education policy analysis archives

A peer-reviewed, independent, open access, multilingual journal

epaa aape

Arizona State University

Volume 18 Number 27 10h of November 2010 ISSN 1068–2341

Distortion or Clarification: Defining Highly Qualified Teachers and the Relationship between Certification and Achievement[1]

Jacob M. Marszalek
Arthur L. Odom
University of Missouri–Kansas City

Steven M. LaNasa
Donnelly College

Susan A. Adler
University of Missouri–Kansas City

Citation: Marszalek, J. M., Odom, A. L., LaNasa, S. M., & Adler, S. A. (2010) Distortion or clarification: Defining highly qualified teachers and the relationship between certification and achievement. *Education Policy Analysis Archives, 18*(27) Retrieved from http://epaa.asu.edu/ojs/837

Abstract: Recent studies of the relationship between teacher preparation pathways and student achievement have resulted in similar statistics but contradictory conclusions. These studies as a group have several limits: they sometimes focus on student-level indicators when many policy decisions are made with indicators at the school-level or above, are limited to specific urban locations or grade levels, or neglect the potential influence of building type, as defined as the grade-levels serviced. Using statewide data from the 2004–2005 school year, we examined the

[1] Accepted under the editorship of Sherman Dorn.

Manuscript received: December 3, 2008
Revisions received: December 9, 2009
Accepted: December 31, 2009

relationships between school-level indicators of student achievement on nationally-normed tests and proportions of alternatively certified teachers, while controlling for building type and other relevant covariates. Our findings indicate that the relationship between teacher preparation and student achievement at the school level depends on whether the building mixes multiple grade levels (e.g., elementary and middle). The implications of Missouri's policy change for research and school improvement are discussed with respect to the current high-stakes testing environment.
Keywords: alternative teacher certification; grade span configuration; educational legislation; politics of education; achievement tests; regression (statistics); robustness (statistics).

Distorsión o aclaración: Definiendo que son "Docentes de Alta Calidad" y la relación entre la titulación y logro académico

Resumen: Estudios recientes de la relación entre las vías de titulación de los docentes y el rendimiento de los estudiantes han tenido como resultado estadísticas similares, pero conclusiones contradictorias. Entendidos en su conjunto estos estudios tienen varios límites: es frecuente que se centren en los indicadores a nivel de los estudiantes, aún cuando muchas de las decisiones de política educativas se realizan usando indicadores a nivel de escuela o distrital, o se limitan a determinadas zonas urbanas o niveles de grado, o no toman en cuenta la influencia potencial del tipo de edificio, según se define los niveles de servicio. Usando datos estatales de Missouri del año escolar 2004-2005, se examinaron las relaciones entre los indicadores a nivel de escuela de el rendimiento de los estudiantes en exámenes (normados) a nivel nacional y la proporción de maestros con certificación alternativas mientras que se controló las variables de el tipo de edificios y otros covariables relevantes. Nuestros resultados indican que la relación entre la preparación de docentes y el logro académico de los estudiantes a nivel de la escuela depende de si el edificio es utilizado para niveles de grado múltiples (por ejemplo, primaria y secundaria). Las consecuencias del cambio de política de Missouri para la investigación y mejora de la escolaridad se discuten con respecto al contexto de la actuales pruebas de consecuencias severas.
Palabras-clave: certificación alternativa de docentes; grados de configuración; legislación educative; política educativa; pruebas de rendimiento; regresión (estadísticas); robustez (estadísticas).

Distorção ou Explicação: Definição de "Professores de alta Qualidade" e a relação entre as qualificações e logro acadêmico

Resumo: Estudos recentes sobre a relação entre as estratégias de preparação de professores e o desempenho dos alunos resultaram em estatísticas similares, mas conclusões contraditórias. Tomados coletivamente, esses estudos têm várias limitações: por vezes concentrando-se em indicadores sobre a produção dos alunos quando muitas das decisões de política educacional são realizadas com base em indicadores da escola ou do distrito; são limitadas a determinadas áreas urbanas ou níveis de ensino; ou não têm em conta a possível influência dos tipos de prédios escolares, tal como definido níveis de serviço. Usando dados oficiais do ano letivo de 2004-2005 foram discutidas relações entre o nível escolar dos indicadores de desempenho dos alunos nos exames nacionais padronizados e a proporção de professores com certificação alternativa, enquanto controlavam variáveis do tipo de edifícios e outras variáveis relevantes. Nossos resultados indicam que a relação entre a preparação dos professores e os resultados dos estudantes na escola depende se o prédio é utilizado para vários níveis de ensino (por exemplo, primário e secundário). As consequências da mudança na política de investigação e melhoria da

escola do Missouri são discutidas em relação ao contexto dos testes atuais para graves consequências.
Palabras-clave: certificação alternativa de professores; graus de configuração; legislação educacional; política educativa; testes de performance; regressão (Estatísticas); robustez (estatísticas).

Introduction

In 2005, the State of Missouri adopted a definition of *highly qualified teachers* mandated by the federal government. Three years before, President George W. Bush had signed the latest reauthorization of the Elementary and Secondary Education Act (ESEA) as the No Child Left Behind Act (NCLB), which defined a highly qualified teacher as having completed a teacher education program and earned a bachelor's degree, thereby obtaining full State certification; being placed in a position which matches his/her area of certification; and not having had certification or licensure requirements waived on an emergency, temporary, or provisional basis. In August 2005, to address teacher shortages, federal policymakers revised the definition of *highly qualified* to include teachers enrolled in alternative certification programs. Under the new definition, *highly qualified* is defined as a teacher who holds at least a bachelor's degree, has demonstrated subject-matter competency in the core academic subject(s) the teacher will be teaching, and is participating in an alternate-route-to-certification program. The definition continues by defining four components of an alternate route: The teacher receives, before and while teaching, high-quality professional development that is sustained, intensive, and classroom-focused to have a positive and lasting impact on classroom instruction; participates in a program of intensive supervision that consists of structured guidance and regular ongoing support for teachers, or in a teacher mentoring program; assumes functions as a teacher for a period not to exceed three years; and demonstrates satisfactory progress toward full certification as prescribed by the state.

This redefinition changed how Missouri considered teachers with a Temporary Authorization Certificate or Special Assignment Certificate (TAC/SAC), a one-year renewable certificate for individuals with a bachelor's degree who are employed by a school district and who complete coursework each year toward their teaching certificate. Under the new definition of a highly qualified teacher, TAC/SAC teachers—or any teacher with just a bachelor's degree—will be considered highly qualified for three years. The policy question is whether this redefinition will change student learning achievement. The answer to this question must address the full spectrum of consequences in this high stakes funding era, the substance of which includes not only student learning at the individual level but also at the school level.

A body of research now exists that supports the premise that good teachers matter to individual student learning (Darling-Hammond & Young, 2002; National Commission of Teaching and America's Future, 1996; Sanders & Horn, 1998) and points to the difference an effective teacher can make even in very challenging circumstances. Teachers are the key to what happens in classrooms (Thornton, 1991, 2005). They make the decisions about what actually gets taught and how it gets taught. They assess what students have learned and what individual needs particular students may have. To use Thornton's (1991, 2005) term, teachers are *curricular-instructional gatekeepers*.

But not every teacher will have a positive impact on student learning. The recent emphasis on improving the learning of all children has raised questions about the preparation teachers need to be effective in classrooms to assume greater importance. These questions have been the concern not only of teacher educators or school administrators but also of politicians and policymakers, as well.

Teacher education today finds itself in the glare of the public spotlight as educators and policymakers seek to determine what characterizes a highly qualified teacher, a teacher who will benefit student learning. Traditionally, state certification of teachers has provided the entry gate through which teachers would be certified as ready to undertake the job. Advocates of professionalism argue that there is a body of research on good teaching and on good practices in teacher education, and they argue that this research can guide the professional oversight of teacher education programs and improve the process of certifying or licensing good teachers (National Commission on Teaching and America's Future, 1996; Sanders & Horn, 1998).

However, others have challenged the need for teacher certification that is based on the satisfactory completion of a university preparation program. Those who take this alternative view argue that good teaching primarily requires strong content knowledge, with the rest learned through apprenticeship while on the job. The advocates of alternative entry pathways into teaching argue that traditional teacher certification programs are obstacles to attracting bright people with strong subject matter backgrounds into teaching (see, for example, Abell Foundation, 2001; Ballou & Soler, 1998; Paige, Stroup & Andrade, 2002).

Thus, a divide appears to separate those who see teaching as specialized work requiring specialized preparation from those who view teaching as something which most academically prepared people could do (Berry, Hoke, & Hirsh, 2004). The literature contains several studies that report on the relationship between certification status and teacher effectiveness in advancing individual student learning achievement, and there is support for both sides. In a review of research on teacher accountability, Wilson & Youngs (2005) examined eight studies reporting on this link. Seven of those found in favor of teacher certification; one did not (p. 611). Six of the studies looked at student achievement in mathematics, two looked at achievement in reading and literacy, and one included data on achievement in science. Goldhaber and Brewer's work (2000, 1997) suggested that teacher content knowledge as indicated by a BA in the field may be more significant than teacher certification in the field. In their 2000 study, students of emergency-certified teachers had the highest gain score. However, this included only 24 emergency-certified teachers out of a sample of 1,201. Furthermore, there was no distinction made among various reasons for emergency certification. In analyzing the same data, Darling-Hammond, Berry, and Thorenson (2001) found that many of the emergency-certified teachers had both teacher preparation (from another state, for example) and content preparation. Darling-Hammond and Youngs (2002) have argued that there is strong evidence to support the assertion that the preparation in pedagogy that preservice teachers receive in traditional certification programs makes a difference. Their review of the literature concluded that there was a significant relationship between certification and student performance at the level of the teacher, the school, the district, and the state (Darling-Hammond & Youngs, 2002).

In a study conducted by Laczko-Kerr and Berliner (2002), the academic performance of the students of regularly certified primary-grade teachers was compared to those who were identified as *under-certified*. In this study, under-certified included emergency-, temporary-, and provisionally-certified teachers, including those who participated in Teach for America. The students of certified teachers outperformed those of under-certified teachers, including those from Teach for America. The authors noted that effect sizes favored students of regularly-certified teachers (Laczko-Kerr & Berliner, 2002).

One criticism of the aforementioned studies is that they rely on aggregated data without accounting for the clustering of individual student scores by teacher or classroom. One recent study that accounted for nested data using multi-level modeling was reported by Boyd, Grossman, Lankford, Loeb, and Wykoff (2006). Although differences were found in student achievement score gains between college-recommended teachers and their alternatively-certified counterparts, the differences were very small for both math and English language arts (ELA) and for all levels of teaching experience. Differences in ELA were statistically significant but perhaps not practically significant; in one example, the difference in achievement scores amounted to about 2.5% of a standard deviation. Differences between the certification groups declined rapidly with years of experience and disappeared after three years. Several important school-level controls were used to prevent potential bias against alternatively certified teachers, who were more likely to be ". . . assigned to schools that have traditionally been difficult to staff" (Boyd et al., 2006, p. 190).

The results of this study suggests minimal importance of the traditional pathway to certification, although the authors admit that the accumulated effects of having several alternatively certified teachers may have a moderate impact on an individual student's achievement. Effect sizes for the differences between traditionally- and nontraditionally-certified teachers were fairly small, statistically significant differences may have been artifacts of the extremely large sample size (about 1 million), and differences disappeared after just a few years. However, one limitation of this study was that it may not have accounted for differences in building type as defined by which grades were serviced (e.g., K-6 as opposed to K-8). Building type may moderate the impact of teacher preparation on student learning, a concept discussed later in this article. Another limitation was that the study did not examine high school students. Furthermore, it was confined to New York City, which may have features significant different from school systems in other cities and states.

A case in point is a study reported by Darling-Hammond, Holtzmann, Gatlin, and Heilig (2005), in which individual student level achievement scores were compared between groups whose teachers had taken different paths to certification. Although the researchers also approached the analysis using multi-level modeling with the same or very similar controls, they came to different conclusions about the role of teacher pathway into the profession. Some effect sizes of teacher certification on individual achievement were of similar size to those reported by Boyd et al. (2006), such as differences of less than 0.10 *SD*. However, other effect sizes reported were larger, between 0.10 and 0.30 *SD*, indicating moderate differences in student achievement. One strength of this study was that it used three different standardized test batteries, including a Spanish-language battery (over 50% of the 35,000 students were Hispanic). However, some limitations were shared with the study of Boyd et al.: high school students were not included, building type was not accounted for, and the results were from a single urban district (Houston).

Both studies paid special attention to the differences between teachers prepared through Teach for America (TFA) and those prepared through college, and arrived at different conclusions. Boyd et al. (2006) concluded that since initial differences between the groups were small and disappeared after three years, there may be no practical advantage to limiting hiring to teachers prepared through the traditional pathway of college education programs. However, Darling-Hammond et al. (2005) concluded that the deficit observed for TFA teachers was large enough to warrant preferential hiring for traditionally prepared teachers. Interestingly, both studies noted the high attrition rate of TFA teachers after three years, and the fact that such turnover may exasperate any negative influence of alternatively certified teachers on student achievement, preventing fair evaluation of the effectiveness of alternatively certified teachers over time (also see Loeb, Darling-Hammond, & Luczak, 2005).

Like Boyd et al. (2006), Kane, Rockoff, and Staiger (2007) examined New York City data, and came to the same conclusions. One difference is that Kane et al. also compared teachers from elementary and middle schools, and while Boyd et al. found a differential effect of teacher certification on test scores between grades 4–5 and grades 6–8, Kane et al. found no difference. However, such a comparison does not take into consideration mixed-grade schools, in which elementary and middle school grades sit in the same school, or middle and high school grades are in the same school. Such combinations may create important differences in educational climate that can have an impact on student learning and teacher effectiveness (e.g., Bedard & Do, 2005; Eccles, Wigfield, Midgley, Reuman, MacIver, & Feldhaufer, 1993). Kane et al. make the same argument as Boyd et al., although more forcefully, that the statistically significant deficit found in scores for students with alternatively-certified teachers do not matter because they have small effect sizes. This is an acceptable conclusion from an administrative point of view, but from the point of view of an individual student (or parent/guardian), it is hard to argue that instruction of even slightly less quality than that given to peers would be acceptable in the current NCLB climate. Additionally, many of the alternatively certified teachers in the studies by Kane et al. and Boyd et al. had had excellent training that was very close to that received in traditional university programs, a threat to validity similar to that previously critiqued by Darling-Hammond et al. (2001). Our concern is not necessarily with well-prepared alternatively certified teachers but with underprepared ones.

Similar to Kane et al. (2007) and Boyd et al. (2006), Constantine, Player, Silva, Hallgren, Grider, and Deke (2009) conducted a study that attempted to account for nested effects but that was limited in scope, addressing in this case grades K-5. The methodology was rigorous, employing random assignment of students to either an alternatively certified teacher or a traditionally certified teacher and a national sample spanning seven states in most geographic regions (most students attended schools in Texas or California). Although no effects of certification pathway on student achievement were found for most groups (the exception being slightly lower math scores for students of alternative teachers in California), one threat to internal validity was a greater amount of mentoring received by alternatively certified teachers and a greater likelihood of mentoring in the second year of teaching than the traditionally certified teachers. Other limitations include: a lack of data on school type, no higher grade levels beyond fifth grade, and limited coverage of the heartland (represented by fewer than 12 schools).

Despite the contradictory evidence, under the previous administration the U.S. Department of Education (2005) argued that positive teacher impact is most closely related to verbal ability and content knowledge of the teacher and that traditional teacher education programs were barriers to attracting teacher candidates with these characteristics. Under current regulations, teachers may be considered highly qualified if they hold a bachelor's degree and demonstrate subject competency. Once hired, teachers who have not followed the traditional path toward teacher certification should receive *high quality* professional development and participate in intensive supervision and support (U.S. Department of Education, 2005). This latter recommendation makes perfect sense, but it leaves open the definition of *high quality*. One explanation for differences in results observed between studies limited to different localities (like New York and Houston) may be differences in definitions of *high quality*.

Another difficulty with this definition of highly qualified teachers is that it is based on the assumption that policy-decisions are based on individual student performance. Because of the current high stakes testing regime enforced by NCLB and many state laws, however, consequences for schools are based on school-level measures of student performance such as mean scores on standardized tests. Efforts to improve schools necessarily are evaluated on the success of changing student scores in aggregate. Although it has been shown that inferences about individual student scores cannot be validly made from school-level variables due to their nested nature (Snijders & Bosker, 1999), it is valid to draw inferences from school-level scores from school-level variables. Unfortunately, the previous studies using multi-level analysis do not model the relationships between school-level achievement scores and school-level variables, such as the percentage of teachers hired from a particular pathway. Darling-Hammond and Youngs (2002) found evidence for such relationships at levels beyond the individual student. Further investigation is needed to guide upper level policy on formulating guidelines for teacher qualifications.

Further investigation is also needed on the potential moderating effect of school atmosphere on the relationship between teacher preparation and student achievement. One area in which differences in school atmosphere manifest themselves is in *building type*, as defined by the grades services in a particular building (e.g., elementary schools generally service grades K-6, and junior high schools grades 7–8). According to Bedard and Do (2005), student outcomes such as learning achievement are a function of peer effect, teacher effectiveness, curriculum, school attributes, and student characteristics. Building types would therefore differ in peer effect and school attributes by virtue of having different mixes of students. For example, sixth-graders are the oldest students in an elementary school serving grades K-6, but they are the youngest students in a middle school serving grades 6–8. One possible mechanism for peer effect is curriculum, as explained by Cook, MacCoun, Muschkin, and Vigdor (2006), who found that discipline problems for North Carolina sixth-graders were greater in middle school than elementary school, and the difference persisted through ninth-grade. Since the curriculum is more fractured and specialized, with different teachers for different subjects, sixth-graders have much less supervision in middle school than in elementary school. This same reasoning was echoed by Bedard and Do (2005).

Teacher effectiveness may also differ, especially if it really does depend on teacher preparation, since 6.1% of middle school teachers are uncertified in contrast to 3.1% and 2.7% in elementary and high school, respectively (Bedard & Do, 2005). Possible mechanisms for the effect of teacher effectiveness were discussed by Eccles et al. (1993) and included differences in teacher discipline practices and teacher self-efficacy, both of which may be affected by the quality of teacher preparation. Eccles et al. surveyed 2500 students as they matriculated from 117 elementary sixth-grade teachers to 134 seventh-grade junior high school teachers in 12 districts in southeastern Michigan. They found that the junior high school teachers tended to approach discipline in a much less developmental way, and felt less efficacious. This has important implications for student performance, since previous research has demonstrated a negative association between transition between school type and academic performance (Berk, 1994).

Thus, several questions become salient given the current literature on the relationship between teacher preparation pathway and student achievement. The first is the influence of alternatively-certified teachers on school-level indicators of learning achievement, which are ultimately used for school assessment in high stakes accountability. The second is that if such an effect is extant, can it be generalized beyond a specific city to the state level, where many policy decisions are made? A third issue is whether any relationship exists between teacher preparation/qualification and school-level learning achievement at the high school level. The fourth question is how much of a moderating effect school type may have when examining the relationship between certification and achievement. And the ultimate question in all of this is whether it is better for educational stakeholders to define highly qualified teachers as they are currently or to return to the earlier more restrictive definition.

Design and Procedures

Data Source

An unusual historical circumstance allows us to collect data to answer these questions. Before 2006, statewide data on school-level characteristics were publicly available through the Missouri Department of Elementary and Secondary Education (DESE) website, including the proportions of teachers in each school with different types of certification. These data, unlike today, do not conflate teachers with alternative certificates with those who have regular certificates in the counts of highly qualified teachers.

The classification and certification standards of Missouri K-12 schools were established by the State Board of Education in 1950. Since then, there have been several revisions. The most recent revisions of standards occurred in April 2000 and December 2002. All districts were required to have an annual written Comprehensive School Improvement Plan (CSIP) designed to direct the overall improvement of educational programs and services. In part, the CSIP contained student performance data, attendance, school finance, and teacher certification report statistics. Each public school in Missouri submitted the core data defined by CSIP beginning in 2002. The data source for the current study was CSIP data reported 2002–2005. CSIP data were available as Excel spreadsheets at http://dese.mo.gov/schooldata/ftpdata.html. Student performance data were retrieved December 5, 2005, and all other data were retrieved September 27, 2006.

Building-Level Student Performance Data

Building- (or school-) level student performance was defined as the median percentile score on the nationally-normed TerraNova Survey. In 2005, the Missouri Assessment Program (MAP) used items from the the TerraNova as the multiple-choice section on its MAP test (Missouri Department of Elementary and Secondary Education [DESE], 2009). Scores on the multiple-choice items were combined with those from additional constructed-response items to form each student's MAP Scale score, which was then assigned one of five Achievement Levels (CTB/McGraw-Hill, 2007). Proportions of students in each Achievement Level were reported for each building, and the proportions were combined by formula to derive a building MAP Index score (CTB/McGraw-Hill, 2007). MAP Index scores were used to assess schools on AYP goals. However, TerraNova percentile score has greater generalizibility, especially outside of Missouri (as each state sets its own proficiency threshold). Also, in our sample median TerraNova percentile scores were highly correlated with the MAP Index and mean MAP Scale scores (bivariate correlation of at least .90 at each grade level). For these reasons, we decided to use Building median TerraNova percentile score as our outcome measure. We chose to examine communication arts and mathematics scores in particular because they were the only tests required in 2005, and far fewer schools reported scores in the other two subject areas (science and social studies).

Evidence for the validity and reliability of TerraNova scores has been well-established (CTB/McGraw-Hill, 2001). The items were developed using a three-parameter logistic item response theory model, and they have been continually evaluated on a national basis using stratified sampling methods (CTB/McGraw-Hill, 2001). Evidence for the reliability of the scores used on the MAP has also been reported in the form of high coefficients of Cronbach's alpha for all grades in both content areas ranging from .89 to .94 (CTB/McGraw-Hill, 2005). For comparison, the college entrance-exam SAT multiple-choice scores range in reliability from .89 to .93 (The College Board, 2009).

Teacher Certification Status

Each district CSIP included a report of the percentage of teachers classified by their certification status. Certification was defined in four different categories: *Regular Certificates* [Life Certificates, Professional Class I Certificates, Professional Class II Certificates, Continuous Professional Certificates, and Provisional Certificates (a provisional certificate was a two-year, non-renewable license for individuals who lacked a few requirements for full certification)]; *Temporary*

Authorization Certificates or Special Assignment Certificates (TAC/SAC) (one-year renewable certificates for individuals with a bachelor's degree who were employed by a school district and completed coursework each year toward a teaching certificate); no certificate (a teacher with a substitute, expired, or no certificate); and *highly qualified teachers* [defined by Sections 1119(a) and 9101(23) of the ESEA in 1965, as reauthorized by NCLB in 2002, which established requirements for the qualifications of teachers who teach a "core academic subject"]. Until August 2005, DESE considered highly qualified teachers those who had each of the following qualifications: obtained full State certification as a teacher, or passed the State teacher licensing examination and hold a license to teach in the State, and may not have had certification or licensure requirements waived on an emergency, temporary, or provisional basis; held a bachelor's degree; and demonstrated subject matter competency in each of the academic subjects in which the teacher teaches, in a manner determined by the State.

The sum of the percentage of teachers with regular certification, no certification, and TAC/SAC certification was 100% per building. (Regular certification and no certification were excluded from this analysis to reduce variance inflation caused by multicollinearity.) The remaining certification category constituted the independent variable, percentage of teachers with TAC/SAC certificates. A possible alternative explanation for the effect of the independent variable on median TerraNova score was the presence of mentoring, or at least other teachers who were highly qualified. Mentoring has been cited as a possible remedy for poor performance by teachers with alternative certificates (Roerhrig, Bohn, Turner, & Pressley, 2008), and as a limitation to the conclusions of Constantine et al. (2009), who found little in the way of negative effects of alternative certification. Therefore, an important control variable was an indicator of the presence of highly qualified teachers, the percentage of courses taught by highly qualified teachers (no other variable for highly qualified teachers was available in the pre-2006 data).

Building Demographics and Other Teacher Characteristics

In addition to student performance and teachers' certification status, we retrieved data on attendance rate (defined as daily average attendance, from http://dese.mo.gov/divimprove/sia/APR.html), percentage of students enrolled in the free and reduced price lunch (FRPL) program, and student-classroom teacher ratio (determined by FTE allocated to instructional class time). Other teacher characteristics included average years of experience and percentage of teachers with a master's degree. We felt that these variables represented viable alternative explanations for variance in median TerraNova score, in part because they had been controlled for in previous studies on the relationship between teacher certification and school achievement (for attendance, Boyd et al., 2006; Darling-Hammond et al., 2005; student-teacher ratio, Darling-Hammond et al., 2005; teacher experience, Boyd et al., 2006); Darling-Hammond et al., 2005; Kane et al., 2007; teacher degree, Darling-Hammond et al., 2005). These variables were the only forms of data on these constructs available in the 2005 data set, and we used them as control variables in all of a our analyses.

Table 1
Summary statistics and bivariate correlations by grade level for communication arts

Variable	Score	MA%	T exper.	S/T	Attend	Lunch	TAC/SAC	HQT	Build.	*M*	*SD*
Grade 3 (N = 1034)											
Median %ile score	—									61.31	11.54
MA %	0.22***	—								48.00	18.70
T experience	0.09**	0.32***	—							12.96	2.89
S/T ratio	0.07*	0.27***	0.01	—						17.48	2.74
Attendance	0.12***	0.06*	0.04	0.08**	—					94.78	2.97
Lunch %	-0.50***	-0.35***	0.01	-0.22***	-0.35***	—				50.94	24.36
TAC/SAC %	-0.15***	-0.20***	-0.16***	-0.05	-0.20***	0.18***	—			0.90	2.26
HQT %	0.12***	0.24***	0.14***	0.22***	0.17***	-0.22***	-0.17***	—		97.79	4.95
Building type	0.00	-0.22***	-0.14***	-0.16***	-0.05	0.08**	0.16***	-0.29***	—	0.07	0.25
Grade 7 (N = 588)											
Median score	—									61.58	11.65
MA %	0.15***	—								39.33	17.15
T experience	0.08*	0.33***	—							12.06	2.73
S-T ratio	-0.18***	0.18***	0.06	—						17.32	4.00
Attendance	0.50***	0.05	0.06	-0.27***	—					92.94	6.54
Lunch %	-0.59***	-0.30***	-0.07	0.07*	-0.48***	—				47.41	20.45
TAC/SAC %	-0.39***	-0.13**	-0.16***	0.16***	-0.53***	0.34***	—			2.55	4.67
HQT %	0.23***	0.31***	0.26***	0.15***	0.20***	-0.29***	-0.11**	—		94.41	6.83
Building type I	-0.03	-0.13**	-0.10**	-0.16***	0.08*	0.21***	-0.04	-0.11**	—	0.13	0.33
Building type II	0.09*	-0.29***	-0.01	-0.35***	0.13**	-0.07*	-0.08*	-0.15***	-0.26	0.33	0.47
Grade 11 (N = 491)											
Median score	—									59.89	9.57
MA %	0.09*	—								41.72	16.31
T experience	0.05	0.37***	—							12.64	2.52
S-T ratio	0.05	0.32***	0.10*	—						18.35	4.64
Attendance	0.50***	-0.13**	-0.05	-0.10*	—					92.74	4.67
Lunch %	-0.55***	-0.32***	-0.05	-0.22***	-0.31***	—				38.72	18.03
TAC/SAC %	-0.34***	-0.01	-0.09*	0.11**	-0.43***	0.23***	—			2.20	3.75
HQT %	0.19***	0.34***	0.22***	0.29***	0.13**	-0.29***	-0.07	—		94.92	5.09
Building type	-0.08*	-0.45***	-0.18***	-0.5***	0.18***	0.31***	-0.05	-0.31***	—	0.38	0.49

*$p < .05$; **$p < .01$; ***$p < .001$.

Table 2

Summary statistics and bivariate correlations by grade level for mathematics

Variable	Score	MA%	T exper.	S/T	Attend	Lunch	TAC/SAC	HQT	Build.	*M*	*SD*
Grade 4 (N = 1031)											
Median %ile score	—									63.36	13.21
MA %	0.23***	—								47.75	18.88
T experience	0.05	0.33***	—							12.92	2.93
S/T ratio	0.09**	0.28***	0.06*	—						17.52	2.90
Attendance	0.24***	0.07*	0.04	0.09**	—					94.78	2.97
Lunch %	-0.5***	-0.34***	0.02	-0.20***	-0.35***	—				51.15	24.39
TAC/SAC %	-0.15***	-0.18***	-0.15***	-0.04	-0.21***	0.18***	—			0.88	2.21
HQT %	0.13***	0.26***	0.17***	0.24***	0.17***	-0.23***	-0.16***	—		97.71	5.14
Building type	-0.08**	-0.23***	-0.16***	-0.16***	-0.05	0.08**	0.13***	-0.28***	—	0.07	0.26
Grade 8 (N = 589)											
Median score	—									63.57	14.00
MA %	0.20***	—								39.48	17.14
T experience	0.15***	0.34***	—							12.09	2.76
S-T ratio	-0.14***	0.16***	0.05	—						17.35	3.98
Attendance	0.43***	0.05	0.06	-0.26***	—					92.88	6.69
Lunch %	-0.65***	-0.30***	-0.08*	0.05	-0.45***	—				47.46	20.57
TAC/SAC %	-0.44***	-0.12**	-0.16***	0.15***	-0.50***	0.34***	—			2.57	4.77
HQT %	0.25***	0.32***	0.27***	0.16***	0.22***	-0.28***	-0.10**	—		94.40	6.86
Building type I	-0.10**	-0.15***	-0.13**	-0.17***	0.09*	0.22***	-0.04	-0.08*	—	0.13	0.34
Building type II	0.12**	-0.29***	0.00	-0.34***	0.11**	-0.07*	-0.10**	-0.16***	-0.27***	0.32	0.47
Grade 10 (N = 493)											
Median score	—									71.97	14.14
MA %	0.10*	—								41.73	16.31
T experience	0.07	0.38***	—							12.67	2.51
S-T ratio	0	0.31***	0.08*	—						18.31	4.66
Attendance	0.54***	-0.10*	-0.06	-0.14**	—					92.80	4.49
Lunch %	-0.54***	-0.32***	-0.05	-0.22***	-0.34***	—				38.73	17.98
TAC/SAC %	-0.41***	-0.02	-0.07	0.08*	-0.40***	0.25***	—			2.25	3.92
HQT %	0.16***	0.32***	0.22***	0.29***	0.14**	-0.28***	-0.05	—		94.92	5.09
Building type	-0.02	-0.44***	-0.16***	-0.51***	0.17***	0.31***	-0.03	-0.30***	—	0.39	0.49

*$p < .05$; **$p < .01$; ***$p < .001$.

Table 3
Summary statistics for percentage of teachers in a building with TAC/SAC certification by content area and grade level

Grade level	*N*	Min.	Mode	Modal freq.	Median	Mean	Max.	*SD*
Communication arts								
3	1034	0	0	.81	0	0.9	18	2.3
7	588	0	0	.69	0	2.6	31	4.7
11	491	0	0	.58	0	2.2	26	3.8
Mathematics								
4	1031	0	0	.82	0	0.9	18	2.2
8	358	0	0	.61	0	2.6	31	4.8
10	493	0	0	.58	0	2.2	27	3.9

Data Analysis and Results

Analytic Rationale

The median TerraNova percentile scores for both communication arts and mathematics had highly skewed distributions. In grades 3 (communication arts) and 4 (mathematics), the skewness statistics were -0.47 ($SE = 0.08$) and -0.28 ($SE = 0.08$) for communication arts and mathematics, respectively, and similar asymmetry was found in the data for grades 7, 8, 10, and 11. Critical ratios with absolute values greater than 3 are considered to be severely non-normal (Kline, 2005), and these were -5.88 and -3.50, respectively. Because no linear transformations were found to provide satisfactory approximations to normality, traditional General Linear Model techniques were eschewed. Instead, a robust regression technique was used, *M*-estimation, which has been shown to be robust for skewed data (Rousseeuw & Leroy, 1987).

Although other approaches to robust regression are available, our set of variables was most suitable for *M*-estimation, which is robust against departures from normality and highly influential outliers in the distribution of the dependent variable. Although some bounded influence estimation approaches such as least trimmed squares estimation are more robust to influential outliers in the independent variables, they are not as efficient as *M*-estimation and are less accurate when normal theory assumptions hold. Hybrid approaches can be as efficient, but they cannot be used with categorical independent variables (e.g., building type) or variables with very many equivalent values, because the algorithms involve iterative resampling (SAS Institute, 2009). For example, in grade 3, 81.4% of the schools had the same percentage of teachers with TAC/SAC certificates, zero (see Table 3 for distributional statistics for percentage of TAC/SAC certificates in each content area and grade level). Another advantage of *M*-estimation is that it allows the use of the full dataset in estimating the regression model by reducing the weight of outliers rather than deleting them, as in bounded influence estimation. We used the *M*-estimation method of PROC ROBUSTREG in SAS for each of our models with the default settings, which provided a 25% breakdown and 95% efficiency (SAS Insitute, 2009).

Besides skewness, independence of observations was another concern. Only data from 2005 was used because the dataset did not allow the assumption of independence of observations between years: school buildings have data on the variables of interest for every year. A second area of concern about independence was grade level. One reason that schools varied in the number of subject areas tested was that they varied in the grade levels taught, and different subject areas are tested at different grade levels. In the area of communication arts, students are tested in grades 3, 7, and 11, and in mathematics, grades 4, 8, and 10. Because some buildings teach both grades 3 and 7 (or 4 and 8), or both grades 7 and 11 (or 8 and 10), they may be represented multiple times in a data set including every grade level. Ultimately, it was decided to model each grade separately in communication arts and mathematics for 2005. This separation also served the purpose of examining the moderating effect of building type on the relationship between teacher certification and school-level achievement.

Communication Arts

Grade 3. The original sample contained 1034 school buildings in which there were third-grade students. An ordinary least squares (OLS) regression model was constructed with grade 3 median TerraNova score as the dependent variable. The independent variables were percentage of TAC/SAC teachers (TAC/SAC) and percentage of courses taught by highly qualified teachers (HQ). Covariates were percentage of teachers with master's degrees, average years of teacher experience, student/classroom teacher ratio, calendar year attendance rate, percentage of students qualified for free/reduced-price lunch (FRPL), and building type. Building type was a dummy variable set up to compare buildings with third-graders but not seventh-graders (coded 0) to buildings with both third- and seventh-graders (coded 1). Bivariate correlations and summary statistics are provided for each grade and subject area in Tables 1 and 2. Collinearity and linearity diagnostics were assessed for each model, and no problems were detected. The model accounted for 27% of the variance in median TerraNova score. A second model was also fitted, in which the interaction of TAC/SAC and building type was added. Although the interaction term was significant at the .01 level, the change in R^2 was small, .01 (see Table 4, which appears with Tables 5–9 at the end of this section).

M-estimation was used to fit a regression model to the data, and the results are reported in Table 4. Because the model uses a different type of estimation that depends on weighting the cases, statistics used in OLS regression such as □ and R^2 cannot be used, but analogous statistics have been developed for robust estimation (SAS, 2009). One such statistic is the robust R^2, which in this case was .23, meaning that 23% of the variance in median TerraNova was explained by the robust model. A robust linear test of the model analogous to F in OLS is □, rho, which is based on the chi-squared distribution. In this case, □ was significant (40.19, $df = 8$, $p < .001$). Because of the number of models planned for testing (communication arts and mathematics models for each grade level, or six total), and the large sample sizes, alpha was set at .01 to control experimentwise Type I error adequately.

Significant effects were found for average years of teacher experience, percentage of students qualified for lunch-program participation, and percentage of TAC/SAC. This pattern was similar to that of the OLS model except for TAC/SAC, which had had no detected statistically significant effect in the OLS model with interactions. In the robust model, TAC/SAC was negatively related to median TerraNova score, accounting for a decrease in building score of 0.47 points for every one percentage-point increase in TAC/SAC certificates with other variables controlled. We fit a second robust model that included the interaction term, and the robust R^2 increased by a small (.01) but statistically significant amount. Buildings that contained both third-and seventh-graders exhibited no evidence of a relationship between TAC/SAC and median TerraNova. However, buildings that contained third-graders but no seventh-graders showed evidence of a decline in median TerraNova with increased percentages of TAC/SAC.

Grade 7. We applied the same analytic procedure to our sample of 588 schools with grade 7 scores in communication arts. The main effects OLS model explained a very large amount of variance in median TerraNova, 43%, and OLS found significant effects for attendance, lunch-program participation, and TAC/SAC percentages with other variables controlled (see Table 5). Lunch-program participation was significant for each of the remaining OLS and robust models, as well, and was also the largest in terms of variance uniquely explained in the OLS models. Attendance and TAC/SAC were significant for most of the remaining models, too, but teaching experience was not significant again (except in the grade 8 robust regression). The presence or absence of the interaction coefficient did not alter these patterns, and it was not significant for the grade 7 communication arts model ($\Delta R^2 = .00$). The interaction term for grade 7 represented a moderating effect of Building Type I (for middle grades), a dichotomous indicator of whether a building had seventh-graders with third-graders (coded 1) or without third-graders (coded 0). Building Type II was a dichotomous indicator of whether a building had seventh-graders with eleventh-graders (coded 1) or without (coded 0). Neither term was a significant main effect, and neither significantly moderated the effect of TAC/SAC.

However, the results differed for the robust regression model. Robust R^2 was .30, but ΔR^2 after adding either interaction term was much less than .01, an amount that was not significant at the .01 level. However, the interaction term for Building Type I by TAC/SAC was significant at the .001 level. In buildings with both seventh- and third-graders, no association (with all other variables controlled) is evident between TAC/SAC and median TerraNova, but in buildings with seventh- but no third-graders, Median TerraNova declines with increased TAC/SAC percentages.

Grade 11. In our sample of 491 schools with eleventh-grade communication arts scores, the OLS model explained 43% of the variance, and revealed significant effects only for attendance and lunch-program participation, but not for TAC/SAC. In this sample, building type referred to buildings with both eleventh- and seventh-graders (coded 1) or buildings with eleventh- but not seventh-graders (coded 0). Its interaction with TAC/SAC was not found to be significant and added very little to the model ($\Delta R^2 = .00$; see Table 6). However, the robust model differed slightly by revealing a significant interaction effect, although it added little to the explanatory power of the main effects model (robust $\Delta R^2 = .01$, □ $= 4.47$, $df = 1$, $p < .01$). Buildings with both seventh- and eleventh-graders experienced a negative relationship between the percentages of teachers with TAC/SAC credentials and median TerraNova, but buildings with just eleventh-graders do not. This result is different from grades 3 and 7 in that the negative association was absent in mixed schools, while it is seen in mixed schools for grade 11 and absent in others.

Building atmosphere may account for the difference in TAC/SAC impact on median TerraNova scores between the building types. High school is arguably the place where subject specialization would have its greatest benefit, and so it may seem reasonable to overlook a lack of pedagogical training to hire a teacher with a great deal of subject area expertise. However, such a compromise might seem untenable for younger students such as seventh-graders, who would seem to require not only subject area expertise but also good pedagogy and more supervision (see for example Bedard & Do, 2005; Cook et al., 2006; Eccles et al., 1993). If a building contains traditional high school grades (grades 9–12) only, pedagogy and supervision and thereby student learning may be less than if the building also contained grade 7. The presence of seventh-graders may require more attention to pedagogy and supervision by all teachers, especially if seventh-graders are allowed to take higher grade-level courses.

Overall, communication arts scores were associated negatively with the percentage of alternatively certified teachers. To put the relationship into perspective, a hypothetical grade 3 example is offered. For a typical building with 100 teachers, one of whom was alternatively certified, the MAP Index score was about 186.0 and the TerraNova score was 60.9. For a similar building with five of its 100 teachers alternatively certified, the MAP Index score was about 183.7 and the TerraNova score was 58.5. Although these deficits seem small at first glance, they can mean the difference in meeting building AYP according to NCLB legislation.

Mathematics

Grade 4. The original sample contained 1031 buildings in which grade 4 mathematics scores were present. The proposed model for grade 4 was exactly like those posited for grades 3 and 11 in communication arts, except that building type indicated buildings with fourth-graders but no eighth-graders (coded 0) or schools with both fourth- and eighth-graders (coded 1). The OLS main effects model explained a large amount of variance, $R^2 = .25$, but lunch-program participation was the only significant coefficient (see Table 7). No significant addition to explanatory power was achieved by adding an interaction term. However, the robust model explained the same amount of variance and revealed a main effect for attendance. Although no additional variance was accounted for by adding an interaction term, the coefficient itself was significant at the .01 level. For mixed schools, no relationship exists between TAC/SAC percentage and median TerraNova with all other variables controlled. However, for schools that included fourth- but not eighth-grade, an increase in TAC/SAC percentage was associated with a decrease in median TerraNova. Grade 4 was the only grade in which evidence for an interaction effect was found at our *a priori* alpha level of .01.

Grade 8. The same procedure was used with the grade 8 building data (N = 589) to explain median TerraNova score. Like grade 7 communication arts, two building type indicators were included in the models. Building Type I indicated either schools with both fourth- and eighth-graders (coded 1) or just eighth-graders (coded 0), and Building Type II indicated either schools with both eighth- and tenth-graders (coded 1) or just eighth-graders and no tenth-graders (coded 0). As for grade 7 communication arts, two interaction terms representing moderating effects of Building Type I and Building Type II on TAC/SAC were tested separately, but both were found insignificant for grade 8 mathematics. For consistency across models, the Building Type I by TAC/SAC term is reported in Table 8 for both the OLS and robust approaches. Both OLS models, with and without the interaction term for Building Type I, explained a very large amount of variance ($R^2 = .50$) in median TerraNova and revealed significant effects for lunch-program participation and TAC/SAC percentage. The robust models explained less variance ($R^2 = .35$), though still a very large amount, but they revealed significant effects not only for lunch-program participation and TAC/SAC but also for attendance and teacher experience. The salient result in this part of the analysis is that TAC/SAC had a significant negative association with TerraNova score even after controlling for the other covariates and influential outliers.

Grade 10. The sample size for grade 10 mathematics scores was 493, with building type was coded as 0 for schools with tenth- but not eighth-graders, and 1 for schools with both. The OLS and robust models had similar results, with both explaining large amounts of variance in median TerraNova (47% and 32%, respectively), and both indicating significant effects for attendance, lunch-program participation, and TAC/SAC (see Table 9). Neither approach indicated main or moderating effects for building type. As with grade 8, the result of main interest was the significant negative association between TAC/SAC and median TerraNova.

Table 4
Parametric and robust linear regression models of grade 3 communication arts median TerraNova score (N = 1034)

	OLS Model 1			OLS Model 2			M-Est. Model 1		M-Est. Model 2	
Variable	*B*	*SE*	□	*B*	*SE*	□	*B*	*SE*	*B*	*SE*
(Constant)	94.01***	12.50		92.73***	12.46		50.81***	11.68	52.31***	11.65
M.A.	0.02	0.02	0.03	0.02	0.02	0.03	0.00	0.02	0.01	0.02
Experience	0.34**	0.12	0.08	0.32**	0.12	0.08	0.36**	0.11	0.35**	0.11
S/T ratio	-0.19	0.12	-0.04	-0.21	0.12	-0.05	-0.21	0.11	-0.24*	0.11
Attendance %	-0.27*	0.11	-0.07	-0.25*	0.11	-0.07	0.15	0.11	0.14	0.11
Lunch program %	-0.24***	0.01	-0.51	-0.24***	0.01	-0.51	-0.25***	0.01	-0.25***	0.01
TAC/SAC certificates	-0.31*	0.14	-0.06	-0.12	0.16	-0.02	-0.47***	0.14	-1.23***	0.30
HQT %	0.04	0.07	0.02	0.03	0.07	0.01	0.11	0.06	0.12	0.06
Building type[a]	2.69*	1.33	0.06	4.64**	1.50	0.10	-2.91*	1.25	-4.99***	1.42
Interaction: TAC/SAC x building				-0.99**	0.36	-0.10			0.95**	0.34
Linear test[b]	46.81***			42.74***			40.19***		36.78***	
R^2	.27			.27			.23[c]		.24[c]	
Standard error of estimate	9.92			9.89			8.65		8.55	
□ R^2				.01					.01[c]	
Test[b] of □ R^2				7.71**					5.63**	

[a]0: Building had no grade above sixth; 1: Building contained both third-grade and seventh-grade. [b]For the OLS model, *F* was used; for the robust model, □ was used. [c]Robust R^2.
*p < .05; **p < .01; ***p < .001.

Table 5

Parametric and robust linear regression models of grade 7 communication arts median TerraNova score (N = 588)

	OLS Model 1			OLS Model 2			M-Est. Model 1		M-Est. Model 2	
Variable	*B*	*SE*	□	*B*	*SE*	□	*B*	*SE*	*B*	*SE*
(Constant)	36.25***	9.25		35.86***			33.30***	8.75	34.51***	8.67
M.A.	0.00	0.03	0.00	0.00	9.28	0.00	0.00	0.02	0.00	0.02
Experience	0.04	0.15	0.01	0.04	0.03	0.01	0.24	0.14	0.23	0.14
S/T ratio	-0.23*	0.11	-0.08	-0.23*	0.15	-0.08	-0.15	0.10	-0.14	0.10
Attendance %	0.33***	0.08	0.18	0.34***	0.11	0.19	0.37***	0.07	0.40***	0.07
Lunch program %	-0.25***	0.02	-0.44	-0.25***	0.08	-0.44	-0.24***	0.02	-0.23***	0.02
TAC/SAC certificates	-0.30**	0.09	-0.12	-0.28**	0.02	-0.11	-0.29**	0.09	-1.18***	0.27
HQT %	0.11	0.06	0.07	0.11	0.10	0.06	0.08	0.06	0.04	0.06
Building Type I[a]	1.49	1.27	0.04	1.83	0.06	0.05	-1.37	1.19	-2.88*	1.31
Building Type II[b]	0.48	0.95	0.02	0.47	1.41	0.02	0.00	0.90	0.06	0.89
TAC/SAC X build. I				-0.17	0.95	-0.02	33.30***	8.75	0.93***	0.28
Linear test[c]	47.60***			42.82***			26.84***		24.70***	
R^2	.43			.43			.30[d]		.30[d]	
Standard error of estimate	8.90			8.91			8.03		7.93	
□ R^2				.00					.00[d]	
Test[b] of □ R^2				0.33					4.00*	

[a]0: Building had no grade below fourth; 1: Building contained both third-grade and seventh-grade. [b]0: Building had no grade above tenth; 1: Building contained both seventh-grade and eleventh-grade. [c]For the OLS model, *F* was used; for the robust model, □ was used. [d]Robust R^2.

**p* < .05; ***p* < .01; ****p* < .001.

Table 6
Parametric and robust linear regression models of grade 11 communication arts median TerraNova score (N = 491)

	OLS Model 1			OLS Model 2			M-Est. Model 1		M-Est. Model 2	
Variable	*B*	*SE*	□	*B*	*SE*	□	*B*	*SE*	*B*	*SE*
(Constant)	3.26	10.18		6.72	10.46		-3.09	9.15	0.53	9.28
M.A.	-0.02	0.03	-0.03	-0.02	0.03	-0.04	-0.01	0.02	-0.01	0.02
Experience	0.17	0.14	0.05	0.17	0.14	0.05	0.16	0.13	0.16	0.13
S/T ratio	-0.01	0.08	-0.01	0.00	0.08	0.00	0.12	0.08	0.15*	0.08
Attendance %	0.67***	0.09	0.32	0.63***	0.09	0.31	0.67***	0.08	0.63***	0.08
Lunch program %	-0.23***	0.02	-0.43	-0.22***	0.02	-0.42	-0.21***	0.02	-0.20***	0.02
TAC/SAC certificates	-0.23*	0.10	-0.09	-0.35*	0.13	-0.14	-0.22*	0.09	0.01	0.13
HQT %	0.03	0.07	0.02	0.03	0.07	0.02	0.06	0.07	0.06	0.07
Building type[a]	-0.35	0.87	-0.02	-0.90	0.95	-0.05	-0.02	0.78	0.75	0.85
Interaction: TAC/SAC x building				0.27	0.19	0.07			-0.41**	0.17
Linear test[b]	45.67***			40.90***			27.06***		24.36***	
R^2	.43			.43			.28[c]		.29[c]	
Standard error of estimate	7.28			7.27			5.73		5.82	
□ R^2				.00					.01[c]	
Test[b] of □ R^2				2.01					4.47	

[a]0: Building had no grade below eighth; 1: Building contained both seventh-grade and eleventh-grade. [b]For the OLS model, *F* was used; for the robust model, □ was used. [c]Robust R^2.
**p* < .05; ***p* < .01; ****p* < .001.

Table 7

Parametric and robust linear regression models of grade 4 mathematics median TerraNova score (N = 1031)

	OLS Model 1			OLS Model 2			M-Est. Model 1		M-Est. Model 2	
Variable	*B*	*SE*	□	*B*	*SE*	□	*B*	*SE*	*B*	*SE*
(Constant)	50.30***	14.24		49.48**	14.24		7.42	13.62	6.87	13.57
M.A.	0.04	0.02	0.06	0.05*	0.02	0.07	0.04	0.02	0.04	0.02
Experience	0.12	0.13	0.03	0.12**	0.13	0.03	0.31*	0.13	0.31	0.13
S/T ratio	-0.15	0.13	-0.03	-0.16	0.13	-0.03	-0.06	0.13	-0.05*	0.13
Attendance %	0.30*	0.13	0.07	0.31*	0.13	0.07	0.66***	0.13	0.68***	0.13
Lunch program %	-0.24***	0.02	-0.45	-0.24***	0.02	-0.45	-0.25***	0.02	-0.25***	0.02
TAC/SAC certificates	-0.25	0.17	-0.04	-0.12	0.19	-0.02	-0.27	0.16	-1.19	0.36
HQT %	-0.04	0.08	-0.01	-0.04	0.08	-0.02	0.01	0.07	0.01	0.07
Building type[a]	-1.27	1.45	-0.03	-0.18	1.61	0.00	0.14	1.40	-1.38	1.55
Interaction: TAC/SAC x building				-0.65	0.42	-0.05			1.06**	0.40
Linear test[b]	44.59***			39.96***			44.48***		40.53***	
R^2	.25			.25			.25[c]		.25[c]	
Standard error of estimate	11.42			11.41			10.03		9.91	
□ R^2				.00					.00	
Test[b] of □ R^2				2.44					2.98[c]	

[a]0: Building had no grade above sixth; 1: Building contained both third-grade and seventh-grade. [b]For the OLS model, *F* was used; for the robust model, □ was used. [c]Robust R^2.

**p* < .05; ***p* < .01; ****p* < .001.

Table 8

Parametric and robust linear regression models of grade 8 mathematics median TerraNova score (N = 589)

	OLS Model 1			OLS Model 2			M-Est. Model 1		M-Est. Model 2	
Variable	*B*	*SE*	□	*B*	*SE*	□	*B*	*SE*	*B*	*SE*
(Constant)	57.99***	9.81		57.45***	9.84		45.27***	9.67	45.82***	9.62
M.A.	0.00	0.03	0.01	0.01	0.03	0.01	0.01	0.03	0.01	0.03
Experience	0.28	0.17	0.06	0.28	0.17	0.05	0.48**	0.16	0.47**	0.16
S/T ratio	-0.21	0.12	-0.06	-0.22	0.12	-0.06	-0.16	0.12	-0.17	0.12
Attendance %	0.12	0.08	0.06	0.13	0.08	0.06	0.28***	0.08	0.29***	0.08
Lunch program %	-0.37***	0.03	-0.54	-0.36***	0.03	-0.54	-0.34***	0.02	-0.34***	0.02
TAC/SAC certificates	-0.57***	0.10	-0.19	-0.54***	0.11	-0.18	-0.56***	0.10	-1.17***	0.31
HQT %	0.14*	0.07	0.07	0.13	0.07	0.06	0.09	0.07	0.09	0.07
Building Type I[a]	0.93	1.42	0.02	1.47	1.57	0.04	-1.59	1.40	-2.67	1.53
Building Type II[b]	1.50	1.07	0.05	1.51	1.07	0.05	-1.09	1.05	-1.10	1.05
Interaction: TAC/SAC and building I				-0.27	0.33	-0.03			0.65*	0.32
Linear test[c]	63.87***			57.52***			35.21***		31.74***	
R^2	.50			.50			.35[d]		.35[d]	
Standard error of estimate	9.99			10.00			8.82		8.92	
□ R^2				.00					.00 [d]	
Test[b] of □ R^2				0.66					2.04	

[a]0: Building had no grade below fifth; 1: Building contained both fourth-grade and eighth-grade. [b]0: Building had no grade above ninth; 1: Building contained both eighth-grade and tenth-grade. [c]For the OLS model, *F* was used; for the robust model, □ was used. [d]Robust R^2.

**p* < .05; ***p* < .01; ****p* < .001.

Table 9

Parametric and robust linear regression models of grade 10 mathematics TerraNova score (N = 493)

	OLS Model 1			OLS Model 2			M-Est. Model 1		M-Est. Model 2	
Variable	*B*	*SE*	□	*B*	*SE*	□	*B*	*SE*	*B*	*SE*
(Constant)	-16.80	14.61		-14.17	15.11		-34.07*	13.39	-31.76*	13.72
M.A.	0.00	0.04	0.00	0.00	0.04	0.00	0.01	0.03	0.01	0.03
Experience	0.39	0.20	0.07	0.39	0.20	0.07	0.42*	0.19	0.42*	0.19
S/T ratio	-0.01	0.12	0.00	0.00	0.12	0.00	0.18	0.11	0.19	0.11
Attendance %	1.06***	0.13	0.34	1.04***	0.13	0.33	1.17***	0.12	1.15***	0.12
Lunch program %	-0.31***	0.03	-0.40	-0.31***	0.03	-0.39	-0.29***	0.03	-0.29***	0.03
TAC/SAC certificates	-0.64***	0.13	-0.18	-0.73***	0.19	-0.20	-0.61***	0.12	-0.47**	0.17
HQT %	-0.01	0.10	0.00	-0.02	0.10	-0.01	0.02	0.10	0.01	0.10
Building type[a]	1.33	1.23	0.05	0.97	1.34	0.03	-1.63	1.12	-1.18	1.22
Interaction: TAC/SAC x building				0.18	0.26	0.04			-0.23	0.24
Linear test[b]	54.31***			48.28***			27.90***		24.97***	
R^2	.47			.47			.32[c]		.32[c]	
Standard error of estimate	10.35			10.35			9.04		8.97	
□ R^2				.00					.00[c]	
Test[b] of □ R^2				0.48					0.63	

[a]0: Building had no grade below ninth; 1: Building contained both eighth-grade and tenth-grade. [b]For the OLS model, *F* was used; for the robust model, □ was used. [c]Robust R^2.

*$p < .05$; **$p < .01$; ***$p < .001$.

Discussion

Grade 3 median TerraNova scores in communication arts for each building were negatively associated with TAC/SAC with all other variables controlled. There was no difference in building type, whether the building also contained seventh-graders or not. However, a significant interaction effect revealed that the effect of TAC/SAC depended on school type, with a negative effect present in elementary schools but not mixed schools. One explanation for this may be that elementary teachers tend to emphasize rote repetitive drill more than higher level thinking (Berk, 1994), unlike more specialized content area teachers in upper grades. Greater emphasis on higher level cognitive learning tasks may be associated with engagement and achievement (Berk, 1994). Such a tendency can be mitigated through appropriate training, and it may be that teachers with TAC/SAC certification as a group are less likely to receive that training. Perhaps being in a mixed school influences elementary teachers to emphasize higher level learning tasks more than they otherwise would have because of more mentoring, collaboration, or sharing of ideas with upper grade-level colleagues.

In grade 7, the picture became more interesting because of the two building type indicators, but resu lts were largely consistent with grade 3. The moderating effect of Building Type I on TAC/SAC indicated that mixed schools again saw no association between TAC/SAC and median TerraNova communication arts score but junior high schools did. Again, we speculate that the intermingling of elementary and middle-school-level teachers accounts for this difference, with more opportunities available for mentoring, collaboration, and the sharing of ideas for pedagogy and classroom management. Another explanation that seems to fit the pattern comes from past research mentioned in the introduction that has shown a negative relationship between school transition and academic performance (Berk, 1994). For example, children tend to perform worse over their primary and secondary academic careers if they have had three major transitions (attending a K-6 school, a 7-9 school, and a 10–12 school) than if they have had two (attending a K-8 school and a 9–12 school; Berk, 1994). Such a transition effect may be masking a negative association between TAC/SAC and median TerraNova in buildings with seventh-grade but not third-grade, or third-grade but not seventh-grade. This explanation also fits the results from grade 4 mathematics scores, which showed another negative association between TAC/SAC and median TerraNova in elementary schools but not in mixed schools.

Grade 11 communication arts scores were found to be negatively associated with TAC/SAC, but that association also depended on building type. However, in grade 11, the association existed for mixed schools and not for high schools, different from grades 3, 4, and 7, in which the association existed for traditional rather than mixed schools. This may reflect a greater emphasis within the mixed schools on supervisory practices due to the presence of early adolescent students and less emphasis in traditional high schools where students are more autonomous. Greater autonomy would be friendlier to the use of student-centered pedagogy, which in turn is more amenable to higher level learning (Lawson, 1995).

Overall, many of the same covariates tended to be significantly related in the same directions to median TerraNova. The covariate chosen to indicate economic level was lunch-program participation, and this factor was consistently the largest contributor of unique explanatory power. Percentage of teachers with TAC/SAC certificates was found to be related to building median TerraNova score at each grade level tested in both communication arts and mathematics. What seems clear is that even with important extraneous variables controlled—such as student-teacher ratio, economic status, teacher experience and education, attendance rate, and building type—the quality of teachers' pedagogical training matters. Admittedly, the effect sizes we found are small (as represented by changes in R^2 for the interactions), but even the smallest difference in scores matters in high stakes testing.[2]

One argument against results based on aggregate data such as these is that they do not reflect a valid measurement of individual student learning. However, the focus of this study is at the building level, where state policy and law are focused and enforced. The consequences of high stakes testing are given out at the building level and evaluated with building level indicators. The current decision of the state of Missouri for all TAC/SAC teachers to be considered highly qualified may be associated with building level achievement score averages, and even a tenth of a point loss may result in the forfeiture of thousands of public dollars.

Regarding our original research questions, several answers have begun to emerge from our results. School-level indicators of learning achievement that are ultimately used for school assessment in high stakes funding decisions do seem to be affected by the level of presence of alternatively-certified teachers, and this is shown for data collected at the state level, a much broader geographic area than a city, where many policy decisions are made. Our data also showed that the relationship exists at the high school level as well as the primary and middle school levels, although it varies depending on whether or not the school is mixed. As stated before, the ultimate question in all of this is whether it is better for educational stakeholders to define highly qualified teachers as they are currently or to return to the earlier more restrictive definition.

Darling-Hammond (2006) noted that one of the most difficult and important questions in education is the relationship between what teachers have learned and student achievement. Our data support the proposition that teachers who first complete teacher education programs and who are placed in teaching positions that correspond to their certification areas have a strong positive influence on student achievement. The data also suggest that teachers who have only a content degree (e.g., bachelor's in History, Biology, Liberal Arts, etc.) and work in the classroom without first gaining full certification may have a negative impact on student achievement. The results do not indicate that content knowledge is unimportant but instead emphasize that content knowledge alone is insufficient in preparing successful teachers. Our data support the increased use of personnel who complete preparation programs, are hired as fully certified teachers, and assigned to their specialty areas.

What distinguishes highly qualified teachers from teachers who begin teaching with preparation in their content fields but little or no teacher preparation coursework? Highly qualified teachers have essential knowledge and skills unavailable to content-only specialists. For example, educational researchers have established through decades of research the importance of student-centered teaching practices, selection and adaption of curriculum to meet student needs and interests, and a focus on students' conceptual understanding and critical thinking rather than knowledge acquisition. Teachers who have participated in approved teacher education coursework understand how students learn and how to facilitate learning. Their work in teacher preparation helps them to make effective decisions about learning outcomes, instructional strategies, assessment, and curriculum. Teachers whose preparation is exclusively in their content fields and general education have only limited knowledge of necessary techniques for successful instruction. They are put into the position of making decisions about instruction, curriculum, and assessment with little background or understanding of the complexity of these issues.

Content-only preparation consigns the framework of effective teaching practices to being defined by personal experiences obtained while earning a degree, or through one's own K-12

---

[2] Sizes of the unique effects of TAC/SAC tended to be larger than those reported by Boyd et al. (2006) and Kane et al. (2007).

experience. Teachers with content-only preparation have only limited knowledge necessary for successful instruction. They tend to teach the way they were taught and to violate some of the most basic instructional principles for effective K-12 learning environments. Teachers who mimic college instruction tend to perpetuate many of the teacher-centered strategies educational researchers know to be ineffectual in facilitating learning and which are inappropriate for the developmental level and desired learning environments described by teacher educators. Teacher-centered instructors inadvertently promote the misconception that good education revolves around knowledge acquisition and the teacher as the source of knowledge (Yager & Penick, 1987).

Implications and Recommendations

One of the most serious implications of this study involves the recent change in the definition of highly qualified teachers by the federal government and adopted by Missouri in 2005. In August 2005, policymakers revised the definition of highly qualified to include teachers enrolled in alternative certification programs to address teacher shortages. Now, highly qualified teachers and TAC/SAC teachers as previously defined are in the same category. As previously indicated, TAC/SAC was defined as a one-year renewable certificate for individuals with a bachelor's degree who were employed by a school district and who completed coursework each year toward their teaching certificate. Under the new definition, teachers with just a bachelor's degree will be defined as highly qualified for three years.

This new definition may have a severe negative impact on schools that have the highest need for genuinely highly qualified teachers. For example, schools with the greatest need for teachers are in high poverty areas, thus increasing the likelihood of hiring teachers with content-only degrees. The overarching result is that schools with the greatest needs will be able to hire in the short term teachers prepared only in content and classify them as highly qualified. While this may address immediate needs, long-term problems will be caused by a spike in the hiring of content-degree-only teachers, a practice that may result in a steady decline in student achievement. From a research and policy perspective, this will further blur the relationship between teacher education and student achievement, making it even more difficult to determine what works.

Two other potential consequences are in the professional environment and in school funding. As pointed out by Kane et al. (2007) and Boyd et al. (2006), teachers taking nontraditional pathways for preparation have much higher rates of turnover during the first three years of employment. If the current definition of highly qualified tempts administrators to apply the bandage of hiring of more of these teachers in schools needing the most help, they may find that the bandage has little sticking power. Although our results suggest that higher proportions of highly qualified teachers under the old definition may help mitigate any negative impact of TAC/SAC teachers, the proposed mechanisms of professional support and mentoring would only work for those teachers that are retained. A constant turnover of TAC/SAC teachers would keep the overall level of teacher preparedness significantly lower within the school (Loeb et al., 2005). We intend to use the 2002–2005 data in a longitudinal analysis to see if this has indeed occurred in the past.

Educational researchers in other states should also seek school data based on the 2002–2005 NCLB definitions of certification and explore their relationships with school-level learning achievement. The results of this study should be replicated to increase the generalizability of our findings as well as to garner the attention of other state policymakers. Because policymakers may hold misconceptions about teacher qualifications and teacher education, continued decision-making based on these largely political perspectives will have serious long term negative consequences for public education in Missouri and the U.S. more generally.

Our data demonstrate that if teacher preparedness is lower, school-level indicators of learning achievement will be lower. Therefore, we strongly recommend that school districts carefully review hiring policies. We also recommend that state departments of education and the federal government revisit teacher certification requirements and definitions. We are unfamiliar with any evidence that finds that teachers who begin teaching with little or no professional preparation are deserving of highly qualified status.

By far the most startling implication will be the way in which this redefinition may affect the student achievement gap discourse. These results show that as previously defined, TAC/SAC certification teachers are not associated with as much school-level student achievement as schools with no TAC/SAC teachers,. The redefinition will serve to mask these teachers as highly qualified when the state's own data show this not to be the case. Most disturbing to future discourse will be the eventual question we can see being raised in various sectors: "Why aren't these children learning if we have highly qualified teachers in their classrooms?" This redefinition is itself symptomatic of a general trend in post-NCLB life, one fraught with gerrymandering, manipulation, and standards-compromising to achieve a seemingly worthy goal, although more frequently an arbitrary one. In the end, we view this decision as one of just many steps taken by those in the policy arena that further shortchanges children who are apparently viewed as those with the least to lose.

Teachers need to know their subject matter and how to teach it. Despite federal guidelines that define highly qualified by subject matter knowledge alone, the evidence suggests that knowledge of pedagogy, of learners, and of the learning process together with a firm content knowledge base is necessary if teachers are to have a positive impact on the achievement of students. The challenge of improving the learning of all teachers is not solved by defining highly qualified teachers as those with strong content knowledge and verbal skills. We do not quarrel that these are important characteristics, but content knowledge and verbal skills alone will not prepare a teacher for the challenges of a classroom. With little preparation for the challenges of teaching, such teachers report more problems and lower self-confidence and sense of efficacy than those who have participated in more traditional preparation programs (Darling-Hammond, Chung, & Frelow, 2002).

There is little evidence that subject matter knowledge trumps all other knowledge or that alternative routes to teaching attract high quality teachers who have a positive impact on student achievement. Yet these assumptions serve as the foundation for policy decisions about admission to the occupation/profession of teaching. Rather than seeking the low-cost solution, if as a nation we really want to assure that no child is left behind, then we need to make different policy decisions. We need the funding to attract and support teacher candidates who are passionate and committed to teaching. Prospective teachers should not be discouraged from receiving strong teacher preparation before being hired simply because they cannot afford to be in school and out of the work force. The support of strong mentoring programs and professional development will enable teachers to continue to grow as professionals. Our evidence suggests that we are sacrificing cohorts of learners while new, underprepared teachers learn how to do the job.

References

The Abell Foundation. (2001, November). *Teacher certification reconsidered: Stumbling for quality*. Baltimore: Author. Retrieved from http://www.abell.org/publications/detail.asp?ID=59

Ballou, D., & Soler, S. (1998, February). *Addressing the looming teacher crunch* (Policy Briefing). Washington, DC: Progressive Policy Institute.

Berry, B., Hoke, M., & Hirsch, E. (2004). The search for highly qualified teachers. *Phi Delta Kappan, 85*, 684–689.

Bedard, K., & Do, C. (2005). Are middle schools more effective? The impact of school structure on student outcomes. *The Journal of Human Resources, 40*(3), 660–682.

Berk, L. (1994). *Child development* (3rd ed.). Boston: Allyn and Bacon.

Berry, B., Hoke, M., & Hirsch, E. (2004). The search for highly qualified teachers. *Phi Delta Kappan, 85*(9), 684–689.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy, 1*(2), 176–216.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

The College Board. (2009). *Test characteristics of the SAT: Reliability, difficulty levels, completion rates.* Retrieved from http://professionals.collegeboard.com/profdownload/2009-Test-Characteristics-of-the-SAT.pdf

Cook, P. J., MacCoun, R., Muschkin, C., & Vigdor, J. (2006, July). *Should sixth grade be in elementary or middle school? An analysis of grade configuration and student behavior* (Working Papers Series SAN06-03). Durham, NC: Duke University Terry Sanford Institute of Public Policy.

Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., & Deke, J. (2009). *An evaluation of teachers trained through different routes to certification: Final report* (Research Report No. NCEE 2009-4043). Washington, DC: Institute of Education Sciences. Retrieved from http://www.mathematica-mpr.com/publications/pdfs/education/teacherstrained09.pdf

CTB/McGraw-Hill. (2005). *Missouri assessment program technical report 2005 supplement.* Jefferson City, MO: Missouri Department of Elementary and Secondary Education. Retrieved from http://dese.mo.gov/divimprove/assess/tech/2005%20TechRpt.pdf

CTB/McGraw-Hill. (2007). *Missouri assessment program: Guide to interpreting results. Communication arts and mathematics: Revised 2007.* Jefferson City, MO: Missouri Department of Elementary and Secondary Education. Retrieved from http://dese.mo.gov/divimprove/assess/2007_gir_manual.pdf

Darling-Hammond, L. (2006). Assessing teacher education: The usefulness of multiple measures for assessing program outcomes. *Journal of Teacher Education, 57*(2), 120–138.

Darling-Hammond, L., Berry, B., & Thoreson, A. (2006). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis, 23*(1), 57–77.

Darling-Hammond, L., Chung, R., & Frelow, F. (2002). Variation in teacher preparation: How well do different pathways prepare teachers to teach? *Journal of Teacher Education, 53*(4), 286–302.

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). Retrieved from http://epaa.asu.edu/epaa/v13n42/

Darling-Hammond, L., & Youngs, P. (2002). Defining "highly qualified teachers:" What does "scientifically-based research" actually tell us? *Educational Researcher, 31*(9), 13–25.

Eccles, J. S., Wigfield, A., Midgley, C., Reuman, D., MacIver, D., & Feldhaufer, H. (1993). Negative effects of traditional middle schools on students' motivation. *The Elementary School Journal, 93*(5), 553–574.

Goldhaber, D., & Brewer, D. (1997). Evaluating the effect of teacher degree level on educational performance. In W. Fowler (Ed.), *Developments in school finance, 1996* (pp. 197–210). Washington, DC: National Center for Education Statistics.

Goldhaber, D., & Brewer, D. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis, 22*, 129–145.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2007). Photo finish: Certification doesn't guarantee a winner. *Education Next, 7*(1), 60–67.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford.

Laczko-Kerr, I., & Berliner, D. C. (2002). The effectiveness of "Teach for America" and other under-certified teachers on student achievement: A case of harmful public policy. *Education Policy Analysis Archives, 10*(37). Retrieved from http://epaa.asu.edu/epaa/v10n37

Lawson, A. E. (1995) *Science teaching and the development of thinking*. Belmont, CA: Wadsworth.

Loeb, S., Darling-Hammond, L., & Luczak, J. (2005). How teaching conditions predict teacher turnover in California schools. *Peabody Journal of Education, 80*(3), 44–70.

Missouri Department of Elementary and Secondary Education. (2009). *Missouri school improvement program: Understanding your annual performance report (APR) 2009–2010: 2009* 4th *cycle APR* (DESE 3341-19 9/03). Jefferson City, MO: Author. Retrieved from http://www.dese.mo.gov/divimprove/sia/dar/understandingyourAPR.pdf

National Commission on Teaching and America's Future (1996). *What matters most: Teaching for America's future*. New York: Report of the National Commission on Teaching & America's Future.

Paige, R., Stroup, S., & Andrade, J. R. (2002).*Meeting the Highly Qualified Teacher Challenge: The secretary's Annual Report on Teacher Quality.* Washington, DC: U.S. Department of Education Office of Postsecondary Education. Retrieved from http://www.ed.gov/offices/OPE/News/teacherprep/AnnualReport.pdf

Roehrig, A. D., Bohn, C. M., Turner, J. E., & Pressley, M. (2008). Mentoring beginning primary teachers for exemplary teaching practices. *Teaching and Teacher Education, 24,* 684–702.

Rousseeuw, P. J., & Leroy, A. M. (1987). *Robust regression and outlier detection*. NewYork: John Wiley and Sons.

SAS Institute (2009). *SAS/STAT User's Guide, Second Edition: The ROBUSTREG Procedure*. Cary, NC: Author. Retrieved from http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/rreg_toc.htm

Sanders, W., & Horn, S. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*(3), 247–256.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* London: Sage Publishers.

Thornton, S. J. (1991). Teachers as curricular-instructional gatekeepers in social studies. In J. Shaver (Ed.), *Handbook of research on teaching and learning in social studies* (pp. 237–248). New York: Macmillan.

Thornton, S. J. (2005). *Teaching social studies that matters: Curriculum for active learning.* New York: Teachers College Press.

U. S. Department of Education, Office of Elementary and Secondary Education. (August, 2005). *Improving teacher quality state grants, ESEA Title II, PartA, Non-regulatory guidance*. Washington, DC: Author.

Wilson, S. M., & Youngs, P. (2005). Research on accountability processes in teacher education. In M. Cochran-Smith & K. M. Zeichner (Eds.), *Studying teacher education: The report of the AERA panel on research and teacher education* (pp 591–643). Washington, DC: American Educational Research Association.

About the Authors

Jacob M. Marszalek
University of Missouri–Kansas City

Arthur L. Odom
University of Missouri–Kansas City

Steve LaNasa
Donnelly College

Susan Adler
University of Missouri–Kansas City

Email: marszalekj@umkc.edu

Jacob M. Marszalek is an assistant professor of educational research and psychology at the University of Missouri-Kansas City. His research interests include educational and psychological measurement, motivation, and program evaluation.

Arthur L. Odom is a professor of education at the University of Missouri-Kansas City. His research interests include teacher professional development, student misconceptions, and inquiry learning models.

Steven M. Lanasa is president of Donnelly College in Kansas City, Kansas. His research interests include college-going decisions, student persistence, and faculty roles.

Susan A. Adler is a professor emeritus of education and former director of teacher education at the University of Missouri-Kansas City. She is currently a visiting professor at the National Institute of Education in Singapore.

education policy analysis archives

Volume 18 Number 27 10h of November 2010 ISSN 1068–2341

Please send errata notes to Gustavo E. Fischman fischman@asu.edu