

Predicting Gain Scores with Hierarchical Linear Models: A Value-Added Approach to Measure Teacher Effectiveness

Bidya Raj Subedi, Ph.D. (Corresponding author)

Department of Research, Evaluation & Assessment, B-240

The School District of Palm Beach County

3300 Forest Hill Blvd., West Palm Beach, Florida 33406

Tel.: 1-561- 969-5810 E-mail: bidya.subedi@palmbeachschools.org

Bonnie Swan, Ph.D.

University of Central Florida, Orlando, Florida

Michael C. Hynes, Ph.D.

University of Central Florida, Orlando, Florida

Received: August 7, 2012

Accepted: August 30, 2013

Published: August 30, 2013

doi:10.5296/jse.v3i3.4187

URL: <http://dx.doi.org/10.5296/jse.v3i3.4187>

Abstract

In this article, we predicted students' mathematics gain scores employing two-level hierarchical linear models (HLM) through value-added approach using data from one of the largest urban school districts in the United States of America. Effects of teacher quality or teacher effectiveness, characterized by teacher's certification in mathematics content area and teacher experience, were measured on students' gain scores. The results showed significant impact on mathematics gain scores due to teacher's content certification and teacher experience at teacher level and pretest scores as well as free and reduced lunch status at student level including cross-level interaction effects of teacher content certification with

student level predictors. We also reported proportions variance explained and d-type effect sizes for teacher level models in order to measure teacher effectiveness.

Keywords: Students' gain scores, Hierarchical linear models, Value-added approach, Teacher effectiveness, D-type effect size

1. Introduction

Students' gain scores can be predicted employing hierarchical linear models (HLM) where student and teacher level data are analyzed simultaneously at level-1 and level-2 models, respectively, incorporating appropriate predictors in the models. This modeling approach allows researchers not only determining the significant predictors at both levels but also measuring the amount of variability and effect size at level-2. For example, we can incorporate significant predictors at teacher level model, defining teacher quality, represented by teacher characteristics such as teacher certification level in related subject area and teacher experience and calculate effect size for teacher level model. The student level predictors such as students' mathematics pretest scores as well as free and reduced lunch status can be included in the model in order to predict gain scores in mathematics. It is also important that we examine the interaction effects between teacher and student level predictors on students' mathematics gain scores. This study aims to predict students' mathematics gain scores due to the effects of potential student and teacher level predictors and estimate the proportions of variance in gain scores lying among teachers, termed as value-added models (VAM) by Rowan, Correnti, & Miller (2002). Based on the proportions of variance, we also compute d-type effect sizes for teacher level models in order to measure teacher effectiveness.

1.1 Teacher Effectiveness and Teacher Quality

Several studies used large data sets to address the relationship between student as well as teacher level factors and student achievement by employing multilevel models. In a review of multilevel studies relating to teacher quality and student achievement, Scheerens and Bosker (1997) found that the differences in student achievement are associated with school (20%) and classroom/teacher level factors (20%), with the remaining difference (60%) at the student level factors (such as socioeconomic status and prior achievement). Rowan et al. (2002) employed a three-level HLM using value-added approach to predict mathematics and reading achievement and annual gains incorporating student, teacher, and school level predictors, respectively, in level-1, level-2 and level-3 models. They allowed variance decomposition among students, classrooms (teachers) and schools in order to measure teacher effectiveness where they mentioned that the purpose of VAM is to estimate the proportions of variance in changes in student achievement lying among classrooms, after controlling for the effects of other confounding variables. Other powerful value-added models (e.g., Jordan, Mendro, & Weerasinghe, 1997; Sanders & Rivers, 1996), that track students' gains over more than one year, have brought about a rethinking among researchers regarding the relative importance of the role of the teacher. Sanders and Rivers' (1996) ground-breaking Tennessee value-added study showed that fifth grade mathematics students matched in performance assigned to ineffective teachers for three years performed dramatically worse than those children assigned to more effective teachers. Jordan et al. (1997) measured the effects of Texas teachers on student achievement and Stronge, Ward, & Grant (2011) examined classroom practices of effective versus less effective teachers based on student achievement gain scores in reading and mathematics.

In the context of the mandates and philosophies of the No Child Left Behind (NCLB) Act in United States of America (U.S.A.), much of what is driving educational reform centers on the premise that teachers matter. For example, by the end of the 2005-2006 school year, states were required, for the first time, to have data collection and reporting mechanisms in place to ensure the ability to publish reports disclosing whether they meet the goal of ensuring all teachers are “highly qualified.” Meeting these standards basically means that teachers must hold a minimum of a bachelor’s degree have state licensure or certification and demonstrate subject-area competence.

Studies reveal that the need for qualified teachers is particularly great in lower-performing schools with higher numbers of low-income and minority students (Allen, 2005; Betts, Rueben, & Danenberg, 2000; Hanushek, Kain, & Rivkin, 2004; Lankford, Loeb, & Wyckoff, 2002; Nye, Konstantopoulos, & Hedges (2004); Sanders & Rivers, 1996; U.S. DOE, 2005) and the problem is even more pronounced in middle schools (Jerald & Ingersoll, 2002).

Evidence is mounting that better teachers can and do make a difference in student achievement (Haycock, 1998; Jordan et al., 1997; Rivkin, Hanushek, & Kain, 2005; Sanders & Rivers, 1996). Still, substantial disagreement exists among researchers as to which teacher qualifications make a difference (Greenberg, Rhodes, Ye, & Stancavage, 2004), and little has been explored on this topic specific to the middle school classroom. For example, Rice (2003) found a serious gap in the knowledge base that still needs to be explored regarding middle school (and elementary) teachers’ effectiveness that is used to guide important teacher policy decisions. In a study related to mathematics achievement using 1996 National Assessment of Educational Progress (NAEP) data, Wenglinsky (2002) found that the effects of classroom practices, when added to those of other teacher characteristics, are comparable in size to those of student backgrounds. They suggested that teachers can contribute as much to student learning in mathematics as the students themselves. In a study measuring the effect of teacher qualification, Croninger, Rice, Rathbun, & Nishio (2007) found potential contextual effects of teachers’ qualifications on student achievement. Harris & Sass (2008) found, in their statewide Florida study, that more experienced teachers were more effective in teaching middle school mathematics. Their research also revealed that only at middle school, did having an advanced degree make significant difference in student achievement in mathematics, compared to other grade and subject matter combinations, where the correlation was either negative, or insignificant.

1.2 Selecting Relevant Predictors

This paper considered appropriate predictors based on substantial studies in past. Darling-Hammond, Holtzman, Gatlin, & Heilig (2005) found that certified teachers consistently produce stronger student achievement gains than do uncertified teachers, and controlling for teacher experience, degrees, and student characteristics, uncertified teachers are less effective than certified teachers. Darling-Hammond (2000) found that measures of teacher preparation and certification are by far the strongest correlates of student achievement in reading and mathematics, both before and after controlling for student poverty and language status. This has also been suggested by others, who found that regular or advanced certification has a major role in significantly impacting student achievement (Clotfelter, Ladd,

& Vigdor, 2007; Klecker, 2008). In a study, Goldhaber and Brewer (2000) found that mathematics teachers who have a standard certification have a statistically significant positive impact on student test scores relative to the teachers who either hold private school certification or are not certified in their subject area.

Researchers in past have used student level predictors in multilevel model by incorporating students' prior achievement and socioeconomic background in the model to predict mathematics and reading achievement (see Rowan et al., 2002; Scheerens & Bosker, 1997). Pituch (1999) measured the interaction effects of student level predictors with those in teacher and school levels. Other studies explored the effect of teacher degree and experience on student mathematics achievement (Ballou and Podgursky, 2000; Goldhaber, and Brewer, 1998; Howley, 1996; Lippman, Burns, & McArthur, 1996; Monk, 1994; Rice, 2003). In a study in California, Shields et al. (2003) found that about 50% of the novice teachers (first or second year) employed at high-poverty schools were under-prepared compared with 30% in low poverty schools. For example in 2007, Also according to the National Science Board Science and Engineering Indicators report (2012), fully certified mathematics teachers were more prevalent in schools with fewer minority students (92% versus 84% for high- versus low-minority); and less poverty (89% versus 81% for high versus less poverty). For seniority, this is especially true in the first few years (Harris & Sass, 2008; Rice, 2010).

This study is conducted in a large urban school district in Florida, U.S.A., where the teacher characteristics are considered to define highly qualified teachers who taught core subjects. In Florida, core subjects include English, reading, language arts, mathematics, science, foreign languages, civics, government, economics, history, geography and most arts.

The purpose of this paper is to predict students' mathematics gain scores due to significant student and teacher level predictors and estimate the proportions of variance in gain scores lying among teachers which is termed as VAM. Further, the d-type effect sizes are computed for teacher level model using the proportions of variance explained through VAM approach and teacher effectiveness is determined based on such effect sizes. The findings of this research will be valuable for evaluators and researchers in K-12 school settings. This study extends prior works of authors (Swan, Dixon, & Subedi, 2005; Swan 2006) in terms of measuring teacher effectiveness in HLM through value-added models.

1.3 Research Questions

The following research questions are answered through this study.

1. What are the significant predictors at student and teacher levels for predicting students' mathematics gain scores?
2. What are the percentages of the variance explained and effect sizes at teacher level for unconditional and conditional models?

2. Methods

2.1 Data

This study used 6,184 students and 253 mathematics teachers from all middle schools in Orange County Public Schools (OCPS), Florida. This urban public school district, with student population of more than 175000, is the 12th largest among 16,000 school districts in

U.S.A. at the time of data collection. Student ethnic distribution for OCPS included approximately 38% white, 28% black, 28% Hispanic, and about 6% other. The distribution for male and female students was 51% and 49% respectively. Sixty-six percent of the teachers were White, 23% Black, 7% Hispanic, and 4% were classified as Asian.

The primary source of data was obtained from teacher and student records of OCPS. This included mathematics NRT-NCE (Norm Referenced Test-Normal Curve Equivalent) portion of the Florida Comprehensive Assessment Test (FCAT) test scores. The reliability of FCAT NRT-NCE test ranged from 0.90 to 0.91 (HumRRO, 2002).

2.2 Variables

2.2.1 Outcome variable

The student mathematics gain scores is used as an outcome measure, which is computed as the difference of school years 2005 and 2004 NRT-NCE scores. The NCE scores for both years ranged from 6.7 to 99 and gain scores ranged from -31.4 to 45.0.

2.2.2 Predictors

The predictors used at student level were student free and reduced lunch status (FRL) and pretest scores. The free and reduced lunch status is a dichotomous variable coded 0 (not participating in free and reduced lunch) and 1 (participating in free and reduced lunch). The 2004 NRT-NCE scores are defined as pretest scores.

The predictors at teacher level were mathematics content certification, advanced mathematics degree, and teacher experience. Both teacher's content certification (1 indicating holding a mathematics content certificate and 0 indicating not holding such certificate) and advanced mathematics degree (1 indicating teacher's advanced degree in mathematics or mathematics education and 0 indicating not holding such degree) are coded as dichotomous variables. Note that advanced degree is defined as master level or higher. Teacher experience is a continuous predictor that measured the number of years taught by a teacher which ranged from 0 to 37. This predictor is used for the teachers who taught those students in regular or advanced mathematics courses for the entire school year 2005 (2004-05) for grades 6-8.

Note that we estimated parameters twice, first without predictors in level-1 and level-2 models (unconditional model) and then with significant predictors in both levels (conditional model). Only significant predictors are incorporated in level-1 and level-2 models and the results tables.

2.3 Model Development

This study employed a two-level HLM in order to predict students' gain scores where student and teacher level data were incorporated in level-1 and level-2 models, respectively. Students' pretest scores as well as free and reduced lunch status are used as level-1 predictors. Teacher's content certification in mathematics, teacher's experience, and teacher's advanced degree in mathematics are used as level-2 predictors.

First, level-1 and level-2 unconditional models, which do not include any predictors, are

developed, and teacher-to-teacher variance in average gain scores is determined (to answer part of research question 2). Then level-1 and level-2 conditional models, which include potential student and teacher level predictors, are developed. The level-2 variance terms are deleted from the model if they were not significant or did not explain more variance in students' gain scores after including these random terms in the model.

Students' mathematics gain scores can be predicted using following level-1 and level-2 unconditional models.

$$(\text{GNMATHACH})_{ij} = \beta_{0j} + r_{ij} \quad (1)$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (2)$$

where β_{0j} is the intercept and γ_{00} is the average mathematics gain scores for teachers. Further, r_{ij} and u_{0j} are the random effect terms at student and teacher level models, respectively. The single equation can be expressed as follows by substituting Equation (2) in Equation (1).

$$(\text{GNMATHACH})_{ij} = \gamma_{00} + u_{0j} + r_{ij} \quad (3)$$

Several assumptions are examined associated with the analysis. The linearity assumption was checked by line graphs that showed the linear relationship between the predictors and students' gain scores. The multicollinearity assumption was checked by bivariate correlation analysis among all possible pairs of predictors. The assumption about residuals' normality was checked by plotting histograms with normal curves of the residuals which showed a normal distribution of both level-1 and level-2 residuals. The scatter diagram for residuals versus predicted values at both level-1 and level-2 models indicated the homoscedasticity. The analysis found no correlations of the predictors with the error terms at their respective levels. Similarly, virtually zero correlations were found between level-1 and level-2 error terms. Thus, the analysis indicated no violations related to HLM analysis.

Further, the level-1 conditional model for predicting mathematics gain scores due to students' pretest scores and free and reduced lunch (FRL) can be expressed as follows.

$$(\text{GNMATHACH})_{ij} = \beta_{0j} + \beta_{1j} (\text{PRETESTSCORES})_{ij} + \beta_{2j} (\text{FRL})_{ij} + r_{ij} \quad (4)$$

where β_{0j} is the intercept, β_{1j} and β_{2j} are slopes or effects of pretest scores and FRL, respectively. The term r_{ij} is the random effect for student i nested in teacher j .

The level-2 conditional model can be given as follows for predicting level-1 coefficients as outcome.

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01} (\text{CERTICONT})_j + \gamma_{02} (\text{TCHREXP})_j + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11} (\text{CERTICONT})_j \\ \beta_{2j} &= \gamma_{20} + \gamma_{21} (\text{CERTICONT})_j \end{aligned} \quad (5)$$

In Equation (5), the term γ_{00} represents the average mathematics gain scores for teachers, γ_{01} is the mean achievement gain difference between those students taught by teachers who hold mathematics content certificates and who do not hold such certificates, γ_{02} is the effect of

teachers' experience on average student mathematics achievement gain, γ_{10} represents average pretest scores-mathematics achievement gain slope/effect, γ_{11} is the interaction effect between teacher's mathematics content certification and students' pretest scores, γ_{20} is average poverty-mathematics achievement gain slope/effect, and γ_{21} is the interaction effect between teacher's mathematics content certification and student FRL.

After substituting equation (5) in (4), the single-equation can be expressed as

$$\begin{aligned} (\text{GNMATHACH})_{ij} = & \gamma_{00} + \gamma_{01} (\text{CERTICONT})_j + \gamma_{02} (\text{TCHREXP})_j + \\ & \gamma_{10} (\text{PRETESTSCORES})_{ij} + \gamma_{11} (\text{CERTICONT})_j * (\text{PRETESTSCORES})_{ij} + \gamma_{20} (\text{FRL})_{ij} + \\ & \gamma_{21} (\text{CERTICONT})_j * (\text{FRL})_{ij} + u_{0j} + r_{ij} \end{aligned} \quad (6)$$

Equation (6) consists of fixed portions (containing γ terms as constants) and random portions (containing u and r terms as variables) of effects.

Considering the level-2 units as teachers and gain scores as an outcome measure, the d-type effect size for teacher level model can be computed using the following formula as provided in Rowan et al. (2002).

$$d = \frac{\sqrt{(\text{Variance in achievement gain lying among teachers})}}{\sqrt{(\text{Total student + teacher variance in student achievement gain})}} \quad (7)$$

The teacher-to-teacher variation (i.e., random effects) and fixed effects in all models are estimated using SAS PROC MIXED procedure (see Singer, 1998). The research questions 1 and 2 are answered based on the information of p-values associated with predictors, magnitude of variance components, and effect sizes.

Since the students are not placed within teachers' classrooms randomly, and student and teacher level data (predictors) incorporated in two separate models provide better estimates of variance and predictors' effects, our best choice of statistical design to measure the effects of teacher and student level predictors involves selecting the HLM technique. According to many researchers, hierarchical models can provide a general framework for this type of analysis (Goldstein, 1995; Raudenbush & Bryk, 2002; Subedi, 2004).

3. Results

Several academic and non-academic predictors at student and teacher levels are found significant in predicting middle school mathematics gain scores. The two research questions were answered based on the two-level HLM analysis of unconditional and conditional models. In order to answer research question 1, the findings of the study are presented in Table 1 which provides the predictors' effect estimates, standard errors, and p-values. At student level, pretest scores ($p < .0001$) as well as free and reduced lunch status ($p < .0001$) are found significant. Similarly, content certification ($p < .01$) and teacher experience ($p < .05$) are found significant at teacher level. We also examined the interaction effects between student and teacher level predictors. The results showed the significant interaction effects of teacher content certification with students' pretest scores ($p < .01$) and free and reduced lunch status

($p < .05$). The effect of student's free and reduced lunch status on mathematics gain scores is found negative, and all other individual as well as interaction effects are found positive.

In order to answer research question 2, the results of the study are presented in Table 2 which provides the percentages of variance explained, p-values, and effect sizes for unconditional and conditional models. The results showed significant teacher-to-teacher variation ($p < .0001$) for both unconditional and conditional models. The unconditional model accounted for 3.7% of variance with an effect size of 0.19 at teacher level, and the conditional model accounted for 4.9% of variance with an effect size of 0.22 at teacher level. These effect sizes are considered as small according to Rowan et al. (2002).

Table 1. Estimation of predictors' effects for predicting mathematics gain scores in conditional model

Predictors	Effect Estimate	Standard Error	p-Value
Pretest scores	0.03	0.001	<.0001
Free and reduced lunch	-2.81	0.253	<.0001
Content certification (Certi.)	1.76	0.632	0.005
Teacher experience	0.42	0.020	0.036
Content Certi. * Pretest scores	0.02	0.001	0.006
Content Certi. * Free and reduced lunch	0.75	0.302	0.013

Table 2. Estimations of variance explained, p-values, and effect sizes for teacher level models for predicting gain scores

Model	Variance Explained	p-value	Effect Size (d-type)
Unconditional (without predictors)	3.7%	<.0001	0.19
Conditional (with predictors)	4.9%	<.0001	0.22

4. Discussion

We predicted students' mathematics achievement gain through unconditional model employing two-level HLM and found significant teacher-to-teacher variability in average mathematics gain scores. This showed a significant contribution of teacher mean gain scores (averaged out for each teacher based on student performance) to students' mathematics gain scores in grades 6 through 8.

Several predictors at student and teacher levels significantly predicted mathematics gain scores for middle school students. The conditional model explored the effect of teacher quality or

teacher effectiveness, represented by teacher's mathematics content certification and teacher experience, on student mathematics gain scores. The study results showed that the effects of middle school teacher content certification and teacher experience as well as the interaction effects of teacher content certification with students' free and reduced lunch status and pretest scores in mathematics were significant. These results are consistent to the findings of Darling-Hammond et al. (2005) as well as Goldhaber and Brewer (2000). This provides us a useful piece of information regarding the importance of teacher characteristics and their interaction effects in predicting students' mathematics gain scores. The significant interaction effects imply that the middle school mathematics gain scores substantially depend on not only whether the students with high or low pretest scores are taught by a teacher holding content certification in mathematics but also whether the rich or poor students are taught by such teacher.

Using conditional VAM, we estimated the proportions of variance in gain scores lying among teachers based on conditional models incorporating significant student and teacher level predictors in the model. Further, we reported the effect sizes to measure teacher effectiveness or "teacher quality" based on the proportions of variance explained at teacher level. The conditional VAM for predicting student mathematics gain scores produced slightly larger percentage of variance explained and effect size than the unconditional model at teacher level although the effect sizes for both models were 'small'. Teacher's mathematics content certification and teacher experience were significant constructs of teacher effectiveness impacting students' mathematics gain scores.

5. Conclusions

This study predicted students' mathematics gain scores incorporating significant predictors at student and teacher levels using the data from one of the largest school districts in the U.S.A. We employed a two-level HLM using unconditional and conditional VAM models. The study determined the significant predictors at student and teacher levels. We also reported the proportions of variance explained and d-type effect sizes at teacher level for both unconditional and conditional models. Using conditional VAM approach to predict students' mathematics gain scores, this study measured teacher effectiveness represented by significant teacher level predictors.

The findings revealed significant effects of teacher's mathematics content certification, teacher experience and the interaction effects of content certification with students' pretest scores as well as participation in free and reduced lunch. We found that the conditional VAM explained more variance and produced slightly larger effect size at teacher level than that of the unconditional model. The findings of this study imply that the teacher effectiveness represented by teacher content certification in mathematics and teacher experience are important factors for predicting mathematics gain scores in middle schools.

This research would benefit the school systems, districts, and the State Departments of Education for teacher and school evaluation purposes. First, given the significant predictors of students' mathematics gain scores, the results will be beneficial since these potential predictors can be controlled to improve students' gain scores and reform schools. Second, researchers and

evaluators can replicate similar predictive models using HLM in order to measure the strength of teacher effectiveness through conditional VAM approach.

This study is limited within middle school level in one of the largest urban school districts in Florida, U.S.A. Future researches are suggested to cover the broader population. From analysis perspective, the models used in this study are limited to two-level and the hypotheses tested are limited to specific number of predictors in the model. Future studies are suggested to extend to three-level HLM incorporating potential school level predictors in the level-3 model.

References

- Allen, M. B. (2005). *Eight questions on teacher recruitment: What does the research say?* Denver: Education Commission of the States. <http://www.ecs.org/html/educationissues/teachingquality/trrreport/home/TeacherRecruitmentRetention.pdf>
- Ballou, D., & Podgursky, M. J. (2000). Reforming teacher preparation and licensing: What is the evidence? *Teachers College Record*, 102(1), 5-27. <http://dx.doi.org/10.1111/0161-4681.00046>
- Betts, J. R., Rueben, K. S., & Danenberg, A. (2000). *Equal resources, equal outcomes? The distribution of school resources and student achievement in California*. San Francisco: Public Policy Institute of California. http://www.ppic.org/content/pubs/report/r_200jbr.pdf
- Clotfelter C. T., Ladd H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. Cambridge, MA: National Bureau of Economic Research. http://www.caldercenter.org/PDF/1001104_Teacher_Credentials_HighSchool.pdf
- Croninger, R. G., Rice, J. K., Rathbun, A., & Nishio, M. (2007). Teacher qualifications and early learning: Effects of certification, degree, and experience on first-grade student achievement. *Economics of Education Review*, 26(3), 312-324. <http://dx.doi.org/10.1016/j.econedurev.2005.05.008>
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8(1), 1-44. <http://epaa.asu.edu/ojs/article/view/392/515>
- Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Heilig, J. V. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives*, 13(42), 1-55. <http://epaa.asu.edu/ojs/article/view/147/273>
- Decker, P.T., Mayer, D.P., & Glazerman, S. (2004). *The effects of Teach for America on students: Findings from a national evaluation*. Princeton, NJ: Mathematica. <http://www.mathematica-mpr.com/publications/pdfs/teach.pdf>
- Goldhaber, D., & Brewer, D. (1998). When should we reward degrees for teachers? *Phi Delta Kappan*, 80(2), 134-138.

Goldhaber, D., & Brewer, D. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, 22(2), 129-145. <http://dx.doi.org/10.3102/01623737022002129>

Goldstein, H. (1995). *Multilevel statistical models* (second edition). New York: John Wiley.

Government Accountability Office (GAO) (2006). *State face challenges measuring academic growth that education's initiatives may help address*. Report to Congressional requesters United States. <http://www.gao.gov/assets/260/250819.pdf>

Greenberg, E., Rhodes, D., Ye, X., & Stancavage, F. (2004). *Prepared to teach: Teacher preparation and student achievement in eighth-grade mathematics*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA. <http://www.air.org/files/AERA2004PreparedtoTeach.pdf>

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why public schools lose teachers. *Journal of Human Resources*, 39(2), 326–354. <http://dx.doi.org/10.2307/3559017>

Harris, D. N., & Sass, T. R. (2008). Teacher training, teacher quality, and student achievement. Working Paper 3. Washington, DC: National Center for Analysis of Longitudinal Data in Education Research, Urban Institute. http://www.caldercenter.org/PDF/1001059_Teacher_Training.pdf

Haycock, K. (1998). *Good Teaching Matters*. How well-qualified teachers can close the gap. *Thinking K-16*, 3(2), 1-14. http://www.edtrust.org/sites/edtrust.org/files/publications/files/k16_summer98.pdf

Howley, C. (1996). Compounding disadvantage: The effects of school and district size on student achievement in West Virginia. *Journal of Research in Rural Education*, 12, 25-32. <http://www.jrre.psu.edu/articles/v12,n1,p25-32,Howley.pdf>

Jerald, C. D., & Ingersoll R. M. (2002). *All talk, no action: Putting an end to out-of-field teaching*. Washington, DC: The Education Trust. <http://files.eric.ed.gov/fulltext/ED468741.pdf>

Jordan, H. R., Mendro, R., & Weerasinghe, D. (1997). Teacher effects on longitudinal student achievement: A preliminary report on research on teacher effectiveness. Paper presented at the National Evaluation Institute, Indianapolis, Indiana. <http://dallasisd.schoolwires.net/cms/lib/TX01001475/Centricity/Shared/evalacct/research/articles/Jordan-Teacher-Effects-on-Longitudinal-Student-Achievement-1997.pdf>

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2006). *What does certification tell us about teacher effectiveness? Evidence from New York City*. A Technical Report. Cambridge, MA: National Bureau of Economic Research. <http://www.nber.org/papers/w12155.pdf>

Klecker B. M. (2008). *Is teacher quality related to eighth-grade mathematics achievement? Evidence from the 2007 NAEP Data*. Paper presented at annual meeting of Mid-South Educational Research Association, Knoxville, TN. <http://www.eric.ed.gov/PDFS/ED503409.pdf>

- Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, 24(1), 37-62. <http://dx.doi.org/10.3102/01623737024001037>
- Lippman, L., Burns, S., & McArthur, E. (1996). *Urban schools: The challenge of location and poverty*. A Technical Report. Washington, DC: U.S. Department of Education. from <http://nces.ed.gov/pubs/96184all.pdf>
- Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review*, 13(2), 125-145. [http://dx.doi.org/10.1016/0272-7757\(94\)90003-5](http://dx.doi.org/10.1016/0272-7757(94)90003-5)
- National Science Board (2012). *Science and Engineering Indicators 2012*. Arlington VA: National Science Foundation (NSB 12-01). <http://www.nsf.gov/statistics/seind12/pdf/seind12.pdf>
- Nye, B., S. Konstantopoulos, & L. Hedges (2004). How Large Are Teacher Effects? *Educational Evaluation and Policy Analysis*, 26(3), 237-257. <http://dx.doi.org/10.3102/01623737026003237>
- Pituch, K. A. (1999). Describing school effects with residual terms: Modeling the interaction between school practice and student background. *Evaluation Review*, 23(2), 190-211.
- Raudenbush, S. W., & Bryk, A. (2002). *Hierarchical linear models in social and behavioral research: Applications and data analysis methods* (second edition). Newbury Park, CA: Sage.
- Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Chicago: Spencer Foundation. http://www.epi.org/publication/books_teacher_quality_execsum_intro/
- Rice J. K. (2010). *The impact of teacher experience: Examining the evidence and policy implications*. Washington, DC: National Center for the Analysis of Longitudinal Data in Education Research, Urban Institute. <http://www.urban.org/uploadedpdf/1001455-impact-teacher-experience.pdf>
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417-458. <http://dx.doi.org/10.1111/j.1468-0262.2005.00584.x>
- Rowan, B., Correnti, R., & Miller, R. R. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the "Prospects" study of elementary schools. Philadelphia, PA: Consortium for Policy Research in Education. http://www.americancivil.literacy.org/sites/default/files/researchreport/791_rr51.pdf
- Sanders, W., & Rivers, J. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville: University of Tennessee, Value-Added Research and Assessment Center. http://www.cgp.upenn.edu/pdf/Sanders_Rivers-TVASS_teacher%20effects.pdf

Scheerens, J., & Bosker, R. (1997). *The foundations of educational effectiveness*. New York: Pergamon.

Shields, P. M., Esch, C. E., Humphrey, D. C., Wechsler, M. E., Chang-Ross, C. M., Gallagher, H. A., Guha, R., Tiffany-Morales, J. D., & Woodworth, K. R. (2003). *The status of the 155 teaching profession 2003*. Santa Cruz, CA: The Center for the Future of Teaching and Learning.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, 24(4), 323-355. <http://www.biostat.jhsph.edu/~fdominic/teaching/bio656/references/singer.1998.pdf>

Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education*, 62(4), 339-355. <http://dx.doi.org/10.1177/0022487111404241>

Subedi, B .R. (2004). A Demonstration of three-level hierarchical generalized linear model applied to educational research. Doctoral dissertation, Florida State University, Tallahassee, Florida, U.S.A. <http://diginole.lib.fsu.edu/cgi/viewcontent.cgi?article=4896&context=etd>

Swan, B. A., Dixon, J. K., & Subedi, B .R. (2005). Middle-school teacher preparation, certification and students' mathematics achievement in a large urban district. A Technical Report. The Multi-University Reading, Mathematics and Science Initiative. Learning System Institute, Florida State University, Tallahassee, Florida, U.S.A.

Swan, B. A. (2006). Middle school mathematics teacher certification, degree level, and experience, and the effects on teacher attrition and student mathematics achievement in a large urban district. Doctoral dissertation, University of Central Florida, Orlando, Florida, U.S.A. [http://www.floridainclusionnetwork.com/Uploads/1/docs/murmsi/Swan%20Dissertation%20\(2006\).pdf](http://www.floridainclusionnetwork.com/Uploads/1/docs/murmsi/Swan%20Dissertation%20(2006).pdf)

Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Educational Policy Analysis Archives*, 10(12). http://www.indiana.edu/~educy520/sec6342/week_07/wenglinsky02.pdf

Copyright Disclaimer

Copyright reserved by the authors.

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).