Exploring Predictors of Highly Effective Teachers Through Multilevel Model

Bidya Raj Subedi, Ph.D.
School District of Palm Beach County,
Florida, U.S.A.

Mark Howard
School District of Palm Beach County,
Florida, U.S.A.

ABSTRACT

This paper explored significant teacher and school level predictors of the status of highly effective teachers based on Florida teacher evaluation system employing a two-level multilevel analysis approach known as hierarchical generalized linear model (HGLM). The analysis used 3,895 teachers who taught students in Kindergarten to Grade 12 and 210 schools from one of the largest urban school districts in United States of America. In this study, teacher and school level data are used at level-1 and level-2 models, respectively. At teacher level, the results showed significant effects of percentage of student level demographic, academic, and disciplinary variables aggregated at teacher level as well as teacher's experience and educational degree level. At school level, percentages of Black and percentage of Hispanic students showed significant effects on the status of highly effective teachers. The percentage of variance explained at school level is found 19.5% with an effect size of 0.44 which is determined as a "medium" size of school effect.

Key words: Highly effective teachers, significant predictors, school variance, school effects, hierarchical generalized model.

INTRODUCTION

It is important to measure the effectiveness of teachers' instructional practice in school systems employing a predictive model. The significant predictors can be explored at teacher and school levels in order to measure teacher effectiveness. For the purpose of this research, the effectiveness of a teacher is defined as the rating of 'highly effective' status of a teacher specified by the State of Florida mandate. The Florida Department of Education has endorsed a mandatory teacher evaluation policy for all school districts in Florida employing different rating components. This study used the teacher evaluation data from school year 2017-18. The final teacher evaluation rating components for the school year 2017-18 included instructional practice (57%), student performance rating (33%), and professional growth (10%). Applying these rating components, teachers were classified in state mandated four evaluation categories: highly effective, effective, needing improvement or developing, and unsatisfactory. The current research is based on the School District of Palm Beach County (SDPBC) data in which we dichotomized the original classification rating scores. Specifically, the classification of teachers is dichotomized into two groups by dividing the classification of "highly effective" teachers in first group and the remaining other categories (i.e., not highly effective) in second group by transforming to a binary outcome measure to generate an appropriate predictive model. The SDPBC is the fifth largest urban school district in Florida and eleventh largest school district in USA.

In instructional practice component of teacher evaluation, the SDPBC used the Marzano's Teacher Evaluation Model (Marzano, 2013) known as Palm Beach Model of Instruction. This model is based on Marzano's extensive research for several decades and identifies a complete set of practices directly associated with improved student performance, structured into four domains that develop teacher expertise. The four domains are: classroom strategies and behaviors, preparing and planning, reflecting on teaching, and collegiality and professionalism. Each of these domain builds on the previous one with direct links to create a causal chain that results in increased learning and achievement for all students. In this component, the instructional practice rating score is calculated using a competency-based scoring system.

The student performance rating component included survey 2 (based on fall semester student data) and survey 3 (based on spring semester student data) students match with the exception of semester long acceleration courses as well as grades 11 and 12 intensive reading courses. The teachers teaching courses in kindergarten to grade 12 are evaluated with a requirement of at least 10 students in related course with both pre and post-test scores. The state English Language Assessment (ELA) and mathematics VAM (Value-Added Model) used the percentage of students meeting or exceed expected scores as post-test measures. The other models used current year's standardized scale scores and pass or fail scores as post-test measures.

In professional growth plan or deliberate practice rating, the four-point scoring rubric for teacher evaluation included growths of level 2 (or rated innovating), level 1, no growth, and element not rated during any observation.

In order to predict the status of highly effective teachers, this study generated a two-level hierarchical research design, known as hierarchical generalized linear model (HGLM) incorporating predictors at teacher and school levels. This paper aims to identify significant predictors at teacher as well as school levels and determine school effect while predicting the status of highly effective teachers.

REVIEW ON PREDICTORS OF HIGHLY EFFECTIVE TEACHERS
Searching Teacher Level Predictors
Most of the teacher level predictors in current study are aggregated from student level predictors. For example, percentage of Hispanic students, one of the predictors used in this study, is an aggregate at teacher level from student's race as a Hispanic student status. Since most of the predictors are derived from student level, our attempt will be to review the effect of predictors, on teacher effectiveness, either at student level or at teacher level.

Past studies show positive effect of student achievement (GPA) on teacher efficacy. In a study related to effectiveness of secondary teachers using a researcher-developed instrument, Ferguson and Womack (1993) found GPA being the best predictor in introductory education courses related to effectiveness. Lucas and Schmitz (1991) found a significant relationship between student-teacher evaluations and both the cumulative GPA and scores on the English portion of the basic subject exam. Moulding, Stewart, and Dunmeyer (2014) found that teacher efficacy scores (that were related to teachher effectiveness) were significantly higher for pre-service teachers in schools with higher student achievement. Significant positive relationships were found between collective teacher efficacy (CTE) scores and student achievement on the grade 8 math, writing, and English tests. Further, significant relationships were found between subscales of CTE with all three tests of student achievement (Tschannen-Moran and Barr, 2004). Past studies found teacher self-efficacy as a better predictor than teachers' attitudes for predicting teachers' willingness to differentiate instruction for gifted students (Caldwell, 2012)

and gifted students received higher intellectual ability ratings than average-ability students (Matheis, Kronborg, Schmitt, & Preckel; 2017).

This research found a negative relationship between the status of highly effective teachers and Black as well as Hispanic students. Mangiante (2011) outlined measures of teacher effectiveness in low-income minority schools in accordance with several factors. Padron, Waxman, and, Rivera (2002) suggested five effective instructional practices for educating Hispanic students. Such practices included culturally-responsive teaching, cooperative learning, instructional conversations, cognitively-guided instruction, and technology-enriched instruction. Based on National Assessment of Educational Progress (NAEP) reading data analysis, Ortiz (1986) found that Hispanic and Black children continue to read at significantly lower levels than Whites. In a research using SDPBC high school data, Subedi and Howard (2017) found significant negative effects of African-American and Hispanic students on reading achievement producing its negative interaction effect with teacher effectiveness.

Past researches revealed student's status of English language learner (ELL) as a barrier in education and student's disability status (ESE) impacting negatively on teacher efficacy. Duran, Dugan, and Weffer (1997) argue that Mexican American language minority (ELL) students are considered to have a (English) language problem that interferes with schooling. The distance between student skills and the educational challenge of the content areas continues to be an obstacle, and the teachers as well as students struggle with this obstacle (Duran et al., 1997). Further, Gersten, Walker, and Darch (1998) found that those teachers with the most effective teaching strategies for low-achieving students tended to report (a) that they tolerate less maladaptive behavior in their classrooms and (b) that they may actively resist placement of handicapped students in their (class)rooms.

Student's misbehavior has been consistently linked to teachers' stress (Yoon, 2002). Past research shows the relationship between student's discipline and teacher's efficacy. Findings in past revealed that disruptive behaviors take time away from instruction, hamper teachers' efforts to sustain a positive learning climate, and consequently weaken teachers' self-efficacy beliefs to effectively deal with individual students (Zee, Jong, & Koomen; 2016). In a research using SDPBC data, Subedi and Howard (2017) found significant negative effect of student suspensions on reading achievement in elementary, middle and high schools showing a negative interaction effect of student suspensions with teacher effectiveness in middle schools.

With experience, many teachers come to oppose permissiveness and take on a more custodial pupil control ideology than they held in their early years of teaching (Willower, Eidell, Hoy; 1988), and indeed, in some schools good control and good teaching are equated (Hoy & Spero, 2005). Wolters and Daugherty (2007) corroborated that aspects of teachers' sense of efficacy were greater for those teachers with more teaching experience, and Hanushek (2005) found a significant effect of teacher experience on teacher quality. Based on the review of 30 studies published within the last 15 years that analyzed the effect of teaching experience on student outcomes in the United States, Kini and Podolsky (2016) found teaching experience positively associated with student achievement gains throughout a teacher's career. Further, the researchers report that gains in teacher effectiveness associated with experience are most steep in teachers' initial years, but continue to be significant as teachers reach the second, and often third, decades of their careers. According to Tschannen-Moran and Hoy (2007), contextual factors such as the teaching resources and interpersonal support available were found to be much more salient in the teachers' self-efficacy beliefs of novice teachers, however, among experienced teachers for whom an abundance of mastery experiences were available,

contextual factors played far less important role in their self-efficacy beliefs. More experienced teachers confer benefits to their colleagues, their students, and to the school as a whole.

Past studies show positive effect of teacher's degree on student academic outcomes. In a review of five studies on teacher effectiveness, Rice (2003) found that teachers who have earned advanced degrees have a positive impact on high school mathematics and science achievement when the degrees earned were in these subjects. In past, Subedi, Reese, and Powell (2015) found a significant positive effect of teacher's degree (education) level on student's high school GPA. The simple truth is that the public education cannot fulfill its mission if students growing up in poverty, students of color and low-performing students continue to be disproportionately taught by inexperienced, under-qualified teachers (Peske & Haycock, 2006).

Searching School Level Predictors

Past research has evidenced that the need for qualified teachers is particularly great in lower-performing schools with higher numbers of minority and low-income students (Allen, 2005).

Sorting students by achievement level exposes minority and poor students to lower quality teachers as well as less resourced classmates and the probability of having a novice teacher is higher for black and Hispanic high school students relative to white students after controlling for prior achievement (Kalogrides and Loeb, 2013). Teachers in the highest minority schools have lower value-added on average, regardless of experience (Steele, Pepper, Springer, and Lockwood; 2015). In Washoe County, Nevada, classrooms with higher concentrations of minority, poor, and low-achieving students were more likely to be taught by teachers with lower evaluation scores (Borman, & Kimball, 2005).

METHODS

Design of the Study

This study predicts the status of highly effective teachers incorporating teacher and school level predictors employing a two-level hierarchical generalized linear model (HGLM). The most appropriate statistical design to measure the school effects involves a HGLM technique because the teachers were not assigned randomly within schools and the predictors incorporated separately at teacher and school level models provide better estimates of variance and predictors' effects (Raudenbush & Bryk, 2002; Goldstein, 1997). In order to determine school level effect sizes, we need variance estimation at teacher and school levels. Past studies show that the estimation of level-1 variance components can be estimated in binary response models (Bryk and Thum, 1989; Finn and Rock, 1997; Goldstein, 1991; Guo & Zhao, 2000; Longford, 1994; McCulloch, 1994). For example, Bryk and Thum (1989) predicted dropout as a binary outcome and estimated variance associated with dropout and Goldstein (1991) adopted a procedure to estimate variance at level-1 multilevel nonlinear model by means of linearization.

Data and Variables

This study used 3895 teachers, who taught students in Kindergarten (K) to Grade 12, and 210 schools from SDPBC. The district-operated elementary, middle and high schools as well as charter schools are included in this study. Based on past literature review, we used potential teacher and school level predictors to predict teacher level outcome measure in this study.

Outcome variable

The status of highly effective teachers teaching in Grades K–12 in SDPBC is used as an outcome measure in this study. The teachers included in this study were originally evaluated and classified in four categories: highly effective, effective, needing improvement or developing,

and unsatisfactory. In the final teacher categories with binary transformation, the highly effective category was coded as 1 and the rest categories were coded as 0.

Teacher Level Predictors

Percentage of high GPA. This predictor is derived from student level predictor, which originally ranged from 0 to 4, aggregated at teacher level as a percentage of students with 2.0 or more GPA for each teacher. This predictor ranged from 0% to 100%.

Percentage of Black. This predictor is the percentage of Black (African-American) students taught by each teacher. This predictor is generated by aggregating student level predictor which was coded as 1 for Black students and 0 for others. This predictor ranged from 0% to 100%.

Percentage of Hispanic. This predictor is the percentage of Hispanic students taught by each teacher. This predictor is generated by aggregating student level predictor which was coded as 1 for Hispanic students and 0 for others. This predictor ranged from 0% to 100%.

Percentage of ESE. This predictor is the percentage of ESE students taught by each teacher. This predictor is generated by aggregating student level predictor which was coded as 1 for ESE students and 0 for others. This predictor ranged from 0% to 100%. The ESE students included those students with following physical or mental disorder conditions: orthopedically impaired, occupational therapy, physical therapy, dual-sensory impaired, autism spectrum disorder, language impaired, traumatic brain injured, deaf or hard of hearing, developmentally delayed, visually impaired, established conditions, emotional as well as behavioral disability, other health impaired, specific learning disability and intellectual disability.

Percentage of ELL. This predictor is the percentage of ELL students taught by each teacher. This predictor is generated by aggregating student level predictor which was coded as 1 for ELL students and 0 for others. This predictor ranged from 0% to 100%. The student ELL status is defined as LY (the student is an English Language Learner and is enrolled in classes specifically designed for English Language Learners) and LF (the student is being followed up for a two-year period after having exited from the ESOL program) codes.

Percentage of Gifted. This predictor is the percentage of gifted students taught by each teacher. This predictor is generated by aggregating student level predictor which was coded as 1 for gifted students and 0 for others. This predictor ranged from 0% to 100%.

Average OSS days. This predictor is the students' OSS days averaged for each teacher. This predictor is generated by computing the average of student level OSS days for each teacher. This predictor ranged from 0 to 10.

Teacher experience. This is a continuous predictor with total number of teaching years in SDPBC. This predictor ranged from less than one year to 37 years.

Teacher's degree level. This is an ordinal variable that ranged from 1 to 4. The values for this predictor is coded as 4 for doctorate degree, 3 for master or post-master degrees, 2 for bachelor degree, and 1 for associate degree or college education below bachelor degree.

School Level Predictors

Percentage of Black. The school percentage of Black students was generated by aggregating student level predictor (which was coded as 0 and 1) at school level. This predictor ranged from 0% to 100%.

Percentage of Hispanic. The school percentage of Hispanic students was generated by aggregating student level predictor (which was coded as 0 and 1) at school level. This predictor ranged from 11.5% to 100%.

Research Questions

The following research questions are addressed through this study:
1. What are the significant predictors at teacher and school levels for predicting the status of highly effective teacher?
2. What are the proportion of variance explained and effect size at school level for predicting the status of highly effective teacher?

Appropriate Model Selection

Since the status of highly effective teacher is a dichotomous outcome measure, the most appropriate model to predict such outcome is a two-level HGLM with teacher and school level data at level-1 and level-2 models, respectively. Past literature revealed that employing such type of model for binary response outcomes (Goldstein, 1991; Raudenbush & Bryk, 2002; Subedi & Powell, 2016). For the purpose of estimating random effects, several studies in past used the estimation of level-1 variance components in binary response model (Bryk and Thum, 1989; Goldstein, 1991; Guo & Zhao, 2000; Longford, 1994; McCulloch, 1994). Bryk and Thum (1989) predicted dropout as a binary outcome and estimated variance associated with dropout. Similarly, Goldstein (1991) adopted a general approach for the estimation of variance (at level-1 model) in multilevel nonlinear model using a linearization.

Several past studies also determined effect sizes for school level employing multilevel models (Goldstein, 1997; Rowan, Correnti, & Miller, 2002; Thomas, Sammons, Mortimore, & Smees, 1997). They determined effectiveness based on effect sizes which were computed using variance of school level model.

Model Development

In order to predict highly effective teachers, this study employed a two-level HGLM where teacher and school data were incorporated in level-1 and level-2 models, respectively. The conditional model was developed incorporating only significant predictors in the model. The fixed and random effects were estimated and presented in Table 1 and Table 2, respectively.

Assuming that Y_{ij} is the status of highly effective (HE) teacher, the log of probability of HE can be predicted by the level-1 conditional model for i^{th} teacher nested in j^{th} school as given by Equation (1).

$$\log(P(Y_{ij}=1)/(1- P(Y_{ij}=1))) = \beta_{0j} + \beta_{1j} (PCTHIGPA)_{ij} + \beta_{2j} (PCTBLACK)_{ij} + \beta_{3j} (PCTHISP)_{ij} +$$
$$\beta_{4j} (PCTESE)_{ij} + \beta_{5j} (PCTELL)_{ij} + \beta_{6j} (PCTGIFTED)_{ij} + \beta_{7j} (AVGOSS)_{ij} +$$
$$+\beta_{8j} (TCHREXP)_{ij} + \beta_{9j} (TCHDEGLVL)_{ij} + e_{ij} \qquad (1)$$

In Equation (1), β_{0j} is the intercept. The coefficients β_{1j}, β_{2j}, and β_{3j}, β_{4j}, β_{5j}, and β_{6j}, β_{7j}, β_{8j}, and β_{9j} are the effects of percentages of high GPA, percentages of Black, Hispanic, ESE, ELL, gifted students, average OSS days, teacher experience and teacher's degree level, respectively. The

term e_{ij} is teacher level random term distributed normally with mean zero and constant variance.

Further, in order to predict the coefficients in Equation (1), the level-2 model can be formulated as below incorporating school level predictors.

$\beta_{0j} = \gamma_{00} + \gamma_{01} (SCHLPCTBLACK)_j + \gamma_{02} (SCHLPCTHISP)_j + u_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30}$

$\beta_{4j} = \gamma_{40}$

$\beta_{5j} = \gamma_{50}$ (2)

$\beta_{6j} = \gamma_{60}$

$\beta_{7j} = \gamma_{70}$

$\beta_{8j} = \gamma_{80}$

$\beta_{9j} = \gamma_{90}$

In Equation (2), γ_{00} represents the average rate of highly effective status of teacher for all schools and u_{0j} represents the random effects at school level with multivariate normal distribution. The coefficients γ_{01} and γ_{02} represent the effects of school percentages of Black and Hispanic students, respectively, on highly effective teacher status.

The following coefficients represent their effects on the predicted probability of highly effective teachers:

- γ_{10} represents the effect of percentages of high GPA,
- γ_{20} represents the effect of percentages of Black students,
- γ_{30} represents the effect of percentages of Hispanic students,
- γ_{40} represents the effect of percentages of ESE students,
- γ_{50} represents the effect of percentages of ELL students,
- γ_{60} represents the effect of percentages of gifted students,
- γ_{70} represents the effect of average OSS days,
- γ_{80} represents the effect of teacher experience,
- γ_{90} represents the effect of teacher's degree level.

The fixed effects (intercepts and slopes) and random effects (variance components) at teacher and school levels are estimated using PROC GLIMMIX procedure in SAS program (Kim, Preisser, Rozier, & Valiyaparambil, 2006; SAS Institute, 2006).

The research question 1 is addressed by estimating fixed effects, γs, and p-values associated with these effects in Equation (2). The research question 2 is addressed by calculating school level effect size using the formula given in Equation (3) below, based on the estimated school level variance term (u_{0j}), as provided by Rowan et al. (2002).

d = √(Variance in teacher effectiveness among schools) / √(Total teacher
+ school variance in teacher effectiveness)   (3)

According to Rosenthal (1994), the variance explained by school can be translated into effect size which is termed as d-type effect size.

RESULTS
Table 1 shows the list of significant predictors (at .05 significance level), based on HGLM analysis predicting highly effective teachers in grades K-12, estimate of effects, standard

errors, and p-values. At teacher level, the effects of percentage of students with high GPA (p = .0314), the percentage of Black (p <.0001), percentage of Hispanic (p = .0044), percentage of ESE (p = .0004), percentage of ELL (p = .0060), percentage of gifted (p = .0067) students, average OSS days (p = .0367), teacher experience (p < .0001), and teacher's degree level (p = .0117) are found significant. At school level, the effects of percentage of Black (p <.0001) and percentage of Hispanic (p=.0127) students are found significant.

The results showed the positive effects of the percentages of students with high GPA, percentage of gifted students, teacher experience and teacher's degree level. The direction of effects of other predictors are found negative.

We also computed the effect size at school level. Table 2 shows the variance components, p-values, percentage of variance explained and d-type effect size (at school level) for predicting the status of highly effective teachers. The proportion of variance explained and effect size are found 19.5% and 0.44, respectively. The effect size is classified as 'medium' representing the moderate strength of school effect while predicting the status of highly effective teachers.

Table 1. Estimation of predictors' effects for predicting highly effective teachers

Predictors	Estimate	Std. Error	p-value
Teacher level			
Percentage high GPA	0.070	0.033	<.0314
Percentage Black	-0.205	0.037	<.0001
Percentage Hispanic	-0.106	0.037	.0044
Percentage ESE	-0.080	0.023	.0004
Percentage ELL	-0.089	0.033	.0060
Percentage Gifted	0.151	0.056	.0067
Average OSS days	-0.029	0.014	.0367
Teacher experience	0.005	0.001	<.0001
Teacher's degree level	0.032	0.013	.0117
School level			
Percentage Black	-0.476	0.086	<.0001
Percentage Hispanic	-0.273	0.110	.0127

Table 2. Estimations of variance explained, p-values, and effect sizes at school level for predicting highly effective teachers

Group	Variance Component	p-value	Variance Explained	Effect size (d-type)
School level	.0410	<.0001	19.5%	0.44 (Medium)
Teacher level	.1697	<.0001		

DISCUSSION

This paper predicted the status of highly effective teachers incorporating teacher and school level predictors employing multilevel models. As supported by past researches, higher GPA as well as student achievement impacted positively on teacher efficacy or teacher effectiveness (Ferguson and Womack, 1993; Lucas and Schmitz, 1991; Moulding, Stewart, and Dunmeyer, 2014; Tschannen-Moran and Barr, 2004). Intuitively, these findings imply that there will be increased teacher effects if the students in their classrooms have high academic achievement. Further, percentage of gifted students correlated positively with teacher effectiveness supporting Caldwell (2012) and Matheis et al. (2017). Without further reasoning, it can be perceived that gifted students support instructional effectiveness by producing their good academic results.

The results of this study which showed negative effects on the status of highly effective teachers due to percentages of Black and Hispanic students (aggregated at teacher level) are supported by past research works (Duran et al.; 1997; Subedi & Howard, 2017). Further, the average out-of-school suspension (OSS), which is a part of student's discipline, impacted negatively on the status of highly effective teachers. This result is supported by Yoon (2002), which showed that the student's misbehavior is consistently linked to teachers' stress. Further, Zee et al. (2016) reported that disruptive behaviors hamper teachers' efforts to sustain a positive learning climate, and weakening teachers' self-efficacy beliefs to effectively deal with individual students. Student's statuses of ELL and ESE influenced negatively on the status of highly effective teachers which is supported by several researches. Past study show that ELL minority students are considered to have an English language problem that interferes with schooling (Duran et al., 1997). According to Gersten et al. (1998), teachers with the most effective teaching strategies for low-achieving students tended to report that they tolerate less maladaptive behavior in their classrooms and they may actively resist placement of handicapped students in their classrooms.

This study found a positive effect of teacher experience on the status of highly effective teachers. Intuitively, the experience in teaching helps teachers gain more mastery in teaching skills and, consequently, this would increase their self-efficacy. Hanushek (2005) found a significant effect of teacher experience on teacher quality. According to Wolters and Daugherty (2007), the aspects of teachers' sense of efficacy were greater for those teachers with more teaching experience. With experience, teachers come to oppose permissiveness and take on a more custodial pupil control ideology than they held in their early years of teaching (Willower et al., 1988). Similarly, teacher's degree produced a positive effect on the status of highly effective teachers which is consistent with past research findings that show significant positive effect of teacher's degree either on student achievement or teacher's qualification. In a review of five studies on teacher effectiveness, Rice (2003) found that teachers who have earned advanced degrees have a positive impact on high school mathematics and science achievement when the degrees earned were in these subjects. Subedi, Reese, and Powell (2015) found a

significant positive effect of teacher's degree level on student's high school GPA. According to Peske and Haycock (2006), the public education cannot fulfill its mission if students growing up in poverty, students of color and low-performing students continue to be disproportionately taught by under-qualified teachers.

Further, two school level predictors, school percentages of Black and Hispanic (minority) students, produced significant negative effects on the status of highly effective teachers. These results are supported by Allen (2005) which found that the need for qualified teachers is particularly great in lower-performing schools with higher numbers of minority. Similarly, Kalogrides and Loeb (2013) found that the probability of having a novice teacher is higher for Black and Hispanic high school students relative to white students after controlling for prior achievement. Borman and Kimball) (2005) reported that the classrooms with higher concentrations of minority students were more likely to be taught by teachers with lower evaluation scores, and further, Steele et al. (2015) corroborated that teachers in the highest minority schools have lower value-added on average, regardless of experience.

School Effects
The school effect in this study is computed by the effect size at school level model while predicting the status of highly effective teachers. The effect size, based on the percentage of variance explained at school level, is determined to be 0.44, which is classified to be 'medium' according to Rowan et al. (2002) implying moderate strength of school effect to predict the status of highly effective teachers.

This paper applied the approach to determine effect size for dichotomous outcomes using variance among schools in highly effective teachers. As a method, this study extended the approach of Rowan et al. (2002) computing effect sizes for level-2 model in binary response models. For this purpose, we assumed the computation of level-1 variance in generalized linear model supported by several past studies (Goldstein, 1991; Guo & Zhao, 2000; Kim et al., 2006).

CONCLUSIONS
This study predicted the status of highly effective teachers employing a multilevel model incorporating significant predictors at teacher and school levels. We explored the substantial predictors of highly effective teachers using the data from one of the largest school districts in USA. Several demographic, academic, and disciplinary predictors were found significant at teacher and school levels. We also determined the effect size at school level, which is found a "moderate" size of effect, based on the percentage of variance explained.

This paper offers several implications for educational policy makers. Given the exploration of significant predictors at teacher and school levels for predicting the status of highly effective teachers, an intervention is recommended by controlling such predictors for related teachers in those schools with low percentage of highly effective teachers. Such intervention will help excel those teachers who were not able to be classified as "highly effective" teachers. Consequently, this would help the schools and districts increase the percentage of highly effective teachers.

Employing a two-level HGLM, this study systematically demonstrated a valid approach for computing effect size while predicting a binary outcome that could be beneficial for educational research practitioners. Such a theoretically grounded and empirically evidenced model would be useful for other school districts to measure school effects. This study is limited to two-level models. Therefore, researchers in future are suggested to explore the predictors of highly effective teachers at student, teacher, and school levels employing a three-level HGLM.

References

Allen, M. B. (2005). *Eight questions on teacher recruitment: What does the research say*? Denver, CO: Education Commission of the States. https://files.eric.ed.gov/fulltext/ED489332.pdf.

Borman, G. D. and Kimball, S. M. (2005). Teacher quality and educational equality: Do teachers with higher standards-based evaluation ratings close student achievement gaps? *The Elementary School Journal*, 6(1), 3-20. https://doi.org/10.1086/496904

Bryk, A. S., & Thum, Y. M. (1989). The effects of high school organization on dropping out: An exploratory investigation. *American Educational Research Journal*, 26(3), 353-383. https://doi.org/10.3102/00028312026003353

Caldwell, D. W. (2012). *Educating gifted students in the regular classroom: Efficacy, attitudes, and differentiation of instruction*. Unpublished Doctoral Dissertation. https://digitalcommons.georgiasouthern.edu/etd/822

Duran, B. J., Dugan, T., & Weffer, R. E. (1997). Increasing teacher effectiveness with language minority students. *The High School Journal, 80*(4), 238-246. https://www.jstor.org/stable/40364455?seq=1#page_scan_tab_contents

Ferguson, P., & Womack, S. T. (1993). The impact of subject matter and education coursework on teaching performance. *Journal of Teacher Education*, 4(1), 55-63.

Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology*, 82(2), 221-234. http://dx.doi.org/10.1037/0021-9010.82.2.221

Gersten, R., Walker, H., & Darch, C. (1988). Relationship between teachers' effectiveness and their tolerance for handicapped students. *Exceptional Children*, 54(5), 433-438. https://doi.org/10.1177/001440298805400506

Goldstein, H. (1997). Methods in school effectiveness research. *School Effectiveness and School Improvement*, 8, 369-95.

Goldstein, H. (1991). Nonlinear multilevel models, with an application to discrete response data. *Biometrika*, 78, 45-51.

Guo, G., & Zhao, H. (2000). Multilevel modeling for binary data. *Annual Review of Sociology*, 26(1), 441-462.

Hanushek, E. A., Kain, J. F., O'Brien, D. M., Rivkin, S. G. (2005). *The market for teacher quality*. MA, Massachusetts Avenue Cambridge: National Bureau of Economic Research. doi: 10.3386/w11154

Hoy, A. W., & Spero, R. B. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education*, 21(6), 343-356. https://www.sciencedirect.com/science/article/pii/S0742051X05000193

Kalogrides, D., & Loeb, S. (2013). Different teachers, different peers: The magnitude of student sorting within schools. *Educational Researcher*, 42(6), 304-316. https://doi.org/10.3102/0013189X13495087

Kim H.Y., Preisser, J. S., Rozier, R. G., & Valiyaparambil, J. V. (2006). Multilevel analysis of group-randomized trials with binary outcomes. *Community Dentistry and Oral Epidemiology*, 34(4), 241–251.

Kini, T. & Podolsky, A. (2016). *Does teaching experience increase teacher effectiveness? A review of the research*. Palo Alto, CA: Learning Policy Institute. https://learningpolicyinstitute.org/sites/default/files/product-_files/Teaching_Experience_Report_June_2016.pdf

Longford, N.T. (1994). Logistic regression with random coefficients. *Computational Statistics and Data Analysis*, 17, 1-15.

Lucas, C. J., & Schmitz, C. D. (1991). A comparison of academic and affective variables as predictors of student teaching performance. *College Student Journal*, 25, 518-531.

Mangiante, E. (2011). Teachers matter: Measures of teacher effectiveness in low-income minority schools. *Educational Assessment, Evaluation, & Accountability*, 23(1), 41- 63. doi: 10.1007/s11092-010-9107-x

Marzano, R. J. (2013). *The Marzano teacher evaluation model*. Bloomington, IN: Marzano Research Laboratory. http://www.k12.wa.us/TPEP/Frameworks/Marzano/Marzano_Teacher_Evaluation_Model.pdf

Matheis, S., Kronborg, L., Schmitt, M., & Preckel, F. (2017). Threat or challenge? Teacher beliefs about gifted students and their relationship to teacher motivation. *Gifted and Talented International*, 32 (2), 134-160. https://doi.org/10.1080/15332276.2018.1537685

McCulloch, C. E. (1994). Maximum likelihood variance components estimation for binary data. *Journal of the American Statistical Association*, 89, 330-335.

Moulding, L. R., Stewart, P. W., & Dunmeyer, M. L. (2014). Pre-service teachers' sense of efficacy: Relationship to academic ability, student teaching placement characteristics, and mentor support. *Teaching and Teacher Education*, 14, 60-66. https://doi.org/10.1016/j.tate.2014.03.007

Ortiz, V. (1986). Reading activities and reading proficiency among Hispanic, Black, and White students. *American Journal of Education*, 95(1), 58–76.

Padron, Y. N., Waxman, H. C., Rivera, H. H. (2002). *Educating Hispanic students: Effective instructional practices*. Practitioner Brief #5. Berkeley, CA: Center for Research on Education, Diversity & Excellence. https://files.eric.ed.gov/fulltext/ED499031.pdf

Peske, H. G. & Haycock, K. (2006). *Teaching inequality: How poor and minority students are shortchanged on teacher quality*. Washington, DC: The Education Trust. https://files.eric.ed.gov/fulltext/ED494820.pdf

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, IL: Spencer Foundation. https://www.epi.org/publication/books_teacher_quality_execsum_intro/

Rosenthal, R. (1994). *Parametric measures of effect size*. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), The Handbook of Research Synthesis (pp. 231-244). New York: Russell Sage Foundation.

Rowan, B., Correnti, R., & Miller, R. (2002). What large-scale, survey research tells us about teacher effects on student achievement; Insights from the "Prospects" study of elementary schools. *Teachers College Record,* 104(8), 1525-1567.

SAS Institute (2006). *SAS/STAT 9.1 production GLIMMIX procedure for Windows*. Cary, NC: SAS Institute Inc.

Steele, J. L., Pepper, M. J., Springer, M. G., & Lockwood, J.R. (2015). The distribution and mobility of effective teachers: Evidence from a large, urban school district. *Economics of Education Review*, 48, 86-101. https://doi.org/10.1016/j.econedurev.2015.05.009

Subedi, B. R., & Howard, M. (2017). Multilevel predictors influencing reading achievement: Comparison of teacher effects in elementary, middle and high schools. *Advances in Social Sciences Research Journal*, 4(23), 98-106. DOI:10.14738/assrj.423.3944.

Subedi, B. R., & Powell, R. (2016). Factors influencing college readiness: A multilevel study to measure school effects. *International Journal of Learning, Teaching and Educational Research*, 15(11), 71-86. https://www.ijlter.org/index.php/ijlter/article/view/789/pdf_1

Subedi, B. R., Reese, N., & Powell, R. (2015). Measuring teacher effectiveness through hierarchical linear models: Exploring predictors of student achievement and truancy. *Journal of Education and Training Studies*, 3(2), 34-43. https://doi.org/10.11114/jets.v3i2.666

Thomas, S., Sammons. P, Mortimore, P., & Smees, R. (1997). Stability and consistency in secondary schools' effects on students GCSE outcomes over three years. *School Effectiveness and School Improvement*, 8(2), 169-197.

Tschannen-Moran, M. & Barr, M. (2004). Fostering student learning: The relationship of collective teacher efficacy and student achievement. *Leadership and Policy in Schools*, 3(3), 189-209. https://doi.org/10.1080/15700760490503706

Tschannen-Moran, M., & Hoy, A. W. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and Teacher Education, 23*(6), 944-956. http://dx.doi.org/10.1016/j.tate.2006.05.003

Willower, D. J., Eidell, T. L., & Hoy, W. K. (1967). *The school and pupil control ideology*. Penn State Studies Monograph No. 24. University Park, PA: Pennsylvania State University.

Wolters, C. A., & Daugherty, S. G. (2007). Goal structures and teachers' sense of efficacy: Their relation and association to teaching experience and academic level. *Journal of Educational Psychology, 99*(1), 181-193. http://dx.doi.org/10.1037/0022-0663.99.1.181

Yoon, J. S. (2002). Teacher characteristics as predictors of teacher-student relationships: Stress, negative affect, and self-efficacy. Social Behavior and Personality: An International Journal, 30(5), 485-493. https://doi.org/10.2224/sbp.2002.30.5.485

Zee, M., de Jong, P. F., & Koomen, H. M. Y. (2016). Teachers' self-efficacy in relation to individual students with a variety of social–emotional behaviors: A multilevel investigation. *Journal of Educational Psychology, 108*(7), 1013-1027. http://dx.doi.org/10.1037/edu0000106