Abstract Title Page

Title: How (and How Much) Do Schools Matter? Variation in K – 8th Grade Achievement Trajectories in a National Sample

Authors and Affiliations: Kate Schwartz, MPP, Department of Applied Psychology, New York University; Elise Cappella, PhD, Department of Applied Psychology, New York University; Marc Scott, PhD, PRIISM Center, New York University; Edward Seidman, PhD, Department of Applied Psychology, New York University; Ha Yeon Kim, PhD, Graduate School of Education, Harvard University

Background / Context:

Academic development proceeds at different paces for students. Child and family characteristics play a profound and predictable role in achievement trajectories from kindergarten through 8th grade (see Duncan & Brooks-Gunn, 1997; Magnuson, Duncan, & Kalil, 2006). School contexts – including structural and process characteristics – are theorized to contribute to math and reading achievement as well (Bronfenbrenner & Morris, 1998; Tseng & Seidman, 2007). Additionally, the role of school contexts in student development may be magnified during times of transition when young people are more vulnerable to risk (see Smetana & Campione-Barr, 2002).

For early adolescents attending U.S. public schools, nearly 90 percent are enrolled in a middle grade school (Keaton, 2012), generally with a 6-8 or 7-9 grade configuration. Developmental mismatch hypothesis (Eccles & Midgley, 1989) suggests that the instructional and social contexts of these middle grade schools may not match early adolescent needs for autonomy, competence, and relatedness (Deci & Ryan, 2000). Only a handful of studies explicitly test this hypothesis (Kim et al., 2014), but research on school grade configuration shows academic drawbacks associated with middle grade schools including increased high school drop-out rates (Alspaugh, 1998) and decreased performance on standardized tests (Byrnes & Ruby, 2007; Rockoff & Lockwood, 2010; Schwartz et al., 2011). Propensity score matching analysis in a national sample suggest higher reading (but not math) achievement for 8th grade students in K-8 schools (Kieffer, 2013).

However, few studies of school effects examine longitudinal trajectories of achievement starting at school entry and in a national sample. In addition, little research has examined whether it is the grade configuration or other aspects of school context such as composition and climate that contribute to achievement trajectories. Links between school contexts and student achievement have been found with school size (Lackney, 2004; Schneider, 2002), quality of facilities (e.g., Earthman, 2002), and stress/disorder (Flannery, Sugai, & Anderson, 2009; Mitchell et al., 2010; Warren et al, 2003). At least one well-designed study found no relation between school grade configuration and student achievement when controlling for other school contextual variables (Weiss & Kipnes, 2006). Longitudinal studies are needed with large samples and multilevel analyses to determine if school characteristics predict trajectories of student achievement from kindergarten forward.

Purpose / Objective / Research Question / Focus of Study:

Given early adolescence as a potentially disruptive period in development and school transitions as a time of inherent potential for the promotion of academic achievement, it is critical to identify the school structures and processes that maximize positive outcomes for youth. The current investigation describes math and reading achievement trajectories in a national sample from kindergarten entry through 8th grade, and examines the role of school grade configuration (K-8; 6-8; 7-9) and school context (e.g., composition, professional climate, facilities) on achievement trajectories. We contribute to developmental mismatch theory (Eccles et al., 1989) and systems theory of social settings (Tseng & Seidman, 2007) by examining whether (and which) school contextual characteristics contribute to growth in achievement. The national sample, multilevel analysis, and measurement of school settings lead to increased precision in estimation and enhanced external validity. Findings will inform educational policy regarding the grade configurations of middle grade schools and school intervention to support academic development in the middle years.

Setting:

Data were drawn from the Early Childhood Longitudinal Study, Kindergarten Class 1998-99 (ECLS-K). We include six waves (fall kindergarten, and spring K, 1st, 3rd, 5th, and 8th grade for the majority

of participants) during which students attended 2,998 unique public schools: 702 in wave 1; 733 in wave 2; 871 in wave 4; 1,087 in wave 5; 1287 in wave 6; and 1562 in wave 7. Wave 3 data (fall 1st grade) was not included because only a subsample of participants were assessed. Data were collected from teachers, administrators, parents, and students.

Population / Participants / Subjects:

The study sample consists of the 6,158 students in the ECLS-K who participated in all 6 waves and who attended kindergarten at a regular public school in wave 1. Students who began in public school and moved to private school are included until they transition, at which point their data are censored from analyses. Similarly, students not assessed in any wave due to disability or language minority status are censored from that wave only. Three students who met the inclusion criteria were omitted entirely as a result of non-assessment due to disability in all six waves.

In 1998, the kindergarteners were 50% female with an average age of 5.7 years. They are white (59%), Hispanic (18%), African American (11%), Asian (6%), Pacific Islander/Hawaiian/Native Alaskan/Native American (3%), and multi-racial (2%). Nineteen percent lived below poverty in kindergarten; 22% of families experienced serious financial hardship since the birth of the child.

Research Design:

The ECLS-K employed a longitudinal research design, following a nationally representative sample of 21,260 kindergarteners from the 1998-1999 school year through the 2006-2007 school year, at which time the majority of participants were enrolled in eighth grade.

Data Collection and Analysis:

The ECLS-K collected data via multiple methods to capture a wide range of child outcomes and developmental contexts. These include parent interviews, school administrator and teacher surveys, student self-reports, student records, and direct student assessments. Data were collected from the full sample in fall and spring of kindergarten, as well as spring of what for most students was their 1st, 3rd, 5th, and 8th grade year. Students who were retained or skipped grades were assessed in the same year as other students with their current grade level included in the prediction model.

Outcomes. Assessments of reading and math achievement used a two-stage testing approach based on item response theory (IRT). In the first stage, a common set of questions covering a broad range of difficulty was asked of all participants. In the second stage, a targeted set of questions was asked based on the students' performance in the first stage. The assessment items were drawn from standardized tests in other large-scale studies of youth, including the NAEP, NELS: 88, ELS: 2002, and TAKS. This approach helps to avoid ceiling effects, capture age-appropriate skills, and ensure comparability over time. The IRT estimates of reliability were high (range = .87 - .96 by wave; Tourangeau et al., 2009). The current study used the IRT-derived theta scores, which are comparable across waves of data collection (see Najarian, Pollack, & Sorongon, 2009).

Predictors. School composition predictors include overall school enrollment as well as the percent of student body: at or above grade level in reading; at or above grade level in math; eligible for free lunch; Hispanic; black; and categorized as having limited English proficiency (LEP). These items were assessed in every wave beginning in the spring of kindergarten using administrator surveys.

School climate predictors, reported by administrators or teachers, include average daily attendance rate (ranging from 1 = less than 90% to 5 = more than 98%) – and four scale scores representing quality of school environment. Adequacy of school facilities scale (α = 0.64-0.71) was created from ten items on the presence and adequacy of various facilities (e.g., computer room). School strain scale

($\alpha = 0.63$-0.68) was created from three items assessing the degree to which teacher absences, student absences, and teacher turnover are a problem. Only the teacher turnover item was asked in wave seven. Positive climate scale ($\alpha = 0.67$ - 0.73) was created from the degree to which administrators think parents are active in programs; the community is supportive of the school; there is consensus on expectations in the school community; and order/discipline is maintained. Only the parent involvement item was asked in wave seven. Professional climate ($\alpha = 0.64$ - 0.71) was created from the average responses, across responding teachers in the school, to six questions concerning collegial school spirit, acceptance/belonging in the school environment, continual improvement among staff, parental support of school staff, and school-wide academic standards.

The final predictor is school grade span: coded K-4, K-5, K-6, K-8, 5-8, 6-8, or 7-8 depending on administrator reports of lowest and highest grade served. The small number of schools that served students beyond 8th grade are coded into the closest corresponding category.

Covariates. Covariates include time invariant student characteristics (e.g., race, age, gender, birth weight), time invariant family/household characteristics (e.g., number of children's books at home, mother's age at first birth, indicator of serious financial troubles), time varying student variables (e.g., student changed schools since previous wave, teacher reports of child behavior), and time varying family variables (e.g., SES, household size, parental involvement). Analyses also control for the region and urbanicity of the school attended in that wave.

Missing data. Missing data were imputed in two stages, school- and individual-level, using the mi impute chained equations subroutines in STATA 13. The stages account for the multi-level design. Additionally, the school-level imputation was stratified by school grade span, allowing for a separate model for elementary (K-4, K-5, and K-6) schools; K-8 schools; middle grade (5-8 and 6-8) schools, and junior high (7-8) schools. This was done based on prior research indicating school context varies along these designations (Kim et al., 2014). Such stratification avoids imputing toward the mean values of the largest categories when missing data in schools with less common grade spans. Under this design, 20 datasets were imputed for the missing school-level predictors, including region and urbanicity, within each stratum in models that included other school-level variables and aggregated individual covariates as auxiliary variables. These were then re-combined and merged back with the student-level covariates and outcome variables based on school ID and wave. For each of the 20 datasets, one dataset was imputed at the individual level, propagating the school-level uncertainty to the individual level and yielding 20 imputed datasets. All analyses were completed across all 20 datasets using mi estimate.

Data analysis. Analyses were conducted using the multilevel model estimation routine xtmixed in STATA to capture within- and between-individual heterogeneity (individual growth curves) across time in reading and math achievement. The heterogeneity model was based on likelihood ratio tests for each imputation of the data; the best fitting model for math and reading was determined to be one that included random linear slopes, random intercepts, and a covariance term. The overall average effect of time was modeled using (normative) year in school to account for the fact that data were not collected at equal intervals. A grade-squared term was included in the model to improve fit (non-linear, common growth was detected). Subsequent models will include interaction terms between school grade span configuration / school context and child race, gender, and poverty status.

Findings / Results: First, we describe the math and reading achievement trajectories of a national sample of public school students from kindergarten through 8th grade. Next, we examine whether students' trajectories differ by school grade configuration or school context (composition, climate) controlling for student characteristics. Subsequent analysis will explore whether the link between

school grade configuration / context and achievement differs by race, gender, or poverty status.

Students' achievement trajectories were found to grow more steeply in the early years (reading: $\beta = .616$, $p = .009$; math: $\beta = .584$, $p = .002$) with a reduced growth rate over time, as indicated by the significant squared terms for reading ($\beta = -.049$, $p = .015$) and math ($\beta = -.041$, $p = .010$). Variance of random intercepts was 0.135 and 0.099 respectively and variance of random slopes was 0.001 for both reading and math, with a covariance of -0.011 for reading and -0.004 for math. This suggests students' trajectories converge somewhat over time, and more so for reading achievement, such that those with lower starting points have steeper slopes in their growth curves and vice versa. The between-subject heterogeneity was substantial. Additionally, many student covariates significantly predicted math and reading achievement, including age, gender, race, SES, disability, pre-K, household number of books, mother's age, transferring schools, frequency child reads, internalizing, externalizing, approaches to learning, self-control, and interpersonal skills. This suggests a good selection of individual controls were used and supports our ability to interpret school-level findings.

Controlling for student covariates, we found no difference in reading growth trajectories by region or urbanicity of the school; however, math achievement varied by region, with higher achievement in the South. In terms of our primary research questions, school composition significantly predicted achievement trajectories. Reading achievement was lower for students in schools with a greater percentage of students eligible for free lunch ($\beta = -.0004$, $p = .027$: magnitude is for each percentage point increase) and higher for students in schools where teachers reported a more positive professional climate ($\beta = .020$, $p = .037$). Math achievement, in contrast, was significantly higher in schools with higher average attendance ($\beta = .004$, $p = .015$) and lower at schools with a higher percentage of black students ($\beta = -.006$, $p = .023$) and facilities with a higher adequacy rating ($\beta = -.005$, $p = .003$). Both reading ($\beta = .033$, $p < .001$) and math ($\beta = .024$, $p = .001$) were associated with higher achievement at schools where administrators report more active parents in 8th grade.

School grade configuration was not significantly associated with reading or math achievement once student and family characteristics, and school context variables, were taken into account.

Conclusions:

The current study extends research on the transition to early adolescence and middle grade schools by examining students' achievement trajectories from school entry through 8th grade in a national sample and beginning to disassociate the role of school context and school grade configuration in achievement trajectories. We find that key characteristics of school composition and climate play roles in student achievement beyond individual student predictors. Specifically, school composition contributed to reading and math trajectories and school climate contributed to reading trajectories.

Importantly, once student and school characteristics are considered, school grade configuration does not significantly contribute to students' academic achievement trajectories before or after the middle grade school transition. This finding supports developmental mismatch theory (Eccles et al., 1989) and systems theories of social settings (Tseng & Seidman, 2007) in suggesting that it may be the attributes of the school context – rather than the presence of a school transition – that places youth at risk during the middle grade years. These results illuminate the need to attend to school composition and climate in policies and practices designed to improve academic achievement. In addition, our ongoing analysis that examines differential effects by race, gender, and poverty status will further address which school attributes are most influential for which groups of students. In doing so, this study will inform district- and school-level initiatives to bolster school contexts and support student academic growth across the first nine years of schooling.

Appendices

Appendix A. References

Alspaugh, J.W. (1998). Achievement loss associated with the transition to middle school and high school. *The Journal of Educational Research*, *92*(1), 20-25.

Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. *Handbook of child psychology*, *1*, 993-1028.

Byrnes, V., & Ruby, A. (2007). Comparing achievement between K–8 and middle Schools: A large-scale empirical study. *American Journal of Education*, *114*, 101–135.

Duncan, G. & Brooks-Gunn, J. (Eds.) (1997). *Consequences of growing up poor*. New York: Russell Sage.

Earthman, G. I. (2002). School facility conditions and student academic achievement. *UCLA's Institute for Democracy, Education, & Access*.

Eccles, J. S., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. *Research on Motivation in Education*, *3*, 139–186.

Flannery, K., Sugai, G., & Anderson, C. (2009). School-wide positive behavior support in high school. *Journal of Positive Behavior Interventions*, *11*, 177.

Keaton, P. (2012). Numbers and types of public elementary and secondary education agencies from the Common Core of Data: School year 2010-11 (NCES 2012-325rev). *National Center of Education Statistics, Institute of Education Sciences, US Department of Education*. Washington, DC.

Kieffer, M. J. (2013). Development of reading and mathematics skills in early adolescence: Do k-8 public schools make a difference? *Journal of Research on Educational Effectiveness*, *6*, 361–379. doi:10.1080/19345747.2013.822954

Kim, H.Y., Schwartz, K., Cappella, E., & Seidman, E. (2014). Navigating middle grades: Role of social contexts in middle grade school climate. *American Journal of Community Psychology, 54*, 28-45. doi:10.1007/s10464-014-9659-x

Lackney, J. A. (2005). New approaches for school design. In F. W. English & G. L. Anderson (Eds.), *The Sage handbook of educational leadership: Advances in theory, research, and practice* (pp. 506–537). California, Sage: Thousand Oaks.

Magnuson, K., Duncan, G., & Kalil, A. (2006). The contribution of middle childhood contexts to adolescent achievement and behavior. *Developmental contexts in middle childhood: Bridges to adolescence and adulthood* (pp. 150-172). New York: Cambridge University Press.

Mitchell, M. M., Bradshaw, C. P., & Leaf, P. J. (2010). Student and teacher perceptions of school climate: A multilevel exploration of patterns of discrepancy. *Journal of School Health*, *80*, 271–279. doi:10.1111/j.1746-1561.2010.00501.x

Rockoff, J. E., & Lockwood, B. B. (2010). Stuck in the middle: Impacts of grade configuration in public schools. *Journal of Public Economics*, *94*, 1051-1061. doi:10.1016/j.jpubeco.2010.06.017

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68.

Schneider, M. (2002). Do school facilities affect academic outcomes? *National Clearing House for Educational Facilities*. Washington, DC.

Smetana, J. G., Campione-Barr, N., & Metzger, A. (2006). Adolescent development in interpersonal and societal contexts. *Annual Review of Psychology*, 57, 255-284.

Tseng, V., & Seidman, E. (2007). A systems framework for understanding social settings. *American Journal of Community Psychology*, *39*(3-4), 217-228.

Warren, J. S., Edmonson, H. M., Griggs, P., Lassen, S. R., McCart, A., Turnbull, A., & Sailor, W. (2003). Urban applications of school-wide positive behavior support critical issues and lessons learned. *Journal of Positive Behavior Interventions*, *5*, 80–91. doi:10.1177/10983007030050020301

Weiss, C. C., & Kipnes, L. (2006). Reexamining middle school effects: A comparison of middle grade students in middle schools and k-8 schools. *American Journal of Education*, *112*, 239–272.