Abstract Title Page

Title: Using Early Literacy Profiles of Hispanic English Language Learners to Predict Later Reading Achievement

Author(s): Francis Huang, Karen Ford, and Marcia Invernizzi

Abstract Body

Background / Context:
Spanish speakers comprise 80% of the English language learners (ELLs) in U.S. schools, and many of these children are considered at risk for academic failure, largely due to low levels of literacy achievement (August & Shanahan, 2006). In order to address this problem, we must be able to identify potential reading problems early so that effective, focused intervention can be provided. Previous research has identified phonological awareness, print knowledge, and orthographic knowledge as robust predictors of later reading achievement for both native English speakers and ELLs (Scarborough, 1998; Chiappe, Siegel, & Wade-Woolley, 2002; Lesaux, Koda, Siegel, & Shanahan, 2006). However, much of the research on ELL literacy development has looked at ELLs as a group, without investigating whether there is within-group heterogeneity that affects literacy outcomes. Recent research (Burrow, Cabell, Konold, & Invernizzi, 2010) on Spanish-speaking ELLs, however, has shown that not all ELLs represent one homogenous group in terms of early literacy profiles. Using cluster analysis (Aldenderfer & Blashfield, 1985), Burrow et al. demonstrated that Spanish-speaking ELLs in the fall of kindergarten represented a heterogeneous group that could be classified into four distinct clusters using three broad measures of early literacy abilities: phonological awareness, orthographic awareness, and alphabet knowledge. The four cluster profiles were described as: 1) the highest early literacy group (i.e., high performing across all literacy domains), 2) the average phonological and orthographic awareness with high alphabet knowledge group, 3) the average phonological awareness with low alphabet and orthographic awareness group, and 4) the lowest literacy group (i.e., low performing across all literacy domains).

Purpose / Objective / Research Question / Focus of Study:
Following a cohort of students from fall of kindergarten to spring of first grade, we investigated whether the general cluster profiles identified in Burrow et al.'s (2010) study remained consistent over time. If cluster profiles in the fall of kindergarten held, our goal was to investigate the differences between clusters in terms of their instructional reading levels at the end of the first grade. Because we could expect that the highest literacy cluster would remain significantly higher at the end of the first grade given that early literacy predicts achievement in succeeding grades (Juel, 1988), we investigated the performance of the other three lower performing groups.

We had two research questions:

1) Would research-based cluster profile patterns remain consistent with the current dataset used (i.e., were the clusters generated similar in characteristics based on early literacy skills)?
2) If clusters profiles were consistent, how would the three lower ranked clusters perform based on their instructional reading levels at the end of first grade, while controlling for other student-level variables?

Setting:
The study involved the secondary data analyses from restricted-use datasets from a Mid-Atlantic state.

Population / Participants / Subjects:
A cohort of 1,424 students (49% female, 51% male) from 271 public schools in a Mid-Atlantic state was followed from kindergarten (fall 2008) through the end of first grade (spring 2010). All of the students were identified by their schools as Hispanic and spoke Spanish at home. Mean age of the participants in the fall of kindergarten was 69 months, and 81% of the participants were classified as economically disadvantaged using the free and reduced price lunch (FRPL) eligibility status. In kindergarten, 96% had no diagnosed disability (i.e., intellectual disabilities, speech/language impairment, developmental delay), and 61% were receiving English as a Second Language services (ESL). By the spring of first grade, 93% of participants had no diagnosed disabilities, and 68% were receiving ESL services. Student descriptive characteristics are presented in Table 1. *[insert table 1 about here]*

Intervention / Program / Practice:
All participants were administered, by their classroom teachers, the Phonological Awareness Literacy Screening (PALS) assessment in kindergarten and in first grade. PALS for kindergarteners (PALS K; Invernizzi, Juel, Swank, Meier, 2003) and PALS for grades one through three (PALS 1-3; Invernizzi, Meier, & Juel, 2003) were designed to identify students who are below grade-level expectations in important literacy fundamentals and identifies students who are at risk of reading difficulties and delays (Invernizzi, Justice, Landrum, & Booker, 2004). PALS is a criterion-referenced assessment administered by the teacher. Some tasks are administered individually, and others are administered in small groups (i.e., five or fewer children).

The tasks we examined in the fall of kindergarten were:

Rhyme awareness. The rhyme awareness task is administered to a group of five or fewer students. Students are shown a series of four pictures of high-frequency words appropriate for kindergartners. All are single-syllable words that can represented pictorially (e.g., tire, book, pot, fire). The names of all the pictures are read aloud by the teacher and students are asked to circle the picture that rhymes with the first picture in the series (i.e., tire, fire). There are ten sets of pictures and students can score a maximum of 10 points, one per set of pictures.

Beginning Sound Awareness. In the beginning sound awareness task, students are shown another series of four pictures of high-frequency words appropriate for kindergartners (e.g., fish, mop, bell, fence). The names of all the pictures are read aloud by the teacher and students are asked to circle the picture that has the same beginning sound (i.e., fence) as the first picture in the series (i.e., fish). There are ten sets of pictures and students can score a maximum of 10 points, one per set of pictures.

Alphabet recognition. In the alphabet recognition task, students are individually shown 26 randomly ordered, lowercase letters and are asked to point to and name each letter. Only lowercase letters are used since lower-case letter recognition is highly correlated with upper-case letter recognition (r=.94 in kindergarten), and there is no ceiling effect when lower-case letters are used (Invernizzi, Juel, & Meier, 2003-2009). Students can score a maximum of 26 points.

Letter sound knowledge. Children are individually shown 23 uppercase letters and three digraphs (ch, th, and sh). (Because the letters Q and X cannot be pronounced in isolation, they are not included. The letter M is used as a practice item). Students are asked to touch each letter on the page and say the sound it represents (e.g., S makes /s/). Only the short vowel sound for each vowel is scored as correct and only the hard sounds for C and G are scored as correct. A total of 26 points is possible.

Invented spelling. The spelling task can be administered to groups of five children or fewer. The task was designed to assess student performance on the representation of beginning, middle, and ending speech sounds. Children are asked to write or attempt to write a series of five high-frequency, single-syllable, consonant-vowel-consonant (CVC) words (e.g., fan, pet), which the teacher reads aloud, one at a time. The child receives points for providing a phonetically correct representation for each sound in the word, with an additional point awarded for spelling the word correctly. Students can score a maximum of 20 points on the task.

In the spring of first grade, students were administered PALS 1-3 which consisted of a different set of task from PALS-K. The main outcome of interest for our study was the instructional oral reading level (M=1.46, SD=1.12, min=0, max=6) of the child, which was determined by combined performance on the *PALS 1-3* Word Reading in Isolation and Oral Reading in Context tasks. Instructional oral reading level was calculated as the highest level at which a student can accurately read at least 75% of the words on the Word Recognition in Isolation list and 90% to 97% (85% to 97% at the Readiness or Preprimer level) of the words on the Oral Reading in Context passage. The oral reading levels addressed in this study were Readiness, Preprimer A, Preprimer B, Preprimer C, Primer, and first through sixth grades. Beginning at first grade, a one-point increment represents one grade level in reading.

The tasks we examined in the spring of first grade were:

Word Recognition in Isolation. The Word Recognition in Isolation task consists of eight graded word lists (Preprimer through sixth grade). The student begins by reading the list that corresponds to the benchmark for his or her assigned grade level and then moves up or down to determine the highest list which can be read with 75% accuracy.

Oral Reading in Context. For this task, the student reads a minimum of three graded passages: the passage that corresponds to the highest word list read with 75% accuracy, the one above, and the one below.

Research Design:
Our research consisted of secondary data analyses.

Data Collection and Analysis:
The data for the present study come from merging two state-provided restricted-use datasets that could be used to track early literacy outcomes using PALS scores. We narrowed down our sample based on only those students who were identified as Hispanic and spoke Spanish at home as their primary language.

We had two main steps in our analyses: the first was to perform cluster analyses on our participants in kindergarten to create the student cluster groupings. The second was to compare the three of the four identified clusters to investigate their instructional reading levels at the end of first grade using multilevel modeling.

Clustering strategy
We used five measures of early literacy from the PALS-K to form our clusters: two measures of phonological awareness (a) beginning sound awareness and (b) rhyme awareness; two measures of alphabet knowledge: (a) letter name recognition and (b) letter sound recognition; and one measure of orthographic knowledge (a) invented spelling. Because the measures were on different scales, all measures were converted to *T*-scores with a mean of 50 and a standard deviation of 10, as recommended by several researchers (e.g., Aldenderfer & Blashfield, 1985; McDermott, 1998). Standardization was done so the different measures contribute equally to the similarities among subjects (Romesburg, 1986). A hierarchical-agglomerative clustering procedure using Ward's (1963) method was employed using SAS PROC CLUSTER. Since there are no completely satisfactory methods for determining the number of population clusters (i.e., there is no test of a null hypothesis) for any type of cluster analysis (Bock 1985; Everitt 1979), we determined the number of clusters by examining the cut points in the resulting dendogram (i.e., tree diagram), as was done by Peck (2005), who had a similar goal of creating comparison subgroups that were readily interpretable. We also used a more formal, but still heuristic, method of using the fusion or amalgamation coefficients produced in the tree diagram (Aldenderfer & Blashfield, 1985), and the method is similar to using a "scree plot" in order to determine the number of factors to retain in factor analysis (Overall & Magee, 1992).

After generating the clusters, we described the cluster characteristics based on variables not used to generate the cluster groupings: age in months, economic status, whether the child was disabled, and whether the child was receiving ESL services.

Modeling strategy
We modeled the students' instructional reading level at the end of the first grade (our dependent variable) on the students identified cluster (our independent variable), while holding other student level variables constant (i.e., age in months, disability status, receiving ESL services, and economic disadvantage). Cluster levels were dummy coded with cluster 3 being the reference group. Age in months, our only continuous covariate, was grand mean centered. We used two-level multilevel modeling (Luke, 2004; Raudenbush & Bryk, 2002) using SAS PROC MIXED to generate the appropriate standard errors to account for the students nested in schools. As the clusters, and not the schools, were our primary area of interest, we did not include school-level covariates.

Findings / Results:
Results of the cluster analysis revealed four distinct cluster profiles. The cluster profiles are presented in Figure 1. The cluster profiles mirrored the pattern of early literacy skills found in the Burrow et al. (2010) study, despite using data from another time period. One difference, however, is that our cluster 2 had above average phonological awareness skills (+.25 *SD*) as compared to the average phonological awareness level in cluster 2 in the Burrow et al. study. The

descriptive profiles of each of the clusters are presented in Table 2. *[insert figure 1 and table 2 about here]*

In the reduced sample (excluding cluster 1), the reading level mean was 1.29 (*SD*=1), indicating that by the end of first grade, our participants were reading slightly above grade level. Table 3 shows results from the full multilevel modeling analyses that used reading level as the outcome variable, using clusters 2 through 4 as our independent variable.

Findings of the multilevel modeling results revealed that differences in reading levels between cluster 3 and cluster 4 were not statistically significant, despite cluster 3 having higher phonological awareness skills in fall of kindergarten. In the two lowest performing groups there were no statistically significant differences in alphabet recognition, letter sounds, or invented spelling in the fall of kindergarten. The inclusion of the independent variable and the covariates accounted for approximately 15% of the overall variance. *[insert table 3 about here]*

Cluster 2, on the other hand, had significantly higher reading levels compared to the other two groups. The difference in reading level between cluster 2 and the two lower clusters amounted to an average of 2/3 of a grade level (*b*=.65).

Conclusions:
Our findings support earlier research (Burrow et al., 2010) that demonstrated heterogeneity among Spanish-speaking ELLs in kindergarten. Further, we found that the same clusters that were identified in the Burrow et al. study remained consistent in our sample, even though we used data from another time period. When we examined reading level for the lower three clusters in our sample, we found that clusters 3 and 4, the two lowest clusters, were achieving at approximately the same level in reading at the end of first grade, while the mean reading level for cluster 2 was approximately 2/3 of a grade level higher.

What distinguished cluster 2 from the other two clusters was higher performance on print-related tasks in fall of kindergarten. Although both cluster 2 and cluster 3 had average or above-average phonological awareness in fall of kindergarten, cluster 2 performed significantly better in alphabet knowledge and orthographic knowledge, the two domains made up of print-related skills (i.e., letter name and letter sound knowledge, spelling). This finding is consistent with previous research (Hammill, 2004; Scarborough, 1998) identifying written language skills as the most accurate predictors of later reading achievement, once children have begun formal literacy instruction.

These findings have clear implications for literacy screening and instruction for Spanish-speaking ELLs. Early literacy screening instruments must include measures of alphabet knowledge and orthographic knowledge in order to best identify children at risk for developing reading problems (Scarborough, 1998). In addition, early literacy instruction should focus on print-related skills in order to ensure that ELLs are well-prepared for the task of learning to read (Hammill, 2004).

Appendices

Appendix A. References

Aldenderfer, M., & Blashfield, R. (1984). *Cluster analysis.* Beverly Hills, CA: Sage Publications.

August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth.* Mahwah, NJ: Erlbaum.

Bock, H. (1985). On some significance tests in cluster analysis. *Journal of Classification, 2,* 77-108.

Burrow, L., Cabell, S., Konold, T., & Invernizzi, M. (2010, April). Using cluster analysis to explain heterogeneity in Hispanic English language learners at kindergarten entry. Invited poster session at the International Reading Association Conference, Chicago, IL.

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.

Everitt, B.S. (1979). Unresolved problems in cluster analysis. *Biometrics*, 35, 169 -181.

Hammill, D. D. (2004). What we know about correlates of reading. *Exceptional Children, 70*, 453-468.

Invernizzi, M., Juel, C., Swank, L., & Meier, C. (2003). *Phonological awareness literacy screening: Kindergarten (PALS-K)*. Charlottesville, VA: University Printing Services.

Invernizzi, M., Meier, J., & Juel, C. (2003-2009). *PALS: 1-3 Technical Reference*. Charlottesville, VA: University Printing Services.

Juel, C. (1988). Learning to read and write: A longitudinal study of fifty-four children from first through fourth grade. *Journal of Educational Psychology, 80*, 437-447.

Lesaux, N. K., Koda, K., Siegel, L. S., & Shanahan, T. (2006). Development of literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second language-learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ: Erlbaum.

Luke, D. (2004). *Multilevel modeling.* Thousand Oaks, CA: Sage.

McDermott, P. (1998). MEG: Megacluster analytic strategy for multistage hierarchical grouping with relocations and replications. *Educational Psychology and Measurement, 58,* 677-686.

Overall, J., & Magee, K. (1992). Replication as a rule for determining the number of clusters in hierarchical cluster analysis. *Applied Psychological Measurement, 16,* 119-128.

Peck, L. (2005). Using cluster analysis in program evaluation. *Evaluation Review, 2,* 178-196. doi: 10.1177/0193841X04266335

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). London: Sage.

Romesburg, H. (2004). *Cluster analysis for researchers.* North Carolina: Lulu Press.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B. K. Shapiro, P. J. Accardo & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum.* Timonium, MD: York Press.

Ward, J. (1963). Hierarchical grouping to optimize objective function. *Journal of the American Statistical Association, 58,* 236-244.

Appendix B. Tables and Figures

Table 1

Participant (*n*=1,424) descriptive statistics in kindergarten: fall (in 271 schools)

Characteristics	*n*	*%*	*M*	*SD*
Gender				
Male	723	51		
Female	701	49		
Students with disabilities	57	4		
Students receiving ESL services	864	61		
Economically disadvantaged	1,151	81		
Age (in months)			69.17	4.57

Table 2

Descriptive profiles of clusters in spring of first grade

Characteristic	Cluster 1 *n*=220 (15%)	Cluster 2 *n*=427 (30%)	Cluster 3 *n*=365 (26%)	Cluster 4 *n*=412 (29%)
Female	0.55	0.52	0.50	0.44
Economically disadvantaged	0.75	0.81	0.82	0.82
Disabled	0.04	0.05	0.07	0.10
Receiving ESL services	0.46	0.64	0.76	0.78
Age in months	88.74	87.42	86.85	86.42
Reading level	2.40	1.76	1.06	1.01

Table 3

Unconditional and full multilevel models using reading level as the dependent variable (*n*=1,204)

Variable	Unconditional model			Full model		
	Estimate		*SE*	*Estimate*		*SE*
Intercept	1.32	***	0.04	1.43	***	0.09
Cluster 2[a]				0.65	***	0.07
Cluster 4[a]				-0.05		0.07
Female (1=yes)				-0.08		0.05
Economically disadvantaged (1=yes)				-0.06		0.07
Classified with disability (1=yes)				-0.42	***	0.10
Receiving ESL services (1=yes)				-0.30	***	0.06
Age in months				-0.02	***	0.01
Variance components						
Student level	0.93	***	0.03	0.78	***	0.03
School level	0.07	**	0.04	0.07	**	0.02

Note. SE=Standard errors; ESL=English as a Second Language

[a] Cluster 3 is the reference group.

** *p*<.01, *** *p*<.001.

Figure 1

Cluster profile descriptions based on five early literacy skills (scores are *T*-scores, *M*=50, *SD*=10)

	Beginning sounds	Rhyme awareness	Alphabet recognition	Letter sounds	Invented spelling
Cluster 1	61.3	59.7	60.6	63.7	68.3
Cluster 2	55.3	54.0	57.6	55.6	52.3
Cluster 3	50.5	50.2	43.4	43.1	44.3
Cluster 4	38.1	40.5	42.4	43.0	42.9