Abstract Title Page
Not included in page count.

Title: Class Management and Homogeneous grouping in Kindergarten Literacy Instruction

Author(s): Guanglei Hong, Janette Pelletier, Yihua Hong, Carl Corter

Abstract Body

Limit 5 pages single spaced.

Background/context:

Description of prior research, its intellectual context and its policy context.

Once children have entered kindergarten, they are organized in classrooms for systematic literacy instruction, which may not take into account the diversity in their prior literacy exposure and their social and emotional experiences. Almost every educator would agree that keeping the class in good order and engaging every student in active learning are essential for effective instruction. Interestingly, promoting literacy learning while helping new kindergartners adapt to classroom norms seem to be accomplished with ease by some kindergarten teachers but pose a major challenge in some other classrooms.

In searching for effective strategies for literacy instruction, numerous classroom observations have indicated that adapting instruction to student ability through skillful management of small-group instruction is one of the characteristics of exemplary language arts teaching (Metsala & Wharton-McDonald, 1997; Morrow, Rutzel, & Casey, 2006; Taylor, Pearson, Clark, & Walpole, 1999; Taylor, Peterson, Pearson, & Rodriguez, 2002). However, there has been no consistent evidence with regard to the effects of within-class homogeneous grouping versus no such grouping across empirical studies (Kulik & Kulik, 1987; Slavin, 1987, Lou, Abrami, Spence, Poulsen, Chambers, and d'Apollonia, 1996; Lou, Abrami, & Spence, 2000). Almost all the previous studies on ability grouping focus on students' learning outcomes. In theory, homogeneous grouping may create an impact on students' academic achievement through changing their general classroom behaviors along dimensions such as manageability by the teacher or self-regulation by the student. Although the trend of poor academic performance for children who engage in disruptive behaviors in school is well-established (e.g., Tremblay, Masse, Perron, LeBlanc, Schwartzman, & Ledingham, 1992; Tremblay, Pihl, Vitaro, & Dobkin, 1994), there are important gaps in our knowledge. Specifically, there are contrasting theoretical arguments with regard to whether homogeneous grouping minimizes rather than exacerbates classroom disruptions, or whether grouped instruction in general reduces rather than increases individual students' problem behaviors. In this study we investigate grouping effects on class management and student behaviors. These outcomes are important in their own right because developing kindergartners' understanding of social norms and school conventions and fostering their capability of self-regulation is an important objective in the first year of schooling.

Reasoning from the basic rationale for homogeneous grouping, we suspect that grouping will improve class behavior by engaging a large number of students in instruction adapted to their individual level. Within a homogeneous group, instructional activities and materials may be geared to the children's overlapping "zones of proximal development" (ZPD) in the Vygotskian framework. Such adaptation is expected to provide the cognitive advantages of working on emergent skills and understanding as well as the motivating aspects of working with appropriate levels of challenge (Ames, 1992). In contrast, a mismatch between the learning tasks and students' current ability is often inevitable during whole-class instruction especially for high-achieving and low-achieving students. As a result, some students may lose motivation and become bored and some may seek peer or adult attention by acting out (Borman, 1978; Greenwood, Horton, & Utley, 2002; Jones, 1948; Schumm, Moody, & Vaughn, 2000; Tombari & Borich, 1999).

From a developmental point of view, however, homogeneous grouping may lead to behavior problems in the classroom. Although young children up to the age of 7 tend to accept school conventions as the way the social world is supposed to be, they appear to take little interest in peer transgressions against school conventions (Nucci & Nucci, 1982; Nucci, 2006). Hence teachers in primary grades cannot rely on the children to regulate peer violations of rules and procedures in small-group settings. Within-class grouping will require the teacher to simultaneously organize and monitor activities in different groups and to assist transitioning between groups. These management tasks pose a major challenge especially to kindergarten teachers who fail to establish instructional routines when the school year starts (Rimm-Kaufman, La Paro, Downer, & Pianta, 2005). According to Grant and Rothenberg's (1986) repeated observations in nine first-grade and second-grade classrooms that adopted reading ability groups, in almost all these classes, teacher time with students in a reading group was frequently interrupted by children outside the group or by the teacher to discipline or give aid to children outside the group.

A related question is whether class manageability at the beginning of the kindergarten year moderates the effect of homogeneous grouping on students' literacy growth. Presumably, individual students' behavioral problems and the teacher's lack of skills may both contribute to management difficulties in a class. Some researchers argued that rules and procedures become fully operative over the first weeks of school (Carter & Doyle, 2006; Doyle, 1986; Emmer, Evertson, & Anderson, 1980). Once the routines are well established, children are able to demonstrate competence while engaged in literacy tasks (Neuman & Roskos, 1997). Grouping could be unproductive in classes that fail to set up good routines from early on. Disruptive behaviors typically result in a shift in focus from academic content to discipline (De Martini-Scully et al, 2000). Classes that spend more time on management issues generally leave students with less time to engage in academic learning and will consequently undermine academic achievement (Shinn, Ramsey, Walker, Stieber, & O'Neill, 1987). Earlier studies have found that the grouping effects on kindergartners' literacy growth depend on the amount of literacy instruction time and the intensity of grouping (Hong & Hong, 2009). More importantly, there is evidence for differential effects of time-by-grouping treatments according to students' prior ability level. Specifically, homogeneous grouping benefits medium-ability students only when there is abundant literacy instruction time. There is a detrimental effect of high-intensity grouping for low-ability students across all literacy subdomains when instructional time is limited. The negative effect is seen even for high-ability children in the higher level domains of word recognition and comprehension (Hong, Corter, Hong, & Pelletier, under review). We reason that in classes with management difficulties, grouping under low instruction time could be especially detrimental to medium and low ability students. In classes where instruction can proceed with good order, grouping is more likely to be successful and its positive effects on medium-ability students as detected in the earlier study may become even more pronounced in these settings.

Purpose / objective / research question / focus of study:
Description of what the research focused on and why.

The purpose of this study is two-fold. Firstly we will examine, given the amount of time allocated to literacy instruction, whether homogeneous grouping helps improve class manageability over the kindergarten year and whether individual students' externalizing problem behaviors will decrease in tandem. Secondly, we will investigate whether the joint effects of

homogeneous grouping and instruction time on the literacy growth of students at different ability levels depend on class manageability at the beginning of the kindergarten year.

Setting:
Description of where the research took place.

We selected data for this study from the Early Childhood Longitudinal study, Kindergarten cohort (ECLS_K). The data were collected by the US National Center for Education Statistics in fall 1998 and spring 1999.

Population / Participants / Subjects:
Description of participants in the study: who (or what) how many, key features (or characteristics).

The ECLS-K data set contains repeated observations of a nationally representative sample of 21,260 children attending 3,197 kindergarten classes in more than 1,000 schools. The kindergarten literacy assessments, administered at the beginning and the end of the year and equated on a vertical scale, covered five literacy learning subdomains of letter recognition, beginning sounds, ending sounds, sight words, and comprehension of words in context (National Center for Education Statistics, 2002a). In addition, ECLS-K researchers surveyed the parents and teachers of each sampled child in fall 1998 and spring 1999, and collected data from school administrators in spring 1999. At the beginning and the end of the year, teachers were asked to rate every child's social-emotional development including externalizing problem behaviors—i.e., the frequency with which a child argues, fights, gets angry, acts impulsively, and disturbs ongoing activities (National Center for Education Statistics, 2002b). We used multiple sources of information in the data to identify a child's reading and language ability status relative to other children in the same class at kindergarten entry. The overall proportions of children at the low, medium, and high ability levels were approximately 25%, 61%, and 14%, respectively.

Intervention / Program / Practice:
Description of the intervention, program or practice, including details of administration and duration.

The teacher reports provide rich information about kindergarten classrooms and children's experiences in literacy instruction (Denton, West, & Walston, 2003). From teacher information in spring 1999, we computed a measure of instructional time allocated to reading and language arts. Approximately 50% of the classes were placed in the "high reading time" category because they spent more than one hour per day on reading and language arts instruction. The rest of the classes were in the "low reading time" category. To measure the intensity of homogeneous grouping, we computed a ratio of the time for ability-grouped literacy instruction to the total amount of literacy instruction time per week. In our sample, a third of the kindergarten classes never used homogeneous ability grouping; 40% of the classes used low-intensity grouping; and the remaining 27% of classes used grouping on a more intensive basis.

Every kindergarten teacher rated the degree of class misbehavior in both fall 1998 and spring 1999. For each time point, we categorized kindergarten classes into three levels of manageability on the basis of teacher perception—classes that misbehave frequently and hard to handle, classes that misbehave occasionally, and classes that behave well. To cross-validate, we also computed the proportion of sampled kindergarten children in a class whose externalizing problem behavior scores were above a disruptive threshold.

Research Design:
Description of research design (e.g., qualitative case study, quasi-experimental design, secondary analysis, analytic essay, randomized field trial).

We conduct secondary analysis of data from a national survey that employed a longitudinal design. If a kindergarten class has a possibility of adopting each of the six treatment conditions defined by literacy instruction time and intensity of grouping, we define the causal effect of one treatment relative to another as the difference between the respective potential outcomes for a child. We allow every child's potential outcome value associated with each instructional treatment to be a function of the treatment setting characterized by class composition, teacher characteristics, and school context (Hong, 2004; Hong & Raudenbush, 2006). Given the student and teacher composition of a class, we assume that every kindergartener in the class will have six potential literacy outcomes at the end of the year corresponding to the six possible treatments.

The first part of our study approximates a factorial design with two levels of instructional time and three levels of intensity of grouping. The second part of the study represents an aptitude-treatment-setting interaction design in which the experimental factors are crossed by three levels of student prior ability and three levels of class manageability. We evaluate the effects of instructional treatments as a function of class manageability for each subpopulation of children defined by their relative ability in a class.

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.

We used marginal mean weighting through stratification to reduce selection bias in analyzing the large-scale non-experimental data. MMW-S is a causal inference method recently applied in epidemiological and educational research (Hong & Hong, 2009; Hong, Corter, Hong, & Pelletier, under review; Huang, Frangakis, Dominici, Diette, & Wu, 2005). This non-parametric method is advantageous in evaluating the causal effects of concurrent multi-valued treatments. In essence, the MMW-S method stratifies a sample on the basis of the propensity score for a certain treatment and then computes the weight for the treated units in each stratum. Under the weak ignorability assumption, the weighted treatment group provides a consistent estimate of the average potential outcome associated with the treatment for the population represented by the analytic sample. The same procedure can be applied to each treatment group one at a time. Recent simulation results (Hong, under review) have revealed that, in typical applications in which a nonlinear or non-additive propensity score model is misspecified as a linear and additive one, MMW-S estimates of treatment effects display a much higher level of robustness when compared with estimates obtained through inverse-probability-of-treatment weighting (IPTW) (Robins, 2000.

In the first part of our investigation, we examined the time-by-grouping effects on class manageability in spring 1999, controlling for the availability of a teaching aide in the class and class manageability in the fall. Subsequently, we assessed the time-by-grouping effects on individual externalizing problem behaviors in spring 1999 through analyzing a three-level model with students at level 1, classes at level 2, and schools at level 3. We adjusted for treatment selection by applying the marginal mean weight at the class level and made additional adjustment for student externalizing problem behaviors in the fall.

In the second part of our analysis, we defined subpopulations by class manageability in the fall as well as by a student's relative prior ability in class. We applied the subpopulation-specific marginal mean weight at the student level. The weighting strategy allowed us to evaluate the time-by-grouping effects on student literacy outcomes within each subpopulation and to compare the treatment effects across the subpopulations.

Findings / Results:
Description of main findings with specific details.

According to our preliminary results, classes that were hard to manage tended to have a relatively high proportion of disruptive children and tended to be taught by inexperienced teachers. When kindergarten classes spent more than one hour per day on literacy instruction, high-intensity grouping appeared to be more effective than no grouping in alleviating class misbehavior in the spring especially for classes that had been hard to manage in the fall. We did not detect any grouping effect on class manageability when literacy instruction time is no more than one hour per day. These results held regardless of the availability of teaching aide. Furthermore, there was evidence for a significant effect of high reading time in combination with high-intensity grouping in reducing kindergartners' externalizing problem behaviors. We are currently in the middle of conducting the second part of our analysis.

Conclusions:
Description of conclusions and recommendations based on findings and overall study.

The above findings lend support to the theoretical argument that providing students with stimulating learning tasks tailored to each ability level along with adequate time for practice and understanding may improve their compliance with school conventions. As Carter and Doyle have argued (2006), rules and procedures are "important but secondary dimensions of classroom management" because they cannot substitute or replace the management of curricular content and pedagogical activities (p.385). Even in kindergarten classes that failed to immediately establish good routines at the beginning, students nonetheless showed growing interest in literacy activities and disruptive behaviors subsided within homogeneous groups when the teachers committed a relatively large amount of time to literacy instruction. These results coincide with the empirical evidence from an earlier study (Hong & Hong, 2009) showing a positive effect of homogeneous grouping on students' literacy growth under high reading time and a lack of grouping effect under low reading time. Putting these pieces together, we are increasingly convinced that, during the kindergarten year, student learning as well as their class behavior will likely be optimized when students receive a substantial amount of literacy instruction time in combination with adaptive instruction through homogeneous grouping.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261-271.

Borman, K. M. (1978). Social control and schooling: Power and process in two kindergarten settings. *Anthropology and Education Quarterly, 9*, 38-53.

Carter, K. & Doyle, W. (2006). Classroom management in early childhood and elementary classrooms. In Carolyn M. Evertson and Carol S. Weinstein (Eds.), *Handbook of Classroom Management: Research, Practice, and Contemporary Issues* (pp. 373-406). Mahwah, NJ: Lawrence Erlbaum.

De Martini-Scully, B., Bray, M. A., & Kehle, T. J. (2000). A packaged intervention to reduce disruptive behaviors in general education students. *Psychology in the Schools, 37*(2), 149-156.

Denton, K., West, J., & Walston, J. (2003). *Reading - Young Children's Achievement and Classroom Experiences* (NCES 2003-070). U.S. Department of Education, NCES. Washington, DC: U.S. Government Printing Office.

Doyle, W. (1986). Content representation in teachers' definitions of academic work. *Journal of Curriculum Studies, 18*, 365-379.

Edmund T. Emmer, Carolyn M. Evertson, & Linda M. Anderson. (1980). Effective classroom management at the beginning of the school year. *Elementary School Journal, 80*(5), 219-231.

Grant, L., & Rothenberg, J. (1986). The social enhancement of ability differences: Teacher-student interactions in first- and second-grade reading groups. *Elementary School Journal, 87*(1), 29-50.

Greenwood, C. R., Horton, B. T., & Utley, C. A. (2002). Academic engagement: Current perspectives in research and practice. *School Psychology Review, 31*(3), 328-349.

Hong, G. (2004). "Causal Inference for Multi-Level Observational Data with Application to Kindergarten Retention," unpublished Ph.D. dissertation, University of Michigan, School of Education.

Hong, G. (Under review). *Marginal mean weighting through stratification: Adjustment for selection bias in multi-level data.* Submitted to *the Journal of Educational and Behavioral Statistics.*

Hong, G., Corter, C., Hong, Y., & Pelletier, J. (Under review). *Reading instruction time and within-class homogeneous grouping in kindergarten: Who will benefit? Who will suffer?* Submitted to *the American Educational Research Journal.*

Hong, G., & Hong, Y. (2009). Reading instruction time and homogeneous grouping in kindergarten: An application of marginal mean weighting through stratification. *Educational Evaluation and Policy Analysis, 31*(1), 54-81.

Hong, G., & Raudenbush, S. W. (2006). Evaluating kindergarten retention policy: A case study of causal inference for multi-level observational data. *Journal of the American Statistical Association, 101*(475), 901-910.

Huang, I-C., Frangakis, C., Dominici, F., Diette, G. B., and Wu, A. W. (2005). Approach for risk adjustment in profiling multiple physician groups on asthma care. *Health Services Research, 40*, 253-278.

Jones, D. M. (1948). An experiment in adaptation to individual differences. *Journal of Educational Psychology, 39*, 257-273.

Kulik, J. A., & Kulik, C. L. (1987). Effects of ability grouping on student achievement. *Equity and Excellence, 23*, 22-30.

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research, 66*(4), 423-458.

Lou, Y., Abrami, P. C., & Spence, J. C. (2000). Effects of within-class grouping on student achievement: An exploratory model. *Journal of Educational Research, 94*(2), 101-112.

Metsala, J. L., and Wharton-McDonald, R. (1997). Effective primary-grades literacy instruction equals balanced literacy instruction. *The Reading Teacher, 50*(6), 518-521.

Morrow, L. M., Rutzel, D. R., & Casey, H. (2006). Organization and management of language arts teaching: Classroom environments, grouping practices, and exemplary instruction. In Carolyn M. Evertson and Carol S. Weinstein (Eds.), *Handbook of Classroom Management: Research, Practice, and Contemporary Issues* (pp.559-582). Mahwah, NJ: Lawrence Erlbaum.

National Center for Education Statistics. (2002a). *Early Childhood Longitudinal Study-Kindergarten class of 1998-99 (ECLS-K), Psychometric Report for Kindergarten Through First Grade*, NCES 2002-05, by Donald A. Rock and Judith M. Pollack, Educational Testing Service, Elvira Germino Hausken, project officer. Washington, DC.

National Center for Education Statistics. (2002b). *User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Code Book*, NCES 2002-135. Washington, DC.

Neuman, S. B., & Fischer, R. (1995). Task and participation structures in kindergarten using a holistic literacy teaching perspective. *Elementary School Journal, 95*(4), 325-337.

Nucci, L., & Nucci, M. S. (1982). Children's social interactions in the context of moral and conventional transgressions. *Child Development, 53*. 403-412.

Nucci, L. (2006). Classroom management for moral and social development. In Carolyn M. Evertson and Carol S. Weinstein (Eds.), *Handbook of Classroom Management: Research, Practice, and Contemporary Issues* (pp.711-734). Mahwah, NJ: Lawrence Erlbaum.

Rimm-Kaufman, S. E., La Paro, K. M., Downer, J. T., & Pianta, R. C. (2005). The contribution of classroom setting and quality of instruction to children's behavior in kindergarten classrooms. *The Elementary School Journal, 105*(4), 377-394.

Robins, J. M. (2000). Marginal structural models versus structural nested models as tools for causal inference. In M. E. Halloran & D. Berry (Eds.), *Statistical models in epidemiology, the environment, and clinical trials* (pp. 95-134). New York: Springer.

Schumm, J. S., Moody, S. W., & Vaughn, S. (2000). Grouping for reading instruction: Does one size fit all? *Journal of Learning Disabilities, 33*(5), 477-488.

Shinn, M., Ramsey, E., Walker, H., Stieber, S., & O'Neill, R. (1987). Antisocial behavior in school settings: Initial differences in an at-risk and normal population. *The Journal of Special Education, 21*, 69-84.

Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence systhesis. *Review of Educational Research, 57*(3), 293-336.

Taylor, B. M., Pearson, P. D., Clark, K. E., & Walpole, S. (1999). Beating the odds in teaching all children to read (Ciera Report No.2-006). Ann Arbor, MI: Center for the Improvement of Early Reading Achievement.

Taylor, B. M., Peterson, D. S., Pearson, P. D., & Rodriguez, M. C. (2002). Looking inside classrooms: Reflecting on the "how" as well as the "what" in effective reading instruction. *The Reading Teacher, 56*(3), 270-279.

Tombari, M., & Borich, G. (1999). Academic motivation. In Authors (Eds.), *Authentic assessment in the classroom: Application and practice*. Upper Saddle River, NJ: Prentice Hall.

Tremblay, R. E., Masse, B., Perron, D., LeBlanc, M., Schwartzman, A., & Ledingham, J. (1992). Early disruptive behavior, poor school achievement, delinquent behavior, and delinquent personality: Longitudinal analysis. Journal of Consulting and Clinical Psychology, 60, 64-72.

Tremblay, R. E., Pihl, R. O., Vitaro, F. & Dobkin, P. L. (1994). Predicting early onset of male antisocial behavior from preschool behavior. Archives of General Psychiatry, 51, 732-739.

Appendix B. Tables and Figures

Not included in page count.