STUDY DESIGN FOR EVALUATING THE YALE-NEW HAVEN TEACHERS INSTITUTE

Ellen Eliason Kisker
Twin Peaks Partners, LLC

July 2006

CONTENTS

Introduction....1
Research Questions....2
Key Features of the Study Design....2

Sample Design....4
Study Period....4
Study Populations....4
Teachers....4
Students....5
Target Sample Sizes and Power Estimates....5

Data Collection Plan....6
Data Sources....7

Data Analysis Plan....8
Data Quality Assessment....8
Descriptive Analyses....8
Multivariate Analyses to Estimate the Effects of Institute Participation on Teacher Outcomes....10
Subgroup Analyses....11
Multivariate Analyses to Estimate the Effects of Institute Participation on Student Outcomes....12
Model When Student Data Cannot Be Linked to Teacher Data....12
Model When Student Data Can Be Linked to Teacher Data....14
Subgroup Analyses....16
Models to Explore Variations in Intensity and Timing of Teachers Institute Participation....16
Interpreting the Analysis Results....17

Notes and References....18

Appendix....19
Tables....19
School Records Data Collection....23
Student-Level Data....23
Teacher-Level Data....23
School Data....24

INTRODUCTION

Teachers Institute seminars, led by university faculty, offer public school teachers opportunities to increase their knowledge and skills in what they teach. Topics are suggested by teachers through teacher Representatives who are responsible for canvassing their colleagues and soliciting ideas for seminar topics. The seminars, which meet over a period of no less than three months, focus on increasing content knowledge, improving writing and oral presentation skills, facilitating interactions between teachers in different grades and schools, and supporting teachers in developing substantial curriculum units. Teachers receive a stipend in recognition of the time they are giving to improve teaching and to cover their expenses of participation.

Research suggests that teacher quality is an important factor in student performance. Teachers Institute seminars aim to improve student performance by improving teacher quality. Quality teachers are defined as those who "really know their subjects, not just 'how' to teach; ... have good basic writing, math, and oral presentation skills; ... expect their students to achieve; ... are enthusiastic about teaching; and... can motivate even highly disadvantaged students to learn." Thus, Teachers Institute seminars are designed to improve teacher quality in the following ways:

- By focusing on content areas that teachers themselves have identified as areas they want to know more about for their teaching, participation in Teachers Institute seminars is expected to help teachers deepen their knowledge of the subjects they teach.
- By guiding each teacher in preparing a substantial curriculum unit based on research on a topic chosen by the teacher and informed by the seminar, and by encouraging discussions of these units among teachers in the seminar, participation in Teachers Institute seminars is expected to improve teachers' writing and oral presentation skills.
- Because teachers will later use the curriculum units they developed themselves, participation in Teachers Institute seminars is expected to provide teachers with curriculum materials that they are more strongly motivated to teach and that more effectively motivate students to learn.
- Through their service as teacher Representatives or seminar Coordinators, teachers are expected to develop leadership capabilities that may lead to further career development.
- Because teachers attend seminars with teachers from other schools and disciplines, Teachers Institute seminar participation is expected to promote the development of teacher networks that may offer new opportunities for learning about what other teachers are doing in their classrooms and help establish supportive peer relationships among teachers.

RESEARCH QUESTIONS

This retrospective evaluation is designed to provide evidence of the effectiveness of the Yale-New Haven Teachers Institute seminars. The overarching research questions to be addressed by the study are:

1. What impacts does teacher participation in Institute seminars have on
   a. student performance?
   b. teacher careers?
2. Do impacts vary among teachers with different characteristics?
3. How do teachers perceive that Institute participation has affected them? How do they perceive it has affected their students?

KEY FEATURES OF THE STUDY DESIGN

Ideally, the impacts of the Teachers Institute would be assessed using a randomized control trial in which teachers would be randomly assigned to a group that could participate in the Teachers Institute seminars or a group that could not. Such a research design would provide the most irrefutable evidence possible for the effectiveness of the Institute in improving the targeted outcomes of students and teachers.

Certain features of the Teachers Institute seminars make an experimental evaluation design difficult. By design, the Teachers Institute plans each year to offer a sufficient number of seminars to meet the demand for them. More importantly, teacher choice is an important element of the Teachers Institute. Specifically, teachers choose their seminars, and in cases of oversubscription to a seminar, committees of teacher Representatives make decisions about which teachers should be admitted to which seminars.

Beyond these challenges, the proposed study of the effects of the Yale-New Haven Teachers Institute seminars will be retrospective, relying primarily on school records data, possibly supplemented by a survey of teachers. It is not possible, of course, to go back in time to randomly assign teachers to conditions.

The best alternative design is a nonexperimental comparison group design in which outcomes for teachers participating in Institute seminars (Institute Fellows) and their students are compared to outcomes for a comparison group of nonparticipating teachers and their students. Ideally, comparison teachers would be selected from the same population as the Institute Fellows and would be similar to them in terms of key characteristics, including pre-intervention levels of the outcome variables if possible. With this design, the possibility of selection bias cannot be entirely eliminated; however, the estimates of the effects of Teachers Institute participation will be better than those derived from comparisons of participating teachers and all other teachers.

The estimates of program effects obtained from the outcomes analyses will be complemented by information collected from the participating teachers about their percep-

tions of the effects seminar participation has had on their teaching skills, resources, and attitudes, as well as their students' responses to the new curriculum units.

This design document focuses on outcomes analyses based on school records data (and potentially new survey data collected from treatment and comparison teachers). Other descriptive and qualitative information about the Teachers Institute seminars and the New Haven Public Schools and teachers may be important for interpreting the results of the outcomes analyses.

SAMPLE DESIGN

STUDY PERIOD

The analyses will focus on outcomes during a five-year period, from 1999-2000 to 2004-2005.[1] Some outcomes, such as teacher retention, pertain to a period of time, and the sample should include all teachers who began the period. Thus, the sample should include all eligible teachers who taught in New Haven Public Schools during the period from 1998-1999 to 2004-2005.

STUDY POPULATIONS

Teachers

The relevant population of teachers for the evaluation includes all New Haven Public School teachers who were identified as eligible for Teachers Institute seminars during the period from 1998-1999 to 2004-2005. Eligible teachers are teachers of grades or subjects qualifying them for potential seminar participation who were identified by teacher Representatives during the study period.

For the main analyses, the treatment group will consist of the eligible teachers during the study period who had ever participated in at least one Teachers Institute seminar. The group of approximately 410 New Haven teachers who ever participated in Teachers Institute seminars can be identified from Teachers Institute records assembled in October of each school year, updated in 2004 to reflect actual Institute seminar participation during the 2004-2005 school year.

The comparison group will include the remaining eligible teachers during the study period who had never participated in Teachers Institute seminars. Because the comparison teachers are from the same population as treatment teachers (that is, from the New Haven Public Schools) and already include only teachers identified as potential participants by teacher Representatives, further matching using propensity scoring procedures is not likely to be necessary. When the data for the teachers are obtained, the comparability of treatment and comparison teachers according to their observed characteristics can be examined, and if substantial differences are evident, propensity score matching to select a more similar subset of teachers for the comparison group can be considered.

An alternative approach to forming the treatment and comparison groups will result in lower power to detect significant effects but may be less affected by selection bias. The groups would be formed from the 410 teachers who ever participated in a Teachers Institute seminar. For each school year, the treatment group would include all teachers who have participated in a Teachers Institute seminar in the current or

prior years, and the comparison group would include teachers who participated in a Teachers Institute seminar in later years but had not yet done so in the current year. Under this approach, the comparison group is likely to be more similar to the treatment group in unobservable characteristics, because these teachers eventually decide to participate in a Teachers Institute seminar. This approach can be applied until later school years when a large proportion of the teachers have participated and too few teachers remain in the comparison group.

Students

Ideally, the students to be included in analyses of student outcomes in each year would include the subset of students taught by treatment and/or comparison teachers during the period since the last measurement of the outcome. Identifying this student population would require data linking students to the teachers they had.

If the available school records data do not contain sufficient information for linking students to teachers, then the student population necessarily includes the entire population of students in schools where eligible teachers worked.

TARGET SAMPLE SIZES AND POWER ESTIMATES

For analyses based on records data, there is no reason not to use data for all teachers on the list of potential seminar participants. If a survey can be conducted to collect new data, it may be necessary for cost or practical reasons to select a random subsample of the potential seminar participants for survey data collection. Table 1 indicates the level of precision that can be expected for teacher outcomes with comparison samples of various sizes.

Table 2 indicates the level of precision that can be expected for student outcomes if student data cannot be linked to data for the teachers the student had. Table 3 estimates the level of precision that can be expected for student outcomes if student data can be linked to data for the teachers the student had.

DATA COLLECTION PLAN

The data needed to address the research questions are summarized in Table 4. To address the research questions, measures of both teacher and student outcomes must be developed. The teacher outcomes available from school records that will be examined include:[2]

- **Retention** — whether or not a teacher teaching in New Haven Public Schools during the previous school year was still teaching in the New Haven Public Schools two years later.[3] Additional measures of retention (such as retention in the current school) can also be explored.
- **Promotion** to a leadership position — whether or not a teacher teaching in New Haven Public Schools during the previous school year was promoted to a leadership position
- **Awards** — whether or not a teacher teaching in New Haven Public Schools during the previous school year received an award or recognition
- **Attendance** — number of days present

Additional teacher outcomes may be available if a teacher survey is conducted. These outcomes may include measures of:

- **Motivation** — degree of motivation to teach
- **Communication** — extent of communication and collaboration with other teachers in the district
- **Teaching Practices** — strategies used to motivate students to learn and curricula used
- **Career Plans** — degree of commitment to teaching in New Haven Public Schools

The student outcomes that will be examined include:

- **Achievement Test Scores** — scores on standardized tests in key subject areas, as well as indicators of low scores (for example, scoring more than one standard deviation below the mean) and high scores (more than one standard deviation above the mean). Raw scores will be converted to Normal Curve Equivalents for the analyses.
- **Grades** — for grades in which students receive letter grades, the student's grade point average
- **Awards and Recognition** — whether or not a student received an award or recognition during the last year
- **Attendance** — number of days present
- **Grade Retention** — whether or not a student was held back in his or her current grade (this may also be examined in terms of graduation to the next grade or school level)

In addition to the outcomes data, data on teacher and student background characteristics are needed. These variables are important for establishing the comparability of treatment and comparison teachers and students initially and will be used as covariates in the statistical models.

DATA SOURCES

The primary data for these analyses are the school records data for teachers and students that have been obtained or will be requested from the New Haven Public Schools for previous school years in the study period. Appendix B describes the range of data to be considered in the data request to the school district.

These data could be supplemented by survey data collected from treatment and comparison teachers (most likely, a subsample of comparison teachers). The survey could collect additional retrospective data on teacher characteristics that could influence outcomes, as well as measure additional outcomes, such as teacher practices and strategies for motivating students, curricula used, and communication with and support from other teachers during a recent period. Because the ability of teachers to recall and accurately report on changing characteristics and practices is likely to be limited, the survey would necessarily need to focus on eligible teachers teaching in the district during the most recent school years (2005-2006, and maybe also 2004-2005).

The survey mode and procedures would depend on the resources available. The least expensive approach to conducting the survey would be to distribute self-administered questionnaires to teachers in school or by mail, asking them to return the completed survey by a deadline. To achieve an acceptable return rate of the questionnaires, incentives would need to be offered to teachers for returning a completed questionnaire. If sufficient resources are available, calling teachers who do not return a completed questionnaire to complete the questionnaire with them by telephone could enhance the survey response rates and may be critical for obtaining a reasonable response rate (at least 70 percent).

Several recent reviews have examined research comparing the results of experimental analyses and nonexperimental comparison group methods using the same data, to identify nonexperimental methods that yield results with the smallest biases. These reviews suggest that simply comparing outcomes of the treatment and comparison groups, controlling for other differences between the two groups in a regression model, often leads to impact estimates with the smallest bias (Bloom, Michalopoulos, and Hill 2005; Glazerman, Levy, and Myers 2002). Thus, the analytic approach for the Teachers Institute is based on such models. Because the data will be nested (students nested within classrooms nested within schools), the analyses will be conducted using hierarchical linear models that take into account the structure of the data.

DATA ANALYSIS PLAN

DATA QUALITY ASSESSMENT

Prior to developing analytic models, it is important to assess the quality of the data. These analyses should assess the completeness of the data and examine the extent of data problems. Analyses of the completeness of the data will assess whether all teachers/students are represented in the data and whether all data items are available for each teacher/student. If significant amounts of data are missing, then methods for imputing data and analytic methods that can handle missing data need to be considered.

Because the sample of teachers will be developed as described above in the discussion of the sample, a comparison of a list of teachers represented in the dataset with the list of teachers in the sample can provide an estimate of the completeness of the teacher data. Determining whether all students are represented in the dataset may be more challenging, requiring comparisons of the number of student data records with enrollment numbers by school (and by class, if student data can be linked with teacher data).

It is also important to examine the data for inaccuracies that could affect the analytic results. For example, are there inconsistencies in the values of related items in the data, or are there values that are out of range and clearly erroneous? All data problems should be examined and reasonable solutions considered, including the possibility of consulting district staff to resolve problems with school records data.

DESCRIPTIVE ANALYSES

Before estimating multivariate models, it is useful to conduct descriptive analyses to provide context for and suggest possible refinements to the more complex analyses discussed in the next section. These analyses may also reveal further data issues that should be resolved before more complex analyses are conducted.

Descriptive analyses of aggregate trends in program participation and outcomes provide an important starting point for more complex multivariate analyses designed to isolate, insofar as possible, the effects of Teachers Institute seminar participation on teacher and student outcomes. Useful school-level descriptive analyses include analyses of:

- Average student achievement and attendance outcomes by grade and year
- Student mobility between schools and into and out of the school district
- Average teacher retention, promotion, and attendance rates by year and during the full study period

- Average proportion of teachers who have participated in Teachers Institute seminars, average proportion of teachers who have participated in multiple seminars, and average proportion of teachers who have participated in a seminar within the last 3 years, by year

An examination of trends in these school averages and correlations between the concentration of Fellows and teacher and student outcomes will provide a broad-brush look at what has been happening during the study period. If the correlations are not positive, this is not necessarily an indication that Teachers Institute seminar participation has not had a positive influence. It may indicate that other factors, such as changes in the composition or distribution of students, may have had a greater influence, and that these need to be considered in the multivariate analyses.

For teachers, useful descriptive analyses include:

- Analyses of the characteristics and previous outcomes of Teachers Institute Fellows and comparison teachers
- Analyses of patterns of seminar participation by treatment group teachers over time
- Analyses of trends in outcomes of treatment and comparison teachers over time, including retention, promotions, awards, and attendance
- Analyses of patterns of participation and outcomes over time for key subgroups of teachers and schools

Comparing the characteristics and previous outcomes of Fellows and comparison teachers is important for assessing the initial equivalency of the two groups. Any observed initial differences between the treatment and comparison groups should be included in the statistical models. If there are substantial initial differences between the groups, creating a better matched comparison group of teachers through propensity score matching should be explored. The comparisons may provide insights into potential biases in comparisons of outcomes of treatment and comparison teachers that may remain even after observed differences are controlled.

Analyses of patterns of seminar participation over time may be helpful for interpreting patterns of effects of being a Fellow over time. For example, if during the study period a significant proportion of Fellows participated in additional seminars and these reinforce and strengthen teaching further, then that might account for any observed growth in the effects of being a Fellow over time.

For students, useful descriptive analyses include:

- Analyses of demographic characteristics of students of treatment and comparison teachers by year, if the available data permit student and teacher data to be linked (again, to understand any observed differences between the treatment and comparison groups that may be useful for interpreting estimates of program effects on students and assessing potential biases in these estimates)

- Analyses of patterns of student exposure to treatment teachers over time
- Analyses of patterns of outcomes of treatment and comparison students by year
- Analyses of patterns of student achievement and attendance outcomes over time for key subgroups of students, teachers, and schools

MULTIVARIATE ANALYSES TO ESTIMATE THE EFFECTS OF INSTITUTE PARTICIPATION ON TEACHER OUTCOMES

Because the study design is not an experimental design, simple comparisons of outcomes for Fellows and comparison teachers cannot be relied on to provide unbiased estimates of the effects of participation in Teachers Institute seminars. To minimize the potential biases in these estimates, it is important to conduct multivariate analyses that control for as many other observed differences among teachers as possible.

Because data for all eligible teachers in all schools in the district will be included in the analyses and the results will not be generalized beyond New Haven Public Schools, a fixed effects model is appropriate for assessing the influence of being a Teachers Institute Fellow on teachers' outcomes in New Haven, controlling for other observed differences between teachers and schools. In this model, the outcome of teacher j in school s, Y_{js}, is a function of the teacher's total years of teaching experience, gender, minority racial/ethnic status, number of years teaching at the school, subject(s) taught, grade level(s) taught, Teachers Institute seminar participation, the concentration of Fellows in the school, and the school:[4]

$$Y_j = b_0 + b_1GENDER_j + b_2MINORITY_j + b_3DEGREE_j + b_4TEACHYRS_j + b_5SCHDUR_j + b_6SUBJECT_j + b_7GRADE_j + b_8FELLOW_j + b_9PCTFEL_j + b_{10}EXP_j + b_{11}AVGCLASS_j + b_{12}PCTNLSP_j + b_{13}PCTMIN_j + e_j$$

where:

Y_j = the outcome of teacher j at the end of the school year

$GENDER_j$ is the teacher's gender

$MINORITY_j$ is an indicator variable equal to 1 if the teacher is a member of a minority ethnic group

$DEGREE_j$ is an indicator variable equal to 1 if the teacher has attained a master's degree or higher

$TEACHYRS_j$ is the total number of years teacher j in school k had been teaching as of the beginning of the school year

$SCHDUR_j$ is the number of years teacher j had taught at his/her current school by the beginning of the school year

$SUBJECT_j$ is a set of dummy variables indicating the subjects that teacher j taught during the school year

$GRADE_j$ is a set of dummy variables indicating the grade level(s) that teacher j taught during the school year

$FELLOW_j$ is an indicator of whether the teacher j was a Teachers Institute Fellow by the beginning of the school year

$PCTFEL_j$ is the proportion of teachers in the school in which teacher j taught who had participated in a Teachers Institute seminar by the beginning of the school year

EXP_j = the average number of years of teaching experience of teachers in teacher j's school

$AVGCLASS_j$ = average class size in teacher j's school

$PCTNLSP_j$ = the percentage of students in teacher j's school who are qualified for a free or reduced-price lunch

$PCTMIN_j$ = the percentage of students in teacher j's school who belong to minority racial or ethnic groups

The coefficient b_8 is an estimate of the effect of being a Teachers Institute Fellow (in other words, a member of the treatment group) on the outcome, after controlling for other observed differences among teachers and schools. The coefficient b_9 is an estimate of the additional benefit that both treatment and comparison teachers in a school may experience from having a higher concentration of Teachers Institute Fellows teaching in their school. In interpreting these coefficients as effects, it is important to keep in mind that these coefficients may be biased estimates of the effects of being a Fellow and working in a school with a higher proportion of Fellows, if Fellows differ from other eligible teachers in other unobserved characteristics that are not controlled in the model.

With the exception of attendance, the teacher outcomes are dichotomous variables, and the models will be estimated using logistic regression methods. The models of outcomes from school records will be estimated separately for each school year in the study period with data for teachers who had started teaching in the district by the beginning of that school year. Thus, it will be possible to examine how the effects of being a Fellow varied over time (the effects may grow, given that some Fellows will have participated in additional Teachers Institute seminars as time goes on, but this growth may be attenuated by the inclusion of additional new teachers in the analyses for later years).

The models of outcomes from survey data, if such data are collected, will be estimated for the most recent school year in the study period and complement the results for school records outcomes.

Subgroup Analyses

To examine whether the effects of participation in Teachers Institute seminars differ among teachers with different characteristics and in schools with different characteristics, the descriptive analyses and models described above can be estimated

using data just from teachers with the teacher characteristics of interest or schools with characteristics of interest. For example, the effects of being a Fellow can be estimated separately for new teachers (for example, those with 3 or fewer years of teaching experience) and more experienced teachers, and for elementary, middle, K-8, and high schools. Teacher subgroups that will be explored include those defined by highest degree attained, major field of study in college, years of teaching experience, years teaching in the current school, grade levels taught, and subjects taught. School subgroups that will be explored include those defined by school type, percentage of children eligible for free or reduced-price school lunch, and percentage of children from minority racial and ethnic groups.

MULTIVARIATE ANALYSES TO ESTIMATE THE EFFECTS OF INSTITUTE PARTICIPATION ON STUDENT OUTCOMES

The final, best model for estimating the effects of Teachers Institute seminar participation on student outcomes will depend on the nature of the data obtained from the school district and the results of analyses testing alternative model specifications. The models specified below illustrate the kinds of models that will be investigated.

Model When Student Data Cannot Be Linked to Teacher Data

If student data cannot be linked with data for teachers that the student had, a two-level hierarchical linear model may provide the best possible estimates of the effects of teacher participation in Teachers Institute seminars on student achievement. In this analysis, the sample of students necessarily includes all students in schools where eligible teachers taught.

If the data obtained from New Haven Public Schools do not permit linking student data to teacher data, the best possible models are ones that take into account the clustering of students within schools and explore the effects of participation in Teachers Institute seminars in terms of the aggregate number of teachers in the relevant grades who have participated in Teachers Institute seminars. Separate models will be estimated for each outcome in each year during the study period.

At the first level, the outcome (achievement test score, award, attendance) of student i in school k is a function of the student's previous outcome and demographic characteristics. For achievement test outcomes, for example:

$$SCORE_{ik} = a_{0k} + a_{1k}PRESCORE_{ik} + a_{2k}AGE_{ik} + a_{3k}GENDER_{ik} + a_{4k}RACE_{ik} + a_{5k}NSLP_{ik} + a_{6k}SCHMOVE_{ik} + r_{ik}$$

where:

$SCORE_{ik}$ = the achievement test score of student i in school k

$PRESCORE_{ik}$ = the previous achievement test score of student i in school k
AGE_{ik} = the age of student i in school k
$GENDER_{ik}$ = the gender of student i in school k
$RACE_{ik}$ denotes a set of dummy variables indicating the race/ethnicity of student i in school k
$NSLP_{ik}$ = whether or not student i in school k is eligible for free or reduced-price school lunches
$SCHMOVE_{ik}$ = whether or not student i in school k moved between schools since the previous testing occasion

In this model, estimated with previous test scores and demographic characteristics centered around their grand mean, a_{0k} is the adjusted mean test score in school k after controlling for differences in pretest scores and differences in student demographic characteristics.

At the second level, each of the coefficients in the student-level model is specified as nonrandomly varying and is a function of school characteristics and aggregate characteristics of teachers in the grades between the previous and current test scores:[5]

$$a_{pk} = b_{p0} + b_{p1}PCTFEL_k + b_{p2}EXP_k + b_{p3}AVGCLASS_k + b_{p4}PCTNLSP_k + b_{p5}PCTMIN_k \quad (p=0 \text{ to } 5)$$

where:
$PCTFEL_k$ = the proportion of teachers in the grades between the previous and current test scores in school k that are Teachers Institute Fellows
EXP_k = the average number of years of teaching experience of teachers in the grades between the previous and current test scores in school k
$AVGCLASS_k$ = average class size in school k (in the grades between the previous and current test scores, if available)
$PCTNLSP_k$ = the percentage of students in the school that are qualified for a free or reduced-price lunch
$PCTMIN_k$ = the percentage of students in school k who belong to minority racial or ethnic groups

For students who moved between schools within the district between testing occasions, the characteristics of the destination school will be used in the school-level model.

School effects are not considered random in this specification, because the results will not be generalized to schools beyond those represented in the study. The model will test, however, whether school effects vary according to school characteristics and in particular, whether a higher proportion of Teachers Institute Fellows in the relevant grades is associated with enhanced growth in student achievement.

The estimated coefficient b_{01} is an estimate of the difference in adjusted mean test scores associated with an increase of one percent in the proportion of teachers in the relevant grades who are Fellows. The estimated coefficient b_{11} is an estimate of the difference in the rate of growth in student achievement between the two testing points associated with an increase of one percent in the proportion of teachers in the relevant grades who are Fellows. In interpreting these coefficients as effects, it is important to keep in mind that the coefficients may be biased estimates of the effects of being a Fellow, if Fellows differ from other eligible teachers in other unobserved characteristics that are not controlled in the model.

Model When Student Data Can Be Linked to Teacher Data

If the data provided permit linking students to the teacher(s) they had, then a three-level hierarchical model of the following form can be estimated. This model can be estimated using data only for students taught by eligible teachers. Separate models will be estimated for outcomes in each grade in each year during the study period.

At the first level, the student outcome is a function of the previous level of the outcome and a set of student demographics. In the case of achievement test scores, for example:

$$SCORE_{ijk} = a_{0jk} + a_{1jk}PRESCORE_{ijk} + a_{2jk}AGE_{ijk} + a_{3jk}GENDER_{ijk} + a_{4jk}RACE_{ijk} + a5_{jk}NSLP_{ijk} + a_{6jk}SCHMOVE_{ijk} + r_{ijk}$$

where:

- $SCORE_{ijk}$ = the achievement test score of student i of teacher j in school k
- $PRESCORE_{ijk}$ = the previous achievement test score of student i of teacher j in school k
- AGE_{ijk} = the age of student i of teacher j in school k
- $GENDER_{ijk}$ = the gender of student i of teacher j in school k
- $RACE_{ijk}$ denotes a set of dummy variables indicating the race/ethnicity of student i of teacher j in school k
- $NSLP_{ijk}$ = whether or not student i of teacher j in school k is eligible for free or reduced-price school lunches
- $SCHMOVE_{ijk}$ = whether or not student of teacher j in school k moved between schools since the previous testing occasion

In this model, estimated with previous test scores and demographic characteristics centered around their grand means, a_{0jk} is the adjusted mean test score in the classroom of teacher j at school k after controlling for differences in pretest scores and differences in student demographic characteristics.

At the second level, each of the coefficients in the student-level model is viewed as a function of teacher characteristics:[6]

$$a_{pjk} = b_{p0k} + b_{p1k}FELLOW_{jk} + b_{p2k}EXP_{jk} + b_{p3k}POSTGRAD_{jk} + u_{jk} \quad (p=1 \text{ to } 6)$$

where:

$FELLOW_{jk}$ = whether or not teacher j in school k was a Teachers Institute Fellow

EXP_{jk} = the number of years of teaching experience of teacher j in school k

$POSTGRAD_{jk}$ = whether or not the teacher has education beyond a bachelor of arts degree

Teacher/classroom effects are specified as random in this specification. The model will test whether teacher effects vary according to teacher characteristics and in particular, whether being a Teachers Institute Fellow is associated with enhanced student achievement. The estimated coefficient b_{00k} is the mean achievement in school k after adjusting for differences in observed teacher characteristics. The estimated coefficient b_{01k} is the average difference in adjusted mean test scores associated with a teacher in school k being a Fellow. The estimated coefficient b_{11k} is the average difference in the rate of growth in achievement between the previous and current testing periods associated with a teacher in school k being a Fellow.

Because the results of the study will not be generalized beyond the schools in the study, the school effects can be specified as nonrandomly varying. Thus, the third-level model can be specified as:[7]

$$b_{pqk} = c_{pq0} + c_{pq1}PCTFEL_k + c_{pq2}EXP_k + c_{pq3}AVGCLASS_k + c_{pq4}PCTNLSP_k + c_{pq5}PCTMIN_k \quad (p=0 \text{ to } 5,\ q=0 \text{ to } 3)$$

where:

$PCTFEL_k$ = the proportion of teachers in the grades between the previous and current test scores in school k that are Teachers Institute Fellows

EXP_k = the average number of years of teaching experience of teachers in the grades between the previous and current test scores in school k

$AVGCLASS_k$ = average class size in school k (in the grades between the previous and current test scores, if available)

$PCTNLSP_k$ = the percentage of students in the school that are qualified for a free or reduced-price lunch

$PCTMIN_k$ = the percentage of students in school k who belong to minority racial or ethnic groups

The estimated coefficient c_{001} is the average difference in the adjusted mean test score for a school associated with an increase of one percent in the proportion of teachers in the school in the relevant grades who are Fellows. The estimated coefficient c_{011} is the average difference in the effect of being a Fellow associated with an increase of one percent in the proportion of teachers in the school in the relevant grades who are Fellows. The coefficient c_{111} is the difference in average rate of growth of achievement associated with being a Fellow in a school where the pro-

portion of teachers in the school in the relevant grades who are Fellow increased by one percent.

Students who remain within New Haven Public Schools during the period between the previous and current testing occasions should have data for both the previous and current tests and can be included in the analyses. Students who enter the school district after the previous testing occasion or leave the district before the current testing occasion will not have outcome data for both time points and cannot be included in the analyses. Thus, the results of the analyses will apply only to students who attend school in the district for the full period.

Because teachers who are not Fellows are encouraged to use the curriculum resources developed by Fellows, it is possible and even likely that the teaching of comparison teachers and the achievement of their students may be improved by the Teachers Institute seminars. Curriculum units developed by Fellows are deposited in all schools, disseminated by school Representatives and Contacts, and made available on the Yale-New Haven Teachers Institute Web site. These potential "spillover" effects will be captured partially by the inclusion of PCTFEL in the school-level model.

Subgroup Analyses

To examine whether the effects of participation in Teachers Institute seminars on student outcomes differ among teachers and schools with different characteristics, the models described above can be estimated using data just from teachers with the teacher characteristics of interest (in the models using data that can be linked for students and teachers) or schools with characteristics of interest. The analyses will explore the same subgroups of teachers and schools described earlier for the analyses of teacher outcomes.

Models to Explore Variations in Intensity and Timing of Teachers Institute Participation

Beyond the effects of ever participating in a Teachers Institute seminar, it is of interest to investigate whether the effects decay or grow over time and whether participating in multiple seminars enhances the effects. One approach to exploring these questions would be to add or substitute one or more variables characterizing the timing and/or intensity of participation in the models described above. The descriptive analyses of patterns of seminar participation can be used to guide the specification of variables characterizing participation patterns.

Estimates of the effects of patterns of seminar participation from these models are subject to greater potential selection bias than the simpler models examining the effects of being a Fellow, because each participation group is compared to the full comparison group in these models. It is likely that within the group of Fellows,

teachers who chose to participate multiple times and teachers who participated more recently differ systematically in unobserved ways from those who participated only once or who participated longer ago. For example, even after controlling for other measured differences among teachers, teachers who chose to participate multiple times may have been more highly motivated teachers with greater curiosity and energy, and the estimated effect of participation in multiple seminars may reflect these unmeasured differences rather than effects of multiple seminar participation.

Propensity scoring methods can be explored to account for selection bias in the models. These methods match comparison group teachers to Fellows based on observable characteristics, assuming that if the distributions of observable characteristics are similar, then the distributions of unobservable characteristics should also be similar. Then, for example, average outcomes for Fellows who participated multiple times can be compared with the average outcomes for their matched comparison group teachers to estimate the effects of participating multiple times, and the average outcomes for Fellows who participated only once can be compared with the average outcomes for their matched comparison group teachers to estimate the effects of participating once in a Teachers Institute seminar. The estimates of effects of participating multiple times can be compared with the estimates of effects of participating once to get some indication of the additional benefits that may be associated with continued participation in seminars.

If a teacher survey is conducted, it may be possible to collect additional data on teacher characteristics that may be associated with both decisions to participate in multiple seminars and outcomes. These variables can be used to reduce the potential biases due to unobserved characteristics and to improve the matching in the propensity scoring analyses.

INTERPRETING THE ANALYSIS RESULTS

The analyses described above will provide estimates of the effects of participation in Teachers Institute seminars on teachers and students. Descriptive analyses of data collected by the Teachers Institute from participating teachers about their perceptions of the usefulness and effectiveness of the seminars and their perceptions of the effects on students of using the new curriculum units provide additional information about the possible nature and magnitude of the effects of seminar participation. To the extent that these analyses paint consistent pictures of the effects of Teachers Institute seminar participation, the analyses together strengthen confidence in the findings of either analysis individually. To the extent that the analyses offer inconsistent information about program effects, they may raise questions that need to be addressed in a prospective evaluation.

NOTES AND REFERENCES

NOTES

1. The study period may need to be adjusted, depending on what data the school district can provide.

2. The list of teacher and student outcomes that will be examined may need to be refined depending on the data actually received from New Haven Public Schools.

3. To be eligible, teachers must be assigned New Haven classes for the following year in which they can teach the curriculum units they develop during the Teachers Institute.

4. The independent variables in this model will be informed by the results of the descriptive analyses, and the final models may include different independent variables.

5. Additional school characteristics may be added to the model if they are available.

6. If during the time between the previous achievement test score (PRESCORE) and the achievement score (SCORE), the student was taught by multiple eligible teachers, the teacher characteristics can be specified as the average teacher characteristic experienced by the student.

7. Additional school characteristics may be added if they are available.

REFERENCES

Bloom, Howard S., Charles Michalopoulos, and Carolyn J. Jill. "Using Experiments to Assess Nonexperimental Comparison-Group Methods for Measuring Program Effects." In Howard S. Bloom, editor. *Learning More from Social Experiments: Evolving Analytic Approaches*. New York, NY: Russell Sage Foundation, 2005.

Glazerman, Steven, Dan M. Levy, and David Myers. *Nonexperimental Replications of Social Experiments: A Systematic Review*. Princeton, NJ: Mathematica Policy Research, Inc., 2002.

Raudenbush, Stephen W., and Anthony S. Bryk. *Hierarchical Linear Models: Applications and Data Analysis Methods, Second Edition*. Thousand Oaks, CA: Sage Publications, 2002.

Snijders, Tom, and Roel Bosker. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Thousand Oaks, CA: Sage Publications, 1999.

APPENDIX

TABLES

TABLE 1: MINIMUM DETECTABLE DIFFERENCES — TEACHER OUTCOMES

Outcome Proportion	p	Variance sigma**2	Potential Sample Size: Fellows nt	Potential Sample Size: Comparison Teachers nc	Response Rates RR	Minimum Detectable Difference MDD	(a) MDD with Regression Adjustment
.1 / .9	0.1	0.09	410	200	0.8	0.081	0.064
.2 / .8	0.2	0.16	410	200	0.8	0.108	0.086
.3 / .7	0.3	0.21	410	200	0.8	0.124	0.098
.4 / .6	0.4	0.24	410	200	0.8	0.132	0.105
.5 / .5	0.5	0.25	410	200	0.8	0.135	0.107
.1 / .9	0.1	0.09	410	400	0.8	0.066	0.052
.2 / .8	0.2	0.16	410	400	0.8	0.088	0.070
.3 / .7	0.3	0.21	410	400	0.8	0.101	0.080
.4 / .6	0.4	0.24	410	400	0.8	0.108	0.086
.5 / .5	0.5	0.25	410	400	0.8	0.110	0.087
.1 / .9	0.1	0.09	410	600	0.8	0.060	0.048
.2 / .8	0.2	0.16	410	600	0.8	0.080	0.064
.3 / .7	0.3	0.21	410	600	0.8	0.092	0.073
.4 / .6	0.4	0.24	410	600	0.8	0.098	0.078
.5 / .5	0.5	0.25	410	600	0.8	0.100	0.080
.1 / .9	0.1	0.09	410	800	0.8	0.057	0.045
.2 / .8	0.2	0.16	410	800	0.8	0.076	0.060
.3 / .7	0.3	0.21	410	800	0.8	0.087	0.069
.4 / .6	0.4	0.24	410	800	0.8	0.093	0.074
.5 / .5	0.5	0.25	410	800	0.8	0.095	0.075

The calculations assume a one-tailed test, alpha = .05, power = .80, and a regression r-squared of .2.

Example (see the first row of the highlighted section of the table): Suppose the outcome is whether or not a teacher received an award or special recognition in the past year. Suppose that the probability of receiving an award in the absence of the Teacher Institute is .10. If the probability that an Institute teacher received an award or special recognition in the past year is .157 or higher, the difference is statistically significant. If the difference is adjusted to take into account other differences between the two groups of teachers, then an increase to .145 or higher would be statistically significant.

TABLE 2: MINIMUM DETECTABLE DIFFERENCES — STUDENT OUTCOMES

(Student data not linked to teacher data)

Outcome	p	Number of Schools	Number of Students per School	Design Effect Due to Clustering	Variance sigma**2	Potential Sample Size: Students of Fellows nt	Potential Sample Size: Students of Comparison nc	Response Rates RR	Minimum Detectable Difference MDD	(a) MDD with Covariates
6th grade test scores in normal curve equivalents										
		23	65	2.34	443.52	598	897	0.8	4.73	4.23
Proportion of 6th grade students scoring above a threshold										
.1 / .9	0.1	23	65	2.34	0.09	598	897	0.8	0.067	0.039
.2 / .8	0.2	23	65	2.34	0.16	598	897	0.8	0.090	0.052
.3 / .7	0.3	23	65	2.34	0.21	598	897	0.8	0.103	0.060
.4 / .6	0.4	23	65	2.34	0.24	598	897	0.8	0.110	0.064
.5 / .5	0.5	23	65	2.34	0.25	598	897	0.8	0.112	0.066
8th grade test scores in normal curve equivalents										
		23	65	2.34	443.52	598	897	0.8	4.73	4.23
Proportion of 8th grade students scoring above a threshold										
.1 / .9	0.1	23	65	2.34	0.09	598	897	0.8	0.067	0.039
.2 / .8	0.2	23	65	2.34	0.16	598	897	0.8	0.090	0.052
.3 / .7	0.3	23	65	2.34	0.21	598	897	0.8	0.103	0.060
.4 / .6	0.4	23	65	2.34	0.24	598	897	0.8	0.110	0.064
.5 / .5	0.5	23	65	2.34	0.25	598	897	0.8	0.112	0.066
10th grade test scores in normal curve equivalents										
		23	65	2.34	443.52	598	897	0.8	4.73	4.23
Proportion of 10th grade students scoring above a threshold										
.1 / .9	0.1	23	65	2.34	0.09	598	897	0.8	0.067	0.039
.2 / .8	0.2	23	65	2.34	0.16	598	897	0.8	0.090	0.052
.3 / .7	0.3	23	65	2.34	0.21	598	897	0.8	0.103	0.060
.4 / .6	0.4	23	65	2.34	0.24	598	897	0.8	0.110	0.064
.5 / .5	0.5	23	65	2.34	0.25	598	897	0.8	0.112	0.066
12th grade test scores in normal curve equivalents										
		14	65	2.34	443.52	364	546	0.8	6.06	5.42
Proportion of 12th grade students scoring above a threshold										
.1 / .9	0.1	14	65	2.34	0.09	364	546	0.8	0.086	0.050
.2 / .8	0.2	14	65	2.34	0.16	364	546	0.8	0.115	0.067
.3 / .7	0.3	14	65	2.34	0.21	364	546	0.8	0.132	0.077
.4 / .6	0.4	14	65	2.34	0.24	364	546	0.8	0.141	0.082
.5 / .5	0.5	14	65	2.34	0.25	364	546	0.8	0.144	0.084

The calculations assume a one-tailed test, with an alpha of .05, power of .80, an intraclass correlation of .07, and proportion of variance explained of .2.

TABLE 3: MINIMUM DETECTABLE DIFFERENCES — STUDENT OUTCOMES

(Student data linked to teacher data)

Outcome	p	Number of Schools	Number of Students per School	Design Effect Due to Clustering	Variance sigma**2	Potential Sample Size: Students of Fellows nt	Potential Sample Size: Students of Comparison nc	Response Rates RR	Minimum Detectable Difference MDD	(a) MDD with Covariates
6th grade test scores in normal curve equivalents										
		70	20	1.53	443.52	560	840	0.8	3.94	3.53
Proportion of 6th grade students scoring above a threshold										
.1 / .9	0.1	70	20	1.53	0.09	560	840	0.8	0.056	0.041
.2 / .8	0.2	70	20	1.53	0.16	560	840	0.8	0.075	0.054
.3 / .7	0.3	70	20	1.53	0.21	560	840	0.8	0.086	0.062
.4 / .6	0.4	70	20	1.53	0.24	560	840	0.8	0.092	0.066
.5 / .5	0.5	70	20	1.53	0.25	560	840	0.8	0.094	0.068
8th grade test scores in normal curve equivalents										
		70	20	1.53	443.52	560	840	0.8	3.94	3.53
Proportion of 8th grade students scoring above a threshold										
.1 / .9	0.1	70	20	1.53	0.09	560	840	0.8	0.056	0.041
.2 / .8	0.2	70	20	1.53	0.16	560	840	0.8	0.075	0.054
.3 / .7	0.3	70	20	1.53	0.21	560	840	0.8	0.086	0.062
.4 / .6	0.4	70	20	1.53	0.24	560	840	0.8	0.092	0.066
.5 / .5	0.5	70	20	1.53	0.25	560	840	0.8	0.094	0.068
10th grade test scores in normal curve equivalents										
		70	20	1.53	443.52	560	840	0.8	3.94	3.53
Proportion of 10th grade students scoring above a threshold										
.1 / .9	0.1	70	20	1.53	0.09	560	840	0.8	0.056	0.041
.2 / .8	0.2	70	20	1.53	0.16	560	840	0.8	0.075	0.054
.3 / .7	0.3	70	20	1.53	0.21	560	840	0.8	0.086	0.062
.4 / .6	0.4	70	20	1.53	0.24	560	840	0.8	0.092	0.066
.5 / .5	0.5	70	20	1.53	0.25	560	840	0.8	0.094	0.068
12th grade test scores in normal curve equivalents										
		45	20	1.53	443.52	360	540	0.8	4.92	4.40
Proportion of 12th grade students scoring above a threshold										
.1 / .9	0.1	45	20	1.53	0.09	360	540	0.8	0.070	0.051
.2 / .8	0.2	45	20	1.53	0.16	360	540	0.8	0.093	0.068
.3 / .7	0.3	45	20	1.53	0.21	360	540	0.8	0.107	0.077
.4 / .6	0.4	45	20	1.53	0.24	360	540	0.8	0.114	0.083
.5 / .5	0.5	45	20	1.53	0.25	360	540	0.8	0.117	0.085

The calculations assume a one-tailed test, with an alpha of .05, power of .80, an intraclass correlation of .07, and proportion of variance explained of .2.

TABLE 4: SUMMARY OF RESEARCH QUESTIONS, DATA NEEDS, AND ANALYTIC APPROACHES

Research Questions	*Data Needs and Sources*	*Analytic Approach*
What impacts does teacher participation in Institute seminars have on teachers' careers?	School district records data on teacher retention, attendance, promotions, and special recognition for Institute Fellows and comparison teachers Survey of Institute and comparison teachers about their career progress, attitudes toward teaching and learning, strategies, and expectations of students Background information on teachers from records and surveys	Multivariate analyses of the effects of Institute participation on teacher outcomes, controlling for other differences between teachers.
What impacts does teacher participation in Institute seminars have on student performance?	School records data on student performance for students of Institute Fellows and comparison teachers (if possible) or for students at schools and in grades in which Fellows and comparison teachers taught Survey of participating teachers about student responses to new curriculum units and other changes in teaching strategies Survey of students of participating teachers	Multivariate analyses of the effects of Institute participation on student outcomes, controlling for other differences between students and teachers. Analyses will use hierarchical linear models that take into account the structure of the school records data.
How do teachers and students perceive that Institute participation has affected them?	Survey of participating teachers about their seminar experiences and perceived usefulness of program participation in areas associated with teacher quality perceived impacts on their careers, and perceived student responses to new curriculum units and changed teaching practices (if any)	Descriptive analyses of survey data collected from Institute participants

SCHOOL RECORDS DATA COLLECTION

Although a request for school records data has already been initiated in New Haven, it would be useful at this point to review what data have been received and follow up systematically to determine what additional data may be obtainable and clarify what is needed.

An accompanying spreadsheet summarizes the data that have been received to date. Below are questions about the data received and data that might be available. Following these questions are three tables that systematically lay out data that would be very useful for analysis (these could be considered wish lists for student, teacher, and school data).

Student-Level Data

- It would be very helpful to document what standardized testing is done in the New Haven Public Schools, and to identify any changes that have occurred during the past ten years. What standardized achievement tests are administered in which grades? Has this been the same for each of the last ten years? If not, what other tests were administered in previous years? How far back in time are data on standardized test scores retained? Once this information is known, the completeness of the data already received can be assessed, and a new, specific request for remaining test score data can be made. See the Test Schedule tab in the accompanying spreadsheet, which reflects the data already obtained. Are any other standardized tests administered, or are tests administered in any grades not indicated in the table?
- The tables below list other data that would be valuable for the analyses. For those that have not yet been obtained, are they available? Exactly what data are available?
- Are records maintained in such a way that students can be linked to the teachers they had in each grade? For example, is it possible to obtain data indicating the teachers that individual students had during each semester? For how far back in time is this information available?
- For the analyses, it is highly desirable to link student data from multiple sources, such as test scores from multiple years. This requires that the student data files contain identifying information that is common across sources. Ideally, each file would include a permanent student ID number that could be used to merge data across data files. Are District students assigned permanent, unique ID numbers, and can these be included in all the data files?

Teacher-Level Data

- The original data request relied on school district staff to select a sample of teachers who had not been Fellows. It might be simpler for the district to pro-

vide data on all teachers, and that would give you more options for defining a sample of comparison teachers for the analysis. Is it possible for the district to provide data for all teachers? For how many years back can these data be provided?

- Similarly, linking teacher data from multiple sources, such as data indicating the grades and subjects taught each year, would be much more reliable if teachers have permanent ID numbers that could be used to link data for the same teacher from multiple data files. (Unless, of course, cumulative records for each teacher are kept together and can be provided in a single data file.)
- No teacher-level data have been provided yet. The data listed in the table below would be valuable for the analysis if they are available.

School Data

- Most of these data are available from public use datasets. The table notes which items can be obtained from the Common Core of Data and from the Connecticut Department of Education Web sites for varying numbers of years back.
- Does the District maintain individual student-level and teacher-level data files that are used to develop the aggregate school and District data provided to the State? If so, would it be possible to provide these data files to us? Alternatively, for many items, it appears that the State has student-level data files — if so, is there a way for the District to provide access to these files?

SCHOOL-LEVEL INFORMATION (DATA AGGREGATED ACROSS STUDENTS)

Format and Time Frame: Aggregate data, broken down by grade; Data from the 1999-2000 school year and continuing through the 2004-2005 school year

Information Desired	***Available from School District?*** *(Mark E if available in electronic form and P if available in paper form.)*	***Available from Common Core of Data?*** *(Varying number of years back)*	***Available from CT State Department of Education Web Site***	***Notes*** *(e.g., record data definition or additional information about availability)*
Information about the Data				
Dates or time periods for which data are available				
School Identifiers (to link data files)		X	X	
School name		X	X	
School level/grades offered				
Whether school is a magnet school				
Whether school is a vocational school			X	
Enrollment				
Number of students enrolled		X*		
Percent of students enrolled in bilingual and ESL services			X	SSP***
Percent of students enrolled in special education services			X	SSP***
Percent of students enrolled in compensatory education			X	SSP***
Percent of students enrolled in a talented and gifted program			X	SSP***
Student mobility			X**	SSP***
Student stability				
Student participation in extracurricular activities				
Aggregate Student Demographics				
Percent of students by gender		X	X	
Percent of students by ethnicity		X	X	SSP***, enrollment Xcel file
Percent of students with non-English home language			X	SSP***
Percent of students eligible for free or reduced price lunch		X	X	SSP***
Aggregate Student Standardized Test Scores				

* Total, by grade, by race/ethnicity, and by gender

** Percent of students above entry-level-grade who attended this school the previous year

*** Strategic School Profile

Information Desired	***Available from School District?*** *(Mark E if available in electronic form and P if available in paper form.)*	***Available from Common Core of Data?*** *(Varying number of years back)*	***Available from CT State Department of Education Web Site***	***Notes*** *(e.g., record data definition or additional information about availability)*
General Achievement test			X Participation in PSAT (grades 10 and 11)	Can be downloaded into files that can be read into Xcel.
Reading or English test			CMT Reading (grades 4, 6, 8) SAT Verbal (graduating seniors)	Substantial school-level data on participation, scores, and % achieving mastery are available online through interactive reports, and many can be downloaded into files that can be read into Xcel.
Math test			CMT Math (grades 4, 6, 8) SAT Math (graduating seniors)	Substantial school-level data on participation, scores, and % achieving mastery are available online through interactive reports, and many can be downloaded into files that can be read into Xcel.
Science test				
Social Studies or History test				
Writing test			CMT Writing (grades 4, 6, 8)	Substantial school-level data on participation, scores, and % achieving mastery are available online through interactive reports, and many can be downloaded into files that can be read into Xcel.
Advanced Placement (AP) test participation and performance			X	

Information Desired	***Available from School District?*** *(Mark E if available in electronic form and P if available in paper form.)*	***Available from Common Core of Data?*** *(Varying number of years back)*	***Available from CT State Department of Education Web Site***	***Notes*** *(e.g., record data definition or additional information about availability)*
Aggregate Student Behavioral Indicators				
Number of disciplinary referrals				
Percent of students with disciplinary referrals				
Percent of students attending school				
Percent of students absent from school or class without permission				
Number of suspensions				
Number of expulsions				
Percent of students tardy				
Number or percent of students receiving awards and prizes, by type of award/prize				
Aggregate Student Achievement/Progress				
Percent of students who were retained or held back after previous year			X	SSP
Graduation rates				
% graduates who enter college				
Aggregate Teacher Characteristics				
Number of full-time teachers		Number of FTE teachers	X (# certified teachers)	SSP
Number of part-time teachers				
Percent of teachers minority race/ethnicity			X	SSP
Average number of years of experience			X (experience in CT)	SSP
% master's degree or above			X	SSP
% trained as mentors, assessors, cooperating teachers			X	SSP
Average days absent due to illness or personal time			X	SSP
% certified staff assigned to the same school in previous year			X	SSP
School Characteristics				
Number of instructional days or hours per year			X (hours per year)	SSP
Estimated hours of instruction per year by subject area			X	SSP
Number of classrooms				
Student:teacher ratio		X		
Average class size			X (grades K, 2, and 5)	SSP
Per pupil expenditure				
Title I eligibility, participation		X	X	

SCHOOL-LEVEL INFORMATION (DATA AGGREGATED ACROSS STUDENTS)

Format and Time Frame: Individual teacher data, with student data aggregated to the teacher level: Data from the 1999-2000 school year and continuing through the 2004-2005 school year

Information Desired	***Available from School District?*** *(Mark E if available in electronic form and P if available in paper form, and indicate the years for which the data are available.)*	***Available from School?*** *(Mark E if available in electronic form and P if available in paper form, and indicate the years for which the data are available.)*	***Notes*** *(e.g., record data definition or additional information about availability)*
Identifiers (to permit linking data files)			
School			
Name			
Teacher ID			
Social Security Number			
Teacher Demographics			
Birth date/age			
Race/ethnicity			
Teacher Background			
Highest degree completed			
Field of study in college			
Field of study in graduate school (if applicable)			
Certifications received			
Year began teaching			
Year began teaching in New Haven Public Schools			
Class Characteristics			
Average class size			
Subjects taught			
Grades taught			
Demographic characteristics of students			
Attendance and Recognition			
Attendance (number of days present)			
Whether left New Haven Public Schools, and if so, reason left			
Awards and special recognition			
Promotions (to leadership roles, not just formal employment promotions)			
Teacher Attitudes			

STUDENT-LEVEL INFORMATION (DATA FOR EACH STUDENT SEPARATELY)

Information Desired	***Available from School District?*** *(Mark E if available in electronic form and P if available in paper form, and indicate the years for which the data are available.)*	***Available from School?*** *(Mark E if available in electronic form and P if available in paper form, and indicate the years for which the data are available.)*	***Notes*** *(e.g., record data definition or additional information about availability)*
Identifiers (to permit linking data files)			
School			
Name			
Student ID			
Social Security Number			
Student Demographics			
Date of birth			
Gender			
Ethnicity			
English language proficiency			
Special education identification			
Talented/gifted identification			
Eligibility for free or reduced-price lunch			
Retained/held back in previous grades			
Date of entry into current school			
Student Standardized Test Scores			
General Achievement test			
Reading or English test			
Math test			
Science test			
Social Studies or History test			
Writing test			
Student Grades/Academic awards/Progress			
Reading grade			
Math grade			
Science grade			
Social Studies or History grade			
Behavior/Motivation/Effort grade			
Academic awards and prizes received			
Whether graduated			
Whether admitted to college			
Attendance and Behavior			
Number of times tardy			
Number of times absent from school or class without permission			
Number of suspensions			
Number of expulsions			
Number or percent of days present			
Whether participated in extracurricular activities, by type of activity			