Principals' Perceptions of the Causes of Teacher Ineffectiveness in Different Secondary Subjects

By Bruce Torff & David Sessions

The performance of the teachers in our nation's schools has long been a concern among educators, parents, and policymakers, but recent educational reform initiatives have put issues of teacher quality in the spotlight. For example, the *No Child Left Behind* Act of 2001 requires states to ensure that all classrooms be staffed by a "highly qualified" teacher. Alternative teacher-certification programs have proliferated, with the goal of increasing the supply of teachers and enhancing their quality. It has also been suggested that teacher certification be deregulated so that college graduates who lack course work in education could qualify for a teaching certificate (Ballou & Podgursky, 1998; Finn, 1999; Finn & Madigan, 2001; Gross, 1999; Hess & Finn, 2004; Hirsch, 1996; Kanstoroom & Finn, 1999; Kramer, 1991; Podgursky, 2005; Ravitch, 2000; Sykes, 1995). The question arises as to the efficacy of these initiatives for enhancing teacher quality.

Bruce Torff is an associate professor in the Department of Curriculum and Teaching and David Sessions is an associate professor in the Department of Information Technology and Quantitative Methods, both at Hofstra University, Hempstead, New York.

Such a determination requires that the fundamentals of teaching competence be identified, since a shortfall in any of these fundamentals constitutes a potential cause of teacher ineffectiveness and thus a threat to teacher quality. The fundamentals of teaching competence can be categorized as encompassing *content knowledge* (expertise in the subject being

taught, also known as "subject-matter knowledge") and *pedagogical knowledge* (expertise in teaching strategies and tactics, typically taught in teacher-education courses) (Shulman, 1987).

Arguments as to which of these fundamentals comprises the greater threat to teacher quality are abundant in the literatures in psychology and education, among other places (Ball, 1990; Darling-Hammond, 2005; Darling-Hammond & Bransford, 2005; Finn, 1999; Finn & Madigan, 2001; Gross, 1999; Hess & Finn, 2004; Hirsch, 1996; Kanstoroom & Finn, 1999; Kramer, 1991; Ma, 1999; Null & Ravitch, 2006; Podgursky, 2005; Ravitch, 2000; Sykes, 1995). But since determining whether content knowledge or pedagogical knowledge more often causes teacher ineffectiveness is a question best resolved empirically, teacher performance must be evaluated. This evaluation has proven contentious, and it seems clear that foolproof means to assess teacher performance remain elusive (Brandt, 1996; Copland, 2001; Darling-Hammond, 1986; Ellet & Teddie, 2003; Frase & Streshly, 1994; Goldstein, 2004; Haefle, 1993; Peel & Inkson, 1993; Sullivan, Mousley, & Gervasoni, 2000; Wayne & Youngs, 2003; Wise, et al, 1984). As described below, teacher performance can be assessed in at least two ways: using measures of student achievement, typically test scores; and using evaluations made by principals, the administrators who supervise teachers in schools.

Assessing Teacher Performance through Student Achievement

Some commentators recommend that teacher performance be evaluated using student-achievement measures (typically standardized test scores) (e.g., Finn, 1999). Critics of this approach (e.g., Darling-Hammond, 2005) have suggested that test scores may be misleading since they are dependent upon factors other than teacher performance (e.g., students' prior experience, students' family and community environment, and teachers' class assignments—the proportion of higher- and lower-achieving classes to which a particular teacher is assigned). Accordingly, in a study with two cohorts, only 8% and 2% (respectively) of the variance in students' test scores was statistically attributable to teacher-related variables (Hill, Rowan, & Ball, 2005). Moreover, test scores reflect student performance on the single day the test is given, which may not accurately reflect students' level of competence.

At the same time, given the widespread emphasis on testing in modern schools, it seems important to examine the research on the influence of teachers' content knowledge and pedagogical knowledge on student achievement. One strand of this work suggests that teachers' content knowledge is positively associated with student achievement (Ballou & Podgursky, 1999; Ma, 1999; Podgursky, 2005). For example, Hill, Rowan, and Ball (2005) studied 2963 first- and third-grade students and 699 teachers in 15 states, with data collected primarily in urban school districts engaged in instructional-improvement initiatives. Results indicated that "teachers' knowledge of mathematics for teaching" (i.e., knowledge of math as it applies specifically to teaching, as opposed to applications in science or engineer-

ing) evinced a positive association with students' test scores in math. The variable "mathematical knowledge for teaching" could be interpreted as encompassing pedagogical knowledge as well as content knowledge, insofar as applying mathematical knowledge to teaching involves pedagogical skills such as lesson planning and lesson implementation. However, the researchers' conclusion that their results call for "policy initiatives to improve students' mathematical achievement by improving teachers' mathematical knowledge" leaves little doubt that they see content knowledge as the pressing teacher-quality issue in math. At the same time, the researchers did not assess pedagogical knowledge, so comparison of the relative contribution of content knowledge and pedagogical knowledge to the variation in student achievement was not attempted.

A different strand of research underscores the impact of teachers' pedagogical knowledge on students' academic achievement (for a review see Darling-Hammond et al., 2005). Laczko-Kerr and Berliner (2002) compared recently hired undercertified and certified teachers (N=293) from five low-income school districts who were matched on a number of variables, resulting in 109 pairs of teachers whose students completed a variety of tests. On tests of reading, math, and language arts, students of certified teachers (i.e., ones who had completed extensive pedagogical coursework) outperformed students of teachers who were undercertified. Obtaining similar findings, Goldhaber (2000) investigated how teachers' certification status affected the performance of 6210 secondary students' performance on a math examination. Results indicated that students of teachers with public school certification outperformed the students of teachers who either held private-school certification or were uncertified in math. Monk (1994) found that teacher-education coursework had a significant positive association with student achievement in math and science, and Isbell (2002) found that students with noncertified teachers scored below students with fully certified teachers in tests of reading in a large urban school district in California. Whereas these studies indicate that pedagogical knowledge is positively associated with student outcomes, they do not address the question of whether deficiencies in content or pedagogy are more often the culprit when teachers' work is ineffective.

The available test-score research suggests that teachers' content knowledge and pedagogical knowledge both appear to be positively associated with student outcomes, but which has the greater effect remains in dispute. Moreover, these studies rest on the assumption that student-achievement data provide a nonproblematic method for evaluating teacher performance, and this assumption is not without its critics, as noted (Darling-Hammond, 1986; Darling-Hammond & Snyder, 2000; Flowers & Hancock, 2003; Kelly, 2004; Kupermintz, 2003; Matsumura, 2002; Nagy & Moorhead, 1990; Neill, 1999; Peterson, 2000). Even if the student-achievement studies are assumed to be valid, they do not directly address the issue of teacher failure. With content and pedagogy each contributing to the variance in student outcomes, it is possible that teachers sometimes fail due to deficiencies in content

knowledge, and other times due to shortfalls in pedagogical knowledge, and still other times to both. The student-achievement data leave it unclear whether content or pedagogy more often causes teachers' work to be ineffective.

Assessing Teacher Performance through Principals' Evaluations

An alternative strategy for identifying the causes of teacher failure is to consult with principals, the administrators charged to evaluate teachers in schools. Principals' evaluations have been criticized as inherently subjective (Ambady & Rosenthal, 1993; Blumberg, 1980; Fant & Stevens, 1991; Frase & Streshly, 1994; Kerrins & Cushing, 2000; Machell, 1995; Stodolsky, 1984), open to bias based on gender and physical attractiveness (Ambady & Rosenthal, 1993; Lee, Smith, & Cioci 1993; Rinehart & Young, 1996), and limited by the fact that principals typically lack content knowledge in secondary subjects other than the ones in which they are certified.

Other educators have argued, however, that principals are suitable candidates to consult on issues of teacher quality (Beerens, 2000; Frase & Hetzel, 1990; Peterson, 2000). To begin with, principals review the test scores of students in each teacher's classroom (Nagy & Moorhead, 1990). Principals observe teachers' classroom performance, gaining first-hand knowledge of their work—including preparation of lesson plans and materials, lesson implementation, and ability to establish rapport with students and handle classroom-management challenges (Darling-Hammond & Snyder, 2000; Denner, Salzman, & Bangert, 2001; Ferrell, 1992). Principals receive feedback about teacher performance from students, parents, department heads, and administrators. Finally, principals typically possess the educational background required to assess teacher performance; they are usually former teachers with personal experience in classroom teaching, and techniques for supervision and evaluation of teachers are part of administrative training and licensure.

Imperfect as principals' evaluations may be, they provide school-level managers' perspective on the causes of teacher ineffectiveness. With commentators concerned about content-knowledge deficiencies suggesting that "school-level managers are in the best position to know who teaches well and who teaches badly" (The Fordham Foundation, 1999, p. 9), it seems appropriate to ask these same school-level managers for their views on the causes of teacher ineffectiveness. Moreover, use of principals' evaluations allows direct comparison of the frequency with which principals judge content knowledge and pedagogical knowledge [to contribute to teacher ineffectiveness.

Torff and Sessions (2005) conducted survey research in which 242 secondary-level principals in high- and low-performing schools in New York State evaluated the threats to teacher quality posed by deficiencies in content knowledge and in four components of pedagogical knowledge. These components included (1) *lesson-planning skills* (preparation of appropriate learning experiences prior to an instructional period), (2) *lesson-implementation skills* (effective execution of planned learning experiences during an instructional period), (3) *ability to establish rapport with*

students (adequate human relations and communications skills), and (4) *classroom-management skills* (ability to successfully keep students on task and attentive).

Results indicated that the most frequent perceived causes of teacher ineffectiveness were deficiencies in components of pedagogical knowledge—three in-class components that entail student-teacher interaction (*lesson-implementation skills*, *ability to establish rapport with students*, and *classroom-management skills*), followed by the out-of-class component *lesson-planning skills*. Deficiencies in *content knowledge* were the least common perceived cause. Moreover, principals in high- and low-performing schools produced similar ratings (except that *lesson-planning skills* were rated as more problematic in low-performing schools), indicating that causes of teacher ineffectiveness varied little across schools that differed in socioeconomic status. These results run counter to claims that deficiencies in content knowledge constitute the greater threat to teacher quality, especially in low-performing schools (Finn, 1999; Finn & Madigan, 2001; Gross, 1999; Hess & Finn, 2004; Hirsch, 1996; Kanstoroom & Finn, 1999; Kramer, 1991; Ravitch, 2000; Sykes, 1995). The results also suggest that, in principals' judgment, programs and policies that emphasize content knowledge and deemphasize pedagogical knowledge fail to address the most common underlying causes of problems of teacher quality.

Differences across Subjects in Causes of Teacher Ineffectiveness

The results reported by Torff and Sessions (2005) do not include an assessment of the extent to which causes of teacher ineffectiveness differ across secondary subjects (English, mathematics, science, social studies, and languages other than English). Such a subject comparison seems needed, given the high level of concern over teacher quality in math and science (e.g., Finn & Kanstoroom, 2000; Kanstoroom, 1999; National Commission on Mathematics and Science Teaching for the 21st Century [USA], 1998). Research has shown that classes in math and science, as with classes in other secondary subjects, are often taught by teachers who appear to lack appropriate preparation in content knowledge (i.e., lacking in appropriate content coursework; Ingersoll, 2001, 2004, 2005). "Out of field" class assignments as such have been decried by a diverse group of educators, educational researchers, and policy makers (e.g., Darling-Hammond, 2000, 2005; Ingersoll, 2005; Kanstoroom & Finn, 1999; Murnane, et al., 1991).

But it does not necessarily follow that properly-assigned math and science teachers are more often vexed by deficiencies in content knowledge than deficiencies in pedagogical knowledge. Nor does it necessarily follow that content deficiencies are more problematic in math and science than in other subjects. It remains unclear how threats to teacher quality vary across secondary subjects. This analysis was not attempted in previous research on principals' perceptions of causes of teacher ineffectiveness (Torff & Sessions, 2005). The previous research ought to be replicated and also extended to analyze threats to teacher quality broken out by subject. These objectives are addressed in the research reported in this article.

Method

Survey Development

The research employed a survey adapted from one developed for previous research (Torff & Sessions, 2005). The survey used in the prior study was designed to tap aspects of teacher quality derived from teachers' guides created by administrators at 20 school districts in New York State. Each guide described the knowledge and skills expected to be demonstrated by teachers in a particular district. The guides charged teachers to demonstrate, among other things, the five components of teacher quality described above: *content knowledge*; *lesson-planning skills; lesson-implementation skills*; *ability to establish rapport with students*; and *classroom-management skills*. The first teacher-quality component refers to *content knowledge,* and the last four components are varieties of *pedagogical knowledge*. These components have been widely cited as essential to teacher efficacy (e.g., Berg, 2003; Berliner, 1987, 1992, 1994; Darling-Hammond. 2005; Darling-Hammond & Bransford, 2005; Darling-Hammond, Wise, & Klein, 1999; Shulman, 1987; Steeves & Brown, 2000; Stone, 2002).

The question is raised concerning the extent to which these components of teaching expertise are separate or overlapping. It seems reasonable to assert, for example, that a lack of content knowledge could influence the efficacy of a teacher's lesson plans, or that a weak lesson plan could influence lesson implementation. Similarly, Ingersoll (2004) has suggested that content knowledge shortfalls (as indicated by out-of-field class assignment) may have an impact on classroom management. However, low correlations (and insignificant ones) presented later in Table 2 indicate that for the principals surveyed in this research, there was considerable separation of the five components. These correlations support the use of the five components as meaningfully distinct (although not entirely discrete) aspects of teaching expertise.

The five components were incorporated into a survey on which respondents rated the frequency with which each component has been a significant factor when teachers' work was judged to be ineffective. Ratings were made using four-point scales (1=very rarely, 2=seldom, 3=sometimes, and 4=frequently). The more frequently a component was rated as a cause of teacher ineffectiveness, the greater the threat to teacher quality posed by that component.

In the modified version of the survey used in this study (Figure 1), principals produced five sets of ratings–one for each secondary subject (English, math, science, social studies, and languages other than English). Five independent variables were added, including three covariate measures (age, years of experience as a classroom teacher, and years of experience as an administrator). Also assessed were two effects (categorical variables that may be associated with a dependent variable): gender and educational attainment (master's degree; master's degree plus 30 credits; doctoral degree).

Figure 1
Survey instrument

When a teacher's classroom work is ineffective, it is because the teacher (circle one for each statement):

	Subject	Very Rarely	Seldom	Some-times	Frequently
fails to demonstrate needed content knowledge (teacher does not exhibit suitable expertise in the subject being taught)	English	1	2	3	4
	Math	1	2	3	4
	Social Studies	1	2	3	4
	Science	1	2	3	4
	LOTE *	1	2	3	4
fails to write effective lesson plans (teacher does not prepare appropriate learning experiences prior to an instructional period)	English	1	2	3	4
	Math	1	2	3	4
	Soc. St.	1	2	3	4
	Science	1	2	3	4
	LOTE	1	2	3	4
fails to implement lesson plans skillfully (teacher does not execute planned learning experiences effectively during an instructional period)	English	1	2	3	4
	Math	1	2	3	4
	Soc. St.	1	2	3	4
	Science	1	2	3	4
	LOTE	1	2	3	4
fails to establish sufficient rapport with students (teacher does not demonstrate adequate human relations and communication skills)	English	1	2	3	4
	Math	1	2	3	4
	Soc. St.	1	2	3	4
	Science	1	2	3	4
	LOTE	1	2	3	4
fails to maintain satisfactory classroom discipline (teacher does not successfully keep students on task and attentive)	English	1	2	3	4
	Math	1	2	3	4
	Soc. St.	1	2	3	4
	Science	1	2	3	4
	LOTE	1	2	3	4

* LOTE = languages other than English (e.g., Spanish)

Participants and Procedure

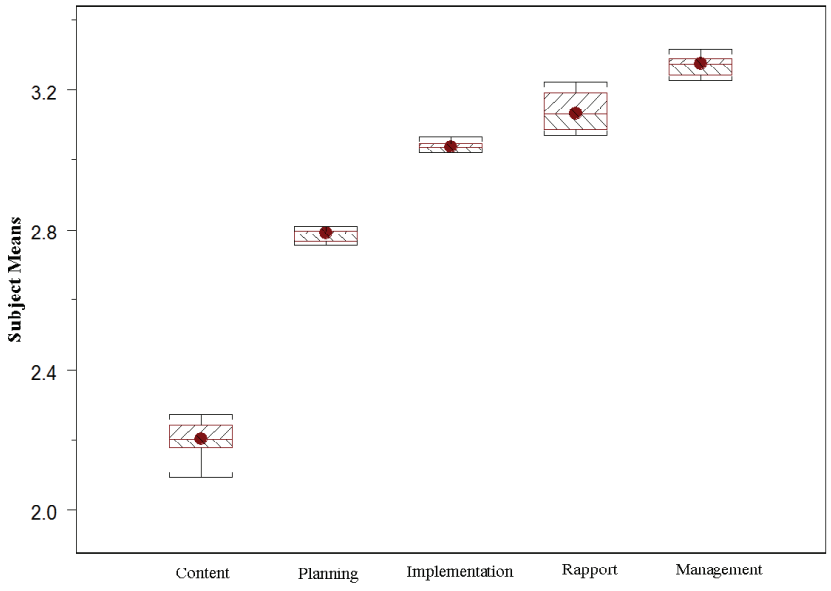
Three hundred fifty secondary schools in New York State and Michigan were selected at random using lists of all the schools in each state (175 schools per state). Principals at these schools received by mail a survey and a postage-paid return envelope, along with instructions indicating that there were no "correct" answers to the survey questions and that all responses were confidential. Data were entered into a spreadsheet for statistical analysis using SAS 9.1.

Results

The survey was completed and returned by 251 principals, yielding a response rate of 72%. As shown in Table 1, the sample included 183 men (72.9%) and 68 women (27.1%). The participants averaged 49.17 years of age (SD=8.64) with 12.19 years of teaching experience (SD=6.48) and 12.87 years of experience in school administration (SD=8.01). The sample included 49 principals (19.5%) with a master's degree, 169 principals (67.3%) with a master's degree plus 30 credits, and 33 principals (13.2%) with a doctoral degree. The demographic characteristics of the respondents were highly comparable to those reported in a prior study of principals' perceptions of causes teacher ineffectiveness (Torff & Sessions, 2005); given such consistency across independent samples, the data were considered to be suitable for further analyses.

The box-and-whisker plots depicted in Figure 2 show distinct separation of subject means between components of teaching expertise. In light of this separation, data analysis was carried out in two phases. To begin with, a MANCOVA was conducted to investigate relationships among the 25 dependent variables (ratings for five components in each of five subjects) while controlling for the impact of the

Figure 2
Box-and-Whisker Plots of Secondary Subject Means, within Teaching Components



Note. Content = content knowledge. Planning = lesson-planning skills. Implementation = lesson-implementation skills. Rapport = ability to establish rapport with students. Management = classroom-management skills.

Table 1
Descriptive Statistics

Categorical Independent Variables		n	(%)
Gender:	Male	183	(72.91)
	Female	68	(27.09)
Education:	Master's	49	(19.52)
	Master's +30	169	(67.33)
	Doctorate	33	(13.15)

Continuous Independent Variables	n	Mean	Std. Dev.	Min	Median	Max
Age	251	49.17	8.64	31.00	51.00	69.00
Teaching Experience	251	12.19	6.48	.00	11.00	31.00
Administrative Experience	251	12.87	8.01	1.00	12.00	40.00

Dependent Variables	n	Mean	Std Err	Std. Dev.	Min	Median	Max
Content Knowledge	248	2.22 *	.06	.89			
English	250	2.11	.06	.92	1.00	2.00	4.00
Math	250	2.26	.06	.99	1.00	2.00	4.00
Social Studies	250	2.20	.06	.91	1.00	2.00	4.00
Science	250	2.30	.06	.99	1.00	2.00	4.00
For. Language	248	2.22	.06	.95	1.00	2.00	4.00
Lesson Planning	248	2.78 *	.06	.95			
English	250	2.81	.06	.98	1.00	3.00	4.00
Math	250	2.79	.06	1.01	1.00	3.00	4.00
Social Studies	250	2.80	.06	.98	1.00	3.00	4.00
Science	250	2.76	.06	1.00	1.00	3.00	4.00
For. Language	248	2.75	.06	.98	1.00	3.00	4.00
Implementation	248	3.02 *	.05	.85			
English	250	3.05	.06	.87	1.00	3.00	4.00
Math	250	3.00	.06	.91	1.00	3.00	4.00
Social Studies	250	3.03	.06	.87	1.00	3.00	4.00
Science	250	3.00	.06	.90	1.00	3.00	4.00
For. Language	248	3.02	.06	.88	1.00	3.00	4.00
Rapport with Students	254	3.13 *	.06	.87			
English	251	3.12	.06	.92	1.00	3.00	4.00
Math	251	3.21	.05	.86	1.00	3.00	4.00
Social Studies	251	3.06	.06	.96	1.00	3.00	4.00
Science	251	3.18	.06	.91	1.00	3.00	4.00
For. Language	250	3.08	.06	.93	1.00	3.00	4.00
Classroom Mgmt.	254	3.26 *	.05	.84			
English	251	3.28	.06	.89	1.00	4.00	4.00
Math	251	3.22	.06	.92	1.00	3.00	4.00
Social Science	251	3.26	.05	.86	1.00	3.00	4.00
Science	251	3.23	.06	.88	1.00	3.00	4.00
For. Language	250	3.30	.05	.81	1.00	3.00	4.00

* Obtained by averaging responses across subjects, within a teacher-quality component.

covariates (age, years of experience as a classroom teacher, and years of experience as an administrator) and effects (gender and educational attainment). The results of these comparisons of conditional means are discussed in detail below; in general, the covariates and effects had little influence. Given these circumstances, a second analysis was undertaken with Hotelling's T^2 (employing unconditional means), to clarify discussion of the study's results and facilitate the replication of prior research findings (Torff & Sessions, 2005).

Table 2 presents the minimum and maximum correlations between subjects (within and across components) and their associated p-values (under the null hypothesis of zero correlation). These correlations indicate substantial separation among the components, as noted. In both the MANCOVA and Hotelling's T^2 models, the residuals were consistent with the assumptions underlying linear modeling. No significant outliers were detected. Comparisons of means were conducted using multiple comparisons procedures based on simultaneous confidence interval methods, including Bonferroni's method. All effect sizes reported below are measured as partial eta-squares.

The overall MANCOVA tests for significance showed that the variables *age*, *educational attainment*, *gender*, and *years of experience as an administrator* were statistically significant at a .05 level. However, closer inspection of the influence of these covariates and effects on the means of the dependent variables revealed that only *educational attainment* showed any systematic impact. This impact was restricted to the components *lesson-planning skills* and *lesson-implementation skills*. These two components were rated as a slightly more frequent cause of teacher ineffectiveness by principals with a higher level of educational attainment (relative to the ratings for

Table 2
Minimum and Maximum Correlations (with p-values) across Subjects, within and across Teaching Components

	Content Knowledge	Lesson Planning Skills	Lesson Implementation	Rapport with Students	Classroom Management Skills
Content Knowledge	.72 - .92 (.00)	.30 - .50 (.00)	.20 - .41 (.00)	.15 - .27 (.00)	.27 - .37 (.00)
Lesson-Planning Skills		.83 - .93 (.00)	.29 -.50 (.00)	.13 - .33 (.00 - .05)	.53 - .70 (.00)
Lesson-Implementation Skills			.87 - .95 (.00)	.13 - .38 (.00 - .05)	.3 - .49 (.00)
Rapport with Students				.81 - .97 (.00)	.52 - .70 (.00)
Classroom-Management Skills					.87 - .98 (.00)

these components produced by principals with less formal education). However, the effect sizes were small, with partial eta-squared statistics ranging from .05 to .11. (Effect sizes less than .20 have been categorized as "small," between .20 and .40 as "moderate," and above .40 as "large" [Cohen, 1988]; in this case, larger partial eta-squared statistics indicate a greater separation between variables.) All other partial eta-squares associated with the covariates and effects were exceedingly small (<.03). As reported below, differences across subjects within a component were dominated by differences between components (which included moderate-to-large effect sizes). Results as such indicate that the within-component effects associated with the variable *educational attainment* were insufficiently strong to merit inclusion in further analyses. Moreover, in this study, partial eta-squared statistics of .10 or less are treated as negligible, since they indicate an inconsequential contribution to the variance in variables of interest. The covariates and effects thus had little meaningful influence on the subject-by-component dependent variables.

Consequently, unconditional means of the dependent variables were compared using Hotelling's T^2 procedure (Littell, Stroup, & Freund, 2002). Three sets of significance tests are relevant: (1) comparisons of component means for aggregated subjects; (2) comparisons of component means within individual subjects; and (3) comparisons of subject means within individual components.

Comparisons of Component Means for Aggregated Subjects

Comparisons of the causes of teacher ineffectiveness for the aggregated data of the five secondary subjects are presented in Table 3. All comparisons of means differed at a .05 level of significance except one (*lesson-implementation skills* versus *ability to establish rapport with students*). Although the variable *classroom-management skills* was reported as the most frequent cause of teacher ineffectiveness, negligible effect sizes indicate that *classroom-management skills*, *lesson-implementation skills,* and *ability to establish rapport with students with students* were comparably frequent causes. The variable *classroom-management skills* was rated as more often a cause of teacher ineffectiveness relative to *lesson-*

Table 3
Comparisons of Component Means for Aggregated Secondary Subjects: Significance Levels and Effect Sizes

	Planning	Implementation	Rapport	Management
Content	.00 (.25)*	.00 (.40)	.00 (.42)	.00 (.53)
Planning		.00 (.06)	.00 (.09)	.00 (.19)
Implementation			.12 (.01)	.00 (.05)
Rapport				.00 (.03)

* p-value (partial eta-squared)

Note. Content = content knowledge; Planning = lesson-planning skills; Implementation = lesson-implementation skills; Rapport = ability to establish rapport with students; Management = classroom-management skills.

planning skills (with a small effect size of .19). The least frequent perceived cause of teacher ineffectiveness was *content knowledge*, which was separated from *lesson-planning skills* with a moderate effect size of .25 and from the student-interactive pedagogical variables (*classroom-management skills*, *ability to establish rapport with students,* and *lesson-implementation skills)* with large effect sizes of .53, .42, and .40, respectively.

Comparisons of Component Means within Individual Subjects

The question arises regarding the extent to which these results obtain in the five secondary subjects analyzed individually (Table 4). In each of these subjects,

Table 4
Comparisons of Component Means within a Secondary Subject: Significance Levels and Effect Sizes

	Planning	Implementation	Rapport	Management
English:				
Content	.00 (.29)*	.00 (.41)	.00 (.43)	.00 (.57)
Planning		.00 (.06)	.00 (.07)	.00 (.16)
Implementation			.39 (.00)	.00 (.05)
Rapport				.00 (.04)
Math:				
Content	.00 (.17)	.00 (.32)	.00 (.40)	.00 (.42)
Planning		.00 (.05)	.00 (.12)	.00 (.14)
Implementation			.00 (.03)	.00 (.04)
Rapport				1.00 (.00)
Social Studies:				
Content	.00 (.25)	.00 (.38)	.00 (.34)	.00 (.53)
Planning		.00 (.06)	.00 (.05)	.00 (.16)
Implementation			.74 (.00)	.00 (.05)
Rapport				.00 (.08)
Science:				
Content	.00 (.16)	.00 (.33)	.00 (.34)	.00 (.42)
Planning		.00 (.06)	.00 (.12)	.00 (.17)
Implementation			.01 (.03)	.00 (.05)
Rapport				.31 (.00)
LOTE:				
Content	.00 (.23)	.00 (.37)	.00 (.36)	.00 (.52)
Planning		.00 (.07)	.00 (.08)	.00 (.21)
Implementation			.42 (.00)	.00 (.08)
Rapport				.00 (.09)

* p-value (partial eta-squared)

Note. Content=content knowledge; Planning=lesson-planning skills; Implementation=lesson-implementation skills; Rapport=ability to establish rapport with students; Management=classroom-management skills; LOTE=languages other than English.

respondents produced the same pattern of threats to teacher quality as in the aggregated-subject results reported above. Although several pairwise comparisons were statistically significant at the .05 level, only three sets of differences produced meaningful effect sizes.

First, in all subjects, *content knowledge* was rated as a less frequent cause of teacher ineffectiveness relative to the pedagogical variables, with effect sizes ranging from .16 to .57. When *content knowledge* was compared to *lesson-planning skills*, the obtained effect sizes were small-to-medium (ranging from .16 to .29). The separation between the variables *content knowledge* and *lesson-planning skills* was somewhat less pronounced in math and science (with small effect sizes of .17 and .16, respectively) relative to moderate effect sizes obtained in languages other than English, social studies, and English (.23, .25, and .29 respectively). When *content knowledge* was compared to the student-interactive pedagogical variables, the effects were stronger: *classroom-management skills* produced large effect sizes ranging from .42 to .57; *ability to establish rapport with students* yielded moderate-to-large effect sizes ranging from .34 to .43; and *lesson-implementation skills* produced moderate-to-large effect sizes ranging from .32 to .41.

Second, in all subjects the variable *classroom-management skills* was rated as a more frequent cause of teacher ineffectiveness relative to the variable *lesson-planning skills.* Effect sizes were generally small, ranging from .14 to .21.

Third, in math and science, but not in the other subjects, there was meaningful separation of the variables *lesson-planning skills* and *ability to establish rapport with students.* The latter was judged to be the more frequent cause of teacher ineffectiveness, but the effect was weak, with a partial eta-squared statistic of .12 in both math and science.

Comparisons of Subject Means within Individual Components

In general, these data indicate that differences across teaching components were substantial, whereas differences across subjects within a teaching component were not. But that does not preclude the possibility of differences across subjects within individual components.

Table 5 presents differences across subjects for each separate teaching component. Although various subject differences were statistically significant (at a .05 level), the obtained effect sizes were negligible, ranging from .01 to .08. Since no meaningful differences were found, subjects could not be readily rank-ordered along any of the components assessed in this study. In particular, no effect size exceeded .06 for subject differences in *content knowledge,* a variable of considerable interest in the debate on threats to teacher quality. Principals' perceptions of the causes of teacher ineffectiveness were similar in math and science as in other secondary subjects.

Table 5
Comparisons of Secondary-Subject Means Within a Component: Significance Levels and Effect Sizes

	Math	Social Studies	Science	LOTE
Content Knowledge:				
English	.00 (.04)*	.00 (.04)	.00 (.06)	.01 (.03)
Math		.06 (.01)	.23 (.01)	.02 (.02)
Social Studies			.01 (.03)	.60 (.00)
Science				.01 (.03)
Lesson Plan:				
English	.59 (.00)	.70 (.00)	.18 (.01)	.13 (.01)
Math		.80 (.00)	.37 (.00)	.37 (.00)
Social Studies			.27 (.00)	.27 (.01)
Science				.86 (.00)
Implementation:				
English	.07 (.01)	.01 (.01)	.12 (.01)	.23 (.01)
Math		.37 (.00)	.99 (.00)	.54 (.00)
Social Studies			.42 (.00)	.81 (.00)
Science				.48 (.00)
Rapport:				
English	.02 (.02)	.01 (.03)	.07 (.01)	.09 (.01)
Math		.00 (.08)	.18 (.01)	.00 (.06)
Social Studies			.00 (.08)	.20 (.01)
Science				.00 (.05)
Management:				
English	.00 (.03)	.52 (.00)	.02 (.02)	.37 (.00)
Math		.01 (.03)	.28 (.00)	.00 (.05)
Social Studies			.00 (.05)	.16 (.01)
Science				.00 (.03)

* p-value (partial eta-squared)

Note. Content=content knowledge; Planning=lesson-planning skills; implementation=lesson-implementation skills; Rapport=ability to establish rapport with students; Management=classroom-management skills; LOTE=languages other than English.

Discussion

With issues of teacher quality in the spotlight, it has been suggested that teachers of mathematics and science too often lack content knowledge in the subjects they teach (Finn, 2000; Kanstoroom, 1999; National Commission on Mathematics and Science Teaching for the 21st Century [USA], 1998; Ravitch, 2000). Accordingly, research is needed to determine whether teacher ineffectiveness in these subjects is more frequently caused by deficiencies in content knowledge or in pedagogical knowledge, and whether teachers of mathematics and science are more often content-deficient relative to other teachers. Research as such requires that teacher performance be assessed, but this assessment has proven contentious (Berk, 1990;

Brandt, 1996; Copland, 2001; Darling-Hammond, 1986; Ellet & Teddie, 2003; Frase & Streshly, 1994; Goldstein, 2004; Haefle, 1993; Lavely, Berger, & Follman, 1996; Machell, 1990; Peel & Inkson, 1993; Sullivan & Mousley, 2000; Wayne & Youngs, 2003; Wise, et al, 1984; Worthen & Sanders, 1973). Use of principals' evaluations to assess teacher performance, while hardly foolproof, has the advantage of providing school-level managers' perspective on whether content knowledge or pedagogical knowledge constitutes the more frequent perceived cause of teacher ineffectiveness in secondary schools, especially in mathematics and science (Beerens, 2000; Darling-Hammond & Snyder, 2000; Denner, Salzman, & Bangert, 2001; Ferrell, 1992; Frase & Hetzel, 1990; Peterson, 2000).

Research was conducted in which 251 principals responded to a survey asking them to rate the frequency with which teacher ineffectiveness has, in their experience, been caused by five components of teacher quality, including *content knowledge* and four types of pedagogical knowledge (*lesson-planning skills* and three skills requiring interaction with students—*lesson-implementation skills, ability to establish rapport with students,* and *classroom-management skills*). When data from the five secondary subjects were aggregated, perceived threats to teacher quality could be ranked on three levels, with deficiencies in student-interactive pedagogical skills (especially classroom management) as most problematic, deficiencies in lesson-planning skills in the middle, and deficiencies in content knowledge as least likely to cause teacher ineffectiveness. These results are highly similar to ones obtained in previous research (Torff & Sessions, 2005).

The present study also examined causes of teacher ineffectiveness broken out by secondary subject (English, mathematics, science, social studies, and languages other than English). But very little variation across subjects was found, yielding the conclusion that principals judged the causes of teacher ineffectiveness to be similar in the five secondary subjects. Hence, the aggregated-subjects results obtained in this study and in prior research did not appear to be an artifact of strong effects associated with individual subjects. Teacher ineffectiveness in mathematics and science, as in other subjects, was rated as more often produced by deficiencies in pedagogical knowledge than deficiencies in content knowledge. And teacher ineffectiveness was rated as no more likely to stem from content deficiencies in mathematics and science compared to other subjects.

Results as such suggest that principals do not agree with claims that content deficiencies comprise the exigent threat to teacher quality in schools (Finn, 1999; Gross, 1999; Hess & Finn, 2004; Hirsch, 1996; Kanstoroom & Finn, 1999; Kramer, 1991; Ravitch, 2000; Sykes, 1995). The results also suggest that principals do not judge content deficiencies to be the exigent teacher-quality problem in math and science or to cause teacher ineffectiveness more frequently in these subjects than in other ones, as has been asserted (Finn & Kanstoroom, 2000; Kanstoroom, 1999; National Commission on Mathematics and Science Teaching for the 21st Century [USA], 1998).

The results have implications for contemporary initiatives to recruit and retain the workforce of teachers. As noted, alternative certification programs have proliferated, based on the belief that alternative routes to teacher certification have potential to increase both teacher quantity and quality. It has also been suggested that teacher certification be deregulated, such that teachers can gain certification if they have a bachelor's degree in a content area but no course work in pedagogy (Ballou & Podgursky, 1998; Finn, 1999; Finn & Kanstoroom, 2000; Kanstoroom & Finn, 1999; Podgursky, 2005).

Programs and policies as such may increase the quantity of teachers, particularly in subjects with a teacher shortage (such as math and science). However, according to principals, the quality of the teacher workforce will not likely be enhanced, at least from the standpoint of reducing the incidence of teacher ineffectiveness. Alternative-certification programs typically provide scant training in pedagogy, often a brief course in the summer preceding the initial year in the classroom (compared to the full year or more of pedagogical training, including student teaching, usually required in traditional teacher-preparation programs). Moreover, proposals to deregulate teacher certification portend to employ teachers with no pedagogical training. With deficiencies in pedagogical knowledge rated as causing more incidents of teacher ineffectiveness, programs and policies that are long on content and short on pedagogy may not reduce the frequency of these incidents.

Employing teachers with inadequate pedagogical knowledge may exacerbate problems of teacher quality, by contributing to teacher attrition. Research indicates that teacher shortages are attributable less to recruitment than retention (Darling-Hammond, 2005; Darling-Hammond & Bransford, 2005; Ingersoll, 2001, 2003, 2004, 2005). "After five years, 40 to 50 percent of all beginning teachers have left teaching altogether [and] the 'best and brightest' among new teachers…are the most likely to leave" (Ingersoll, 2004, pp. 144). These departing teachers often cite pedagogical challenges, especially in classroom management, as their main reason for leaving (Darling-Hammond, 2005; Ingersoll, 2003, 2004). Accordingly, more than two-thirds of the participants in Teach for America leave teaching after their initial two-year commitment – an attrition rate far higher than that of traditionally prepared teachers. Increasing the teacher workforce by adding individuals with minimal or nonexistent pedagogical knowledge seems likely to put upward pressure on attrition rates, further destabilizing schools.

The results of the study reported in this article also have implications for professional-development programs provided for practicing teachers. It has been suggested that professional development programs in math and science be focused exclusively on content knowledge (Hill, Rowan, & Ball, 2005; Kanstoroom, 1999). This suggestion follows from the concern that content-knowledge deficiencies comprise the greatest threat to teacher quality in these subjects. But such concern seems misplaced, according to the principals surveyed for the research reported in this article. Professional-development initiatives focused on content knowledge

apparently miss the target, by providing additional training where it is least needed while failing to address the components of pedagogical knowledge that principals regard as the main causes of ineffective teaching. These initiatives ought to be devoted at least part of the time to the pedagogical skills pressed into action when teachers interact with students—*classroom-management skills*, *ability to establish rapport with students*, and *lesson-implementation skills*.

The finding that principals view deficiencies in pedagogical knowledge as the most frequent cause of teacher ineffectiveness could be seen as a sign of weak performance on the part of the teacher-preparation programs charged to teach this knowledge. Initiatives to reform these programs have gained momentum of late (e.g., Cochran-Smith & Zeichner, 2005; Darling-Hammond, 2005; Darling-Hammond & Bransford, 2005; Darling-Hammond & Sykes, 2003; Wilson, Floden, & Ferrini-Mundy, 2001). But it is unsurprising that deficiencies in pedagogical knowledge were rated as the more frequent cause of teacher ineffectiveness. This knowledge has little in common with the academic background that teachers develop over more than 15 years in school. In this time, teachers acquire academic skills that are consistent with content-knowledge preparation but have little in common with the pedagogical challenges of teaching, especially where interaction with students is concerned. Being a good student in a biology class, for example, does little to help the student develop the classroom-management techniques and other pedagogical skills a biology teacher needs.

Accordingly, the results presented here could be interpreted to suggest that teacher-certification programs and policies place additional emphasis on pedagogical knowledge, especially the components of this knowledge rated as the most frequent causes of teacher ineffectiveness. The results could also be construed as suggesting that preservice and inservice teacher-education programs better prepare prospective teachers to meet the challenges of interacting with students. Teachers might well be required to complete rigorous training in both content knowledge and pedagogical knowledge, with appropriate emphasis on the interactive skills identified by principals as the most exigent threats to teacher quality.

Limitations and Future Research

The study's participants were employed in New York State and Michigan, and the causes of teacher ineffectiveness may differ elsewhere. This study was focused on secondary principals, but the causes of teacher ineffectiveness (and principals' perceptions of same) may vary between elementary and secondary schools. Principals' judgments of teacher performance are subjective and potentially biased and thus may differ from the actual causes of teacher ineffectiveness (Amadi & Rosenthal, 1993; Blumberg, 1980; Fant & Stevens, 1991; Kerrins & Cushing, 2000; Lee, Smith, & Cioci 1993; Rinehart & Young, 1996; Stodolsky, 1984). In particular, since principals are not always experts in the subjects taught by the teachers they supervise (typically having expertise in but one of the secondary

subjects), they may be less attentive to deficiencies in content knowledge than in pedagogical knowledge; such a possibility raises the need for additional research in which content-knowledge experts are surveyed.

Additionally, research is needed to examine how principals' evaluations of teachers' content knowledge and pedagogical knowledge relate to various measures of student outcomes, including test scores, performance assessments, and students' ratings of teacher performance. Research on principals' views of the antecedents of teacher quality might also encompass their views concerning what makes teachers' work *most* effective. The factors that lead to optimally effective teaching may differ from the factors associated with teacher ineffectiveness.

Finally, it has been suggested that the constituents of teaching expertise include, in addition to content knowledge and pedagogical knowledge, a third dimension called *pedagogical content knowledge* (Shulman, 1987). This term refers to the knowledge of how the content in a discipline can be marshaled for teaching purposes—to make disciplinary ideas accessible and comprehensible to students. Future research might well examine principals' perceptions of the role played by pedagogical content knowledge in both teacher ineffectiveness and teacher effectiveness.

These studies would be particularly illuminating were they to compare teachers who participated in a traditional teacher-education program and teachers who received alternative certification. Shedding light on the antecedents of teacher quality, this research has potential to inform teacher-certification policy, enhance techniques in teacher evaluation, and improve practices in teacher education.

References

Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality & Social Psychology, 64*, 431-441.

Ballou, D., & Podgursky, M. (1998). The case against teacher certification. *The Public Interest*, Summer 1998, 17-29.

Ballou, D., & Podgursky, M. (1999). Teacher training and licensure: A layman's guide. In M. Kanstoroom & C. Finn (Eds.), *Better teachers, better schools.* Washington, DC: Fordham Foundation.

Beerens, D. (2000). *Evaluating teachers for professional growth*. Thousand Oaks, CA: Corwin Press.

Berg, J. H. (2003). *Improving the quality of teaching through National Board Certification: Theory and practice*. Boston, MA: Christopher-Gordon Publishers.

Berk, R. A. (1990). Limitations of using student-achievement data for career-ladder promotions and merit-pay decisions. In J. V. Mitchell, Jr., S. L. Wise, & B. S. Plake (Eds), *Assessment of teaching: Purposes, practices, and implications for the profession* (pp. 261-306). Mahwah, NJ: Lawrence Erlbaum Associates.

Berliner, D. C. (1987). In pursuit of the expert pedagogue. *Educational Researcher, 15,* 5-13.

Berliner, D. (1992). The nature of expertise in teaching. In F. Oser, A. Dick, & J. Patry (Eds.),

Effective and responsible teaching (pp. 227-248). San Francisco: Jossey Bass.

Berliner, D. (1994). Expertise: The wonders of exemplary performance. In J. N. Mangieri & C. C. Block (Eds.), *Creating powerful thinking in teachers and students: Diverse perspectives* (pp. 161-186). Fort Worth, TX: Holt, Reinhardt, & Winston.

Blumberg, A. (1980). *Supervision and teachers: A private cold war* (2nd edition). Berkeley, CA: McCutchan.

Brandt, R. (1996). On a new direction for teacher evaluation. *Educational Leadership, 53*, 30-33

Cochran-Smith, M., & Zeichner, K. (2005). *Review of research in teacher education.* Washington, DC: American Educational Research Association.

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

Copland, M. A. (2001). The myth of superprincipal. *Phi Delta Kappan. 82,* 528-33.

Darling-Hammond, L. (2005). *A good teacher in every classroom: Preparing the highly qualified teachers our children deserve.* San Francisco: Jossey-Bass.

Darling-Hammond, L., & Bransford, J. (Eds.). (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do.* San Francisco: Jossey-Bass.

Darling-Hammond, L (1986). A proposal for evaluation in the teaching profession. *Elementary School Journal, 86,* 531-551.

Darling-Hammond, L. (2000). *Solving the dilemmas of teacher supply, demand, and quality*. New York: National Commission on Teaching and America's Future.

Darling-Hammond, L., & Snyder, J. (2000). Authentic assessment in teaching in context. *Teaching and Teacher Education, 16,* 523-545.

Darling-Hammond, L., & Sykes, G. (2003). Wanted: A national teacher supply policy for education: The right way to meet the "highly qualified teacher" challenge. *Education Policy Analysis Archives, 11,* 1-57.

Darling-Hammond, L., Wise, A. E., & Klein, S. (1999*). A license to teach: Raising standards for teaching.* San Francisco: Jossey-Bass.

Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Vazquez Heilig, J. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives, 15,* 1-51 .

Decker, P., Mayer, D. & Glazerman, S. (2004). *The effects of Teach for America on students: Findings from a national evaluation.* Princeton, NJ: Mathematica Policy Research.

Denner, P., Salzman, S., & Bangert, A. (2001). Linking teacher assessment to student performance: A benchmarking, generalizablility, and validity study of the use of teacher work samples. *Journal of Personnel Evaluation in Education, 15,* 287-307.

Ellett, C., & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education, 17,* 101-128.

Fandt, P., & Stevens, G. (1991). Evaluation bias in the business classroom: Evidence relating to the effects of previous experiences. *Journal of Psychology: Interdisciplinary & Applied, 25,* 469-477.

Ferrell, B. (1992). Lesson plan analysis as a program evaluation tool. *Gifted Child Quarterly, 36,* 23-26.

Finn, C. (1999). Foreword. In M. Kanstoroom & C. Finn (Eds.), *Better teachers, better schools.* Washington, DC: Fordham Foundation.

Finn, C., & Kanstoroom, M. (2000). Improving. empowering, dismantling. *The Public Interest, 140,* 64-74.

Finn, C., & Madigan, K. (2001). Removing the barriers for teacher candidates. *Educational Leadership, 58*, 29-32.

Flowers, C., & Hancock, D. (2003). An interview protocol and scoring rubric for evaluating teacher performance. *Assessment in Education: Principles, Policy, & Practice, 10,* 161-168.

The Fordham Foundation. (1999). The teachers we need and how to get more of them. In M. Kanstoroom & C. Finn (Eds.), *Better teachers, better schools.* Washington, DC: Fordham Foundation.

Frase, L. E., & Streshly, W. (1994). Lack of accuracy, feedback, and commitment in teacher evaluation. *Journal of Personnel Evaluation in Education, 8,* 47-57.

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

Goldstein, J. (2004). Making sense of distributed leadership: The case of peer assistance and review. *Educational Evaluation & Policy Analysis, 26,* 173-197.

Gross, M. (1999). *The conspiracy of ignorance: The failure of American public schools.* New York: Harper Collins.

Haefele, D. L. (1993). Evaluating teachers: A call for change. *Journal of Personnel Evaluation in Education, 7,* 21-31.

Hess, F., & Finn, C. (Eds.). (2004*). Leaving no child behind?: Options for kids in failing schools.* New York: Macmillian.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42,* 371-406.

Hirsh, E. D. (1996). *The schools we need and why we don't have them.* New York: Doubleday.

Ingersoll, R. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal, 37,* 499-534.

Ingersoll, R. (2003). *Who controls teachers' work? Power and accountability in America's schools.* Cambridge, MA: Harvard University Press.

Ingersoll, R. (2004). Revolving doors and leaky buckets. In C. Glickman (Ed.), *Letters to the next president: What we can do about the real crisis in public education* ***(pp. 141-150).*** New York: Teachers College Press.

Ingersoll, R. (2005). The problem of underqualified teachers: A sociological perspective. *Sociology of Education, 78,* 175-178.

Isbell, L. L. (2002). Teacher certification and teacher efficacy as correlates of student achievement. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 63*(5-A), pp. 1791.

Kanstoroom, M. (1999). Qualified math and science teachers: Rethinking the regulatory approach. Testimony prepared for delivery to the Committee on Science, U.S. House of Representatives, August 4.

Kanstoroom, M., & Finn. C. (Eds.) (1999). *Better teachers, better schools.* Washington, DC: Fordham Foundation.

Kelly, S. (2004). Are teachers tracked? On what basis and with what consequences. *Social*

Psychology of Education, 7, 55-72.

Kerrins, J., & Cushing, K. (2000). Taking a second look: Expert and novice differences when observing the same classroom teaching segment a second time. *Journal of Personnel Evaluation in Education, 14,* 5-24.

Kramer, R. (1991). *Ed school follies: The miseducation of America's teachers.* New York: Free Press.

Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee value added assessment system. *Educational Evaluation & Policy Analysis, 25,* 287-298.

Laczko-Kerr, I., & Berliner, D. C. (2002, September 6). The effectiveness of "Teach for America" and other under-certified teachers on student academic achievement: A case of harmful public policy. *Education Policy Analysis Archives, 10*, http://epaa.asu.edu/epaa/v10n37/

Lavely, C., Berger, N., & Follman, J. (1996). Contemporary process-product teacher evaluation research relationships. *Journal of Instructional Psychology, 23,* 293-296.

Lee, V. E., Smith, J., & Cioci, M. (1993). Teachers and principals: Gender-related perceptions of leadership and power in secondary schools. *Educational Evaluation & Policy Analysis, 15,* 153-180.

Littell, R. C., Stroup, W. W., & Freund, R. J. (2002). *SAS for linear models.* Hoboken, NJ: Wiley Interscience.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States.* Mahwah, NJ: Lawrence Erlbaum Associates.

Machell, J. (1995). The teacher evaluation environment: An examination of attributes related to teacher growth. *Journal of Personnel Evaluation in Education, 9,* 259-273.

Matsumura, L. C., Garnier, H., Pascal, J., & Valdes, R. (2002). Measuring instructional quality in accountability systems: Classroom assignments and student achievement. *Educational Assessment, 8,* 207- 229.

Monk, D. H. (1994). Subject matter preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review, 13*, 125-145.

Murnane, R. J., Singer, J. D., Willett, J. B., Kemple, J. J., & Oben, R. J. (1991). Who will teach? Policies that matter. Cambridge, MA: Harvard University Press.

Nagy, P., & Moorhead, R. (1990). Administrative response to classroom testing data: A problem solving perspective. *Alberta Journal of Educational Research, 36,* 18-34.

National Commission on Mathematics and Science Teaching for the 21st Century [USA]. (1998). *Before it's too late: A report to the nation from the National Commission on Mathematics and Science Teaching for the 21st Century [USA]*. Washington, DC: United States Department of Education.

Neill, M. (1999). Stop misusing tests to evaluate teachers. *Social Education, 63,* 330-32.

Null, J. W., & Ravitch, D. (Eds.) (2006). *Forgotten heroes of American education: The great tradition of teaching teachers.* Greenwich, CT: Information Age Publishing.

Peel, S. & Inkson, K. (1993). High school principals' attitudes to performance evaluation: Professional development or accountability? *New Zealand Journal of Educational Studies, 28,* 125-141.

Peterson, K. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices*. Thousand Oaks, CA: Corwin Press.

Podgursky, M. (2005). Teaching licensing in U.S. pubic schools: The case for 49 years of

age simplicity and flexibility. *Peabody Journal of Education, 80,* 15-43.

Ravitch, D. (2000). *Left back: A Century [USA] of failed school reforms.* New York: Simon & Schuster.

Rinehart, J. S., & Young, I. P. (1996). Effects of teacher gender and principal gender on ratings of teacher performance. *Journal of Personnel Evaluation in Education, 10,* 313-323.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*, 1-22.

Steeves, K. A., & Browne, B.C. (2000). *Preparing teachers for National Board Certification: A facilitator's guide.* New York: Gilford Publications.

Stodolsky, S.S. (1984). Teaching evaluation: The limits of looking. *Educational Researcher, 13*, 11-18.

Stone, R. (2002). *Best practices for high school classrooms: What award-winning secondary teachers do.* Thousand Oaks, CA: Corwin Press.

Sullivan, P., Mousley, J., & Gervasoni, A. (2000). Caution: Classroom under observation. *Asia-Pacific Journal of Teacher Education, 28,* 247-260.

Sykes, C. (1995). *Dumbing down our kids: Why American children feel good about themselves but can't read, write, or add*. New York: St. Martin's Press.

Torff, B., & Sessions, D. (2005). Principals' perceptions of the causes of teacher ineffectiveness. *Journal of Educational Psychology, 97,* 530-537.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73,* 89-122.

Wilson, S., Floden, R., & Ferrini-Mundy, J. (2001). *Teacher preparation research: Current knowledge, gaps, and recommendations.* Seattle, WA: University of Washington Center for the Study of Teaching and Policy.

Wise, A., Darling-Hammond, L., McLaughlin, M., & Berstein, H. (1984-85). Teacher evaluation and teacher professionalism. *Educational Leadership, 4,* 28-33.