent class membership?**

The present study also determined if the cognitive processes related to language, reading, and math measures vary as a function of latent class (also see Swanson et al., 2020). For example, deficits in the phonological system (phonological storage) have been attributed to RD in English (e.g., Stanovich & Siegel, 1994) and Spanish (e.g., González & Valle, 2000). More recent studies have suggested that executive processes related to WM, are also significantly related to L2 reading, as well as and math performance (e.g., Swanson et al. 2015; 2019a,b). This executive system reflects controlled attention because the information to be recalled is presented in the context of competing information. Thus, we predict that processes related to the executive system (WM, inhibition) and/or the phonological storage system (STM, naming speed) play a unique role in predicting a latent class of children at risk for learning disabilities in reading and/or math. Thus, of interest is whether the influence of phonological processes on determining risk is more apparent in the younger grades, whereas processes related to the central executive controlling system of WM play an important role in predicting risk in the older grades. When applied to EL children, Swanson et al. (2004) found significant cross-language transfer in Grade 1 children on L1 STM and WM measures and L2 English word identification. Additional studies found that growth on measures of Spanish vocabulary, reading, STM, and WM accounted for significant variance in 1) predicting growth in English reading (e.g., Swansonet al. , 2006) and 2) growth on Spanish measures of naming speed, STM, and working memory predicted L2 math performance (e.g., Swanson et al., 2019b).

To the best of our knowledge, only a few studies we are aware of have addressed these aforementioned issues using latent class and latent transition analysis on an array of language, achievement and cognitive measures as in this study (Grimm, Solari, Gerber, Nylund-Gibson & Swanson, 2019; Swanson et al., 2016). Our earlier study (Swanson et al., 2016), although limited in focusing primarily on EL children reading difficulties, did find a distinct latent class of children at risk for reading difficulties at year 1 that were also exhibiting difficulties three years later. There was an extremely low incidence of children identified as good readers in wave 1, being identified as at risk for reading difficulties in the later testing waves. More critically, EL children at risk for reading difficulties in Wave 1 were unlikely to move out of the risk category in the later testing waves.

There are several limitations in this previous study that we hope to address. First, the previous study was limited by focusing on EL children at risk for reading difficulties as defined on English measures of reading. Also, the sampling of children was isolated to children selected from classrooms where reading instruction was presented in English only. If risk status for some EL children is related to learning disabilities, which in turn is assumed to have a biological base (see footnote 1), then it must be shown that the disabilities in reading/or math do not reside in just one language system. Second, the current study considers the classification of children at risk for learning disabilities based on performance on English and Spanish normed referenced measures of vocabulary, reading and mathematics. The previous study relied only on calculation scores based on norms for monolingual children and therefore we cannot determine if the risk factors related to reading difficulties were independent or comorbid with other achievement areas. Furthermore, the earlier study included only English and Spanish receptive vocabulary measures (Peabody Picture vocabulary tests), and tests of English and Spanish expressive vocabulary were not included in the classification of risk. Thus, determining whether at-risk status among EL children reflected comorbid achievement difficulties and/or expressive language difficulties and/or isolated deficits could not be determined. Also, because learning disabilities may be considered as a multifaceted process, there may be deficits in both language systems that are exhibited primarily on high order processes (math word problems, reading comprehension), rather than isolated to basic processes (calculation, word identification). Finally, the previous study did not analyze the transitional probabilities between certain grades, which usually are considered a pivotal point in determining later risk. For example, the "fourth-grade slump" (a drop in scores after grade 3) has been considered a risk factor for the later performance in reading (e.g., see Etmanskie, Partanen, & Siegel, 2016, for review).

In summary, the present study determines whether a latent class emerges related to risk for learning disabilities in reading and/or math risk among EL children and whether this latent class is transient or stable across testing periods. The classification of EL children at risk was based on performance below the 16th percentile on norm-referenced measures of reading, math, and vocabulary in both Spanish and English. A further refinement in the sample selection of children who are at risk includes: (a) establishing that such children's academic difficulties were not due to general language learning difficulties (e.g., Bishop & Snowling, 2004) or general intellectual difficulties (e.g., Ferrer, Shaywitz, Holonan, & Shaywitz, 2010) and (b) separating children at risk for RD and/or MD from children with comorbid difficulties, such as those with Attention-Deficit/Hyperactivity Disorder (ADHD; e.g., Boada, Willcutt, & Pennington, 2012; Snowling, 2012). Thus, measures of fluid intelligence and behavioral ratings of attention were also included in the classification battery. Measures of attention were included because children with ADHD are viewed as having primary processing difficulties that are distinct from children with learning disabilities (e.g., Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005). Further, children with learning disabilities are assumed to be in the normal range of intelligence and this represents a distinct category of children from those with general intellectual challenges (e.g., Giofrè, Toffalini, Altoè, & Cornoldi, 2017).

In general, we predict finding children at risk for RD and/or MD, children not at risk who were proficient in both languages (English and Spanish), and children not at risk who were more proficient in their second language than their first language (because their educational experience was in L2).

Method

Participants

This study was approved by the Institutional Review Board at the University of New Mexico, Project title:Math problem solving growth and working memory growth in English languge learners, IRB protocol: 995581-9. The study included a sequential-cohort design in which children in Grades 1,2,3 were assessed in Years 1 and 2, creating three cohorts (Cohort 1: Grades 1-2, Cohort 2: Grades 2-3, Cohort 3: Grades 3-4). Two hundred sixty-seven (*N*=267) students in grades 1 (*n* =118), 2 (*n* =90), and 3 (*n* =59) from two large urban school districts in the southwest United States participated in this study. [3] The sample was part of a four year federally funded longitudinal study assessing cognitive growth among EL children (Swanson et al., 2019a). Only those children who were tested in reading and math at Waves 1 and 2 in both English and Spanish were included in the analysis. One-hundred percent of the children in the sample participated in a full or reduced Federal lunch program and were drawn from neighborhoods with high Hispanic/Latino representation.[4] The children in this study were designated as ELs by their school and were selected from 31 elementary classrooms at Wave 1.

The first year sample included 131 boys and 136 girls who returned signed consent forms. School records indicated children's primary home language was Spanish (> 90%). All children were selected from dual language classrooms in which instruction was in both English and Spanish. No significant differences in gender representation emerged across the grades, χ^2(df=2, N= 267) = 4.12, p = .13.

Measures Used for Identifying Latent Classes

The study included group and individual administrations of a battery of tests. The series of tests were counterbalanced into one of four presentation orders. No Spanish and English versions of the same test (except for the Expressive One-word Picture Vocabulary Test, Spanish-Bilingual Edition; Brownell, 2001) were presented simultaneously. All participants were administered both English and Spanish versions of each measure by bilingual graduate students and staff researchers. Because the norm-referenced measures for establishing the latent classes are commercially available, along with information on their validity and reliability, they are only briefly reviewed here. Additional detail was provided below for the experimental cognitive measures.

Vocabulary: receptive and expressive. The Peabody Picture Vocabulary Test (PPVT) (Dunn & Dunn, 2007) was used as a measure of receptive vocabulary and was administered in English. In this task, children were presented with four pictures and were asked to select the picture that matched the word read aloud in English. The Test de Vocabulario en Imágenes Peabody (TVIP) was also administered**.** This measure is similar to the PPVT (Dunn & Dunn, 1981) in the presentation and administration, except that words were read aloud in Spanish (Dunn, Lugo, Padilla & Dunn, 1986). The Expressive One-Word Picture Vocabulary Test - Spanish-Bilingual Edition (EOWPVT-SBE: Brownell, 2001) was used as a measure of English and Spanish expressive vocabulary. The sample KR_{20} (reliability) for the receptive and expressive vocabulary measures at Wave 1 were .97 and .96 for English, and .91 and .95 for Spanish measures, respectively. The sample KR_{20} for the receptive and expressive vocabulary measures at Wave 1 were .96 and .96 for English, and .93 and .95 for Spanish measures, respectively.

Reading: word identification and passage comprehension. The Woodcock-Muñoz Language Survey-Revised (WMLS-R) established a norm-referenced reading level in English and Spanish (Woodcock, Muñoz-Sandoval & Alvarado, 2005). The WMLS-R Spanish and English Word Identification and Passage Comprehension subtests were administered. The sample KR_{20} for the word identification and comprehension subtests at Wave 1 were .95 and .86 for English, and .90 and .87 for Spanish measures, respectively. The sample KR_{20} for the word identification and comprehension subtests at Wave 2 were .92 and .82 for English, and .92 and .87 for Spanish measures, respectively.

Math: calculation and word problem-solving. The Calculation and Applied Math Problem-Solving subtests from the Woodcock-Johnson III (Woodcock, McGrew, Schrank, & Mather, 2007) were administered for the English presentation. The Calculation and Problemas Aplicados subtests from the Batería III Woodcock-Muñoz (Muñoz-Sandoval, Woodcock, McGrew, & Mather, 2005) were administered to establish norm-referenced math levels in Spanish. Both of these subtests were individually administered and assessed children’s early mathematical operations (e.g., counting, addition, and subtraction) through practical problems. The sample KR_{20} for the Calculation and Applied Problems subtests at Wave 1 were .79 and .81 for English, and .71 and .87 for Spanish measures, respectively. The sample KR_{20} for the

Calculation and Applied Problems at Wave 2 subtests were .79 and .81 for English, and .71 and .87 for Spanish measures, respectively.

Fluid intelligence and attention. Fluid intelligence was assessed by administering the Raven Colored Progressive Matrices test (RCMT, Raven, 1976). The RCMT is commonly used to tap fluid intelligence because of its brevity in administration, as well as its high correlation with other nonverbal intelligence measures that are assumed to tap reasoning, thinking, or the ability to acquire new knowledge (referred to as Fluid Intelligence). The sample KR_{20} was .96 for Wave 1, and .80 for Wave 2. Attention was assessed by administering the Conners' Teacher Ratings Scales-Revised: Short Form (CTRS–R:S; Conners, 1997). The CTRS-R:S is used to evaluate students' problem behaviors by obtaining ratings from teachers. The homeroom teacher was selected for each child and was asked to complete the CTRS–R:S. The primary measure for this study was the Attention-Deficit/Hyperactivity index.

Cognitive Measures Used for Determining Correlates of Latent Class Membership

The cognitive measures assumed to be related to the latent classification assessed the storage of phonological information (short-term memory and naming speed) and executive processing (central executive of working memory, visual-spatial working memory, and inhibition). The convergence of the measures for the English and Spanish versions was established in an earlier study (see Swanson et al., 2015; 2019a, for further discussion), and a full description of each cognitive measure is provided in Swanson et al. (2015, 2019a).

Phonological storage. Three short-term memory (STM) tasks were administered to capture the storage of phonological information: the forward digit span task, the word span task, and the phonetic memory span task. To assess the children's forward digit span task, The Forward Digit Span subtest of the Wechsler Intelligence Scale for Children-Third Edition

(WISC-III; Wechsler, 1991) was adopted. More specifically, this subtest assessed children's STM because it was assumed that forward digit spans involved a subsidiary memory system (the phonological loop). The Word Span task was previously used by Swanson et al. (2004) and assessed the children's ability to recall increasingly large word lists (a minimum of two words to a maximum of eight words). The Phonetic Memory Span task assessed the children's ability to recall increasingly large lists of nonsense words (e.g., des, seeg, seg, geez, deez, dez) ranging from two to seven words per list. The sample KR_{20} for Digit Span, Word Span, and Phonetic Span at Wave 1 were .54, .80, and .71 for English, and .66, .80, and .77 for Spanish measures, respectively. The sample KR_{20} for Digit Span, Word Span, and Phonetic Span at Wave 2 were .50, .73, and .69 for English, and .35, .74, and .68 for Spanish measures, respectively.

Naming speed. The Comprehensive Test of Phonological Processing's (CTOPP; Wagner, Torgesen, & Rashotte, 2000) Rapid Letter Naming and Rapid Digit subtests were administered to assess speed in recalling numbers and letters in an English and Spanish version. The sample Cronbach alpha reliabilities for the letters and numbers subtests at Wave 1 were .95, .94 for English measures, and .95 and .92 for Spanish measures, respectively. The sample Cronbach alpha reliabilities for the letters and numbers subtests at Wave 1 were .92, .89 for English, and .91 and .88 for Spanish measures, respectively.

Executive processing. Three complex span measures (tasks that included both a process and storage question) and an updating task were administered. The Conceptual Span, Listening Sentence Span, Digit Sentence Span, and Updating task were administered in English and Spanish to capture the executive component of WM (tasks described in detail in Swansonr et al., 2015). The WM tasks required children to hold increasingly complex information in memory, while simultaneously responding to a question about the task. Because WM tasks were assumed to tap a measure of controlled attention referred to as updating, an experimental updating task was also administered. The sample KR_{20} for Conceptual Span, Listening Span, Digit Sentence Span, and Updating were .84, .88, .66, .83 for English, and .82, .88, .61 and .76 for Spanish measures, respectively. The sample KR_{20} for Conceptual Span, Listening Span, Digit Sentence Span, and Updating at Wave 2 were .66, .80, .56, .87 for English and .62, .75, .40 and .85 for Spanish measures, respectively.

Visual-spatial working memory. This component of WM was measured using two tasks (see Swanson et al., 2004 for review of these tasks). The Visual Matrix task assessed the participants' ability to remember visual sequences within a matrix. The Mapping and Directions Span task assessed whether the children could recall a visual-spatial sequence of directions on a map with no labels. The sample KR_{20} for the Visual Matrix and Mapping/Directions measures at Wave 1 were .96 and .79, respectively. The sample KR_{20} for the Visual Matrix and Mapping/Directions measures at Wave 2 were .96 and .79, respectively. The instructions for these measures were provided in both English and Spanish, but the measure was only administered once.

Inhibition. The Random Number and Random Letter Generation Tasks were administered to assess inhibition. Children were first asked to write, as quickly as possible, numbers (or letters) in a non-random sequential order to establish a baseline. They were then asked to write numbers as quickly as possible, out of order, in 30 seconds. Scoring included an index for randomness, information redundancy, and percentage of paired responses to assess the tendency of participants to suppress response repetitions. The sample KR_{20} for English letter and number measures at Wave 1 were .73 and .76, and for Spanish letters and numbers were .85 and

.74, respectively. The sample KR_{20} for English letter and number measures at Wave 2 were .68 and .69, and for Spanish letters and numbers were .77 and .73, respectively.

Cut-off Point

The manifest variables (vocabulary, reading, math, fluid intelligence, and attention) used to determine discrete groups were dummy coded as reflecting a normative score at or below the 16th percentile (1 = at or below the 16th percentile, 2 = above the 16th percentile). The 16th percentile (an 85 standard score) was based on the normative data in the test manual from the standardized vocabulary, math, reading, and fluid intelligence measures. The Conners scale was in T-scores, with high scores representing higher levels of inattention. Therefore, the 16^{th} percentile on the Conners scale was a T-score of 60.

Procedures

Ten bilingual graduate students or research assistants trained in test administration tested all participants in their schools. Each child was tested individually and in small groups of 10 to 15 students after informed consent was obtained for participation. For each testing wave, children participated in two sessions of individual testing that lasted 30 to 60 minutes and two group-testing sessions that lasted approximately 60 minutes and occurred for two consecutive days. One of six presentation orders, related to the individually administered tasks (WM, STM, phonological processing, and reading), was randomly assigned to each child to control for order effects. Also, the presentation orders of Spanish and English tests were counterbalanced across all children. For the group-administered tests, the presentation order of English and Spanish measures for each type of task were also counterbalanced across small groups.

Statistical Analysis

Latent class analysis. Before computing the estimations for the LTA (e.g., Collins & Lanza, 2010), latent class analysis (LCA) was conducted at each time period to evaluate the model fit. Also, because LCA is an exploratory analysis, a series of models were computed varying the number of latent classes between one and seven (Nylund et al., 2007; see Masyn, 2013, for a comprehensive review). A combination of statistical indicators and substantive theory were used to decide on the best fitting model. Statistical model comparisons included likelihood ratio tests: the Lo-Mendell-Rubin Test (LMR) and the Bootstrap Likelihood Ratio Test (BLRT). Both statistical procedures compared the improvement between neighboring class models (i.e., comparing models with three vs. four classes, and four vs. five, etc.) and provided *p*-values. *P*-values were used to determine if there was a statistically significant improvement in fit for the inclusion of one more latent class. A non-significant *p*-value for a K-class indicated that the previous K-class with a significant *p*-value fit the data better. The models with different numbers were compared using information criteria (i.e., Bayesian Information Criteria-BIC, Akaike Information Criteria-AIC, and Adjusted BIC). Lower values on these fit statistics indicated a better model fit. Among the information criterion measures, the Adjusted BIC and BIC are generally preferred, as is the BLRT for statistical model comparisons (Nylund et al., 2007; Nylund-Gibson & Choi, 2018). It is important to note, as indicated by Nylund-Gibson and Choi (2018) "fit indices often do not all point to a single solution, thus the recommended procedure for exploring and deciding on the number of classes is to jointly consider statistical fit indices, substantive interpretability and utility, and classification diagnostics, which help to illuminate how well the classes are classifying and differentiating among the individuals considered "(p. 443). We primarily used SAS (Lanza, Dziak, Huang, Xu & Collins, 2011) software to examine the manifest variables and determine the number of latent classes as well as perform the multilevel logistic modeling (see below). The Mplus (Muthén & Muthén, 2012) software was used to compute the likelihood ratio using the Lo-Mendell-Rubin Test (LMR) and the Bootstrap Likelihood Ratio Test (BLRT).

Nestedness. At least 50 level-2 observations (classrooms in this case) are needed to assure that estimated parameters are unbiased (Maas & Hox, 2005). The number of clusters in our sample (i.e., 31 classes at Waves 1 and 2,) was not sufficient to use in a multilevel latent class analysis (LCA). However, because Wave 1 yielded the largest number of clusters, we conducted a two-level latent class analysis with the categorical variables (see Muthén and Muthén's Mplus User's Guide 1998-2010 in 2012, pages 360 to 366). In this situation, it is recommended to consider entropy values when there is a multilevel structure to the data (see Kaplan & Keller, 2011, p. 53) because the preferred index (Bayesian Information Criterion, BIC) may underestimate the number of classes.

Latent Transition Analysis. Latent transition analysis (LTA) was used after the optimal number of LCA had been selected at each time point. The analysis utilized the PROC LTA procedure in SAS version 9.3 (Lanza et al., 2011). A three-step method was used to determine the LTA model (e.g., Asparouhov & Muthén, 2014; Morin, Bujacz, & Gagne, 2018; Nylund-Gibson, Grimm, Quirk, & Furlong, 2014). The first step examined whether the same number of latent classes could be identified across two testing waves (i.e., configural similarity). Several authors have suggested fitting the covariate after determining latent models (Asparouhov & Muthén, 2014) so as not to alter the model. Grade level and gender were entered into the analysis when predicting latent classes in the subsequent logistic models. The second step integrated the two retained solutions (one at each time point) in a single LTA model, allowing for the estimation of transition probability solutions estimated across the two testing waves. We constrained the latent classes to be equal across the two test waves to examine the transitioning from one class to another over time. In applications of LTA, full measurement invariance is assumed for practical reasons, because it ensures that the number and structure of classes are the same across time and allows for a straightforward interpretation of transition probabilities. From these analyses, the following sets of parameters were estimated: latent class membership probabilities at Time 1 (δ delta parameters), probabilities of transitions between latent classes over time (tau τ parameters), and item response probabilities conditional on latent class membership and time (ρ rho). Because of missing data on some of the measures (i.e., Fluid intelligence), latent class models were specified using the full information maximum likelihood estimation procedure (see Collins & Lanza, 2010, for a rationale). Multiple LTA models were conducted to identify an adequate model fit of the data.

Multilevel logistic model. A multilevel logistic model, via SAS PROC GLIMMIX software (SAS Institute, 2010), was used to analyze cognitive differences between latent classes. The reference group was the latent class comprised of average achievers in reading and math. Of interest was whether the odds of being identified as being within a particular latent class, when compared to average achievers, increased as a function of performance on the cognitive measures. Cognitive measures were reduced to latent constructs based on an earlier study (Swanson et al., 2019a). Latent scores were computed by multiplying the z-score of the target variable by the standardized factor, loading weight based on the total sample (see Nunnally & Bernstein, 1994, p. 508, for calculation procedures). Latent variables were specified as indicators of speed (naming speed for numbers and letters), inhibition (random generation of numbers and letters), STM (Digit Forward Span, Word Span, and Phonetic Span), executive processing (Conceptual Span, Listening Span, Digit Sentence Span, Updating), and visual-spatial WM (Visual Matrix, Mapping & Directions).

The equation for estimating the LCs for the unconditional model was:

$$\eta_{ij} = \beta_{0j} + \beta_{1j} X_{ij} \quad \text{(Eq. 1)}$$

Equation 1 represented a simple level-1 model with one student-level predictor, where η_{ij} represented the log odds of reflecting a latent class other than an average achiever for student i in classroom j, β_{0j} is the intercept or the average log odds of being designated at risk in classroom j, ij is a student-level predictor for student i in classroom j, and β_{1j} represents the slope associated with X_{ij}, showing the relationship between the student-level variable and the log odds of not being designated at risk. It is important to note that unlike the hierarchical linear model used to analyze the total sample, this model has no error variance at level-1 (see Snijders & Bosker, 1999; pp. 225-227). As the effect of the student-level predictor was modeled as fixed or constant across classrooms, this was represented as a random intercept-only model.

$$\eta_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} W_j + u_{0j} \quad \text{(Eq. 2)}$$

The combined level-1 and level-2 model represented the log odds of being designated as a latent class at risk for student i in classroom j (η_{ij}) at a typical classroom ($\gamma 00$), at the student-level ($\gamma 10\ X$ij) and classroom-level predictor ($\gamma 01 W$j), as well as the classroom-level error [$u0$j, $u0$j ~ $N(0,\tau_{00})$].

Results

Classification Variables

The means and standard deviations for the classification variables for Waves 1 and 2, as a function of the total sample and the three cohorts, are reported in Table 1. Also reported are difference scores (Wave 2 minus Wave 1 performance) on the classification measures. As shown in Table 1, the negative difference scores showed that some measures (e.g., reading) decreased in normative values from Waves 1 to 2. Substantial declines (negative difference score > - 5.00) on normative tests in the total sample occurred on measures of English and Spanish math and measures of Spanish reading comprehension. The manifest variables that showed substantial declines on the normed manifest variables (negative difference score > - 5.00) as a function of each cohort were as follows: English and Spanish reading for Cohort 1 (grades 1 to 2), Spanish receptive vocabulary for Cohort 2 (grades 2 to 3), and English receptive vocabulary for Cohort 3 (grades 3 to 4).

Latent Class Analysis

Model fit. Prior to reporting our results related to LTA, LCA models were computed on the total sample for Waves 1 and 2. Of the indices reported in Table 2, the BIC and the adjusted sample BIC is generally preferred, as is the BLRT for statistical model comparisons (Nylund et al., 2007). The BIC was lower for the LC4 model than the LC5 model and LC4 yielded adequate sample proportionality and item probabilities that were more easily interpreted. An additional consideration was the interpretability of the classes, as well as the size of the smallest class. The lowest adjusted BIC values emerged for the LC4 model. Both the LMR and BLRT yielded non-significant *p*-values for the solution of the latent class five (LC5) model, and significant *p*-values for the solutions of three latent class models, indicating that the LC4 model provided an excellent fit to the data. Thus, the LC5 model did not represent an improvement of the LC4 model, and the LC4 model was an improvement to the LC3 model. The BIC and Sample adjusted BIC values are shown in bold in Table 2.

As shown in Table 2, the entropy value (Wave 1) for the four latent class models was .77 and .85 for Wave 2. These two values are considered an acceptable value (Nylund et al., 2007).

In terms of sample proportion, the four-class model represented a relatively large number of participants (e.g., proportionally was greater than 5%) within each latent class. Masyn (2013) suggested that class proportion values can be considered (i.e., "assign meaning to the classes," p. 559) when determining the number of latent classes. Consistent with Nylund-Gibson and Choi (2018) we also inspected the "elbow" of point of "diminishing returns" in determining the model fit (e.g., small decreases in the IC (information criteria) for each additional latent class) (p. 443-444). Thus, for classification purposes, as well as our predictions, the four latent classes were considered appropriate for the interpretation of the data.

Transition Analysis

Transition probabilities were estimated for the total sample to identify the rate of change or stability over time for the latent class groups. Although the LC4 model was supported at Waves 1 and 2 separately (establishing configurable similarity), we determined if the LC4 model in LTA was a good fit for a transition model. A measurement variance model was computed that yielded BIC values of 3927.53, 3958.01, and 4092.20 for the LC3, LC4, and LC5 models in LTA, respectively. Because of the elbow between the LC3 and LC4 models, one could argue that the LC3 model provided a more parsimonious fit than the LC4 model. However, when we examined moving from the LC4 to the LC3 model, we saw that some classes (LC2 and LC4) were being combined, while the classes in the middle of the overall data (.00) distribution (i.e., balanced bilingual-average achievers) remained almost unchanged. This is not an unusual finding. As stated by Ryoo and colleagues (2018), "model parsimony is not necessarily the goal of LTA. The goal of LTA, particularly when it is used in an exploratory vein, is to describe the data by identifying classes within the population of observations in the data (p.31)." Thus, our selection of the LC4 model provided a more nuanced description of the data.

To further analyze class stability, however, it is useful to constrain each element of the matrix of ρ parameters at Wave 1 to be equal to its corresponding element at Wave 2, which has the effect of imposing measurement invariance across time (Kam, Morin, Meyer, & Topolnytsky, 2016; Nylund-Gibson et al., 2014). This analysis forced the change process to be stationary (i.e., forces individuals transitioning between classes to have the same probability level of change across the two-time points). This comparison allowed us to check whether the latent classes have the same meaning across the two testing waves and three cohorts. The objective of researching measurement invariance is to find the lowest level of "inequivalence" possible that fits the data well.

The left side of Table 3 shows the transition probabilities of the latent classes (τ parameters) for the transitional model. The probabilities reflected membership in the same latent class model at two consecutive measurement times. Transition probabilities off the diagonal reflected the likelihood of one latent class status group at Wave 1 transitioning to a different latent class status group at Wave 2. For example, for the total sample, the LC1 status group (balanced bilingual-average achievers) at Wave 1 yielded an 87% estimated chance of maintaining their status at Wave 2. In contrast, the LC2 status group (unbalanced bilingual-average achievers) at Wave 1 had only a 46% chance of being in the same group at Wave 2, suggesting that the item estimates were highly unstable. In contrast, the LC3 status group (at risk for learning disabilities) and the LC4 status group (English dominant) estimates were stable at Waves 1 and 2 (1.0 and 95%, respectively).

To unpack the transitions that may be occurring across cohorts, the findings were further analyzed by dividing the analysis into the grade-level cohorts. As shown on the left side of Table 3, the LC2 status group (unbalanced bilingual-average achievers) classification was the least stable for Cohort 3 (grades 3 to 4, 6%) with approximately 94% of the sample falling into the LC4 status group (English dominant) at Wave 2. Those children who were in the LC1 group (balanced bilingual-average achievers (LC1) that were most likely to fit the category of ELs at risk for late learning disabilities was Cohort 2 (grade 2 to grade 3, 24%) and Cohort 3 (grade 3 to grade 4, 15%).

The right side of Table 3 also shows the membership probabilities (proportion of the sample of a particular latent class at Waves 1 and 2). The largest proportional representation of the total sample occurred for children in the LC1 status group (balanced bilingual-average achievers) and the LC2 status group (unbalanced bilingual-average achievers) at Wave 1, and the LC1 and LC4 (English dominant) status groups at Wave 2. The results show that membership probabilities increased dramatically from time 1 (Wave 1) to time 2 (Wave 2) for the LC4 status group.

Labeling of latent class groups. The means and standard deviations on the manifest variables as a function of the four latent class status groups at each time point are reported in Appendix A. As shown in Appendix A, there were 14 manifest variables for both Waves 1 and 2. To simplify the labeling of the latent class groups, the frequency of standard scores at or below 85 were identified. None of the mean scores for Waves 1 or 2 were at or below an 85 standard score for the LC1 status group. Thus, since English and Spanish vocabulary, math, and reading were in the average range, this group was tentatively labeled as balanced bilingual-average achievers. Except for Spanish expressive language, all mean scores were at or above 85 for the LC2 status group. Both the LC1 and LC2 status groups were designated as average achievers since reading and math scores were in the average range. This later group (LC2) was labeled as unbalanced bilingual-average achievers because their expressive Spanish vocabulary was below 85. As shown in Appendix A, the normative achievement and vocabulary scores for the LC1 status group were in the average range (85 to 115 standard score). Thus, we maintained the label for the LC1 status group as balanced bilingual-average achievers. In contrast, we labeled the LC2 status group as an unbalanced bilingual-average achievers because their English vocabulary scores exceeded their Spanish vocabulary scores, but achievement scores (reading and math) in both English and Spanish were in the average range. That is, their achievement scores were in the same average range as the LC1 status group, but exceeded the LC1 status group in English vocabulary.

For the LC3 status group, six of 14 scores at Wave 1 and seven of 14 scores at Wave 2 were at or below a standard score of 85. Low scores (at or below 85) emerged on English and Spanish measures of vocabulary, math, and reading. Fluid intelligence and attention scores were in the normal range. Because the academic performance was low across both language systems, but fluid intelligence was in the average range, the LC3 status group was labeled as at risk for learning disabilities. For the LC4 status group, two of 14 scores at Wave 1 and three of 14 scores at Wave 2 were in the low range. The low scores were on Spanish measures of vocabulary and reading comprehension. Because scores were primarily low in Spanish vocabulary and reading comprehension, the LC4 status group was labeled as English dominant. However, because these descriptions merely reflected cut-off scores on normative measures, a further focus considered those items that yielded high probabilities of reflecting each latent class.

Magnitude differences. To facilitate the labeling of four latent class status groups, the magnitude of differences on the manifest variables as a function of latent class status groups was computed. Using Cohen's (1988) criteria for medium (.50) to large (.80) effect sizes (ESs), ESs at or greater than .50 are shown in bold for performance at Waves 1 and 2 in Table 4. The higher performance of the LC1 status group (balanced bilingual-average achievers) when compared to the LC2 status group (unbalanced bilingual-average achievers) at Wave 1 were on measures of Spanish vocabulary, Spanish math problem solving and Spanish reading comprehension. In contrast, the LC2 status group yielded higher performance relative to the LC1 status group on measures of English vocabulary at Wave 1 and English vocabulary, English word identification, and fluid intelligence. In general, the differences between the LC1 and LC2 status groups were that the LC1 status group yielded higher performance on Spanish measures, whereas the LC2 status group yielded higher performance on English measures.

Relative to average achievers (the LC1 and LC2 status groups), the latent class model that appeared to yield low achievement scores in both languages, or who were academically at risk, was the LC3 status group (at risk for learning disabilities). The LC1 status group exceeded (ESs > .50) performance of the LC3 status group on all 14 measures at Waves 1 and 2. Except for performance on Spanish vocabulary measures, the LC2 status group exceeded the LC3 status group on all measures at Waves 1 and 2. Also, the LC4 status group (English dominant) exceeded the LC3 group on nine of 14 measures at Wave 1 and 11 of 14 measures at Wave 2.

Likewise, the LC1 status group exceeded (ES s > .50) the LC4 status group on four of the 14 measures at Wave 1 and five of the 14 measures at Wave 2. In contrast, the LC4 status group exceeded the LC1 status group on measures of English vocabulary at Wave 1 and English vocabulary, classroom attention rating and fluid intelligence at Wave 2. The LC2 status group exceeded the LC4 status group on measures of Spanish calculation and Spanish reading at Wave 1, as well as measures of Spanish vocabulary, Spanish calculation, Spanish reading, and classroom attention at Wave 2.

Taken together, we considered the LC3 status group as children at potential risk for learning difficulties. While the majority of comparisons between the latent class groups revealed language-specific differences (Spanish higher performance for the LC1 status group and English higher performance for the LC2 and LC4 groups), the LC3 status group revealed both English and Spanish differences on measures of vocabulary, math, and reading. However, not all measures played a major role in predicting latent class membership. The next analysis computed probability estimates for each of the manifest variables that were conditional on latent class membership.

Item probabilities. Item probabilities based on the LC4 transition model are shown in Table 5. Shown are the probabilities (rho estimates) for performance *below* the cut-off threshold of the 16th percentile (85 standard score) on the manifest variables. These rho estimates reflected the latent class abilities of the given item-response, conditional on the given latent-class membership. To facilitate discussion, and because there is no set standard for determining meaningful probabilities, item latent class abilities above 60% were selected and these values are shown in bold. That is, probabilities above .60 indicated "at-risk" status for that particular manifest variable. The results show that all of the parameters of vocabulary, reading, math and attention for the LC1 status group were below .50. Thus, consistent with the previous descriptive analysis, the LC1 status group was labeled as a balanced bilingual average achiever. Except for Spanish expressive language, all of the parameters for the LC2 status group were below .50. For lack of a better term, this group was again labeled as unbalanced, or English dominant (stronger English than Spanish vocabulary), bilingual-average achievers. The LC3 status group showed high risk on measures of English language, English problem solving, and English reading comprehension; the LC4 status group showed low Spanish reading comprehension and language.

The LC3 status group was again labeled as at risk for learning disabilities, and the LC4 group was labeled as English dominant (relatively lower Spanish proficiency).

Correlates of Latent Classes

Demographic information. Appendix A shows the distribution of the sample for gender and cohort representation as a function of latent class status. Significant differences emerged as a function of the latent class status group on gender representation, χ^2(df=2, *N*= 267) = 10.40, *p* = .014, and cohort representation, χ^2(df=2, *N*= 267) = 36.93, *p* < .001. As shown in Appendix A, a higher representation of males emerged in the LC3 status group when compared to the other latent class status groups. The results also show that a high representation for Cohort 1 (grade 1 to 2) occurred in the LC1 and LC2 status groups relative to the other cohorts.

Latent measures of cognition. Further analysis determined those variables external to the classification measures that played a significant role in predicting latent class membership at Wave 2. We assumed that separate cognitive processes would predict latent class membership. Because we were not interested in the variance related to individual cognitive tasks, but what was common amongst the observed variables, as well as controlling for measurement error and enhanced reliability, latent measures served as predictors in the analysis. Previous analyses tested the categorization of the variables (i.e., working memory [WM], short-term memory[STM], naming speed, and inhibition) and provided a good fit to the data in a previous study (Swanson et al., 2015; 2019a). The confirmatory factor model was computed for measurement purposes (latent variables control for measurement error as different variables have different weightings on a construct) and also for practical reasons: some constructs (e.g., WM, STM) included several tasks. Overall, the factors included STM (span=nonword or pseudoword, real words, digit forward, digit backward), naming speed (letters, numbers), the executive component of WM (conceptual span, sentence span, listening span, updating), and inhibition (random generation= letters and numbers). For the present study, the model provided an acceptable fit to the data at Wave 1 (CFI = .92, RMSEA = .044; 90% CI: .039 to .048). Because we focused on two language systems (English and Spanish), measures that tapped the same constructs, but were assessed in both English and Spanish, were not combined into a single factor (however, see Swanson, Orosco, & Lussier,., 2012, for a second-order analysis). This separation by language system was maintained to assess whether bilingual proficiency may play an important role in latent class status. Based on the loadings, scores were computed by multiplying the *z*-score of the target variable by the standardized coefficients based on the total sample at Wave 1. The mean z-scores related to each factor as a function of the four latent class models are reported in Appendix A.

Effect sizes. To facilitate the interpretation of the LC status group differences on the English (E-) and Spanish (S-) cognitive measures, ESs of moderate and high magnitude (ESs > .50) were again considered meaningful. Table 6 shows ES comparisons on 18 (nine at Wave 1 and nine at Wave 2) cognitive measures (E-STM, S-STM, E-Speed, S-speed, E-Inhibition, S-Inhibition, E-Executive WM, S-Executive WM, & visual-WM) as a function of LC. Of particular interest was identifying the cognitive measures that separated children at risk for learning disabilities (the LC3 status group) from the other latent classes.

As shown in Table 6, at Wave 1 the LC3 status group (at risk for learning disabilities) group yielded low performance (absolute ES value .50 or greater) relative to the LC1 status group (balanced bilingual-average achievers) and the LC2 status group (unbalanced bilingual-average achievers) on English measures of STM, Speed and WM. At Wave 2 the LC3 status group yielded low performance relative to the LC1 and LC2 status groups on measures of

English and Spanish measures of STM and WM. Relative to the LC4 status group (English dominant), the LC3 status group yielded low performance on measures of English STM, English naming speed, and English WM at Waves 1 and 2.

The results in Table 6 also show that no meaningful ESs (ES > .50) emerged between the LC2 and LC4 status groups. These findings may account for the higher representation of the LC2 status group among the younger cohorts (Cohort 1) and the decreasing representation among the older cohorts (Cohorts 2 and 3). The ES findings, as well as the cross-sectional patterns in the data, suggested that as children increased in age decreases in Spanish reading comprehension occurred.

Generalized Linear Polytomous Model

A multilevel logistic model determined the cognitive variables that uniquely predicted the LC status groups at Wave 2. The LC4 status group (English dominant) was the reference group. The estimates for the multilevel logistic unconditional and conditional models in predicting the odds of being classified in one of the latent classes, when compared to the LC4 status group as a function of variables external to the classification (cognitive variables), are shown in Table 7. Thus, the three intercept values for the LC1 (balanced bilingual-average achievers), LC2 (unbalanced-average achievers), and LC3 (at risk for learning disabilities) status groups shown in Table 7 are the status group comparisons to the LC4 status group.

Table 7 shows the intercepts for the unconditional means model. The unconditional model was assumed to have no error at level-1 (Snijders & Bosker, 1999). That is, the level-1 residual follows a logistic distribution with a mean of 0 and a variance of 3.29 (Snijders & Bosker, 1999, p. 227). Thus, only the intercept variance is reported. The cross-sectional intercept for both Waves 1 and 2 would not converge. The intercept variance for wave 1 was .94 and .80

for wave 2. Thus, we used the Wave 1 intercept variance because it yielded a larger intraclass correlation than Wave 2. The intraclass correlation was computed as .22 (.94/.94 + 3.29) for Wave 1, suggesting that approximately 22% of the variability was accounted for by children nested in classrooms at Wave 1, leaving approximately 78% of the variability to be accounted for by the latent measures (or other unknown factors). As shown in Table 7, the three intercepts for the unconditional model indicated a significant amount of variability in the log odds of being classified as one of the two latent class models relative to the LC4 model. As shown, the LC3 model was significantly lower in intercept values when compared to the LC4 model, whereas the LC1 model yielded significantly higher intercepts than the LC4 model. No differences were found between the LC2 and LC4 models in the unconditional means model.

Table 7 also shows a comparison of five conditional models (i.e., Models 1, 2, 3, 4, and 5) when entering the cognitive variables into predictions of LC status. Model 1 tested whether gender and cohort would eliminate the intercept differences between latent classes. Model 2 tested whether the log-odds were significant when predictor variables in Wave 1 for cognitive measures were entered into the model. Model 3 tested whether the log-odds were improved when a reduced model was entered into the analysis. The reduced model entered only those variables found significant in the full model (Model 2). Models 3 and 4 entered only variables tested at Wave 2. These two models (Models 3 and 4) followed the same pattern of entry as Models 1 and 2 (full model followed by the reduced model).

For all the conditional models, we have three estimates of the intercept but only one slope associated with the covariates. Thus, the cognitive covariates remained *constant* across the logits/intercepts within each model. This allowed for the interpretation that the increase in log-odds of falling into a latent class (e.g., LC1, LC2, or LC3) versus the LC4 resulted from a one-unit increase in the covariate, holding the other covariates constant across all intercepts. Holding the covariates constant was also done since no predictions were made related to covariates interacting with the intercepts. Thus, the covariates estimated the relationship between all of the intercepts in the log odds of being within a particular latent class, relative to the reference category (LC4).

Model 1 entered measures of gender and cohort status. Both gender and cohort (grade level) variables played a significant role in predictions of latent class. However, none of these significant covariates eliminated intercept differences. In contrast to the unconditional model, a significant intercept advantage emerged for LC2 when compared to LC4, indicating that changes in gender and grade representation influenced the predictions of latent class status.

Model 2 entered Wave 1 cognitive measures into the model. Three intercepts were significant, as were the measures of Spanish naming speed and English and Spanish WM. The reduced model (Model 3) entered only cognitive measures found significant in the full model. When compared to Model 2, Model 3 provided a parsimonious fit to the data. That is, Model 2 yielded a less parsimonious fit to the data (Deviance=639.83) than Model 3 (Deviance=649.51), χ^2 (df =9) =9.69 (649.51 – 639.83), $p > .05$. Also, the AIC and BIC indices were lower for the reduced than full model, suggesting that Model 3 had a better fit to the data.

Model 4 entered all Wave 2 cognitive variables. Two intercept values for LC1 and LC3 were significant, suggesting that the students in the LC4 underperformed the students in the LC1, but outperformed the LC3 status group. No intercept differences emerged between LC2 and LC4. The significant covariates in Model 4 were measures of gender, grade level, English and Spanish naming speed, Spanish WM, and the interaction for the inhibition measures. The significant covariates were entered into the reduced model (Model 5). Model 5 (reduced model), provided a more parsimonious fit to the data when compared to Model 4, χ^2 (df =9) =11.47(628.84 – 617.37), $p > .05$. When comparing the fit indices (AIC and BIC), values were lower for the Wave 2 than the Wave 1 model.

In summary, the results suggest that the executive component of WM and naming speed played an important role in contributing unique variance to intercept differences among the latent classes. As shown across the models for the covariates at Waves 1 and 2, an increase in the log-odds of falling into a latent class (e.g., LC1, LC2, or LC3) versus LC4 resulted in a one-unit increase in the covariates (gender, grade level, Spanish naming speed, Spanish working memory) holding the others covariates constant across all intercepts. Interestingly, some expected predictors (e.g., visual-spatial WM) based on the literature, did not yield reliable estimates. These findings may be related to the sample size within each of the LCs. Unfortunately, there is no published standard minimum number of manifest variables or sample size within the LC literature to consider whether we adequately represented the LCs and/or have an appropriate sample size for the LCA and LTA analysis (however, see Dziak, Lanza, & Tan, 2014, for a discussion).

Discussion

The main purpose of this study was to identify EL children at risk for learning disabilities, within a heterogeneous sample, that varied in first and second language proficiency and academic skills. The LC3 status group (at risk for learning disabilities) showed average intelligence and attention, but difficulties in performance across both English and Spanish measures of reading and math. The results also yielded two additional outcomes. First, 100% of children that were identified as at risk for learning disabilities, maintained their latent class status from Waves 1 to Wave 2. Interestingly, 13% of students in the LC1 status group (balanced bilingual-average achievers) transitioned into the LC3 status (learning disability) group at Wave 2. These children with late-emerging learning disabilities were particularly apparent in children transitioning from grades 2 to 3 (24%) and grades 3 to 4 (15%). We interpret these findings as suggesting that a transition in the language of instruction (Spanish to English) in the upper grades may have played a critical role in these outcomes. This inference will be discussed in detail below.

Second, children at risk for learning disabilities were separated from the other LC status groups on executive processing measures. As shown in Table 6, this separation included performance on WM measures within their first language and second language systems. The results supported the notion that children at risk for learning disabilities maintain difficulties across the two testing waves, but what differentiates them from the other latent groups is weak performance on measures of executive processing (WM) measures (Swanson et al., 2006; 2016). Below we addressed the three questions that directed this study:

1. **Can a latent classification of children at risk for learning disabilities be identified among a heterogeneous sample of EL children?**

The results showed that a latent class group (LC3) emerged for children at risk in reading and math in both language systems, and this latent class had a low probability of transitioning into average achieving latent classes. More specifically, these latent class membership probabilities were estimated at two testing waves, one year apart. These probabilities reflected the proportion of children expected to belong in each latent class at each time period. As shown in our analysis (see Table 3), the latent class membership probabilities for the latent class of EL children at risk for learning disabilities was approximately 20% of the total sample at Wave 1, and this probability increased to 25% at Wave 2.

2. Does membership in a latent class of EL children at risk change over time?

An answer to this question was partly addressed in our response above. However, to answer this question more thoroughly, transition probabilities were calculated that reflected the probability of transitioning from a particular latent class at time t to another latent class at time $t + 1$. Together, these probabilities reflected the amount of change in the latent class over time. As shown in Table 3, there was an 87% probability that balanced bilingual-average achievers would maintain their status at Wave 2, but only a 46% probability that unbalanced bilinguals would maintain their status at Wave 2. There was a 95% probability that the LC4 status group (English dominant) would emerge as English dominant (relatively lower Spanish proficiency) at Wave 2. More importantly, a lack of balance between English and Spanish proficiency would underlie that instability noted in LC2. As shown in Table 3, approximately 54% of students in the LC2 status group at Wave 1 emerged as the LC4 status group at Wave 2.

Because of the cohort-sequential nature of the study design, these transitional probabilities were also investigated within each cohort. Table 3 shows that the probability of transitioning into the risk class (the LC3 status group) was most likely to occur for balanced bilingual-average achievers (24%) in Cohort 2 (Grade 2 to 3). For Cohorts 1 and 3, the transition from Wave 1 to Wave 2 indicated a high probability (9% and 15%, respectively) of the balanced bilingual-average achievers in this sample moving into the risk category (the LC3 status group). Thus, there appears to be a critical period during the third grade that plays an important role in determining late-emerging risk class. The greatest dispersion (children moving to another LC status group at Wave 2) at Wave 1 occurred for Cohort 2 (grades 2 to grade 3).

3. Do specific cognitive measures predict latent class membership?

The findings suggest that performance on general cognitive measures, such as WM and naming speed, were associated with latent class membership. The candidate most often referred to in the literature as potentially underlying higher cognitive performance in bilingual children is executive processing. Our results are consistent with these findings. We inferred that the increasing importance of the executive component of WM occurs because of the increasing emphasis on more complex academic tasks, such as reading comprehension, in the later grades. Thus, our findings were interpreted as suggesting that EL children at risk for reading and/or math experience difficulties when drawing on a WM system that is independent of their problems in phonological storage, naming speed, and inhibition processes. We assumed that L2 (English) achievement depends on WM, which not only takes into account the storage of phonological information for later retrieval, but also controls attention. Thus, if children have adequate WM resources across both Spanish and English language systems, then the execution of various fundamental processes (such as word encoding, lexical access, syntactic and semantic analysis, etc.) does not deplete the limited resource pool as much as it does for EL children with weaker WM skills. In general, our findings are consistent with several studies that have established significant relationships between WM and achievement within and across language systems in EL children (e.g., Swansonet al., 2004; 2019a).

Although the focus of this study was on determining whether children at risk for LD could be identified among EL children and whether this latent class was stable across testing waves, an unexpected finding was that we were able to identify children who may be at risk for late-emerging learning disabilities. We were surprised to find that some children who are relatively proficient in both Spanish and English during the first testing wave would be at risk on achievement measures at the second testing wave. Our best inference on this finding was that the language of instruction in the older grades may have placed a heavy emphasis on English, rather than allowing for the access of information from both English and Spanish language systems. As English learners continue to develop competencies in their second language, there is often a decline in their L1, or Spanish language skills (also see Gottardo, Collins, Bacium & Gebotys, 2008; for the difference in instructional strategies across grade 1 and 2). Indeed, early signs of language loss, or attrition, in English learners may present as a decrease in L1 vocabulary skills, increased use of general/nondescriptive words, increased grammatical errors in L1, decreased ability to handle L1 language tasks, or to develop academic language skills in their L1 (Guiberson, Barrett, Jancosek, & Yoshinaga Itano, 2006). When the Spanish language is no longer emphasized or supported in the classroom, the opportunities for maintaining and continuing to build L1 diminish. Indirect evidence for this outcome was the high incidence of English dominant achievers in Wave 1 (LC2) transitioning to low Spanish proficiency (LC4) at the second testing wave.

Our sampling reflected children who were sequential bilinguals (L2 follows L1 development) and therefore may not reflect bilingualism when two languages are learned simultaneously (e.g., Sabourin & Vīnerte, 2015). Also, our geographic region (e.g., studies conducted in the southwest U.S. versus Canada) may limit the generalizability of the findings. Likewise, we use normative measures and therefore it is important to note that some of the standardized assessments do not report the language background of their norming sample or are based primarily on a monolingual English speaking population. As indicated by one reviewer, when comparing second language (L2) speakers' performance using the standardized scores, it is relative to others who speak a single language. Given that there is research showing that Spanish-English bilingual children may show different developmental trajectories in both languages (e.g., Hoff, 2017), it is possible to observe 'lower' standard scores as children age, reflecting the slower growth. Thus, to compensate for some of these limitations we used a conservative cut-off for determining risk (< 1 standard deviation) rather than cut-offs commonly used with monolingual children (e.g., 25 percentile, 35^{th} percentile, see Swanson, 2018 for review).

When comparing the LC1 status group (balanced bilingual-average achievers) with the LC3 status group (at risk for learning disabilities), we also note that the two groups were comparable in English inhibition (ES= -.04) at Wave 1. This suggests that the two status groups may have experienced similar difficulties accessing the English language system and/or inhibiting the Spanish language system, which in turn may have put some balanced bilingual-average achievers in the at-risk group. Therefore, some vulnerability in executive processing among sequential bilinguals may have played a role in later performance (cf. Kieffer & Christodoulou, 2019). Regardless of these inferences on our part, it is important to note that the majority of studies finding higher performance in achievement and cognitive processing in bilingual children have focused on children who learned L1 and L2 simultaneously. Therefore, the higher academic performance of sequential bilinguals (who learn their L1 first, then L2 later) with different levels of language proficiency needs further research.

Comparison of Studies

How do results compare to our previous work (Swanson et al., 2016) on latent transition analysis of EL children at risk? What clearly separated the two studies was that the earlier study found a latent class of children with reading difficulties, whereas **no** unique reading difficulties or disabilities latent class occurred for the current study. The current study found a latent class of children at risk in reading and math that cut-across both language systems. What was shared between the two studies, however, was the stability of the at-risk classification. The earlier work defined risk based on English measures of reading below the 25th percentile. For the total sample in grades 1 to 3 in this earlier study, children defined at risk for reading difficulties in Wave 1 yielded a 100% chance of being labeled as at-risk at Wave 2 (grades 2, 3 and 4). Likewise, in the current study, children who were defined children at risk (below the 16th percentile on both Spanish and English measures of reading and math) also had a 100% chance of maintaining their risk status at wave 2. Thus, whether defined within the English system (Swanson et al. 2016) or in both English and Spanish as in the current study, EL children the risk for learning disabilities across testing waves was stable.

The studies varied considerably in identifying children with emergent learning disabilities. In the earlier study, only children in the low attentive/English and Spanish reading comprehension group at Wave 1 had a 9% chance of transitioning into the at-risk latent group (children at risk for reading difficulties) at Wave 2. This latent class group was characterized by low classroom attention and low reading comprehension, but high skills in sight word recognition at wave 1. More importantly, this earlier study found that no average readers in wave 1 were identified as at risk at Wave 2. Such was not the case in the current study.

In the current study, thirteen percent (13%) of children labeled as bilingual average achievers at wave 1 (referred to as LC 1) were label at risk for learning disabilities at wave 2. At wave 1, as a group, these children yielded English and Spanish vocabulary, math, and reading scores in the average range (> 85). However, the incidence of late-emerging risk among these wave 1 average readers occurred across all cohorts: 9% for Cohort 1 (grade 1 to grade 2), 24% for Cohort 2 (grade 2 to grade 3), and 15% for Cohort 3 (grade 3 to grade 4). Thus, by using both language systems that focused on additional performance measures of risk beyond reading and receptive language, we were able to identify several children who were average achievers at wave 1, but exhibited risk status at wave 2.

We have three explanations related to the differences in findings related to identifying children with late-emerging learning disabilities. First, vacillation between English and Spanish academic instruction may have played a primary role. The language of instruction in the earlier work was English, whereas the current study sampled children with academic instruction in Spanish in grades 1 and 2 and Spanish-English instruction in grades 3-4. Thus, one can infer that if the language of instruction stays the same as in the earlier study, EL children who are average readers are less likely to be late-emerging children with reading difficulties. However, as shown in the current study, late-emerging risk status occurred among all cohorts (whether the instruction was Spanish, English or mixed English-Spanish). Second, by broadening the range of manifest variables in the current study we were able to identify comorbid difficulties (reading and math on both basic (word identification, calculation) and high order [reading comprehension and mathematical problem solving] measures across both language systems. Thus, because the current study included a broader range of manifest variables we were better able to show that later risk for a learning disability cut-across both language systems. Finally, the cognitive measures that reflect risk status varied as a function of classification criteria (see Swanson et al., 2018; for discussion of this issue in the domain of mathematics). In the earlier study, the late-emerging group (referred to in the study as low attention and comprehension) shared comparable performance with the at-risk group on several cognitive measures. The only cognitive measures that appeared significant between current risk and later risk status was within the L1 system. Consistent with earlier work (Swanson et al., 2004), the two risk groups could only be differentiated in terms of risk status on their proficiency on measures of WM and phonological processing within their **first** (Spanish) language system. In contrast, children at risk for learning disabilities and the late-emerging learning disabilities in the present study exhibited risk factors in both language systems. The status group (at risk for learning disabilities) group yielded low performance relative to the LC1 status group (balanced bilingual-average achievers) on English measures of STM, Speed and WM at wave 1, but exhibited deficits in both English and Spanish measures of STM and WM at wave 2. The important point is that if the classification of EL children with learning disabilities is to be manifested across both language systems, then similar cognitive difficulties should be reflected in both languages.

Implications

Taken together, these results provide support for the notion that children initially identified as at risk for LD reflect a highly stable latent class across two testing waves within this particular EL sample. The results also show that the incidence of children moving into the risk group (LC3) were children identified as balanced bilingual average achievers in Cohort 2 (grades 2 to 3). That is, children with average achievement and comparable language development in both English and Spanish were more likely than their unbalanced bilingual, English dominant (LC2) average achieving counterpart to move into the risk group.

Before outlining further implications of our findings, however, it is important to note that all children were Spanish speakers, but not all children were fluent Spanish readers. This has been a consistent finding in earlier studies (e.g., Swanson et al., 2004; 2006; 2016). As noted in Table 1, Spanish passage comprehension standard scores were substantially lower than English passage comprehension scores, clearly suggesting that the EL children in our sample were not truly biliterate. However, we assumed that the students reflected a representative sample. That is, except for Spanish comprehension measures, the normed results for English reading, fluid intelligence, and math skills approximated a normal distribution found with monolingual samples (*M* = 100, *SD* = 15). Given these qualifications for our study, we provide four implications to the current literature below.

First, the results indicate that distinct latent classes emerge using the 16th percentile as an "a priori" cut-off score for determining risk for learning disabilities. However, it is important to note that the same or other latent classes may have emerged with other cut-off scores. Thus, we have not shown that the identification of latent classes validates a specific cut-off point; rather the results suggested that the measures were able to identify subgroups to the cut-off for which they were applied (Swanson et al., 2016; 2018). Although the 16^{th} percentile or below across multiple years has been used as an "a priori cut-off point" to identify children at risk, the issue as to whether other cut-off scores yield similar latent classes of EL children at risk and/or discrete or identifiable groups has not been established.

Another issue with cut-off scores being related to reading and/or math is if 'proficiency' is better understood as a continuous variable, rather than 'performance' above or below a specific cut-off point. However, the vast majority of studies on children at risk for reading and math difficulties have used the dichotomization of normed-referenced achievement measures as a means to study children classified as at risk. In terms of common cut-off score designations for risk, a norm-referenced score (e.g., < 25th percentile) is commonly used to designate reading and/or math difficulties, whereas at least a full standard deviation (< 85 standard score) or more is used to identify disabilities. Thus, we used the 16^{th} percentile or a normative score < 85 as a designation for children at risk for learning disabilities. No doubt, categorizing data is sometimes not recommended when compared to analyzing continuous measures, because creating discrete variables from continuous variables has been shown to increase type I error, weaken reliability, and decrease power (MacCallum et al., 2002). Also, some studies suggest that the identification of children at risk in reading and math disabilities is best viewed as fitting on a continuum (e.g., Stanovich, 1988). However, what we have shown is that latent subgroups do emerge within this continuum.

Finally, we were able to identify those variables associated with EL children at risk for learning disabilities. The primary implication of our findings is that performance on both Spanish and English measures of achievement and cognition is associated with differentiating children at risk for learning disabilities from other latent classes. The measures that were used to assess children at risk were in both their first and second languages. For example, as shown in Table 6, both L1 measures (Spanish WM) and L2 measures (English WM) played a major role in predicting latent classes at Wave 2. The LC3 group (children at risk for learning disabilities) were found to be the least likely to change class and could be differentiated from the other latent classes on measures of WM.

Limitations

There are at least three limitations to this study. First, as useful as LCA is to determine meaningful patterns within the data, identifying children at risk commonly requires setting a cut-off point based on normative data, and therefore there may be artifacts in the psychometric measures selected. Although the measures in our study are commonly used, a selection of alternative measures may likely yield different results.

Second, the findings from this study should be interpreted with caution because LTA has not at present provided procedures for conducting a power analysis. For example, the small sample size for Cohort 3 may have potentially yielded more identifiable items in the areas of vocabulary and cognition with a larger sample.

Finally, we have an absence of intervention information. Although all children in this study participated in reading and math instruction in the dual language classroom, we had no information as to whether some were receiving additional instruction outside of the classroom. Because school records showed that all children had learned English as an L2, this allowed us, however, to explore the cognitive processes accessed regarding L1 Spanish versus L2 English, as well as transfer to L2 reading and math problem-solving that have not been explored in previous studies. Thus, our study is limited to discussing the stability of the risk classification and not whether a particular intervention program could later influence the classification of children at risk (also see Swanson et al., 2006, 2015; 2016 for discussion).

Summary

In summary, this study yielded three important findings. First, latent classifications of children at risk could be identified among a sample of ELs. Four latent classes emerged across two testing waves: balanced bilinguals-average achievers, unbalanced bilinguals-average achievers, children at risk for learning disabilities, and English dominant children (children relatively low in Spanish proficiency). Second, latent class membership probabilities for the risk class at Wave 1 (LC3) remains stable over time. Such was not the case for the groups that varied in language acquisition (Spanish vocabulary). Children low in expressive Spanish (LC2) language increased their chances overtime of decreasing their performance on the Spanish reading comprehension measure. That is, several L2 children in Wave 1 emerged with the L4 status group in Wave 2. Finally, children at risk could be identified on measures external to the classification. Children identified as at risk for learning disabilities (LC3) performed poorly on cognitive measures of English STM and WM when compared to the other latent classes.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders, 5th edition*. Washington, DC: American Psychiatric Association.

Artiles, A. J., Rueda, R., Salazar, J. J., & Higareda, I. (2005). Within-group diversity in minority disproportionate representation: English language learners in urban school districts. *Exceptional Children, 71*(3), 283-300. doi: 10.1177/001440290507100305

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using M-Plus. *Structural Equation Modeling, 21*(3), 329-341. doi:10.1080/10705511.2014.915181

Asparouhov, T., & Muthén, B. (2015). Residual associations in latent class and latent transition analysis. *Structural Equation Modeling, 22*(2), 169-177. doi:10.1080/10705511.2014.935844

August. D., & Hakuta, K. (1997). *Improving schooling for minority-language children: A research agenda.* Washington, DC: National Academy Press.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63,* 1-29. doi: /10.1146/annurev

Baddeley, A. D., & Logie, R. H. (1999). The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance* and executive control (pp. 28-61). Cambridge, U.K.: Cambridge University Press.

Barnes, M. A., Clemens, N. H., Fall, A., Roberts, G., Klein, A., Starkey, P., Flynn, K. (2020). Cognitive predictors of difficulties in math and reading in pre-kindergarten children at high risk for learning disabilities. *Journal of Educational Psychology, 112*(4), 685-700. doi:10.1037/edu0000404

Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology*, *65*(4), 229-235. doi:10.1037/a0025406

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

Boada, R., Willcutt, E. G., & Pennington, B. F. (2012). Understanding the comorbidity between dyslexia and attention-deficit/hyperactivity disorder. *Topics in Language Disorders, 32*(3), 264-284. doi:10.1097/TLD.0b013e31826203ac

Bonifacci, P., Giombini, L., Bellocchi, S., & Contento, S. (2011). Speed of processing, anticipation, inhibition and working memory in bilinguals. *Developmental Science, 14*(2), 256-269.

Brandenburg, J., Klesczewski, J., Schuchardt, K., Fischbach, A., Büttner, G., & Hasselhorn, M. (2017). Phonological processing in children with specific reading disorder versus typical learners: Factor structure and measurement invariance in a transparent orthography. *Journal of Educational Psychology, 109*(5), 709-726. doi:10.1037/edu0000162

Branum-Martin, L., Fletcher, J. M., & Stuebing, K. K. (2013). Classification and identification of reading and math disabilities: The special case of comorbidity. *Journal of Learning Disabilities, 46*(6), 490-499.

Brownell, K. (2001). *Expressive One-Word Picture Vocabulary Test* (3rd Edition). New York: Academic Therapy Publications.

Bumgarner, E., Martin, A., & Brooks-Gunn, J. (2013). Approaches to learning and Hispanic children's math scores: The moderating role of English proficiency. *Hispanic Journal of Behavioral Sciences, 35*(2), 241-259. doi:10.1177/0739986312473580

Catts, H. W., Compton, D., Tomblin, J. B., & Bridges, M. S. (2012). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology, 104*(1), 166-181. doi: 10.1037/a0025323

Chiappe, P., Siegel, L. S., & Wade-Woolley, L. (2002). Linguistic diversity and the development of reading skills: A longitudinal study. *Scientific Studies of Reading, 6*, 369-400. doi:10.1207/S1532799XSSR0604_04

Chiappe, P., & Siegel, L. S. (2006). A longitudinal study of reading development of Canadian children from diverse linguistic backgrounds. *The Elementary School Journal, 107*(2), 135-152. doi:10.1086/510652

Child, A. E., Cirino, P. T., Fletcher, J. M., Willcutt, E. G., & Fuchs, L. S. (2019). A cognitive dimensional approach to understanding shared and unique contributions to reading, math, and attention skills. *Journal of Learning Disabilities, 52*(1), 15-30. doi:10.1177/0022219418775115

Cirino, P. T., Fuchs, L. S., Elias, J. T., Powell, S. R., & Schumacher, R. F. (2015). Cognitive and mathematical profiles for different forms of learning difficulties. *Journal of Learning Disabilities, 48*(2), 156-175. doi:10.1177/0022219413494239

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: L. Erlbaum Associates.

Collins, L.M. & Flaherty, B.D. (2002). Latent class model for longitudinal data. In J.A. Hageneers & L. McCutcheon (Eds. *) Applied latent class analysis* (p.287-303). Cambirdige, UK:Cambirdge

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis with applications in the social, behavioral, and health sciences*. New Jersey: Wiley and Sons.

Compton, D. L., Fuchs, D., Fuchs, L. S., Elleman, A. M., & Gilbert, J. K. (2008). Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. *Learning and Individual Differences, 18*(3), 329-337. doi: 10.1016/j.lindif.2008.04.003

Conners, C. K. (1997). *Conners' Rating Scales-Revised: Technical manual.* North Tonawanda, NY: Multi-Health Systems.

Cooper, R. P. (2016). Executive functions and the generation of "random" sequential responses: A computational account. *Journal of Mathematical Psychology, 73*, 153-168. doi:10.1016/j.jmp.2016.06.002

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review, 26*(2), 197-223. doi:10.1007/s10648-013-9246-y

Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. Working papers on bilingualism, No. 9. Retrieved from http://files.eric.ed.gov/fulltext/ED125311.pdf

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49,* 222-251. doi:10.3102/00346543049002222

David, C. V. (2012). Working memory deficits in math learning difficulties: A meta-analysis. *International Journal of Developmental Disabilities, 58*(2), 67-84. doi:/10.1179/2047387711Y.0000000007

de Abreu, P. M. J.. (2011). Working memory in multilingual children: Is there a bilingual effect?

Memory, 19(5), 529-537.doi:10.1080/09658211.2011.590504

de Abreu, P. M. J., & Gathercole, S. E. (2012). Executive and phonological processes in second language acquisition. *Journal of Educational Psychology, 104*(4), 974-986. doi:10.1037/a0028390

Dunn, L. M., & Dunn, L. M. (1981). *The Peabody Picture Vocabulary Test-4.* NY: Pearson.

Dunn, L. M., & Dunn, L. M. (2007). *The Peabody Picture Vocabulary Test-4.* NY: Pearson.

Dunn, L. M., Lugo, D. E., Padilla, E. R., & Dunn, L. M. (1986). *Test de Vocabulario Imágenes Peabody.* Circle Pines, MN: American Guidance Service.

Dziak, J. J., Lanza, S. T., & Tan, X. (2014). Effect size, statistical power, and sample size requirements for bootstrap likelihood ratio test in latent class analysis. *Structural Equation Modeling, 21*(4), 534-552. doi:10.1080/10705511.2014.919819

Etmanskie, J. M., Partanen, M., & Siegel, L. S. (2016). A longitudinal examination of the persistence of late emerging reading disabilities. *Journal of Learning Disabilities, 49*(1), 21-35. doi:10.1177/0022219414522706

Farnia, F., & Geva, E. (2019). Late-emerging developmental language disorders in English-speaking monolinguals and English-language learners: A longitudinal perspective. *Journal of Learning Disabilities, 52*(6), 468-479. doi:10.1177/0022219419866645

Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K., & Shaywitz, S. E. (2010). Uncoupling of reading and IQ over time: Empirical evidence for a definition of dyslexia. *Psychological Science, 21*(1), 93-101. doi:10.1177/0956797609354084

Fletcher, J. M., Stuebing, K. K., Morris, R. D., & Lyon, G. R. (2014). Classification and definition of learning disabilities: A hybrid model. In H. L. Swanson, K. R. Harris & S. Graham (Eds.), *2nd ed.; handbook of learning disabilities (2nd ed.)* (2nd ed. ed., pp. 33-

50) NY: The Guilford Press.

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Schatschneider, C., & Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology, 98*, 29-43. doi:10.1037/0022-0663.98.1.29

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2012). Mathematical cognition deficits in children with learning disabilities and persistent low achievement: A five-year prospective study. *Journal of Educational Psychology, 104*(1), 206-223. doi:/10.1037/a0025398

Geary, D. C., Nicholas, A., Li, Y., & Sun, J. (2017). Developmental change in the influence of domain-general abilities and domain-specific knowledge on mathematics achievement: An eight-year longitudinal study. *Journal of Educational Psychology, 109*(5), 680-693. doi:/10.1037/edu0000159

Giofrè, D., Toffalini, E., Altoè, G., & Cornoldi, C. (2017). Intelligence measures as diagnostic tools for children with specific learning disabilities. *Intelligence, 61*, 140-145. doi:10.1016/j.intell.2017.01.014

Georgiou, G. K., Tziraki, N., Manolitsis, G., & Fella, A. (2013). Is rapid automatized naming related to reading and mathematics for the same reason(s)? A follow-up study from kindergarten to grade 1. *Journal of Experimental Child Psychology, 115*(3), 481-496. doi: 10.1016/j.jecp.2013.01.004

González, J. E. J., & Valle, I. H. (2000). Word identification and reading disorders in the Spanish language. *Journal of Learning Disabilities, 33*(1), 44-60. doi:10.1177/002221940003300108

Gottardo, A., Collins, P., Baciu, I., & Gebotys, R. (2008). Predictors of grade 2 word reading and vocabulary learning from grade 1 variables in spanish-speaking children: Similarities and differences. *Learning Disabilities Research & Practice, 23*(1), 11-24. doi:10.1111/j.1540-5826.2007.00259.x

Guiberson, M. M., Barrett, K. C., Jancosek, E. G., & Yoshinaga Itano, C. (2006). Language maintenance and loss in preschool-age children of Mexican immigrants: Longitudinal study. *Communication Disorders Quarterly*, *28*(1), 4-17. doi:10.1177/15257401060280010601

Grimm, R.P., Solari, E.J., Gerber, M.M., Nylund-Gibson, K., & Swanson, H.L. (2019). A cross-linguistic examination of heterogeneous reading profiles of Spanish-speaking bilingual students. *The Elementary School Journal, 120,* 109-131.

Hemphill, F.C., & Vanneman, A. (2011). Achievement Gaps: How Hispanic and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress (NCES 2011-459). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Hoff, E. (2018). Bilingual development in children of immigrant families. *Child Development Perspectives, 12*(2), 80-86. doi:10.1111/cdep.12262

Kam, C., Morin, A. J. S., Meyer, J. P., & Topolnytsky, L. (2016). Are commitment profiles stable and predictable? A latent transition analysis. *Journal of Management, 42*(6), 1462-1490. doi:10.1177/0149206313503010

Kaplan, D., & Keller, B. (2011). A note on cluster effects in latent class analysis. *Structural Equation Modeling, 18*(4), 525-536. doi:10.1080/10705511.2011.607071

Kieffer, M. J. (2011). Converging trajectories: Reading growth in language minority learners and their classmates, kindergarten to grade 8. *American Educational Research Journal, 48,* 1187-1225. doi:10.3102/0002831211419490

Kieffer, M. J., & Christodoulou, J. A. (2019). Automaticity and control: How do executive functions and reading fluency interact in predicting reading comprehension? *Reading Research Quarterly,* doi10.1002/rrq.289

Kieffer, M. J., & Thompson, K. D. (2018). Hidden progress of multilingual students on NAEP. *Educational Researcher, 47*(6), 391-398. Retrieved from https://search.proquest.com/docview/2315229412?accountid=14521

Lanza, S. T., Dziak, J. J., Huang, L., Xu, S., & Collins, L. M. (2011). *Proc LCA and Proc LTA users' guide* (Version 1.2.7). University Park: The Methodology Center, Penn State. Retrieved from: http://methodology.psu.edu.

Lesaux, N. K., & Harris, J. R. (2017). An investigation of comprehension processes among adolescent English learners with reading difficulties. *Topics in Language Disorders, 37*(2), 182-203. doi10.1097/TLD.0000000000000120

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. doi: 10.1007/s11145-005-4713-6

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2013). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review,* doi:10.3758/s13423-013-0565-2

Lohman, D. F., & Gambrell, J. L. (2012). Using nonverbal tests to help identify academically talented children. *Journal of Psychoeducational Assessment, 30*(1), 25-44. doi: 10.1177/0734282911428194

Lohman, D. F., Korb, K. A., & Lakin, J. M. (2008). Identifying academically gifted English-language learners using nonverbal tests: A comparison of the Raven, NNAT, and CogAT. *Gifted Child Quarterly, 52*(4), 275-296. doi: 10.1177/0016986208321808

Lonigan, C. J., Allan, D. M., Goodrich, J. M., Farrington, A. L., & Phillips, B. M. (2017). Inhibitory control of Spanish-speaking language-minority preschool children: Measurement and association with language, literacy, and math skills. *Journal of Learning Disabilities, 50*(4), 373-385. doi: 10.1177/0022219415618498

Maas, C., & Hox, J. (2005). Sufficient sample sizes in multiple regression analysis. *Methodology, 1,* 86-92. doi:10.1027/1614-2241.1.3.85

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*(1), 19-40. doi:10.1037/1082-989X.7.1.19

Mancilla-Martinez, J., Hwang, J. K., Oh, M. H., & McClain, J. B. (2019). Early elementary grade dual language learners from Spanish-speaking homes struggling with English reading comprehension: The dormant role of language skills. *Journal of Educational Psychology,* doi10.1037/edu0000402

Mann Koepke, K., & Miller, B. (2013). At the intersection of math and reading disabilities: Introduction to the special issue. *Journal of Learning Disabilities, 46*(6), 483-489. doi:10.1177/0022219413498200

Mancilla-Martinez, J., & Lesaux, N. K. (2017). Early indicators of later English reading comprehension outcomes among children from Spanish-speaking homes. *Scientific Studies of Reading, 21*(5), 428-448. doi10.1080/10888438.2017.1320402

Masyn, K. (2013). Latent class analysis and finite mixture modeling. In T. Little (Ed.), *The Oxford handbook of quantitative methods in psychology* (Vol. 2, pp. 375-393). Oxford, UK: Oxford University Press.

McCardle, P., Keller-Allen, C., & Shuy, T. (2008). Learning disability identification. In E. Grigorenko (Ed.), *Educating Individuals with Disabilities* (pp. 137-164). New York, NY: Springer

Morgan, P. L., & Farkas, G. (2016). Are we helping all the children that we are supposed to be helping? *Educational Researcher, 45*(3), 226-228. doi:/10.3102/0013189X16644607

Morin, A. J. S., Bujacz, A., & Gagné, M. (2018). Person-centered methodologies in the organizational sciences: Introduction to the feature topic. *Organizational Research Methods, 21*(4), 803-813. doi:10.1177/1094428118773856

Muñoz-Sandoval, A. F., Woodcock, R. W., McGrew, K. S., & Mather, N. (2005). Bateria III Woodcock- Muñoz. Rolling Meadows, IL: Riverside.

Murphy, M.M., Mazzocco, M.M.M., Hanich, L.B., & Early, M.C. (2007). Cognitive characteristics of children with mathematics learning disability (MLD) vary as a function of the cutoff criterion used to define MLD. *Journal of Learning Disabilities*, 40(5), 458-478.

Muthén, B. (2006). The potential of growth mixture modeling. *Infant and Child Development, 15*(6), 623-625. doi:10.1002/icd.482

Muthén, B., & Muthén, L. K. (2012). Mplus:User's Guide. Los Angeles ,CA. Muthén & Muthén.

National Assessment of Educational Progress (2011). Achievement gap: How Hispanics and white students in public schools perform in mathematics and reading on the national assessment of educational progress. Washington DC: US Department of Education.

National Assessment of Educational Progress (2017). The condition of education (update 2017) Washington DC: US Department of Education.

National Assessment of Educational Progress (2019) NAEP Mathematics and Reading Assessments: Highlighted Results at Grades 4 and 8 for the Nation, States, and Districts Washington DC: US Department of Education.

National Center for Education Statistics (2020) Common Core of Data (CCD) "Local Education Agency Universe Survey. Washington DC: U.S. Department of Education

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd Ed.). NY: McGraw-Hill.

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, 14*(4), 535-569. doi:10.1080/10705510701575396

Nylund-Gibson, K., & Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science, 4*(4), 440-461. doi:10.1037/tps0000176

Nylund-Gibson, K., Grimm, R., Quirk, M., & Furlong, M. (2014). A latent transition mixture model using the three-step specification. *Structural Equation Modeling, 21*(3), 439-454. doi:10.1080/10705511.2014.915375

Peña, E. D., Bedore, L. M., & Kester, E. S. (2016). Assessment of language impairment in bilingual children using semantic tasks: Two languages classify better than one. *International Journal of Language & Communication Disorders, 51*(2), 192-202. doi:10.1111/1460-6984.12199

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

Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016a). A meta-analysis of mathematics and working memory: Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology, 108*(4), 455-473. doi:/10.1037/edu0000079

Peng, P., Namkung, J., Fuchs, D., Fuchs, L., Patton, S., Yen, L., Compton, D. L., Zhang, W. J., Miller, A. & Hamlett, C. (2016b). A longitudinal study on predictors of early calculation development among young children at risk for learning difficulties. *Journal of Experimental Child Psychology, 152*, 221-241. doi:/10.1016/j.jecp.2016.07.017

Prevoo, M. J. L., Malda, M., Mesman, J., & van IJzendoorn, M. H. (2016). Within- and cross-language relations between oral language proficiency and school outcomes in bilingual children with an immigrant background: A meta-analytical study. *Review of Educational Research, 86*(1), 237-276. doi:10.3102/0034654315584685

Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology*, *97*(2), 246-256. doi:10.1037/0022-0663.97.2.246

Raven, J. C. (1976). *Colored Progressive Matrices*. London, England: H. K. Lewis & Co. Ltd.

Ryoo J. H., Wang , C., Swearer, S. M., Hull, M., & Shi, D. (2018) Longitudinal Model Building Using Latent Transition Analysis: An Example Using School Bullying Data. Frontiers in Psychology, 9:675. doi: 10.3389/fpsyg.2018.00675

Sabourin, L., & Vīnerte, S. (2015). The bilingual advantage in the stroop task: Simultaneous vs.

Early bilinguals. *Bilingualism: Language and Cognition, 18*(2), 350-355. doi:10.1017/S1366728914000704

SAS Institute (2010). *SAS/STAT software: Changes and Enhancements through release 9.3*. Cary, NC: SAS Institute Inc.Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally
achieving and subtypes of learning disabled children. *Child Development, 60*(4), 973-980. doi:10.2307/1131037

Snijders, T., & Bosker, R. (1999). *Multilevel modeling: An introduction to basic and advanced multilevel modeling.* Thousand Oaks, CA: Sage Press.

Snowling, M. J. (2012). Changing concepts of dyslexia: Nature, treatment and comorbidity. *Journal of Child Psychology and Psychiatry, 53*(9), e1-e3. doi:10.1111/j.1469-7610.2009.02197.x

Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Learning Disabilities, 21*(10), 590-604, 612. doi:10.1177/002221948802101003

Stanovich, K. E., & Siegel, L. (1994). Phenotypic performances profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Education Psychology, 86*(1), 24-53. doi:10.1037/0022-0663.86.1.24

Swanson, H.L. (2020). Learning disabilities as a working memory deficit: A revised model. In Sperling, R., Martin, A., & Newton, K. *Handbook of Educational Psychology & Students with Special Needs* (pp. 19-51). NY: Rutledge

Swanson, H. L. (2012). Cognitive profile of adolescents with math disabilities: Are the profiles different from those with reading disabilities? *Child Neuropsychology, 18*(2), 125-143. doi:10.1080/09297049.2011.589377

Swanson, H. L., & Fung, W. (2016). Working memory components and problem-solving accuracy: Are there multiple pathways? Journal of Educational Psychology, 108(8), 1153-1177. doi:10.1037/edu0000116

Swanson, H. L., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 100*(2), 343-379. doi:10.1037/0022-0663.100.2.343

Swanson, H. L., Kong, J. E., & Petcu, S. D. (2019a). Individual differences in math problem solving and executive processing among emerging bilingual children. *Journal of Experimental Child Psychology, 187*, 25. doi:10.1016/j.jecp.2019.06.006

Swanson, H. L., Kong, J., & Petcu, S. (2019b). Math difficulties and working memory growth in English language learner children: Does bilingual proficiency play a significant role? *Language, Speech, and Hearing Services in Schools, 49*(3), 379-394. doi:10.1044/2018_LSHSS-17-0098

Swanson, H. L., Kong, J., Petcu, S.D., Pimental, M.F. A (2020). Can Difficulties in Language Acquisition and Specific Learning Disabilities Be Separated Among English Learners? *Exceptional Children, 86,* 293-309. doi.org/10.1177/0014402919893931

Swanson, H. L., Kudo, M., & Guzman-Orth, D. (2016). Cognition and literacy in English language learners at risk for reading disabilities: A latent transition analysis. *Journal of Educational Psychology, 108*(6), 830-856. doi:10.1037/edu0000102

Swanson, H. L., Olide, A. F., & Kong, J. E. (2018). Latent class analysis of children with math difficulties and/or math learning disabilities: Are there cognitive differences? *Journal of Educational Psychology, 110*(7), 931-951. doi:10.1037/edu0000252

Swanson, H. L., Orosco, M. J., & Lussier, C. M. (2012). Cognition and literacy in English language learners at risk for reading disabilities. *Journal of Educational Psychology, 104*(2), 302-320. doi:10.1037/a0026225

Swanson, H. L., Orosco, M. J., & Lussier, C. M. (2015). Growth in literacy, cognition, and working memory in English language learners. *Journal of Experimental Child Psychology, 132*, 155-188. doi:10.1016/j.jecp.2015.01.001

Swanson, H. L., Orosco, M. J., Lussier, C. M., Gerber, M. M., & Guzman-Orth, D. (2011). The influence of working memory and phonological processing on English language learner children's bilingual reading and language acquisition. *Journal of Educational Psychology, 103*(4), 838-856. doi:10.1037/a0024578

Swanson, H. L., Sáez , L., & Gerber, M. (2006). Growth in literacy and cognition in bilingual children at risk or not at risk for reading disabilities. *Journal of Educational Psychology, 98*(2), 247-264. doi:10.1037/0022-0663.98.2.247

Swanson, H. L., Sáez, L., Gerber, M., & Leafstedt, J. (2004). Literacy and cognitive functioning in bilingual and nonbilingual children at or not at risk for reading disabilities. *Journal of Educational Psychology, 96*(1), 3-18. doi:10.1037/0022-0663.96.1.3

Vermunt, J. K. (2007). Growth models for categorical response variables: Standard, latent-class, and hybrid approaches. In K. van Montfort, J. Oud & A. Satorra (Eds.), *Longitudinal models* *in the behavioral and related sciences; longitudinal models in the behavioral and related sciences* (pp. 139-158) Lawrence Erlbaum Associates Publishers, Mahwah, NJ.

Vukovic, R. K., & Lesaux, N. K. (2013). The language of mathematics: Investigating the ways language counts for children's mathematical development. *Journal of Experimental Child Psychology, 115*(2), 227-244. doi: 10.1016/j.jecp.2013.02.002

Wagner, R., Torgesen, J., & Rashotte, C. (2000). *Comprehensive Test of Phonological Processes.* Austin TX: Pro-ED.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children-Third Edition*. San Antonio, TX: Psychological Corporation.

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-Deficit/Hyperactivity disorder: A meta-analytic review. *Biological Psychiatry, 57*(11), 1336-1346. doi: 10.1016/j.biopsych.2005.02.006

Woodcock, R. W., McGrew, K. S., Schrank, F.A. , & Mather, N. (2007). Technical manual. Woodcock-Johnson III Normative Update. Rolling Meadows, IL: Riverside Publishing.

Woodcock, R. W., Muñoz-Sandoval, A. F. & Alvarado, C. G. (2005). *Woodcock-Muñoz Language Survey*. Itasca, IL: Riverside Publishing.

Yeung, S. S. (2018). Second language learners who are at-risk for reading disabilities: A growth mixture model study. *Research in Developmental Disabilities, 78*, 35-43. doi:10.1016/j.ridd.2018.05.001

Footnotes

[1]Current categories of learning disabilities include specific disabilities in reading and math [see IDEA reauthorization, 2004, Sec. 300.8(c)(10)]. In contrast, the new American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) does not use the term "learning disabilities" or related terms (e.g., dyslexia, math disabilities, reading disabilities). The new DSM-5 (2013) uses the term "specific learning disorder in reading" or "specific learning disorder in math" and assumes that such disorders have a neurological/biological base. The broad category of a learning disorder in reading and/or math captures a sample of children previously referred to in the literature as having reading disabilities (RD) and/or math disabilities (MD). In general, researchers use the terms such as specific reading disabilities and/or math disabilities to identify children at risk who are of average intelligence, but performance is below a certain percentile (e.g., 8^{th}, 16^{th}, 25^{th} percentile) on a norm-referenced standardized reading and/or math measure (e.g., Cirino et al. 2015; Siegel & Stanovich, 1994). In terms of research, it is not uncommon to find children with normal intelligence defined as having a reading disability, but also experience serious difficulties in math and vice versa (e.g., Mann Koepke & Miller, 2013). This is because both reading and math draw upon or overlap with similar cognitive processes (e.g., Child et al., 2019; Swanson, 2012, 2020).

[2] Traditionally, as indicated above, children at risk for RD and/or MD are operationally defined by performing below a cut-off point on a norm-referenced achievement measure [studies vary from the 16^{th} to 25^{th} percentile on norm-referenced standardized achievement measures (e.g., Murphy et al., 2007; Vukovic & Lesaux, 2013; Swanson et al., 2006)]. Although commonly used to identify children at risk and in need of intervention, there is debate as to

whether a cut-off score on normative data can be used to identify children at risk for learning disabilities. We suggest that our dichotomization of continuous variables on the normative achievement and language measures is valid if it can be shown that a mixture of hidden groups underlies the cut-off point. Although highly critical of the dichotomization of continuous variables in general, the seminal article by MacCallum, Zhang, Preacher, and Rucker (2002) indicated there maybe two possible legitimate uses of dichotomization. One that applies to us is that they (MacCallum et al., 2002) state "..Corresponding dichotomization of quantitative scale and the analysis of group differences simply must be supported by compelling results from taxometric analyses (p.38). By taxometric they meant techniques such as mixture modeling, cluster analysis" (see page 34). Thus, our rationale for the dichotomization of our normative data was because we hypothesized that a distinct latent class of children at risk for learning disabilities would emerge.

[3]Four large elementary urban schools from two large metropolitan areas participated in this study. The sample at Year 1 was also described in Swanson et al., (2020). The sample was identified by the school district as English learners. Two of the elementary public schools in sample yielded, according to state reports, the lowest percentage in reading and math scores proficiency within the state. Minority (Hispanic/Latino) enrollment was 95% of the student body which was higher than the state average. Also, the current study included two urban charter schools also with a high Hispanic/Latino (> 95%) representation. State reports indicated that one of the charter schools at the time of testing (2017-2018) reported that only 35% of children were proficient in reading and 29% were at proficient in math. A state report on the second charter school had also indicated that only 33% of the elementary children were proficient in reading and 29% proficient in math on state measure.

[4]In addition, because children were concurrently being tested on state measures at the time of our year 2 study, we were unable to administer to several measures year 2 to children. For example, several children were not administered the Spanish Word Identification test from the Language-Survey Revised (Woodcock, Muñoz-Sandoval, & Alvarado, 2005) or the English Applied Problems subtest from the Woodcock-Johnson (Woodcock, McGrew, Schrank, & Mather, 2007), thereby reducing our sample to compared children in both year 1 and year 2. The year 1 study included a larger sample, yielding five latent classes (average achievers, poor achievers, reading disabled, English language learners, Spanish Dominant Achievers) that varied in language and achievement scores. The probability estimates indicated that 10% of the year 1 sample (N=391) was at risk for learning disabilities (below cut-off score), and approximately 40% sample reflected a second language acquisition group, not at risk for academic difficulties. The design of the year 1 study was cross-sectional instead of a longitudinal study and therefore the stability (latent transition) of children at risk for learning disabilities could not be assessed. An inspection of achievement measures in Year 1 also showed that children who performed relatively higher on those achievement measures were not participating in the Federal Lunch program (2% of the total sample) when compared to the rest of the sample. Thus, because we're interested in the stability of the children at risk for learning disabilities, while controlling as much as possible the SES classification (participating in the Federal lunch program), only children with vocabulary and achievement measures administered in both years 1 and 2 were selected for the latent transition analysis.

Table 1
Normative Scores for Total Sample for Year 1 (Wave 1) and Year 2(Wave 2

	Total Sample			**Cohort 1**			**Cohort 2**			**Cohort 3**		
	N	M	SD	N	M	SD	N	M	SD	N	M	SD
Year 1 (Wave1)												
Language												
E-PPVT	267	92.29	18.34	118	92.36	18.46	90	90.5	18.16	59	94.87	18.38
E-Express	266	103.45	28.34	118	103.63	31.26	89	102.34	27.98	59	104.74	22.49
S-TVIP	267	95.15	16.08	118	94.4	15.6	90	97.9	15.42	59	92.48	17.6
S-Express	267	79.16	20.76	118	79.32	23.83	90	78.39	19.04	59	80.02	16.52
Math												
E-WPS	267	97.79	13.84	118	103.44	10.22	90	94.36	12.75	59	91.71	17.24
E-Calcul	265	115.46	14.09	116	119.58	16.43	90	112.8	9.95	59	111.41	12.44
S-WPS	266	104.2	13.46	118	109	13.45	89	99.91	11.78	59	101.09	12.94
S-Calcul	262	98.51	11.74	116	106.79	5.72	87	93.28	12.35	59	89.96	8.78
Reading												
E-Word	267	104.34	16.21	118	106.68	15.24	90	99.84	17.27	59	106.52	15.2
E-Comp	267	93.56	18.31	118	98.09	19.49	90	89.64	17.3	59	90.5	15.4
S-Word	267	118.82	18.08	118	126.86	18.9	90	112.38	15.26	59	112.57	13.62
S-Comp	267	92.89	16.51	118	99	18.05	90	90.33	13.31	59	84.58	12.85
Attention/Fluid												
Conners	267	50.35	9.46	118	49.55	9.19	90	51.82	10.15	59	49.73	8.75
Fluid Int.	239	101.19	17	108	99.27	19.26	85	102.58	16.07	46	103.11	12.22

Continue Table 1

	Total Sample			Cohort 1			Cohort 2			Cohort 3		
Year 2 (Wave 2)	N	M	*SD*	N	M	*SD*	N	M	*SD*	N	M	*SD*
Language												
E-PPVT	267	91.40	17.69	118	91.93	18.41	90	92.45	17.52	59	88.75	16.5
E-Express	267	107.11	26.14	118	108.88	29.36	90	103.87	23.42	59	108.51	23
S-TVIP	267	91.49	18.32	118	93.04	18.17	90	88.59	17.56	59	92.82	19.5
S-Express	267	78.29	19.65	118	78.52	21.13	90	77.93	19.2	59	78.41	17.41
Math												
E-WPS	267	93.95	14.18	118	97.86	12.42	90	89.87	15.51	59	92.36	13.49
E-Calcul	267	109.27	12.86	118	112.47	11.99	90	106.45	13.51	59	107.17	12.29
S-WPS	267	97.87	12.38	118	99.58	12.17	90	96.4	12.18	59	96.69	12.86
S-Calcul	267	86.79	12.48	118	90.47	13.45	90	85.96	9.7	59	80.69	11.78
Reading												
E-Word	267	103.56	16.42	118	105.34	16.76	90	100.54	17.05	59	104.64	14.23
E-Comp	267	91.77	15.84	118	94.96	14.52	90	88.07	17.33	59	91.03	14.91
S-Word	248	118.81	15.84	111	122.85	15.52	84	115.13	16.7	53	116.17	13.16
S-Comp	264	85.69	14.55	118	91.98	13.79	90	82.58	13.57	56	77.43	11.79
Attention/Fluid												
Conners	263	51.33	9.91	115	51.75	10.03	89	51.86	10.76	59	49.71	8.16
Fluid Int.	267	98.53	16.28	118	99.27	17.74	90	97.58	16.07	59	98.49	13.51

Continue Table 1

Year 2 (Wave 2) – Year 1 (Wave 1)

	Total Sample			Cohort 1			Cohort 2			Cohort 3		
Language	N	M	SD	N	M	SD	N	M	SD	N	M	SD
E-PPVT	267	-0.88	15.07	118	-0.42	14.38	90	1.95	15.64	59	**-6.12**	14.44
E-Express	266	3.64	23.78	118	**5.25**	27.67	89	1.42	21.41	59	3.77	18.17
S-TVIP	267	-3.67	17.97	118	-1.36	17.06	90	**-9.31**	16.77	59	0.34	19.64
S-Express	267	-0.86	15.47	118	-0.8	17.52	90	-0.46	14.55	59	-1.61	12.37
Math												
E-WPS	267	-3.84	12.41	118	**-5.58**	12.36	90	-4.49	11.07	59	0.65	13.53
E-Calcul	265	**-6.22**	14.61	116	**-7.12**	16.15	90	**-6.35**	14.01	59	-4.24	12.15
S-WPS	266	**-6.37**	13.57	118	**-9.42**	13.03	89	-3.63	13.6	59	-4.41	13.55
S-Calcul	262	**-11.78**	11.99	116	**-16.21**	12.21	87	**-7.58**	9.92	59	**-9.27**	11.56
Reading												
E-Word	267	-0.77	11.87	118	-1.34	12.6	90	0.70	11.28	59	-1.88	11.18
E-Comp	267	-1.79	13.71	118	-3.12	14.91	90	-1.56	13.14	59	0.53	11.81
S-Word	248	0.95	16.13	111	-2.81	16.29	84	***4.01***	15.89	53	3.97	14.75
S-Comp	264	**-7.34**	12.97	118	**-7.02**	13.71	90	**-7.75**	13	56	**-7.36**	11.43
Attention/Fluid												
Conners	263	0.99	9.85	115	2.18	8.29	89	0.12	11.25	59	-0.02	10.28
Fluid Int.	239	-0.93	18.46	108	1.66	20.61	85	-3.28	18.2	46	-2.68	12.01

Note. Cohort 1=grades 1-2, Cohort 2=grades 2-3, Cohort 3-grades 3-4; E-English, S-=Spanish, PPVT=Peabody Picture Vocabulary Test, TVIP= The Test de Vocabulario en Imágenes Peabody; Expressive=One-Word Expressive Vocabulary Test, WPS=Word Problem Solving test from WJ or Batería, Calculation=Arithmetic Subtest from WJ or Batería, Word=Word Identification subtest from Woodcock-Muñoz, Comp=Passage Comprehension subtest from Woodcock-Muñoz . Conners=Conners Behavior Rating Scale (T-score), Fluid Int.=Raven Colored Progressive Matrices Test. Bold=absolute values > .5.00.

Table 2

Fit Indices for Seven Latent Class Models for Year 1 (Wave 1) and Year 2 (Wave 2).

Wave	**LC=1**	**LC=2**	**LC=3**	**LC=4**	**LC=5**	**LC=6**	**LC=7**
Year 1							
Log-likelihood:	-1694.08	-1694.08	-1561.12	-1518.46	-1501.85	-1493.7	-1475.17
G-squared:	1043.04	1043.04	777.12	691.81	658.59	642.28	605.22
AIC:	1071.04	1071.04	865.12	809.81	806.59	820.28	813.22
BIC:	1121.26	1121.26	1022.96	**1021.46**	1072.05	1139.54	1186.29
CAIC:	1135.26	1135.26	1066.96	1080.46	1146.05	1228.54	1290.29
Adjusted BIC:	1076.87	1076.87	883.45	**834.4**	837.42	857.36	856.55
Entropy:	1	1	0.78	0.77	0.81	0.78	0.91
Degrees of freedom:	16369	16369	16339	16324	16309	16294	16279
LMR (p-values)		0	0	0.02	0.58	0.11	0.45
Bootstrap (p-values)		0	0	0	0.09	0.24	0.21
Year 2 (Wave 2)							
Log-likelihood:	-1970.34	-1804.37	-1732.63	-1701.5	-1686.83	-1669.44	-1656.61
G-squared:	1429.89	1097.68	954.19	891.93	862.6	827.82	802.15
AIC:	1457.89	1155.68	1042.19	1009.93	1010.6	1005.82	1010.15
BIC:	1508.11	1259.71	1200.03	**1221.57**	1276.06	1325.09	1383.22
CAIC:	1522.11	1288.71	1244.03	1280.57	1350.06	1414.09	1487.22
Adjusted BIC:	1463.72	1167.76	1060.52	**1034.51**	1041.43	1042.9	1053.48
Entropy:	1	0.91	0.83	0.84	0.86	0.88	0.88
Degrees of freedom:	16369	16354	16339	16324	16309	16294	16279
LMR (p-values)		0	0	0.09	0.03	0.40	0.26
Bootstrap (p-values)		0	0	0.02	0.25	0.50	0.17

Note. LC=Latent Class, AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; CAIC and Adjusted BIC corrected for sample size; LMR = Lo-Mendell-Rubin Test; BLRT = Bootstrap Likelihood Ratio Test.

Table 3

Transitional Probabilities and Delta Estimates for Total Sample and by Grade Level

Tau estimates transition probabilities: Year 1 latent status (rows) by Year 2 latent status (columns)

Delta estimates status membership probabilities:

Latent Class	1	2	3	4	Latent Class	Year 1	Year 2
Total Sample							
1	0.87	0	0.13	0	1	0.32	0.28
2	0	0.46	0	0.54	2	0.35	0.16
3	0	0	1.00	0	3	0.20	0.25
4	0	0	0.05	0.95	4	0.13	0.32
Cohort 1					Latent Class		
1	0.86	0	0.09	0.04	1	0.43	0.38
2	0	0.87	0.04	0.09	2	0.42	0.37
3	0	0	1.00	0	3	0.12	0.17
4	0	0	0.16	0.84	4	0.04	0.09
Cohort 2					Latent Class		
1	0.68	0.08	0.24	0	1	0.26	0.18
2	0	0.44	0	0.56	2	0.38	0.2
3	0	0	0.96	0.04	3	0.27	0.32
4	0	0.17	0	0.83	4	0.09	0.3
Cohort 3					Latent Class		
1	0.85	0	0.15	0	1	0.21	0.18
2	0	0.06	0	0.94	2	0.23	0.01
3	0	0	1.00	0	3	0.26	0.29
4	0	0	0	1.00	4	0.31	0.52

Note. LC1=balanced bilingual average achiever, LC2=unbalanced bilingual average achiever-, LC3=at risk for learning disabilities, LC4=English dominant

Table 4

Effect Sizes on Manifest Variables as a Function of Latent Classes

Year 1 (Wave 1)	LC1 vs. LC2	LC1 vs. LC3	LC1 vs. LC4	LC2 vs. LC3	LC2 vs. LC4	LC3 vs. LC4
Language						
E-PPVT	**-0.73**[a]	**0.84**	**-0.85**	**1.59**	-0.07	**-1.76**
E-Express	**-0.98**	**0.52**	**-0.97**	**1.64**	0.13	**-1.66**
S-TVIP	**0.91**	**0.75**	**1.16**	-0.07	0.39	0.42
S-Express	**1.49**	**0.50**	**1.87**	**-1.05**	0.20	**1.37**
Math						
E-WPS	0.01	**1.36**	0.03	**1.27**	0.03	**-1.17**
E-Calcul	-0.13	**0.92**	-0.01	**1.06**	0.18	**-1.12**
S-WPS	**1.10**	**1.57**	**0.82**	0.55	-0.02	-0.46
S-Calculation	-0.18	**0.94**	0.47	**1.05**	**0.67**	**-0.59**
Reading						
E-Word	0.00	**1.42**	0.16	**1.26**	0.14	**-1.26**
E-Comp	-0.32	**1.72**	0.13	**1.94**	0.41	**-1.48**
S-Word	0.49	**1.07**	**1.20**	**0.62**	**0.71**	-0.08
S-Comp	**0.82**	**1.28**	**1.59**	**0.50**	**0.73**	0.15
Attention/Fluid						
Conners	0.06	**-0.77**	-0.47	**-0.78**	-0.49	0.24
Fluid Int.	-0.31	**0.87**	-0.09	**1.17**	0.25	**-1.02**
Year 2 (Wave 2)						
Language						
E-PPVT	**-0.92**	**1.00**	**-0.74**	**2.19**	0.28	**-1.98**
E-Express	**-0.62**	**0.98**	**-0.63**	**1.72**	0	**-1.70**
S-TVIP	**0.83**	**1.04**	**1.39**	0.26	**0.64**	0.38
S-Express	**2.04**	**1.07**	**2.62**	**-0.87**	0.45	**1.34**
Math						
E-WPS	-0.36	**1.62**	-0.25	**1.78**	0.14	**-1.83**

E-Calcul	-0.17	**1.38**	0.12	**1.32**	0.28	**-1.3**
S-WPS	0.06	**1.31**	0.48	**1.22**	0.41	**-0.69**
S-Calcul	-0.23	**1.06**	**0.56**	**1.36**	**0.86**	**-0.57**
Reading						
E-Word	**-0.52**	**1.63**	-0.05	**1.97**	**0.52**	**-1.80**
E-Comp	-0.49	**1.93**	-0.21	**2.35**	0.32	**-2.28**
S-Word	0.02	**1.24**	**0.78**	**1.10**	**0.72**	**-0.51**
S-Comp	**0.62**	**1.54**	**2.19**	**1.13**	**1.84**	0.22
Attention/Fluid						
Conners	0.04	**-0.86**	**-0.65**	**-0.88**	**-0.66**	0.10
Fluid Int.	**-0.64**	**0.61**	**-0.51**	**1.29**	0.18	**-1.19**

Note. [a]a negative effect size under LC1-LC2 means LC1 has a lower performance than LC2;
LC1=balanced bilingual average achiever, LC2=unbalanced bilingual average achiever-, LC3=at risk for learning disabilities, LC4=English dominant; Bold=Effect size .50 or greater. E-English, S-=Spanish, PPVT=Picture Vocabulary Test, Expressive=One-Word Expressive Vocabulary Test, WPS=Word Problem-Solving test from WJ or Batería, Calcul=Arithmetic Subtest from WJ or Batería, Word=Word Identification subtest from Woodcock-Muñoz, Comp=Passage comprehension subtest from Woodcock-Muñoz. Conners=Conners Behavior Rating Scale (T-score), Fluid Int.=Raven Colored Progressive Matrices Test.

Table 5

Latent Class as a Function of Item Probabilities (Rho Estimates)

Rho estimates (item-response probabilities):

Latent Class	LC1	LC2	LC3	LC4
Language				
E-PPVT	0.4084	0.0883	**0.7898**	0.0935
E-Express	0.2779	0.0691	0.5356	0.0311
S-TVIP	0.0345	0.2694	0.3591	**0.7019**
S-Express	0.1046	**0.7733**	0.3805	**0.9709**
Math				
E-WPS	0.095	0.0665	**0.6104**	0.1346
E-Calcul	0.0294	0.0087	0.1502	0.0104
S-WPS	0.0145	0.0209	0.2291	0.1721
S-Calcul	0.1453	0.0000	0.4695	0.3109
Reading				
E-Word	0.0086	0.0519	0.415	0
E-Comp	0.1147	0.076	**0.806**	0.0863
S-Word	0.0515	0.0205	0.207	0.07
S-Comp	0.0543	0.1988	0.5164	**0.7675**
Attention/Fluid				
Conners	0.0866	0.1319	0.3271	0.2698
Fluid Int.	0.3052	0.1522	0.5582	0.1364

Note. LC1=balanced bilingual average achiever, LC2=unbalanced bilingual average achiever-, LC3=at risk for learning disabilities, LC4=English dominant; E-English, S-=Spanish, PPVT=Picture Vocabulary Test, Expressive=One-Word Expressive Vocabulary Test, WPS=Word Problem-Solving test from WJ or Batería, Calcul=Arithmetic Subtest from WJ or Batería, Word=Word Identification subtest from Woodcock-Muñoz, Comp=Passage Comprehension subtest from Woodcock-Muñoz. Conners=Conners Behavior Rating Scale (T-score), Fluid Int.=Raven Colored Progressive Matrices Test. Bold=probabilities .60 or better.

Table 6

Effect Sizes on Cognition Measures as a Function of Latent Class Comparisons

	LC1 vs. LC2	LC1 vs. LC3	LC1 vs. LC4	LC2 vs. LC3	LC2 vs. LC4	LC3 vs. LC4
Year 1 (Wave 1)						
E-STM	0.10	**0.83**	-0.30	**0.75**	-0.38	**-1.01**
S-STM	0.28	0.46	0.10	0.21	-0.17	-0.36
E-speed	0.20	**-0.56**	0.47	**-0.67**	0.26	**0.97**
S-speed	-0.38[a]	-0.23	-0.33	0.08	-0.02	-0.09
E-inhib	-0.30	-0.04	-0.40	0.26	-0.01	-0.36
S-inhib	-0.18	-0.32	**-0.51**	-0.16	-0.35	-0.17
E-WM	-0.29	**0.62**	**-0.66**	**1.02**	-0.42	**-1.28**
S-WM	**0.58**	0.41	0.26	-0.13	-0.28	-0.14
Visual	0.11	0.23	-0.19	0.13	-0.31	-0.41
Year 2 (Wave 2)						
E-STM	-0.39	**0.97**	-0.47	**1.40**	-0.09	**-1.37**
S-STM	-0.01	**0.64**	0.01	**0.65**	0.02	**-0.63**
E-speed	0.49	**-0.67**	**0.62**	**-1.05**	0.14	**1.23**
S-speed	-0.43	-0.43	**-0.57**	-0.11	-0.22	-0.08
E-inhib	0.19	0.35	-0.13	0.16	-0.30	-0.44
S-inhib	0.16	0.31	0.01	0.17	-0.13	-0.28
E-WM	-0.49	**0.77**	**-0.54**	**1.18**	-0.04	**-1.21**
S-WM	0.24	**0.89**	0.48	**0.60**	0.23	-0.35
Visual	-0.29	0.24	-0.28	**0.54**	-0.02	**-0.50**

Note. [a]a negative effect size under LC1-LC2 means LC1 has a lower performance than LC2
LC1=balanced bilingual average achiever, LC2=unbalanced bilingual average achiever-, LC3=at risk for learning disabilities, LC4=English dominant, E-=English,, S-=Spanish. STM=Short-Term Memory, Speed=Naming Speed, Inhib=Inhibition-Random Generation task, WM=Working Memory executive system, Visual=Visual-spatial WM. Bold=effect size at absolute .50 or greater

Table 7

Model Predicting Latent Class from Cognitive Variables from Year 1 (Wave 1) and Year 2 (Wave 2) (N=267)

Model		Unconditional		1		2		3		4		5	
Fixed Effects		Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Intercept	LC1	-1.23***	0.24	-1.27***	0.21	-1.38***	0.22	-1.33***	0.21	-1.10***	0.22	-1.23***	0.2
Intercept	LC2	-0.44	0.23	-0.48*	0.20	-0.54*	0.21	-0.51**	0.19	-0.23	0.21	-0.37	0.19
Intercept	LC3	0.79**	0.23	0.77***	0.2	0.79**	0.21	0.81***	0.2	1.19***	0.22	1.01***	0.2
Gender				-0.27*	0.12	-0.35**	0.13	-0.30*	0.13	-0.29*	0.13	-0.25*	0.13
Grade				-0.54**	0.19	-0.75**	0.21	-0.70***	0.19	-0.76***	0.18	-0.72***	0.18
E-STM						-0.12	0.17			-0.29	0.16		
S-STM						0.14	0.16			0.09	0.15		
Interaction						-0.04	0.11			-0.12	0.11		
E-speed						0.10	0.16			0.34*	0.16	0.52*	0.14
S-Speed						-0.57***	0.17	-0.40***	0.13	-0.62***	0.16	-0.68***	0.14
Interaction						0.18	0.1			0.05	0.09		
E-Inhibition						-0.09	0.14			-0.008	0.14		
S-Inhibition						-0.15	0.15			0.002	0.14		
Interaction						0.04	0.11			-0.26*	0.11	-0.21**	0.1
E-Exec WM						-0.35*	0.19	.43***	0.16	-0.12	0.16		
S-Exec WM						0.34*	0.17	.39***	0.15	0.67***	0.17	.45**	0.14
Interaction						-0.14	0.16			-0.08	0.14		
Visual-spatial WM						0.2	0.1453			-0.23	0.15		
Error Variance		.94***	0.38	.54*	0.26	0.34	0.21	.41*	0.22	0.28	0.19	0.29	0.19
Fit													
Deviance		686.68		673.42		639.83		649.51		617.37		628.84	
AIC		694.68		685.42		677.83		667.51		655.37		648.84	
BIC		700		693.41		703.14		679.5		680.68		662.16	

Note. *p < .05. **p < .01,***p < .001, E-English, S-Spanish, STM=short-term memory or phonological loop, Exec=executive component of working memory.

Appendix A

Demographic Information and Mean Scores for Manifest and Predictor Variables as a Function Latent Classes

Latent Class		**LC1**		**LC2**		**LC3**		**LC4**	
Sample size		74		39		65		89	
Gender		%		%		%		%	
	Female	60		67		40		45	
	Male	40		33		60		55	
Cohorts		%		%		%		%	
	1	61		69		32		24	
	2	24		28		43		37	
	3	15		3		25		35	
Variable		M	*SD*	M	*SD*	M	*SD*	M	
Year 1 (Wave 1)									
Language									
E-PPVT		89.56	15.92	101.2	15.84	76.45	15.39	102.21	14.12
E-expressive		95.17	26.46	121.19	26.36	82.37	22.04	118.11	21.18
S-TVIP		105.47	14.34	93.44	10.75	94.4	15.2	87.87	15.75
S-expressive		95.38	19.65	68.14	15.28	85.99	17.97	65.51	12.19
Math									
E-WPS		101.79	10.02	101.81	11.3	85.88	13.28	101.38	13.18
E-calculation		118.67	13.59	120.26	8.77	104.25	17.87	118.79	7.83
S-problem solving		113.35	9.94	102.54	9.63	96.62	11.45	102.85	14.67
S-calculation		102.6	9.48	104.21	7.84	91.09	14.67	98.14	9.49
Reading									
E-wordID		109.79	12.97	109.8	16.87	90.21	14.67	107.74	13.34
E-compre		99.43	14.06	104.07	15.9	75.15	14.23	97.52	15.78
S-wordID		130.84	17.78	122.36	15.87	111.43	18.63	112.67	12.6

S-compre		105.68	11.47		95.41	14.36		87.44	16.87		85.13	14
Attention/Fluid												
Conners[a]		47.36	7.17		46.97	6.54		54.11	10.32		51.58	10.35
RAVEN		102.82	16.11		107.89	16.12		88.43	17.09		104.14	14.44
Cognition												
STM												
E-STM		0.03	1.36		-0.10	1.21		-1.17	1.53		0.5	1.74
S-STM		0.21	1.6		-0.24	1.55		-0.62	1.98		0.05	1.75
Speed												
E-speed		0.25	1.36		-0.02	1.27		1.24	2.16		-0.31	1.05
S-speed		-0.13	1.3		0.36	1.26		0.23	1.87		0.39	1.8
Inhibition												
E-inhib		-0.17	0.97		0.21	1.69		-0.13	0.99		0.22	0.96
S-inhib		-0.18	0.7		-0.06	0.61		0.05	0.73		0.17	0.68
Executive WM												
E-WM		-0.32	1.19		0.01	0.97		-1.00	0.98		0.5	1.29
S-WM		0.23	1.36		-0.49	0.99		-0.33	1.39		-0.13	1.39
Visual-WM												
Visual		-0.12	1.22		-0.25	1.06		-0.42	1.42		0.11	1.19
Year 2 (Wave 2)		M	*SD*		M	*SD*		M	SD		M	*SD*
Language												
E-PPVT		89.24	16.25		103.79	14.7		74.69	12.4		99.98	13.02
E-expressive		105.53	23.49		119.85	22.03		83.93	20.17		119.77	21.65
S-TVIP		105.04	16.03		92.34	13.86		88.42	15.85		82.1	16.87
S-expressive		98.38	15.99		68.54	11.61		81.2	16.08		63.75	10.34
Math												
E-WPS		96.44	11.09		100.98	15.23		79.63	9.52		99.25	11.51
E-calculation		113.16	9.36		115.07	13.57		97.61	13.07		112	9.38
S-WPS		103.16	10.18		102.54	10.65		89.41	10.8		97.6	12.54

S-calculation		91.75	11.28		93.96	5.92		78.59	13.53		85.51	11.07
Reading												
E-wordID		107.78	12.24		114.48	14.05		85.6	15.03		108.4	10.58
E-compre		96.01	11.33		101.43	10.77		72.32	13.24		98.23	9.81
S-wordID		126.52	8.76		126.36	9.12		108.08	19.81		116.79	14.72
S-compre		98.46	9.84		93.13	5.63		79.20	15		76.49	10.19
Attention/Fluid												
Conners[a]		47.51	7.89		47.16	7.58		54.86	9.32		53.81	11.04
RAVEN		97.06	15.99		106.72	13.41		87.30	15.91		104.36	12.99
Cognition												
STM												
E-STM		0.11	1.36		0.65	1.4		-1.13	1.18		0.79	1.54
S-STM		0.33	1.51		0.34	1.5		-0.65	1.54		0.31	1.5
Speed												
E-speed		-0.57	0.81		-0.95	0.69		0.08	1.12		-1.05	0.74
S-speed		-0.80	0.87		-0.45	0.72		-0.33	1.32		-0.23	1.1
Inhibition												
E-inhib		0.07	0.89		-0.10	0.86		-0.25	0.94		0.2	1.07
S-inhib		0.24	0.63		0.14	0.59		0.02	0.8		0.23	0.7
Executive WM												
E-WM		0.71	1.25		1.42	1.76		-0.18	1.03		1.48	1.56
S-WM		0.62	1.41		0.27	1.61		-0.55	1.21		-0.07	1.45
Visual-WM												
Visual		0.32	1.39		0.72	1.4		0	1.3		0.75	1.63

Note. Cohort 1 (grades 1 to 2), Cohort 2 (grades 2-3) and Cohort 3 (grades 3 to 4). Manifest variables are in Normed Scores (M=100, *SD*=15) and Cognition measures are in Z-scores. E=English, S=Spanish. LC1=balanced bilingual average achiever, LC2= unbalanced bilingual average achiever, LC3=children at risk for learning disabilities, LC4=English dominant, WPS=Word Problem Solving test from WJ or Batería, Calculation=Arithmetic Subtest from WJ or Batería, PPVT=Peabody Picture Vocabulary Test, TVIP= The Test de Vocabulario en Imágenes Peabody, Word=Word Identification subtest from Woodcock-Muñoz, Comp=Passage Comprehension subtest from Woodcock-Muñoz . Conners=Conners Behavior Rating Scale (T-score), Fluid Int.=Raven Colored Progressive Matrices Test. The Test de Vocabulario en

Imágenes Peabody; Expressive=One-Word Expressive Vocabulary Test, WPS=Word Problem Solving test from WJ or Batería, Calculation=Arithmetic Subtest from WJ or Batería, Word=Word Identification subtest from Woodcock-Muñoz, Comp=Passage Comprehension subtest from Woodcock-Muñoz . Conners=Conners Behavior Rating Scale (T-score), Fluid Int.=Raven Colored Progressive Matrices Test. STM=Short-Term Memory or Phonological Loop; Speed=Naming Speed, Inhib=Inhibition or Random Generation Tasks, Exec=Executive component of working memory, WM=working memory. [a]Conners Behavior Rating Scale is a T-score (M=50, *SD*=10).