

Selected Research on Kindergarten Ability Grouping for Reading Instruction and

Influences on Teaching Identity

Christine Ammirati

July 27, 2010

College of William and Mary

CRIN 591, Dr. Gail McEachron

Selected Research on Kindergarten Ability Grouping for Reading Instruction and Influences on Teaching Identity

In the current era of standards and high-stakes testing, kindergarten has transformed from a primarily social transition between home and school focused on “supporting children’s interests and skills” (Graue, 2010, p. 29) into an academic program “more focused on literacy and numeracy” (Graue, 2010, p. 29). Emphasis has shifted from providing a nurturing initiation into schooling to full-day programs with a full complement of content instruction. In this “era of heightened reading expectations and societal demands of all children being able to read by grade 3” (Simmons et al., 2007, p. 344), developing kindergarteners’ reading proficiency is critical.

Researchers generally agree that “regardless of individual differences at kindergarten entry, schools have a mission to promote reading achievement for all students” (McCoach, O’Connell, & Levitt, 2006, p. 339). The means to attain this end is not as clear-cut. Students bring a wide range of abilities to kindergarten. Homogeneous within-class ability grouping is one common instructional technique used to address this diversity. However, this grouping method is the subject of impassioned debate. Proponents argue that within-class reading ability groups are an effective means to increase kindergarteners’ reading performance. Critics maintain that differentiating young children upon kindergarten entry does not improve academic performance in the long term, is often based on assessments of behavior rather than reading performance, and establishes an enduring framework for segregation and sub optimal performance and self-concept.

A preservice kindergarten teacher must be prepared to implement or avoid within-class ability groups as directed by her future school administrators and, more importantly, must

unpack and reflect upon the ongoing debate about this instructional method. Is differentiating a classroom into homogeneous reading groups an effective means of delivering appropriate curriculum and scaffolding individual needs? Or is within-class ability grouping a stigmatizing structure that relegates low proficiency students to static grouping in future classes thus reinforcing low expectations, behavior challenges, and poor academic performance? How does a classroom teacher balance the contradictory perspectives and research findings and fulfill her responsibilities to her young students to optimize the quality of her instruction and maximize the potential of their first elementary school year?

Researchers routinely construct their own research designs with selected discrete samples but in the last decade they have also had the opportunity to access the database of the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K). This nationwide study followed a cohort of children entering their kindergarten year through the spring of their 8th grade year (2007). The 22,782 participating students attended 1,277 private and public, part-day and full-day kindergartens and constituted a nationally representative sample in terms of racial and ethnic backgrounds and socioeconomic status (SES). Parents, teachers, and schools were also participants in the study. Data were collected on “children’s cognitive, social, emotional, and physical development,... and children’s home environment, home educational activities, school environment, classroom environment, classroom curriculum, and teacher qualifications” (National Center for Education Statistics, 2010) and made available for public use.

Instrumentation included social /emotional development assessments, cognitive assessments, physical measures, and parent, teacher, and school questionnaires and surveys (National Center for Education Statistics, 2010; U.S. Department of Health and Human Services, 2010). With regard to homogenous or within-class ability groups, ECLS-K data show that 64% of the 1998-

1999 kindergarten cohort was placed in within-class ability groups for reading instruction (Koch, Steelman, Mulkey, & Catsambis, 2008, p.410).

Determinants and Consequences of Within-Class Ability Groups

Tach and Farkas used the ECLS-K data to investigate the determinants and consequences of placement in within-class reading ability groups in kindergarten and first grade. They hypothesized that placement in kindergarten within-class ability groups is impacted by the students' incoming level of reading and math proficiency as well as preschool behavior problems that are "more commonly observed among males [and] among children from lower socioeconomic status families," (Tach & Farkas, 2005, p. 1049). However, the researchers predicted that when these performance and behavior factors were taken into account, no significant difference would exist in ability group placement between African-American and lower SES students and white and higher SES students. Lastly, they addressed the perception evident in the literature that a teacher's race interacted "with the student's race to determine the teacher's judgment of the student's behavior" (Tach & Farkas, 2005, p. 1058).

Hierarchical linear modeling (HLM) was used to account for the nested nature of the ECLS-K data and the sample was narrowed to include only those students and classrooms where complete data existed at all analysis points (beginning of kindergarten, end of kindergarten, and end of first grade). In all, the sample totaled 11,769 kindergarteners in 2420 classrooms and, in the following year, 10,747 first graders in 3113 classrooms (Tach & Farkas, 2005, p. 1055-56). Limitations of this study include some under-representation of Hispanics due to the exclusion of students with limited English proficiency.

The researchers concluded that the strongest determinant for reading ability group placement was prior reading test scores with secondary influences of "students' prior

mathematics score performance and the teacher's judgment of the student's learning-related behaviors" (Tach & Farkas, 2005, p. 1072). Although these determinants were already established by kindergarten entry, homogenous grouping for reading instruction was found to have a statistically significant effect on behavior and existing differentials were shown to be amplified by the within-class ability group placement in kindergarten. In testing the interplay between the race of teachers and students, the researchers found no statistically significant interactions at either the kindergarten or first grade level (Tach & Farkas, 2005, p. 1072). The researchers found that ability group placement also exacerbated academic performance. That is, while ability groups were found to have an overall significant and positive effect on reading performance, in fact "the reading gains achieved by higher-placed students [were] larger than the relative losses by lower-placed students" (Tach & Farkas, 2005, p. 1072) thus creating an aggregate gain despite the decline in the lower group. Any aggregate gains were then reversed in first grade. The Tach and Farkas research showed a small statistical significance indicating that "on average, the use of ability groups lead to lower achievement in first grade" (Tach & Farkas, 2005, p. 1070). Homogenous ability groups were therefore not demonstrated to improve reading performance.

In a qualitative research study, Koch, Steelman, Mulkey, and Catsambis created vignettes to analyze the impact of behavior and gender on reading group placement. Stating that "early placement decisions may influence, with some permanency, what comes next in a student's academic life" (Koch, et al., 2008, p. 411), the researchers sought to examine the influence of teachers' gender stereotyping in assigning kindergarteners to differentiated ability groups. In the vignettes, each student was specified to have scored within a tight range of average scores on a reading achievement test. The vignettes depicted students displaying either "good," "average,"

or “bad” behavior and designated them as a girl, boy, or a gender-neutral “student”. Unbiased achievement-based placement decisions would objectively assign all students to an average reading group. Placement in the high or low ability groups was inferred to be a result of variations in student behavior and thus a reflection of the evaluator’s stereotyped beliefs.

In three separate experiments, different groups of student volunteers from a large southern university were asked to evaluate student behavior as depicted in a vignette on a five point Likert scale. To control threat to external validity, each evaluator participated in only one experiment and assessed a single randomly assigned vignette depicting one student. Analysis of variance procedures and Tukey’s posthoc procedures were applied to the results.

Serving as a pretest to establish behavior measures and confidence in the vignettes, the first experiment asked 71 evaluators to rate the behavior depicted in a single vignette. Behavior differences were found to be statistically and dramatically different and these results were used to generate one reliable behavioral score. The second experiment expanded on the behavior assessment, asking a separate pool of 225 evaluators to place a child in a high, average, or low reading ability group. Results showed that despite the fact that each student was specified to have an average score, “students exhibiting problematic behaviors, no matter how slight, [were] placed in lower reading groups than their counterparts” (Koch, et al., 2008, p. 419). However, these same results did not support the researchers’ hypothesis that boys’ behavior would be evaluated more severely than girls’ behavior. The third test repeated the second experiment’s vignettes but included gender-neutral subjects rather than designations of “boy” or “girl”. Based on the behavior depicted in the vignettes, the 114 evaluators were asked to infer the gender of the students. This data supported the hypothesis of gender stereotyping. For well-behaved subjects, 57.9% were inferred to be female compared to only 42.1% who were inferred to be male. The

“bad” behavior vignettes resulted in extreme results with only one evaluator inferring that the student was female and “a staggering 97.4 % believ[ing] the naughtiest child [to be] male” (Koch, et al., 2008, p. 421). In a fourth experiment, 17 pre-service teachers served as evaluators and duplicated the second experiment. Again, although each student was specified to have scored in the average range for reading achievement, the evaluators placed students in either high or low ability groups based on their depicted behavior.

The major limitation of this study was the use of student volunteers as evaluators. The research could be strengthened through use of a random sample and made more relevant if such a sample was drawn from a population of classroom teachers. Nonetheless, the researchers presented a compelling example of how innate prejudices concerning behavior override academic assessment.

Using ECLS-K data, Buttaro, Catsambis, Mulkey, and Steelman examined the interaction of school organization and its role in homogeneous ability grouping. The researchers asserted that kindergarten reading groups lay the foundation for middle and high school second-generation segregation, that is, the isolation within schools of ethnic and racial groups (Buttaro, Catsambis, Mulkey, & Steelman, 2010, p. 1301). They used a subsample of ECLS-K data containing both fall and spring data from the kindergarten year to examine how three distinct components of school organization impact the use of ability grouping. The subsample totaled 12,176 students, 2244 teachers, and 768 schools, the outcome variable was measured in number of minutes per week spent in homogeneous reading ability groups, and HLM was employed for statistical analysis.

The “structure” component of school organization is comprised of student sociodemographics and achievement levels as well as material resources, class size, and school

facility, size, and personnel. Structure was found to contribute to the frequency of within-class ability grouping for kindergarten reading instruction. Specifically, within-class ability grouping increased as the minority concentration of a school's student body increased. Use of homogeneous grouping was found to be characteristic of schools rather than of individual classrooms indicating that school administration rather than individual teachers was dictating the use of this instructional method. The researchers found that high-minority schools used within-class ability grouping at a much higher rate than white schools (Buttaro et al., p. 1324) and that "the diversity in the reading skills of incoming students [was] the strongest predictor of the extent to which within-class ability grouping [was] used in kindergarten" (Buttaro et al., p. 1329). In addition, the data revealed an increased rate of grouping in kindergarten classrooms with structural support in the form of paid teacher aides and full-day schedules (Buttaro et al., p. 1324).

However, the study did not support the hypothesis that the other two components of school organization contributed to a higher incidence of within-class kindergarten grouping. Homogeneous ability grouping was not shown to be affected by "politics," defined as administrative attempts to improve school performance on standardized tests and elite parents' lobbying to maintain tracking. "Cultural expectations of how to instruct disadvantaged minority youth with low levels of skills" were also not proven to influence the existence of homogeneous reading groups (Buttaro et al., p. 1311) within a school. The researchers noted that quantitative methods are insufficient to examine these two hypotheses and suggest that "qualitative data and direct observation" (Buttaro et al., p. 1330) is needed to parse the ECLS-K data set.

The Impact of Ability Groupings on Gains in Reading Proficiency

McCoach, O'Connell, and Levitt adopted a narrow focus in examining homogenous within-class ability grouping. They analyzed the specific effects of this instructional method for kindergarteners and “hypothesized a positive relationship between the frequency with which teachers reported using within-class ability grouping and children’s reading growth” (McCoach et al., 2006, p. 340).

The researchers noted that the entire ECLS-K sample measures teachers’ use of within-class ability grouping for reading instruction at just once a week on average and that “30% of the teachers reported that they never use ability grouping” (McCoach et al., 2006, p. 342). From the ECLS-K database, the researchers refined their sample to include only first-time kindergarteners and students who remained in the same school throughout their kindergarten year. In addition, they excluded learning disabled students and those children with limited-English proficiency as well as schools with fewer than 4 students in the study sample. After these adjustments, their sample totaled 10,191 kindergarteners in 620 schools. Analysis of the nested ECLS-K data used HLM to measure item response theory (IRT)-scaled reading assessment scores. Reliability of this kindergarten reading assessment was measured at .93 (fall) and .95 (spring). The researchers also accounted for school-level variables (e.g., attendance data and Title I funding) and kindergarten teacher variables (e.g., use of ability groups and measures of school climate).

McCoach et al. presented results showing “the frequency with which teachers used ability grouping in reading was associated significantly and positively with” reading improvements demonstrated by the entire school (p.342) but they acknowledged that the effect size is small (ES=.25). They also found that “school gains were higher in schools with full-day kindergartens and that schools in which teachers reported greater frequency of use of ability groups tended to have greater average gains in reading scores across the kindergarten year” (p. 343). Over the

course of the kindergarten year, private schools were not shown to have an effect in any of the models studied, girls were seen to attain greater gains in reading score assessments than boys, and “individual SES status was associated positively with gain in early literacy skills” (p. 344). Ability grouping was only one of a multitude of variables that impacted kindergarteners’ reading development.

Limitations of the study lie in the quantitative and distanced nature of the ECLS-K census data itself. Researchers have no direct contact with the sample and therefore have no observations, interviews, or other qualitative data on which to amplify data and explain results. In particular, McCoach et al. were unable to investigate the four elements that they hypothesized would impact the effectiveness of ability grouping: “curricular differentiation ... when the teacher adapts the depth, pace, or difficulty level of a lesson;” (McCoach et al., 2006, p. 340) flexible regrouping of students; grouping placement methodology; and structure and specificity of the groups, that is, what specific reading content (e.g., phonics, comprehension) is being targeted in within-class ability groups. In addition, unlike Tach and Farkas, McCoach et al. examined aggregated kindergarten data and therefore did not surmise that observed improvements in reading scores were the net effect of strong improvement for high-ability groups offsetting the losses of low-ability groups.

Vadasy, Sanders, and Peyton expanded on these findings of positive effects of homogeneous ability groups and on additional studies that demonstrated the effectiveness of code-oriented phonological interventions for high-risk emergent readers. These researchers investigated whether paraeducators could be trained to provide “explicit supplemental one-to-one instruction in alphabetic and phonemic decoding skills” to disadvantaged and low-scoring kindergarteners (Vadasy, Sanders, & Peyton, 2006, p. 509). They propose using lower-cost

paraeducators rather than classroom teachers to address budget constraints and to lessen the workload burdens of classroom teachers. In their study, Vadasy et al. asked whether compared to classroom controls, “paraeducators can effectively supplement [explicit code-oriented and phonemic awareness kindergarten instruction] for high-risk kindergarten students, and whether the effects of this instruction maintain at the 1-year follow-up” (2006, p. 510).

The sample was drawn from nine urban public elementary schools. Midway through the school year, 19 full day kindergarten teachers were asked to identify students “who would benefit from intensive additional reading instruction” (Vadasy, et al., 2006, p. 510). In contrast to some studies reviewed above, students were not excluded due to limited English proficiency or special education status. A total of 75 of the 126 referred students met the researchers’ eligibility criteria based on obtained scores of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) test. Students were then “randomly assigned within schools to supplemental or regular classroom reading instruction” (Vadasy, et al., 2006, p. 510). Attrition was a major validity threat for this study. Eight students were lost in the kindergarten year and a further 23 students were lost by the end of the students’ first grade year. Chi-square analyses of the remaining control and treatment groups “showed no reliable differences” (Vadasy, et al., 2006, p. 512) and all remaining students were included in the results.

The eleven participating paraeducators had varied “educational levels, general tutoring experience, and experience working with kindergarteners” (Vadasy, et al., 2006, p. 511) and had qualifications similar to the new standards for paraprofessionals enacted by No Child Left Behind. To ensure treatment fidelity, the researchers supplemented initial instructional training with an average of 16 observations for each paraeducator. In addition, “to ensure that classroom reading instruction did not systematically vary between treatment and control groups” classroom

teachers completed the Multigrade Inventory for Teachers and were found not to “differ significantly on reading emphasis or curricula” (Vadasy, et al., 2006, p. 511). Trained and independent evaluators administered pretests and posttests at the end of the kindergarten and at the end of first grade. Reliability measures were provided for assessments.

The intervention consisted of “a set of 62 scripted lessons (with 3-4 activities per lesson) with matched decodable texts to work from during their 30-min individual tutoring sessions” (Vadasy, et al., 2006, p. 511). These individual tutorials took place “during the school day, outside the classroom in a quiet nearby school space... 4 days per week, for 18 weeks” (Vadasy, et al., 2006, p. 511). The interventions generally coincided with classroom reading instruction thus creating opportunities for the kindergarten teachers to create within-class high ability reading groups thus “making the intervention, in effect, the differentiated reading instruction for their lower skilled students” (Vadasy, et al., 2006, p. 512).

The study found that the paraeducators were effective in their interventions. Students “demonstrated significant advantages in reading and spelling skills at the end of kindergarten ... [and] the kindergarten treatment group maintained their higher level of performance through first-grade posttest across reading outcomes” (Vadasy, et al., 2006, p. 521). This positive result suggests that less costly intervention can be effective in addressing the critical reading intervention needs of high-risk kindergarteners.

Major limitations of this study include the use of teacher self-reports rather than direct observations about the exact nature of classroom instruction, the effect of subject attrition, and the schools’ use of tutoring for some of the study participants in the first grade. The researchers suggest that “future research on paraeducators instruction should include a treatment comparison design with a trained teacher comparison group” (Vadasy, et al., 2006, p. 524).

Simmons et al. (2007) examined the efficacy of two elements of supplemental intervention for kindergarteners identified to be at high risk for reading problems. Using an experimental research design, they investigated two key variables impacting supplemental reading instruction: the amount of instructional time and the specificity of instructional design, defined as the “blueprint for instruction” (Simmons et al., 2007, p. 333).

The sample was comprised of 116 students from seven elementary schools in the Pacific Northwest who were selected in September of their kindergarten year due to scores at or below the 25th percentile on the DIBELS and confirmation by their teacher that they were at risk for reading difficulties (Simmons et al., 2007, p. 333). Students with significantly limited English proficiency and severe hearing or visual problems were excluded. SES, race, and gender reflected the district population. Students were assigned to one of three intervention groups using a stratified random sampling procedure and each school implemented all three instructional treatments: a 30-minute/highly specified design intervention, a 15-minute/highly specified design followed by a 15-minute intervention, and a 30-minute/moderately specified design intervention. “Interventionists” of certified teachers and educational assistants conducted 108 days of supplemental small-group instruction either before or after the 2.5-hour kindergarten day. Each tutorial group was capped at five students to optimize instruction and behavior management.

The researchers used a pretest-posttest experimental design and commercially produced measurements of five dimensions of early language and literacy. They found that when compared to 15 minutes of highly specified instruction, 30 minutes of highly specified supplemental, small-group instruction resulted in comparable performance in phonemic awareness measures. However, the longer instructional time period was significantly more effective in all other reading and literacy assessment areas, particularly for those kindergarteners

who began school with the lowest reading proficiency. In terms of instructional design specificity, 30 minutes of highly specified supplemental, small-group instruction was found to be comparable to 30 minutes of moderately specified instruction for phonemic awareness proficiency but was significantly more effective in all other reading and literacy measures. Again, these findings were particularly strong with regard to children who scored at the bottom of their kindergarten cohort. Lastly, 30 minutes of even moderately designed instruction proved to be more effective than a highly specified abbreviated session. In this experiment, only spelling outcomes showed significant differences when comparing results of the 15-minute/highly specified design plus 15-minute intervention to the 30-minute/moderately specified design intervention. For all other measures, highly specified instruction was not statistically significant.

This study detailed treatment integrity controls, methodology, justifications, and reported both reliability and validity measures. Limitations include the lack of a control group that received only in-class instruction. In addition, due to the September selection of students for the sample, the researchers may have misidentified some students' true level of reading proficiency, thus skewing the sample by including participants who did not actually require intervention.

Implications for Teaching Identity

Reading recovery programs, individualized education plans, and gifted acceleration are not available to all kindergarteners in need of specialized instruction. The classroom teacher must therefore differentiate instruction for students manifesting a wide array of early literacy and reading comprehension skills. Current research conflicts on the justice and effectiveness of kindergarten within-class ability group for reading instruction. Critics assert that “reading ability group placement in kindergarten ... is likely to be enormously consequential for later schooling

outcomes” (Tach & Farkas, 2005, pg. 1053) and “may deprive [students] of the strong educational foundations necessary for enrolling in advanced coursework and college-bound curricula later in school (Buttaro et al., p. 1328). Moreover, the process of assigning students to these homogenous groups may be compromised by strong cultural biases regarding gender and behavior that supersede accurate assessments of reading proficiency.

Advocates of within-class ability groupings maintain that this classroom structure allows focused remediation for reading skills, a critical intervention for kindergarteners entering school lacking basic emergent literacy skills. The above studies offer some evidence that kindergarteners’ participation in classroom ability groups may be positively associated with growth in early literacy skills and that session length and instructional design of ability groups impact their effectiveness. In addition, Vadasy et al. demonstrated that paraeducators rather than classroom teachers could be trained to implement structured interventional instruction to ability groups.

Kindergarten is certainly becoming more academic, but it remains a critical year of social-emotional development. The author acknowledges the value of structured groups as a means to facilitate focused instruction for some reading skills. However, she is committed to creating a nurturing introduction to school where students discover their own potential, confidence, and positive disposition toward school and where individual needs are scaffolded and reinforced. Children are acutely aware of and sensitive to existing or perceived deficits and within-class groupings could indeed be a source of anxiety or internalized as a reflection on potential.

In reviewing the arguments and contradictory research presented above, the author seeks a middle ground in the debate over ability groups for kindergarten reading instruction. Language

arts is comprised of multiple elements and can be approached in several ways within the classroom. Vocabulary and text comprehension can be addressed in large groups as well as in heterogeneous groups, phonics and phonemic awareness taught in homogeneous ability groups, and fluency advanced through one-on-one interactions with the teacher. Flexibility of grouping is critical and students should not be made to feel that they have been relegated into rigid groupings based on ability level. Therefore, the author would be attentive to frequent restructuring of both heterogeneous and homogeneous groups and to the possibility that student behavior may impact her decision-making. An effective teacher should be flexible and open to a variety of instructional options. Research studies suggest that within-class ability groups might indeed be an effective means of developing reading skills. In advocating for children who require focused instruction to improve their literacy skills, all promising methodologies must be explored. If carefully and selectively implemented and monitored, within-class ability groups may serve an essential immediate need without causing long-term harm. The author will not exclude them from her instructional options but will proceed with care and caution.

References

- Buttaro, A., Catsambis, S., Mulkey, L., & Steelman, L.C. (2010). An organizational perspective on the origins of instructional segregation: School composition and use of within-class ability grouping in American kindergartens. *Teachers College Record*, 112(5), Retrieved from <http://www.tcrecord.org.proxy.wm.edu/LIBRARY/Issue.asp?volyear=2010&number=5&volume=112>
- Graue, E. (2010). Reimagining kindergarten. *Education Digest: Essential Readings Condensed for Quick Review*, 75(7), Retrieved from <http://web.ebscohost.com.proxy.wm.edu/ehost/>
- Koch, P.R., Steelman, L.C., Mulkey, L., & Catsambis, S. (2008). Naughty or nice? Equity, gender and behavior . *Social Psychology of Education: An International Journal*, 11(4), Retrieved from <http://web.ebscohost.com.proxy.wm.edu/ehost/>
- McCoach, D.B., O'Connell, A.A., & Levitt, H. (2006). Ability grouping across kindergarten using an early childhood longitudinal study. *Journal of Educational Research*, 99(6), Retrieved from <http://web.ebscohost.com.proxy.wm.edu/ehost/>
- National Center for Education Statistics, US Department of Education. (2010). *Early Childhood Longitudinal Program (ECLS): Kindergarten class of 1998-1999 (ECLS-K)*. Retrieved from <http://nces.ed.gov/ecls/kindergarten.asp>
- Simmons, D.C., Kame'enui, E.J., Harn, B., Coyne, M.D., Stoolmiller, M., Santoro, L.E., Smith, S.B., Beck, C.T., & Kaufman, N.K. (2007). Attributes of effective and efficient kindergarten reading intervention: An examination of instructional time and design specificity. *Journal of Learning Disabilities*, 40(4), Retrieved from <http://web.ebscohost.com.proxy.wm.edu/ehost/>

Tach, L.M., & Farkas, G. (2005). Learning-related behaviors, cognitive skills, and ability grouping when schooling begins. *Social Science Research*, 35(4), Retrieved from <http://www.sciencedirect.com.proxy.wm.edu/science>

U.S. Department of Health and Human Services, Administration for Children and Families, (2010). *Early childhood longitudinal study-Kindergarten cohort (ECLS-K)*, Retrieved from http://www.acf.hhs.gov/programs/opre/other_resrch/eval_data/reports/common_constructs/com_appb_eclsk.html

Vadasy, P.F., Sanders, E.A., & Peyton, J.A. (2006). Code-oriented instruction for kindergarten students at risk for reading difficulties: A randomized field trial with paraeducator implementers. *Journal of Educational Psychology*, 98(3), Retrieved from <http://web.ebscohost.com.proxy.wm.edu/ehost/>