Teachers’ qualifications and their impact on student achievement: Findings from TIMSS 2003 data for Israel

Ruth Zuzovsky
Center for Science and Technology Education, Tel Aviv University, Tel Aviv, Israel

Data collected as part of the Trends in International Mathematics and Science Study 2003 (TIMSS 2003) in Israel make it possible to validate several assumptions regarding the relationship between certain teacher characteristics and student achievement. We need this validation if we want to take a stance in the debate, occurring in Israel and elsewhere, regarding the nature of the reforms needed in the selection and preparation of teachers, opportunities for professional development, and the reward mechanisms and incentives that affect teachers’ career structures.

INTRODUCTION

International comparative studies of educational achievement have become an important source of information for those involved in educational policymaking. Although many commentators view these studies as “horse races” that focus mainly on the relative position of one country’s attainment to that of others, there are those who appreciate their educational role (Bryk & Hermanson, 1993; Darling-Hammond, 1992, 1994; Kellaghan, 1996) and the opportunity these studies provide to clarify and reassess local policy assumptions. Diane Shorrocks-Taylor (2000, p. 18) sees the benefit of participating in international comparative studies as the challenge these present to existing local policies: “The process of participation requires self-evaluation which in turn may lead to assumptions being questioned and what previously understood only implicitly now being made explicit and so examined in a more critical way.” This article illustrates the relevance of data obtained as part of the Trends in International Mathematics and Science Study 2003 (TIMSS 2003) for local policymaking in Israel.

The policy issue at the heart of this article relates to the need to ensure the presence of "highly qualified teachers in every classroom" and to determine how best to define and prepare these "qualified" teachers. Quality teachers are often seen simply as "good" teachers and are considered to be those who exhibit desirable traits and uphold the standards and norms of the profession. But quality teachers are also considered to be those who bring about "student learning." These teachers are called "effective" (Berliner, 1987, 2005) or "successful" (Fenstermacher & Richardson, 2005). Fenstermacher and Richardson (cited in Berliner, 2005, p. 207) distinguish between *good* teaching and *successful* teaching as follows:

> By "good teaching" we mean that the content taught accords with disciplinary standards of adequacy and completeness and the methods employed are age appropriate, morally defensible and undertaken with the intention of enhancing the learner's competence with respect to content. By "successful teaching" we mean that the learner actually acquires some reasonable and acceptable level of proficiency from what the teacher is engaged in teaching.

Because of psychometric difficulties in assessing teachers by their normative attributes—the logical, the psychological, and (especially) the ethical, which are defined differently across cultures (Alexander, 2000)—the tendency to evaluate teacher qualities on the basis of student performance is given even greater emphasis. With the increased demands for accountability in line with performance standards and with the growing demand for evidence-based policymaking, student achievement is considered an accurate measure of teacher effectiveness and has become a basis for value-added teacher assessment systems (Braun, 2005; McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004; Sanders, 2000; Sanders & Rivers, 1996).

These notions have also found favor in regard to the effectiveness of teacher education systems. After tracing the development and reform of teacher education in terms of the major questions shaping this field of education, Cochran-Smith (2001) argues that "the outcome" question is what currently motivates teacher education research and policymaking. She set down three ways in which the outcomes of teacher education are constructed. One of them, *long-term impact outcomes*, refers to the relationships between teacher qualifications and student learning. Teachers' qualifications encompass teachers' scores on tests and examinations, their years of experience, the extent of their preparation in subject matter and in pedagogy, what qualifications they hold in their area of expertise, and their ongoing professional development. Student learning is taken simply as the gain scores students attain on achievement tests. Cochran-Smith (2001, p. 531) went on to posit the relationship between teacher qualification and student learning as the percentage of variance in student scores accounted for by teacher qualifications when other variables are held constant or adjusted.

In many countries, teacher qualifications that are considered to be related to student learning have become targets of education reform. However, the nature of this reform is under debate. Some perceive the main problem to be the low academic and cognitive level of those who go into the teaching profession and call for policies aimed at attracting more capable candidates through shorter, less regulated alternative routes (Ballou & Podgursky 1997, 1999, 2000; Goldhaber & Brewer, 2000; United States Department of Education, 2002). Others view the problem mainly as the result of inadequate teacher preparation and call for the "professionalization" of teacher education by making it longer, upgrading it to graduate programs, and regulating it through mechanisms of licensure, certification, and promotion aligned with standards (Darling-Hammond, 1998, 1999, 2000a; Darling-Hammond, Berry, & Thorenson, 2001; Darling-Hammond, Chung, & Frelow, 2002; National Commission on Teaching and America's Future, 1996).

The impact of these different approaches on student learning have been explored in several meta-analytic studies based mainly on United States data but also drawing from the databases of other countries (see, in this regard, Darling-Hammond, 1999, 2000b; Greenwald, Hedges, & Laine, 1996; Organisation for Economic Co-operation and Development, 2005; Santiago, 2002; Wayne & Youngs, 2003; Wilson, Floden, & Ferrini-Mundy, 2001). Other relevant studies have drawn more on local sources of data and have been targeted at specific (country-based) policies (Harris & Sass, 2007; Ingersoll, 2003; Wilson, Darling-Hammond, & Berry, 2001). In Israel, too, teacher qualifications have become the target of several recent reforms, such as those announced by different teacher unions (2004), the National Task Force for the Advancement of Education in Israel (Dovrat Committee, 2005), and the Committee of the Commission for Higher Education (Ariav, Olshtain, Alon, Back, Grienfeld, & Libman, 2006). The reforms suggested in Israel are more in line with the advocacy to professionalize teacher preparation. All suggestions thus far envision improving the candidate selection process, upgrading the disciplinary preparation of teachers, opening advanced degree Master of Education (M.Ed) or Master of Teaching (M.Teach) programs, and providing opportunities for professional development.

Given the relatively few studies conducted in Israel on the impact of these recommended policies on student learning, and because of the conflicting results obtained from the many studies conducted elsewhere, the study documented in this article attempted to validate some of the assumptions at the basis of the suggested policies. More specifically, the study re-examined the extent to which advanced academic degrees, majoring in the field of teaching, years of teaching experience, and intensive participation in professional development activities—all assumed to be cardinal teacher qualifications—are indeed positively associated with student achievement in mathematics and science.

LITERATURE REVIEW

This section offers a summary of research findings related to each of the teacher qualifications considered in this study.

Teachers' Formal Education

Findings related to teachers' academic degrees (Bachelor's, Master's, doctorate, and other) are inconclusive. Some studies show positive effects of advanced degrees (Betts, Zau, & Rice, 2003; Ferguson & Ladd, 1996; Goldhaber & Brewer, 1997, 2000;

Rowan, Chiang, & Miller, 1997); others show negative effects (Ehrenberg & Brewer, 1994; Kiesling, 1984). Some researchers maintain that the requirement for teachers to have a second degree raises the cost, financially as well as in time, of teacher education, which may prevent quality candidates from choosing this profession (Murnane, 1996).

Teacher Education in the Subject Matter of Teaching (in-field preparation)

This characteristic is related to the subject-matter knowledge teachers acquire during their formal studies and pre-service teacher education courses. The evidence from different studies is contradictory. Several studies show a positive relationship between teachers' preparation in the subject matter they later teach and student achievement (Darling-Hammond, 1999, 2000b; Goldhaber & Brewer, 2000; Guyton & Farokhi, 1987), while others have less unequivocal results. Monk and King (1994) found both positive and negative effects of teachers' in-field preparation on student achievement. Goldhaber and Brewer (2000) found a positive relationship for students' mathematics achievement but no such relationship for science. Rowan et al. (1997) reported a positive relationship between student achievement and teachers with a major in mathematics. Monk (1994), however, found that while having a major in mathematics had no effect on student achievement in mathematics, having a substantial amount of under- or post-graduate coursework had a significant positive effect on students in physics but not in life sciences.

Ingersoll (2003) considered the widespread phenomenon in the United States of teachers teaching subjects other than those for which they had formal qualifications. His study of out-of-field teaching (as it is known) portrayed a severe situation where 42% to 49% of public Grades 7 to 12 teachers of science and mathematics lacked a major and/or full certification in the field they were teaching (1999/2000 data). In Israel, a recent survey (Maagan, 2007) placed the corresponding percentages even higher for elementary teachers—42% for mathematics and 63% for science (2005/2006 data).

Teacher Education in Pedagogical Studies

The literature shows a somewhat stronger, and more consistently positive, influence of education and pedagogical coursework on teacher effectiveness (e.g., Ashton & Crocker, 1987; Everston, Hawley, & Zlotnik, 1985; Ferguson & Womack, 1993, Guyton & Farokhi, 1987). Some of these studies compare the effect on student achievement of courses in pedagogical subject matter with the effect of courses in the subject matter itself, and present evidence in favor of the former. An example is a study conducted by Monk (1994) related to mathematics achievement. Other studies reveal no impact of education courses on students' achievement (see, for example, Goldhaber & Brewer, 2000, in relation to science achievement).

Duration of Pre-service Education

Despite evidence that five-year programs result in a higher retention rate and career satisfaction of their graduates than do four-year programs (Andrew, 1990), there is no evidence that graduates of the longer programs become more effective teachers. Data collected in TIMSS 2003 in Israel cannot contribute to this consideration, as the information collected on teachers' pre-service education did not differentiate between consecutive teacher preparation programs at universities (one- to two-year programs taken after completion of the first degree in a discipline) and concurrent programs at teachers' colleges (four- to five-year integrated disciplinary and pedagogy programs).

Certification and Licensing Status

Certified teachers are usually those who have graduated from accredited teacher education programs. Some of these teachers are also required to complete an induction program or pass a national teacher examination test in order to obtain a license. There is debate in the USA between those in favor of full certification (Darling-Hammond, 1999; Darling Hammond et al., 2001) and those who argue that students of teachers who hold full certification achieve similarly to those who study under teachers with temporary "emergency" credentials (Goldhaber & Brewer, 2000). These authors also argue that relaxing requirements for certification is a way not only of attracting academically talented college graduates to teaching but also of recruiting a more diverse pool of candidates needed for a diverse student population. The TIMSS 2003 data at hand for Israel prevented examination of this issue, as all participating teachers were fully certified.

Years of Experience

Studies on the effect of teacher experience on student learning have found a positive relationship between teachers' effectiveness and their years of experience, but the relationship observed is not always a significant or an entirely linear one (Klitgaard & Hall, 1974; Murnane & Phillips, 1981). The evidence currently available suggests that while inexperienced teachers are less effective than more senior teachers, the benefits of experience level off after a few years (Rivkin, Hanushek, & Kain, 2000).

The relationship between teacher experience and student achievement is difficult to interpret because this variable is highly affected by market conditions and/or motivation of women teachers to work during the child-rearing period. Harris and Sass (2007) point to a selection bias that can affect the validity of conclusions concerning the effect of teachers' years of experience: if less effective teachers are more likely to leave the profession, this may give the mistaken appearance that experience raises teacher effectiveness. Selection bias could, however, work in the opposite direction if the more able teachers with better opportunities to earn are those teachers most likely to leave the profession.

Participation in Professional Development Activities

Professional development activities can be conducted by many different organizations, in school and out of school, on the job or during sabbatical leave. On these occasions, practicing teachers update their content knowledge and teaching skills so they can meet the requirements of new curricula, consider new research findings on teaching and learning, and adapt to changes in the needs of the student population, and so on. Criticism has been leveled against the episodic nature of these activities and concern expressed that very little is known about what these activities really comprise and involve.

Conclusions in the literature on the relationship between teachers' participation in professional development activities and student outcomes are mixed. Some studies on in-service professional development have found no relationship to student achievement (see, in regard to mathematics and reading, Jacob & Lefgren, 2004). Other studies have found higher levels of student achievement linked to teachers' participation in professional development activities directly related to the area in which they are teaching (see, in regards to mathematics, Brown, Smith, & Stein, 1995; Cohen & Hill, 1977; Wiley & Yoon, 1995; and in regard to language and mathematics, Angrist & Lavy, 2001). Wenglinsky (2000) found a positive correlation between professional development activities aimed at the needs of special education students, and students' higher-order skills and laboratory skills in science. More recently, Harris and Sass (2007) identified what they call the "lagged effect of professional development," that is, the larger effect of teachers' professional development on student outcomes not becoming apparent until three years after the teachers had completed their courses.

The interpretation of the positive effect of participation in teacher professional development activities is not clear cut, as this variable is confounded with other teacher attributes, that is, teachers who participate in these activities are also likely to be more motivated and, usually, more specialized in the subjects they teach.

METHOD

Sample

The sample of teachers who participated in TIMSS 2003 in Israel comprised 371 mathematics teachers and 317 science teachers, who taught about 4,000 students in 149 sampled classes, each class in a different school. It became evident that in only about one quarter of the mathematics classes and about one third of the science classes all students were taught by a single teacher. In the rest of the classes, all students either were taught by more than one teacher, or students in one class were divided into groups, each of which was taught by one, or sometimes more than one, teacher.

In the present study, which examines the relationship between teacher characteristics and student achievement, it was essential to link teachers exclusively to the class or group of students they taught. Thus, a preliminary step in the analysis was identification of these specific learning units. In the process of preparing such a file, the sample of students who participated in the study was first reduced due to missing data on student or teacher variables relevant to this study. The resulting dataset was then used to identify the specific learning unit. Because grouping of students is very common in Israel, and frequently occurs within several same-grade classes, there were occasions when some students, originally studying with same-grade students from other classes, were identified in the sampled classes. These students, who were part of interclass groups and comprised a too small learning unit, were omitted from the analysis. The number of students omitted for this reason in mathematics (where grouping is very common) reached 625; in science, the number of omissions was 277. Table 1 presents the distribution of the remaining learning units according to the number of teachers assigned to teach each such unit (group of students) and the number of students studying in these groups.

Table 1: Learning units by number of teachers teaching them

Type of group according to number of teachers assigned to teach	Mathematics		Science	
	No. of groups	*No. of students*	*No. of groups*	*No. of students*
Groups taught by one teacher	143	2,036	110	1,656
Groups taught by two teachers	12	232	33	683
Groups taught by three (or more) teachers	1	22	11+(2)	298
Total	156	2,290	156	2,637
Missing		625		277
Total		2,915		2,914

We tested the representativeness of the reduced post-exclusions sample by comparing the achievement of students who remained in the analysis with the achievement of students in the sample before the exclusions. The result of a *t*-test showed a small but significant difference (5 to 6 points; 0.05–0.06 of a standard deviation of student score distribution) in favor of the "pre-exclusion" students. We considered this difference educationally meaningless, and the process of excluding students from the interclass groups did not seriously violate the representativeness of the sample. Table 2 presents the comparisons.

Table 2: Achievement of students before and after excluding students in small groups

The comparison groups	Science scores	*t*-value and sig.	Mathematics scores	*t*-value and sig.
Students in the groups left for the analysis	493 (80) *n* = 2,914	3.3**	498 (79) *n* = 2,915	3.9***
Students before excluding those from small groups	498 (77) n = 2,637		504 (78) n = 2,290	

Note: ** = $p \leq 001$, *** = $p \leq 0001$.

Data Source

We used the responses of teachers and principals to questions in the teacher and school questionnaires to determine the independent teacher variables. Six variables came from the teacher questionnaires; the remaining two came from the school questionnaire. Those variables already known to be associated with science and mathematics achievement (see, for example, extensive reviews by Darling-Hammond, 1999; Greenwald et al., 1996; Wayne & Youngs, 2003) describe teachers' academic preparation, their highest level of education, their preparedness for and feelings of readiness to teach subject matter(s), their years of teaching experience, and their participation in professional development activities. The variables that we drew from the TIMSS' questionnaires were:

1. The ethnic affiliation of the teacher (ISRARB), as inferred from the language used by the teacher with his or her students: 0–*Arabic speaking*; 1–*Hebrew speaking*.
2. Teacher's gender (TSEX): 0–*male*; 1–*female*.
3. Seniority as inferred from the number of years of teaching (TAUT), a continuous variable.
4. Teachers' highest levels of education (MA): 0–*up to and including first university degree*; 1–*beyond first university degree*.
5. Teachers' major areas of study in the field they teach (INFLD): 0–*study in areas other than the subject they teach*; 1–*study in the areas they teach*—mathematics or mathematics education, and, in the sciences, at least one of five relevant areas—biology, physics, chemistry, earth sciences, and science education; and 2–*study in at least one of the areas they teach and in addition in science or in mathematics education*.
6. Teachers' feeling of readiness to teach the content of instruction (READY), an index based on teachers' average responses toward a list of topics in mathematics and in science, on a scale of 1–*not feeling ready to teach the topics* to 3–*feeling highly ready to teach the topics*. The index was cut into three categories, indicating low, medium, and high readiness.
7. Extent of participation in professional development activities focusing on **content knowledge** (PDICK), on a scale of 1–*never or once or twice per year* to 4–*more than 10 times per year* (data from school questionnaire).

8. Extent of participation in professional development activities focusing on **pedagogy** (PEDAG), an index describing teachers' participation in four professional development activities: implementation of the national curriculum, improving school's own goals, improving teaching skills, and using ICT, on a scale of 1–*never or once or twice per year* to 4–*more than 10 times per year* (data from school questionnaire). The index was recoded into three categories indicating low participation (1), medium participation (2), and high participation (3).

The dependent variables used were the average estimates of the five plausible scores in mathematics or science.

Analyses

Two main analyses were carried out in this study. The first involved breaking down the group means of student achievement (dependent variable) by the categories of the different teacher variables (independent variables). We used the differences in achievement scores of students taught by teachers characterized by the extreme variable categories as the measures of teachers' effectiveness. The second analysis that we conducted was a multilevel regression analysis using hierarchical linear modeling software (HLM6) (Raudenbush, Bryk, Cheong, & Congdon, 2000), carried out using all five estimates of individual plausible values in mathematics or science.

The models specified for the analysis were two-level models of students nested in groups taught exclusively by one or more teachers. At their first level, the models contained three variables describing student characteristics. The first of these was *Number of books in student's home* (BOOK) on a scale of 1–*up to 10 books* to 5–*more than 200 books*. This variable provides a proxy of the socioeconomic and cultural background of each participating student's home. The second variable was *Student's self-confidence in learning the subject taught* (SCM/S1), a dummy variable derived from an index constructed for the TIMSS 2003 survey (Martin, Mullis, & Chrostowski, 2004). This variable indicated whether a student had *a lot of confidence* (1) or *not* (0). The third variable described the level of education students aspired to complete (ASPIR) on a scale of 1–*finish high school* to 5–*finish university beyond initial level*. At their second level, the models contained the eight above-described teacher variables. In cases where several teachers were teaching the same group of students, we averaged and sometimes rounded the values of their relevant variables.

To avoid problems of multicollinearity and to maximize interpretability, the second-level teacher variables were centered and standardized around their grand mean (Aiken & West, 1991, p. 43). Thus, their regression coefficients represent the change in achievement score points due to an increase by one standard deviation above the standardized mean of the relevant teacher variable.

The regression analyses were carried out separately for the two dependent variables. For each analysis, we specified three models. The first model was a "null" variance component model with no predictors. This model provided estimates of the variance components at each of the model's levels, indicating an upper limit to the explanatory power of the different models specified later. The second model included only the student-level variables; the third model also contained the second-level variables, that is, the teacher variables. The models specified for each dependent variable contained the same predictors at each level.

The most important outputs of this analysis were estimates of the regression coefficients of the predictors that indicate their effect on student achievement (the slopes of the predictor's regression lines). Allowing the coefficients of the first-level variables (student variables) to be modeled as random yields a "slope as outcome" model (Raudenbush & Bryk, 1986) in which the slope (regression coefficient) of a student-level variable is itself regressed over the higher-level teacher variables. This "slope as outcome" model is formally equivalent to an interaction model, indicating the existence of an interaction between a student-level variable, which varies randomly among the second-level units of analysis, and relevant second-level teacher variables.

The significant slope variation revealed for all student-level variables in our analysis justified our decision to look for interactions between these variables and the teacher variables or—to state this another way—to determine if student-level variables buffer the effect of teacher variables on achievement. In this study, we looked for interactions related to only one student-level variable—*Student's aspiration to finish a high level of academic education* (ASPIR). This variable is usually associated with both a student's intellectual capabilities and his or her socioeconomic background. The effects of the two other student variables were specified as fixed effects.

RESULTS

1. Breakdown of Group Achievement Scores by Categories of Teacher Variables

Table 3 presents the group mean achievement and its standard deviation as well as the gaps in the mean achievement of groups of students taught by teachers belonging to the distal categories of each teacher variable. In those cases where there were only a few groups in the distal categories, the categories were collapsed with others to render a more accurate picture. Large gaps indicate that the relevant teacher variable is associated with achievement (i.e., is effective).

Among the mathematics teacher variables, we found the following ones to be the most effective in terms of their association with student achievement.

- *Ethnic affiliation* (ISRARB): Students in groups taught by Hebrew-speaking mathematics teachers achieve more—almost three quarters of a standard deviation of the group's mean mathematics scores (*SD* = .57)—than do students studying in groups taught by Arabic-speaking teachers.
- *Gender* (TSEX): Students taught by female teachers achieve more—about a third of a group standard deviation—than do students taught by male teachers.
- *Seniority* (TAUT): Students studying in groups taught by mathematics teachers with more than 15 years of experience achieve more—by about a half of a group standard deviation—than do students studying in groups taught by mathematics teachers with less experience (five years or fewer).

Table 3: Breakdown of group mean achievement in mathematics and science by the categories of teacher variables

Teacher variables	Mathematics		Science	
	No. of groups 156	*Group mean and SD*	*No. of groups 156*	*Group mean and SD*
Ethnic affiliation (ISRARB)				
0 Arabic-speaking	40	465 (48)	35	464 (39)
1 Hebrew-speaking	116	507 (55)	121	490 (48)
Gap: Hebrew-speaking vs. Arabic-speaking		**42**		**26**
Teacher's gender (TSEX)				
0 Male	30	480 (53)	26	477 (47)
1 Female	113	501 (57)	130	486 (47)
Gap: Female vs. male		**21**		**9**
Second degree (MA)				
0 First degree or less	119	498 (55)	108	481 (45)
1 Second degree or more	37	493 (63)	48	491 (52)
Gap: Second degree vs. first degree		**-5**		**10**
Seniority (TAUT)				
Up to 5 years	55	482 (55)	53	479 (45)
5–15 years	48	450 (56)	53	475 (46)
15+ years	53	507 (58)	50	499 (48)
Gap: High vs. low seniority		**25**		**20**
Participation in content-oriented prof. development (PDICK)				
Never/Once or twice a year	20	485 (57)	18	465 (39)
3–5 times a year	55	493 (49)	49	482 (50)
6–10 times a year	43	507 (62)	46	492 (47)
More than 10 times a year	38	496 (51)	43	487 (47)
Gap: More than 10 times a year vs. never		**11**		**22**
Participation in pedagogically-oriented prof. development (PEDAG)				
Never /Once or twice a year	60	506 (51)	62	488 (42)
3–10 times a year	59	491 (53)	39	486 (42)
More than 10 times a year	37	490 (69)	55	478 (56)
Gap: More than 10 times a year vs. never		**-16**		**-4**
Major in subject-area being taught (INFLD)				
0 no major	10	502 (78)	3	466 (36)
1 majoring in at least one relevant content area	68	496 (55)	51	489 (43)
2 majoring in at least one relevant content area and in ped. content knowledge	78	497 (55)	102	482 (49)
Gap: Major in content and pedagogy vs. no major	**-5**		**16**	
Feeling ready to teach topics included in the list				
1 Low	55	492 (58)	75	477 (44)
2 Medium	32	502 (55)	29	478 (50)
3 High	69	497 (57)	52	498 (48)
Gap: High readiness vs. low readiness		**5**		**21**

We found the rest of the mathematics teacher variables to be less associated with achievement. The mean group achievement of students taught by mathematics teachers who had a second degree or by teachers who had majored in at least one relevant subject area (e.g., mathematics or mathematics education) was slightly lower (about one tenth of a group standard deviation) than that for students studying in groups taught by teachers with a first or no academic degree, and lower again for those taught by teachers without a major in a relevant subject area. Teachers' feelings of readiness to teach mathematics topics seemed to have almost no effect on student outcomes in mathematics.

The results also revealed a slight advantage in achievement (about one fifth of a group standard deviation) in favor of students studying in groups taught by mathematics teachers who had participated frequently during the last year (more than 10 times) in **content-oriented** professional development activities over students of mathematics teachers who had either participated only once or twice a year in these types of activity or who had not participated at all. However, there was no apparent advantage in achievement—and, in fact, there was a slight disadvantage—for students who had intensively participated in **pedagogically-oriented** professional development activities over students of mathematics teachers who had rarely, if ever, participated in this type of professional development.

The picture regarding the achievement of students taught by science teachers was only partially similar. The effects of ethnic affiliation, gender, and seniority for students studying with science teachers were smaller than those detected in groups studying with mathematics teachers. However, the students studying in groups taught by Hebrew-speaking science teachers attained higher scores than the students in groups taught by Arabic-speaking science teachers, but only by about half a standard deviation of the group's mean distribution of science scores ($SD = .47$). Students in groups taught by female science teachers did better, but only moderately so (about a fifth of a group standard deviation) than students in groups studying with male science teachers.

The effect of seniority was almost the same in the two school subjects. Students studying in groups taught by teachers with more than 15 years of experience gained higher achievement scores than students studying in groups taught by teachers with fewer years of experience (about 0.4 of a group standard deviation).

Students taught by science teachers with a second degree or by teachers who had majored in at least one of five relevant subject areas and in **pedagogy** (science education) had achievement levels that were slightly higher (about one fifth to one third of a group standard deviation) than those of students studying in groups taught by science teachers with a first degree or less. Students taught by teachers who had majored in the field they taught also did better than those students taught by teachers who had not majored in the field they were teaching. This pattern contrasts with the pattern that emerged relative to mathematics teachers. There, having a second degree or majoring in the field of study was negatively associated with student achievement.

Students in groups taught by science teachers who felt highly ready to teach science topics outperformed students in groups taught by science teachers who did not feel the same degree of readiness (about half a group standard deviation).

We also found from our analyses that students in groups taught by science teachers who participated more than 10 times per year in content-oriented professional development activities had an achievement level about half a group standard deviation higher than that for students in groups taught by science teachers who had either minimally participated or had not participated at all in such activities. This effect was more profound among the science teachers than among the mathematics teachers. As was the case with the mathematics teachers, students studying with science teachers who intensively participated in pedagogically-oriented professional development activities had a lower level of achievement than students studying with science teachers who had not participated in such activities. This negative effect was less profound among the science teachers than among the mathematics teachers.

2. Multi-Level Regression Analysis

Table 4 sets out the results of the HLM analyses. The upper part of the table presents estimates of the variance components of each of the following models—the "null" or "basic" model with no explanatory variable, the "student-level" model, and the "full" model, which also included the teacher variables as well as a "slope as outcome" part, where the coefficients of the student-academic-aspiration variable were regressed on all teacher variables. Table 4 also presents the added explained variance to the between-group variance (BGV), in percentages beyond that explained by the student model, and the cumulative percentage of BGV explained, indicating the total explanatory power of the models. The lower part of the table presents the intercept of each regression equation, the estimated coefficients of all predictors, and the standard error of measurement.

In the full models, the coefficients of the teacher variables indicating main effects appeared first. Their second appearance indicated the effect of these variables on the slope of student academic aspiration. The statistically significant coefficients in the table reveal the interaction effect between teacher variables and student variables that can explain the differential effect of teacher variables on the achievement of students who differ in their academic aspirations.

As is evident from the table, the percentage of BGV relative to the total variance in outcomes was higher in mathematics (41.7%) than in science (27.4%), suggesting that grouping in mathematics is more likely to be based on ability than is grouping in science. In general, the explanatory power of the specified model was higher for mathematics than for science. Student variables, in both school subjects, explained only a small portion of the BGV (about 6.2% to 6.5%). The added percentage of the BGV that could be explained once we had included the teacher variables in the model amounted to almost 16 in mathematics but only about 6 in science. Also, the ratio between the added explanatory power of the BGV offered by the teacher variables compared to that offered by the student variables was 2.5:1 (15.5: 6.2)

Table 4: Results of the multilevel regression analysis of mathematics and science outcomes: Average of five plausible values

	Mathematics	*N* = 2,290 students in 156 groups		**Science**	*N* = 2,637 students in 156 groups	
	Null model	*Students*	*Teachers*	*Null model*	*Students*	*Teachers*
Variance components						
• Between-group variance	2,829 (41.7%)	2,653 (45.2%)	2,214 (36.8%)	1848 (27.4%)	1,728 (29.6%)	1,631 (21.4%)
• GHFSG1 Slope	–	43	33	–	20	15
• Within-group variance	3,958	3,167	3,173	4,896	4,085	4,082
• Total variance	6,787	5,863	6,020	6,744	5,833	5,728
• % of BSV explained beyond Model 1			16.5			5.6
• % of added BGV explained			15.5			5.2
• Cumulative % of BGV explained		6.2	21.7		6.5	11.7
Final estimation of student var. main effects						
Intercept	501.2 (4.5)***	417.9 (6.9)***	409.7 (12.9)***	491.7(3.8)***	400.1 (6.3)***	394.0 (15.0)***
Book slope #		5.9 (1.2)***	5.8 (1.2)***		6.9 (1.6)***	7.0 (1.6)***
SCM1 slope #		43.4 (2.8)***	43.3 (2.8)***		39.9 (4.4)***	39.9 (4.4)***
ASPIR slope		9.6 (1.2)***	5.4 (2.3)		11.3 (1.1)***	9.2 (3.5)*
MA			-11.6 (14.4)			10.1 (12.9)
ISRARB			36.2 (15.9)*			17.5 (14.3)
T_SEX			-18.8 (15.0)			-12.5 (15.0)
ZTAUT_M/S			4.2 (5.6)			0.1 (5.1)
ZMRDY_M/S			6.2 (5.7)			8.2 (6.6)
ZMPEDG_M/S			-22.4 (9.0)*			-16.1 (7.6)*
ZGPDIK_M/S			20.4 (9.4)*			8.7 (8.4)
ZNSINFLD			-6.7 (6.4)			1.9 (5.6)

Table 4: Results of the multilevel regression analysis of mathematics and science outcomes: Average of five plausible values (contd.)

	Mathematics *N* = 2,290 pupils in 156 students			**Science** *N* = 2,637 students in 156 groups		
	Null model	*Students*	*Teachers*	*Null model*	*Students*	*Teachers*
Final estimation for GHFSG1 slope						
MA			-1.0 (2.9)			-1.8 (2.5)
ISRARB			1.3 (3.1)			-0.3 (2.7)
T_SEX			7.0 (3.0)*			3.1 (3.1)
ZTAUT_M/S			0.6 (1.2)			0.6 (1.0)
ZMRDY_M/S			-1.1 (1.1)			-2.1 (1.5)
ZMPEDG_M/S			3.7 (1.8)*			3.1 (1.4)*
ZGPDIK_M/S			-3.9 (1.9)*			-1.0 (1.7)
ZNSINFLD			0.7 (1.3)			-0.8 (1.1)

Notes:

BGV = Between-group variance

* = $p \leq .05$, -.08, ** = $p \leq .01$, *** = $p \leq .001$

= fixed effect

in mathematics and 0.8:1 (5.2: 6.5) in science. This finding indicates that the set of teacher variables chosen as measures of quality explained more of the BGV in mathematics than they did in science.

The HLM findings also supported the findings of the breakdown analysis. The relationship between teachers' ethnic affiliation and student achievement was found in both school subjects, although it reached significance only in mathematics. We also found, for both school subjects, a significant negative association between frequent participation in pedagogically-oriented professional development activities and student achievement. We furthermore found a positive main association between frequent participation in content-oriented professional development activities and achievement, but this relationship was significant only for mathematics. Teachers' feelings of readiness to teach science or mathematics showed a positive but non-significant relationship with student achievement.

Many of the variables describing teachers' qualifications did not yield similar relationships with student achievement in both school subjects. Mathematics teachers' advanced academic degrees and teachers having a major in the field of teaching seemed to have a negative association with students' outcomes in mathematics. However, in the case of science teachers, these variables showed a positive relationship with students' science outcomes. Teachers' years of experience had a positive, although not significant, relationship with student achievement in the case of mathematics teachers, but there was almost no such relationship in respect of science teachers. Although these relationships were not significant, the contrast is clear.

Teacher's gender was positively associated with student achievement in favor of female teachers, according to the breakdown analysis, and also when fitted as a single predictor in the multilevel analysis. However, when fitted with other teacher-level variables, teacher gender appeared to have a negative relationship with student outcomes, albeit not a statistically significant one, indicating that its effect was probably mediated by other teacher variables.

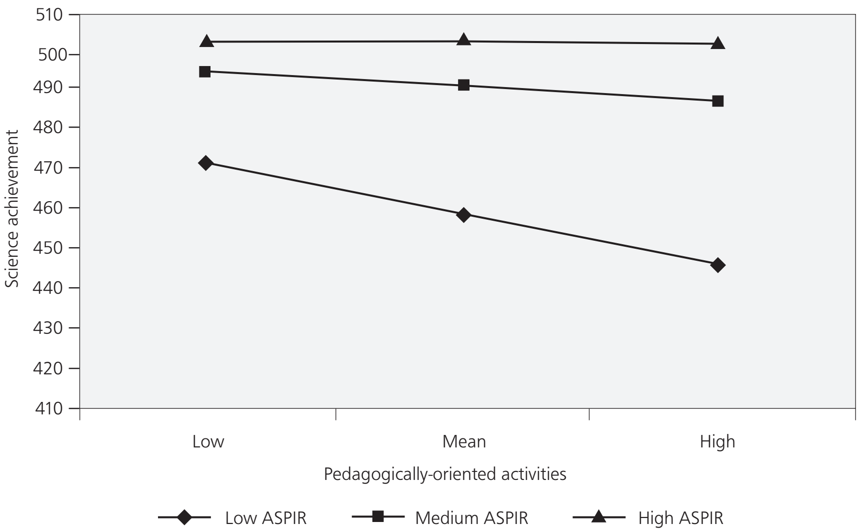
The analyses thus reveal that some of the teacher variables, such as having second degrees, majoring in the field of teaching, participating in content-oriented professional development activities, and seniority, which are taken as indicators of quality and considered desirable teacher qualifications and criteria for remuneration and reward systems, appear not to have a consistent positive association with student achievement in different school subjects. These associations, as will be elaborated in the next section, were also inconsistent for different students.

3. Interactions between Teacher Variables and Student Variables

Interesting significant interactions were found between teachers' participation in professional development activities and students' academic aspirations. The negative relationship between frequent participation in pedagogically-oriented professional activities and student achievement and the positive relationship between frequent participation in content-oriented professional development activities and student achievement were more profound for students with low academic aspirations. This pattern was significant in the case of mathematics but less clear in the case of science.

Plotting predicted mathematics and science group mean scores as a function of frequent participation of the students' teachers (low [minimal], mean, and high [maximal]) in pedagogically-oriented professional development activities (Figures 1 and 2) and plotting predicted mathematics group mean scores dependent on the frequent participation of the students' teachers in content-oriented professional development activities (Figure 3) for three values of students' academic aspiration (low, mean, and high) helps us to visualize and understand the interaction effect between students' characteristics and their teacher's attributes. The plots chosen refer only to the significant interactions found.

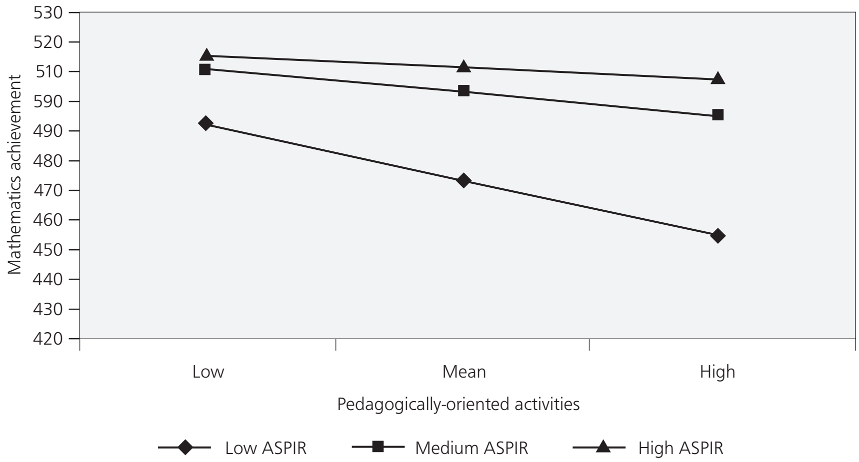
Figure 1: The relationship of frequent participation in pedagogically-oriented professional development activities with the achievement of students with different levels of academic aspiration (ASPIR), science



	Low	Mean	High
Low ASPIR	472.350	459.440	446.530
Medium ASPIR	495.151	491.311	487.471
High ASPIR	504.018	503.705	503.393

When students who score low on the scale of academic aspirations are taught by mathematics or science teachers who frequently participate in pedagogically-oriented professional development activities, they achieve less than when taught by teachers who do not, or only rarely, participate in such activities. This negative association is weak for students with mean academic aspirations and even weaker for students with high academic aspirations. Thus, intensity of participation of both mathematics and science teachers in pedagogically-oriented professional development activities seems to increase the achievement gap between high and low academically-motivated students.

Figure 2: The relationship of frequent participation in pedagogically-oriented professional development activities with the achievement of students with different levels of academic aspiration (ASPIR), mathematics

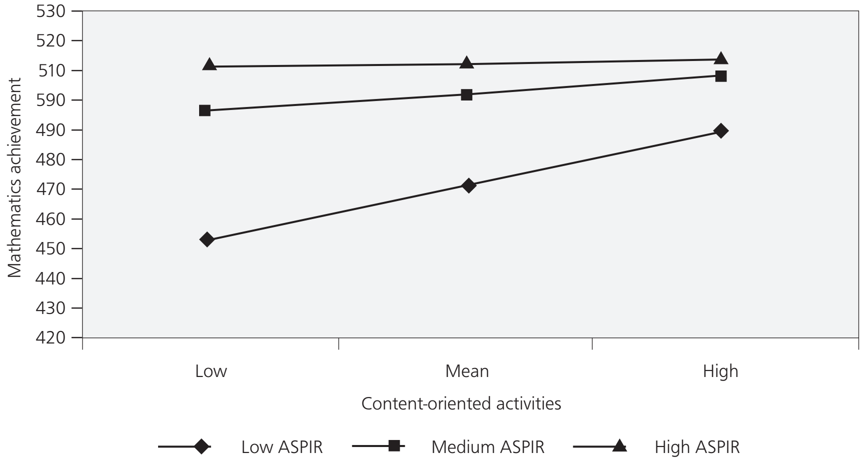


	Low	Mean	High
Low ASPIR	492.850	474.135	455.421
Medium ASPIR	509.915	502.178	494.440
High ASPIR	516.147	512.418	508.689

In contrast, mathematics teachers' frequent participation in content-oriented professional development activities has the opposite association with achievement of students with different academic aspirations. Teachers' frequent participation in content-oriented professional development activity was positively related to achievement of students with low academic aspirations, had a small positive association with the achievement of students with medium academic aspirations, and had almost no association with the achievement of students with medium or high academic aspirations. Thus, intensity of participation in content-oriented professional development activities narrows the gap between high and low academically-motivated students.

A positive interaction effect (significant only in the case of mathematics) was found between level of students' academic aspirations and another teacher variable—gender: achievement of students with high academic aspirations was fostered by female teachers. This interaction effect buffers the negative main association between teacher's gender and student achievement found in the full multilevel regression model.

Figure 3: The relationship of frequent participation in content-oriented professional development activities with the achievement of students with different levels of academic aspirations (ASPIR), mathematics



	Low	Mean	High
Low ASPIR	450.679	467.38	484.087
Medium ASPIR	490.112	495.426	500.739
High ASPIR	504.512	505.666	506.820

DISCUSSION

The results of these analyses are disappointing in terms of providing a clearer indication not only for policymakers in Israel but elsewhere on how best to improve the recruitment, education, and remuneration of teachers. Most of the teacher variables commonly regarded as desired qualifications or as indicators of quality, such as advanced academic degrees, securing a major in the subject being taught, and years of teaching experience that were adopted as reform targets in teacher policy and as criteria for remuneration, had only marginal and statistically non-significant positive relationships with student achievement. These associations, moreover, were inconsistent across the two subject areas and varied according to different student groupings. In interpreting these results, we begin with substantive explanations and later bring in others that are more methodological in nature.

Two variables had opposite (although not significant) influences in the two subject areas. The first was *having an advanced degree* and the second was *majoring in the field of teaching*. Both variables were *positively* but not significantly associated with achievement in science. However, they had a negative (but again not significant) association with achievement in mathematics, a result contrary to other findings (Goldhaber & Brewer, 1997, 2000; Rowan et al., 1997).

Rowan, Correnti, and Miller (2002) reported similar findings relative to teachers having advanced degrees. They found that neither teachers' mathematics certification nor teachers having post-graduate degrees in mathematics improved the achievement of their students (who, in this case, were in the upper grades of schools). Struggling to interpret their findings, the authors suggested that advanced academic training in mathematics somehow interferes with effective teaching, either because it limits time spent on professional pedagogical training or because it produces teachers who somehow cannot simplify and clarify their advanced understanding of mathematics for school students.

A possible explanation for the differential effects on student achievement brought about by the students' mathematics teachers or science teachers having advanced qualifications, a major in the subject of teaching, and a good number of years of experience could be found in the fundamental differences between these two subject areas. Because science is a constantly developing domain, science teachers with advanced and frequently updated education in their field of teaching may have a greater advantage. Mathematics, however, is regarded as a classic domain of knowledge, which means that what is taught in schools tends to be basic, and so teaching experience is more important than updated knowledge of the domain. Indeed, the positive main effect of experience in teaching mathematics confirms this.

The fact that the positive main association between frequent participation in content-oriented professional development activities and student achievement reached significance only for mathematics and the fact that the negative association between participating in pedagogically-oriented professional activities and student achievement reached significance in both mathematics and science could point to initial differences between teachers—those who prefer to upgrade their content knowledge and those who prefer to learn new pedagogical tools. Preference for content might indicate higher teacher qualifications.

The opposite associations were more profound for students with low academic aspirations. Teachers' intensive participation in pedagogically-oriented professional development activities appeared to contribute to outcome inequality among students with different academic aspirations, which usually is associated with level of socioeconomic status. The opposite occurred with respect to teachers' participation in content-oriented professional development activities, in that it appeared to narrow the achievement gap between students with different academic aspirations.

The relationships between the two types of professional development opportunities and the achievement of students differing in terms of their academic aspirations can be interpreted in the following way: highly motivated students (who are usually more able) are less sensitive to teachers' input; in many cases, they can manage on their own. The problems that students with low academic aspirations (usually less able) face require more focus on the content teachers teach, and this cannot be replaced by instructional strategies.

The findings in this study on the positive relationship between content-focused professional development activities and student achievement, at least for mathematics teachers, support policy interventions aimed at providing more such opportunities. The findings in this study on the negative association between participating in pedagogically-focused professional development activities (which are very popular in Israel) and student achievement suggest we should investigate what actually happens in such in-service training and how the knowledge gained in these courses is translated into action in classrooms. Another conclusion from our study is that policies related to recommended reforms in teacher education, rather than being broadly applicable and generic, need to account for the manner in which these effects vary across subject areas and across different types of students.

We also wish to offer some methodological explanations for the findings regarding the ineffectiveness of many teacher variables commonly regarded as relevant. One possible explanation could be the lack of variability in some teacher variables, which results in underestimation of their correlation with achievement. This is the case, for instance, with level of certification, which (according to the TIMSS 2003 data for Israel) did not correlate with student achievement because all teachers in Israel are fully certified. The non-linear effects or the low threshold effect that some teacher variables exhibit is another explanation. For instance, if a teacher has a first academic degree in mathematics, he or she is much more likely to have students gaining the higher scores on achievement tests than if he or she has only a non-university education. However, teachers' further education—for a second academic degree in mathematics—had no relationship with student achievement.

Multicollinearity between teacher variables can also mask the relationship between some of these variables and student achievement. Wayne and Youngs (2003) used the following example to illustrate this problem: analysis may show that having a Master's degree matters, but it is likely that teachers with a Master's degree also have more teaching experience. Thus, what appears attributable to a Master's degree may instead be attributable to experience. Studies, such as this present one, that assess multiple teacher characteristics simultaneously are therefore more reliable. Wayne and Youngs (2003) also posit that "failing to reject the null hypothesis [i.e., that the relationship can occur by chance] does not rule out a relationship. Perhaps the study's sample size was too small—the measurement error too great—to provide statistical confirmation. These issues prevent rejection of the possibility that differences occurred randomly, even if the qualification being studied does influence student achievement" (p. 93). For these reasons, Wayne and Youngs sum up by saying, "Studies may establish that an observed teacher quality indicator matters but cannot convincingly show that an observed teacher quality indicator does not matter" (p. 93).

These explanations provide cautions for policymakers who tend to conclude that teachers' formal education, certification, and procurement of a major in the subject of teaching are not indicators of quality. The teacher qualifications covered by this present study should therefore not be dismissed. However, there is also an obvious need to look for other important and more sensitive indicators of quality. Although not always easy to measure, attributes such as enthusiasm, motivation, charisma, ability to convey ideas clearly, and verbal ability should be added to the models that try to associate teachers' required qualifications with students' achievement outcomes.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* London, Newbury Park, New Delhi: Sage.

Alexander, R. (2000). *Culture and pedagogy: International comparisons in primary education*. Oxford: Basil Blackwell.

Andrew, M. D. (1990). Difference between graduates of 4-year and 5-year teacher preparation programs. *Journal of Teacher Education, 41*, 45–51.

Angrist, J. D., & Lavy, V. (2001). Does teacher training affect pupil learning? Evidence from matched pairings in Jerusalem public schools. *Journal of Labor Economics, 19*(2), 343–369.

Ariav, T., Olshtain, E., Alon, B., Back, S., Grienfeld, N., & Libman, Z. (2006). *Committee for guidelines for teacher education: Report on programs in higher education institutions in Israel*. Jerusalem: Council for Higher Education (in Hebrew).

Ashton, P., & Crocker, L. (1987). Systematic study of planned variation: The essential focus of teacher education reform. *Journal of Teacher Education*, May–June, 2–8.

Ballou, D., & Podgursky, M. (1997). Reforming teacher training and recruitment: A critical appraisal of the recommendations of the National Commission on Teaching and America's Future. *Government Unit Review, 17*(4).

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

Ballou, D., & Podgursky, M. (2000). Reforming teacher preparation and licensing: What is the evidence? *Teacher College Records, 102*(1), 5–22.

Berliner, D. C. (1987). Simple views of effective teaching and a simple theory of classroom instruction. In D. C. Berliner & B. Rosenshine (Eds.), *Talks to teachers* (pp. 93–110). New York: Random House.

Berliner, D. C. (2005). The near impossibility of testing for teacher quality. *Journal of Teacher Education, 56*(3), 205–213.

Betts, J., Zau, A., & Rice, L. (2003). *Determinants of student achievement: New evidence from San Diego*. San Francisco, CA: Public Policy Institute of California.

Braun, H. J. (2005). *Using student progress to evaluate teachers: A primer on value-added models*. Princeton, NJ: Educational Testing Service. Available online at: http://www.ets.org/research/pic.

Brown, C. A., Smith, M. S., & Stein, M. K. (1995). *Linking teacher support to enhanced classroom instruction*. Paper presented at the annual conference of the American Educational Research Association, New York.

Bryk, S., & Hermanson, K. L. (1993). Educational indicator systems and observations on their structure, interpretation, and use. *Review of Research in Education, 19*, 451–484.

Cochran-Smith, M. (2001). The outcomes question in teacher education. *Teaching and Teacher Education, 17*, 527–546.

Cohen, D. K., & Hill, H. (1977). *Instructional policy and classroom performance: The mathematics reform in California*. Paper presented at the annual conference of the American Educational Research Association, Chicago, IL.

Darling-Hammond, L. (1992). Educational indicators and enlightened policy. *Educational Policy*, September, 235–265.

Darling-Hamond, L. (1994). Policy uses and indicators in making education count. *Developing and using international indicators* (pp. 357–378). Paris: Organisation for Economic Co-operation and Development.

Darling-Hammond, L. (1998). Teachers and teaching: Testing hypotheses from a national commission report. *Educational Researcher, 27*(1), 5–15.

Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence* (Research Report R-99-1). Washington DC: Center for the Study of Teaching and Policy, University of Washington.

Darling-Hammond, L. (2000a). How teacher education matters. *Journal of Teacher Education, 51*(3), 166–173.

Darling-Hammond, L. (2000b). Reforming teacher preparation and licensing: Continuing the debate. *Teachers College Record, 102*(1), 5–27.

Darling-Hammond, L., Berry, B., & Thorenson, A. (2001). 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.

Dovrat Committee (Israel) Task Force for the Advancement of Education in Israel. (2005). *National plan for education*. Jerusalem: Government of Israel.

Ehrenberg, R. G., & Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from high school and beyond. *Economics of Education Review, 13*, 1–17.

Everston, C. W., Hawley, W., & Zlotnik, M. (1985). Making a difference in educational quality through teacher education. *Journal of Teacher Education, 36*(3), 2–12.

Fenstermacher, G. D., & Richardson, V. (2005). On making determinations of quality in teaching. *Teachers College Record, 107*(1), 186–213.

Ferguson, R. F., & Ladd, H. F. (1996). How and why money matters: An analysis of Alabama schools. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 265–298). Washington, DC: Brooking Institution.

Ferguson, P., & Womack, S. T. (1993). The impact of subject matter and education coursework on teaching performance. *Journal of Teacher Education, 44*(1), 155–163.

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

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*(2), 129–145.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66*(3), 361–396.

Guyton, E., & Farokhi, F. (1987). Relationships among academic performance, basic skills, subject-matter knowledge, and teaching skills of teaching education graduates. *Journal of Teacher Education, 38*, 37–42.

Harris, D. N., & Sass, T. R. (2007). *Teacher training, teacher quality and student achievement.* Unpublished manuscript, Grant R305M04121 from US Department of Education.

Ingersoll, R. M. (2003). *Out of field teaching and the limits of teachers' policy: A research report.* Washington DC: Center for the Study of Teaching and Police, University of Washington.

Jacob, B. A., & Lefgren, L. (2004). The impact of teacher training on student achievement: Quasi-experimental evidence from school reform efforts in Chicago. *Journal of Human Resources, 39*(1), 50–79.

Kellaghan, T. (1996). IEA studies and educational policy. *Assessment in Education: Principles, Policy, and Practice, 3*(3), 143–160.

Kiesling, H. J. (1984). Assignment practices and the relationship of instructional time to the reading performance of elementary school children. *Economics of Education Review, 3*, 341–350.

Klitgaard, R. H., & Hall, G. R. (1974). Are there unusually effective schools? *Journal of Human Resources, 10*(3), 40–106.

Maagan, D. (2007). *Teaching staff survey in elementary schools (2005/6)*. Jerusalem: Israel Central Bureau of Statistics (in Hebrew)

Martin, M. D., Mullis, I. V. S., & Chrostowski, S. J. (2004). *TIMSS 2003 technical report.* Boston, MA: Boston College.

McCaffrey, D. F., Lockwood, J. R., Koreetz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics, 29*(1), 67–101.

Monk, D. H. (1994). Subject area preparation of secondary mathematics and science: Teachers and student achievement. *Economics of Education Review, 1*3, 125–145.

Monk, D. H., & King, J. (1994). Multilevel teacher resource effects on pupil performance in secondary mathematics and science. In R. G. Ehrenberg (Ed.), *Contemporary policy issues: Choices and consequences in education* (pp. 29–58). Ithaca, NY: ILR Press.

Murnane, R. J. (1996). Staffing the nation's schools with skilled teachers. In E. Hanushek & D. Jorgenson (Eds.), *Improving America's schools: The role of incentives* (pp. 241–256). Washington, DC: National Research Council, National Academy.

Murnane, R. J., & Phillips, B. (1981). Learning by doing, vintage, and selection: Three pieces of the puzzle relating teaching experience and teaching performance. *Economics of Education Review, 1*(4), 453–465.

National Commission on Teaching and America's Future (NCTAF). (1996). *What matters most: Teaching for America's future*. New York: Author.

Organisation for Economic Co-operation and Development (OECD). (2005). *Attracting, developing and retaining effective teachers. Final report: Teachers matter*. Paris: OECD Publishing.

Raudenbush, S. W., & Bryk, A. S. (1986). A hierarchical linear model: A review. *Sociology of Education, 59*, 1–17.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. E., & Congdon, R. (2000). *Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2000). *Teachers, schools and academic achievement* (working paper 6691, revised). Cambridge, MA: National Bureau of Economic Research.

Rowan, B., Chiang, F. S., & Miller, R. J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education, 70,* 256–285.

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

Sanders, W. L. (2000). Value-added assessment from student achievement data: Opportunities and hurdles. *Journal of Personnel Evaluation in Education, 14*(4), 329–339.

Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future academic achievement.* Knoxville, TN: Value-Added Research and Assessment Center, University of Tennessee.

Santiago, P. (2002). *Teacher demand and supply: Improving teaching quality and addressing teacher shortage* (OECD Education Working Paper No. 1). Paris: OECD Head of Publication Service.

Shorrocks-Taylor, D. (2000). International comparisons of pupil performance: An introduction and discussion. In D. Shorrocks-Taylor & E.W. Jenkins (Eds.), *Learning from others* (pp. 13–28). Dordrecht: Kluwer.

United States Department of Education. (2002). *Meeting the highly qualified teachers' challenge*. Washington, DC: Author.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89–122.

Wenglinsky, H. (2000). *How teaching matters: Bringing the classroom back into discussion of teacher quality policy*. Princeton, NJ: Educational Testing Services.

Wiley, D., & Yoon, B. (1995). Teacher reports of opportunity to learn analyses of the 1993 California learning assessment system. *Educational Evaluation and Policy Analysis, 17*(3), 355–370.

Wilson, S. M., Darling-Hammond, L., & Berry, B. (2001). Steady work: The story of Connecticut's reform. *American Educator, 25*(3), 34–39, 48.

Wilson, S. M., Floden, R. E., & Ferrini-Mundy, J. (2001). *Teacher preparation research: Current knowledge, gaps, and recommendations. A research report.* Washington DC: Center for the Study of Teaching and Policy, in collaboration with Michigan State University, University of Washington.