

TO STRENGTHEN TEACHING: An Evaluation of Teachers Institute Experiences



by Rogers M. Smith
University of Pennsylvania

TO STRENGTHEN TEACHING: AN EVALUATION OF TEACHERS INSTITUTE EXPERIENCES

**By Rogers M. Smith
The University of Pennsylvania**

March 2009

CONTENTS

Executive Summary	5
To Strengthen Teaching: An Evaluation of Teachers Institute Experiences	7
I. Teacher Quality and Professional Development Programs.....	7
The Ingredients of Teacher Quality.....	7
The Ingredients of Successful Professional Development Programs.....	8
II. The Teachers Institute Approach to Teacher Quality.....	8
III. Past Evaluations of the Teachers Institute Approach.....	11
The National Demonstration Project and Institute Evaluations.....	12
The Yale National Initiative.....	13
IV. Recent Results.....	13
The Curriculum Units in the Classroom.....	17
Student Responses to Curriculum Units.....	18
Site Variations.....	20
Corollary Benefits.....	21
V. The New Haven Quantitative Evaluation.....	22
Hypotheses.....	22
Results.....	23
VI. Conclusions and Recommendations for Continuing Evaluations.....	24
Notes	27

EXECUTIVE SUMMARY

Recent research indicates that the single most important factor in student performance is *teacher quality*. By common consensus, quality teachers are:

1. Teachers who really *know their subjects*;
2. Teachers with *good basic writing, math, and oral presentation skills*;
3. Teachers with *high expectations of their students*;
4. Teachers who are *enthusiastic about teaching*; and
5. Teachers who *can motivate all students to learn*.

Though the literature on professional development programs is less advanced, researchers agree that most forms of professional development fail to foster teacher quality along these five key dimensions. Most find that high quality professional development programs include:

1. A focus on *content* and on *pedagogy linked to content*;
2. *Active teacher learning*;
3. *Teacher leadership*;
4. *Extended duration*;
5. *Collective participation* by teachers from the same school, grade, or subject;
6. *Alignment* with state and local standards;
7. *Ongoing evaluation*.

For 30 years, the Yale-New Haven Teachers Institute has been developing a unique model for improving teacher quality. The Institute approach includes all of the seven elements of successful professional development programs, five very extensively; and it significantly strengthens teachers in all five of the major dimensions of teacher quality. At its heart are partnerships between institutions of higher education and public schools. Teachers Institutes offer five to seven seminars each year meeting weekly over roughly three months, led by university or college faculty members, on topics that teachers have selected to increase their mastery of what they teach. Institute programs focus on *content* and *pedagogy linked to that content*; *active teacher learning*; *extensive teacher leadership*; seminars of *substantial duration*; the development of curriculum units aligned to *state and local standards*; and they involve *continuing evaluations* and opportunities for *collective participation*, though to varying degrees. In Institute seminars teachers gain more sophisticated content *knowledge*, and they also *enhance their skills* by preparing curriculum units adapting the themes of the seminars for their students. Most teachers are *enthusiastic* about the seminars and the opportunity to teach the units they have written. They *expect more* of the students taking them. And they *succeed in motivating their students* to learn at higher levels.

Annual surveys of teacher participants at all four existing Institutes in the last six years confirm earlier results: teachers drawn from all grade levels and all subject areas participated out of desires to obtain curriculum suited to their needs (84.6% to

89.3% at the four sites), to increase their mastery of their subjects (85.5% to 90.5%), and especially to obtain materials *to motivate their students* (93.2% to 94.7%). Over 96% of all participating teachers rated the Institute seminars overall "moderately" or "greatly" useful. High percentages said the seminars increased their knowledge (87.3% to 93.7%) and raised their expectations of students (87.8% to 94.75%). Though data on unit use are less extensive, they show that after teaching their Institute units, 2/3 of all participants rated them superior to all other curriculum they had used. *Roughly 60% of all participants rated student motivation and attention as higher during these units, producing substantially greater content mastery.*

A five-year quantitative study of the impact of Institute experiences on teachers and students in New Haven shows that the Institute attracts participants roughly equally from each identifiable demographic group: women and men, younger and older teachers, new and experienced teachers, white and African American teachers — with the latter in fact over-represented. Though causality cannot be imputed, *Institute participants also had nearly twice the retention rate of non-participants* in local teaching. There were no statistically significant effects on the only measures of student achievement for which data were available, standardized test scores and grades. This expected result underlines the need for evaluations focused on the specific outcomes Institutes seek to achieve, including heightened student interest and attendance and improved student content knowledge in curricular areas featured in local district standards and goals.

Other research suggests that Institutes also serve to foster teacher leadership, to develop supportive teacher networks, to heighten university faculty commitments to improving public education, and to foster positive partnerships between school districts and institutions of higher education. To confirm, maintain, and improve these impacts, it is advisable for Institutes to continue surveys of teacher participants and to undertake a variety of additional continuing quantitative and qualitative studies of the effects of Institutes on student and teacher outcomes.

TO STRENGTHEN TEACHING: AN EVALUATION OF TEACHERS INSTITUTE EXPERIENCES¹

This report is divided into six sections. Section I summarizes the conclusions of recent research on the ingredients of teacher quality and effective professional development programs. Section II describes the Teachers Institute approach to improving teacher quality. Section III reviews past evaluations of the Teachers Institute model. Section IV reports the results of recent research relying on teacher surveys. Section V describes a quantitative study of the impact of participating in Institute seminars on New Haven teachers and students from the 2000-2001 through the 2004-2005 academic years. Section VI details the report's conclusions and recommendations for continuing evaluations of the Teachers Institute approach.

I. TEACHER QUALITY AND PROFESSIONAL DEVELOPMENT PROGRAMS

In recent years, educational researchers have converged on the conclusion that the best way to help students learn is to *improve teacher quality*.² Scholars also see an "emerging" consensus on the characteristics of professional development programs that can contribute to this goal.³ The mandate of Sec. 1119 of the No Child Left Behind Act for school districts to employ "highly qualified" teachers who receive "high quality" professional development has prompted more thorough data collection by school officials and more rigorous evaluations of teacher quality and of professional development programs. Much remains unsettled, but some features contributing to effective teachers and successful professional development are clear. Programs that contribute to teacher quality, as well as promising strategies for ongoing program evaluations, can be identified with confidence.

The Ingredients of Teacher Quality

Though researchers differ on some important issues, virtually all experts agree on the value of five features. Quality teachers are:

1. Teachers who really know their subjects, not only "how" to teach;⁴
2. Teachers who have good basic writing, math, and oral presentation skills;⁵
3. Teachers who expect their students to achieve;⁶
4. Teachers who are enthusiastic about teaching;⁷ and
5. Teachers who can motivate even highly disadvantaged students to learn.⁸

Two pertinent caveats appear in the literature. First, content knowledge and skills provided by advanced degrees do not always translate into greater ability to convey that content to students, a finding that some researchers interpret as indicating the importance of pedagogical as well as content training.⁹ Second, if over time teach-

ers rely too extensively on content knowledge gained from an assigned standardized curriculum, their teaching may become complacent, unimaginative, and less effective.¹⁰

The Ingredients of Successful Professional Development Programs

Scholars concur that "nationwide, the typical professional development experience is not high quality."¹¹ Many traditional forms of professional development are limited in duration, content, and active learning, leaving teachers uninformed and uninspired.¹² But high quality programs exist, and researchers are gaining knowledge of their key characteristics. Recent studies stress that these programs feature:

1. A focus on content and on pedagogy linked to specific content;¹³
2. "Active" learning involving feedback on teaching strategies and practices;¹⁴
3. Opportunities for teacher leadership;¹⁵
4. Extended duration;¹⁶
5. Collective participation of teachers by school, grade, or department;¹⁷
6. Alignment with state and local standards and curriculum;¹⁸ and
7. Appropriate evaluation methods and objectives.¹⁹

The chief caveats here are first, that due both to costs and program structures, relatively few professional development program evaluations have included randomized controlled trials or rigorous quasi-experimental designs with well-matched participant and non-participant groups.²⁰ Few existing programs, then, can claim to have been validated through those forms of empirical research. Second, the research literature suggests that districts may face a choice "between serving larger numbers of teachers with less focused and sustained professional development or providing higher quality activities for fewer teachers," since "good professional development requires substantial resources" in terms of time, expertise, and dollars.²¹ As the Gates Foundation recently stated, an "effective professional learning community" requires "teachers who work together to improve each other's practice," including "weekly time to work together to meet shared challenges and improve their skills. . . . They need ongoing, job-embedded professional development."²² These needs cannot be met through forms of professional development that feature brief workshops for passive audiences of large numbers of teachers.

II. THE TEACHERS INSTITUTE APPROACH TO TEACHER QUALITY

Since 1978, the Yale-New Haven Teachers Institute has been developing a unique approach to improving teacher quality, one now employed by affiliated Teachers Institutes in three other cities: Pittsburgh, Houston, and Philadelphia. Each Institute provides professional development designed to help teachers face the challenges of educating effectively in low-income communities with high proportions of racial and ethnic diversity. In New Haven, where 45 schools serve over 20,000 students, 70% of the students come from families receiving school meal financial assistance,

and 87% are African American or Hispanic. Teachers from all schools are eligible to participate in Institute seminars. In Pittsburgh, 65 schools serve 28,000 students, 68% financially eligible for school lunch aid. Roughly 57% are African American. The Teachers Institute of Pittsburgh has drawn participants from most schools in the district.²³ In Houston, the nation's seventh largest school district includes over 300 schools and serves 202,000 students, 58% Hispanic, 30% African American, many with limited English proficiency, with 78% of all students eligible for meal assistance. The Houston Teachers Institute has Representatives in roughly 30 schools and draws extensively from those schools, though all Houston public school teachers are eligible to participate. In Philadelphia, where 281 schools serve over 167,000 students, with 80% eligible for meal assistance and the same percentage either African American or Hispanic, the Teachers Institute of Philadelphia targets 25 schools in the South and West Regions, representing over 18,000 students. Each Institute makes the curriculum units written by the teachers who participate in its seminars, designated "Fellows," available to all interested teachers in their districts and beyond, in electronic and print formats.

Though these Institutes cannot and have not solved all the daunting educational problems in these settings, data obtained through annual teacher surveys and school district records indicate that each Institute embodies most of the seven features that characterize successful professional development programs, particularly a focus on content and content-linked pedagogy; active learning; teacher leadership; extended duration; and alignment with state and local standards. The Institutes also provide opportunities for collective participation and all engage in continuing evaluation, though with some variations. At each site the Institute approach has significantly strengthened teachers in regard to all five ingredients of teacher quality noted above: content knowledge; basic communication and calculation skills; expectations of students; teacher enthusiasm; and capacities to motivate students.

At the heart of the Teachers Institute approach is a special sort of partnership between institutions of higher education and public schools. Each Institute consists of a Director and staff; a network of Teacher Representatives in the local public schools; and a community of higher education faculty members. Each year, the Institutes offer weekly seminars over several months, led by university or college faculty members, on topics that the teachers, through their Representatives, have selected to enhance their knowledge of what they teach. The seminars sometimes include teams of participants from a single school, but most also combine teachers from different schools teaching different subjects at different grade levels, who all wish to learn more about topics in history, literature, politics, chemistry, biology, art, or other areas, in order to improve their teaching.

In these seminars teachers gain more sophisticated content knowledge, and they also enhance their writing and oral skills by preparing substantial curriculum units

that they discuss with their fellow teachers in the seminars, then teach to their students. These units of roughly fifteen pages or more are based on research into a topic inspired by the seminar and assisted by the faculty member, but adapted to the teachers' courses and students. The units often represent the first substantial research and writing project that teachers have undertaken since college. They contain a narrative describing the unit's goals and strategies and several sample lesson plans. By presenting their units in progress to their fellow teachers, seminar participants receive useful pedagogical feedback as they also improve their oral communication skills. Roughly half of each seminar's time is devoted to improving teachers' content knowledge, half to developing effective teaching strategies for their students. Both faculty and teachers are paid stipends, in recognition that all are giving time and expertise to improve teaching and curriculum.

The Institute approach thus involves a stress on improving teachers' *content knowledge* and on the development of *pedagogical strategies tied to the content* of their units. Teachers are *active learners* in these seminars, receiving feedback from their peers and often trying out the units with their students as they prepare them. Teachers play *leadership roles* in every stage of the program, including designation of seminar topics, approval of faculty proposals for seminars on those topics, approval of teacher participants, and assistance in coordinating seminar activities and program schedules throughout the year. The seminars are of substantial *duration*, involving a minimum of 26 hours in session, while Fellows also meet with seminar leaders, research seminar topics, and write their units over many more hours. All units are explicitly *aligned with state and local curricular standards* and goals. The seminars involve a mix of *collective and individual* forms of teacher participation, though they are weighted toward individual participation.²⁴ And each site engages in *continuing evaluation*, though for reasons discussed below, chiefly through means of teacher surveys and qualitative studies, not randomized controlled trials or quasi-experiments.

The Institutes are especially well-suited to insure that teachers do not learn advanced content in isolation from attention to how it can be conveyed effectively, and that they do not passively receive content from curriculum prepared by others. Instead, they research and write curriculum themselves, challenged and aided by their peers to insure that it is exciting for students and teachers alike. The Institute approach does embrace the alternative of "higher quality activities for fewer teachers." Each Institute offers seminars for roughly 50 to 80 teachers per year. But through dissemination of their units, Institutes can impact far more teachers than they enroll; and over time, significant percentages of teachers in particular schools and regions can be direct participants. In New Haven, nearly 600 teachers had been Institute seminar participants by 2005, including 32% of the teachers then at work in New Haven high schools; 25% of the middle-school teachers; and 14% of the elementary-school teachers.²⁵

III. PAST EVALUATIONS OF THE TEACHERS INSTITUTE APPROACH

Since the founding of the Yale-New Haven Teachers Institute in 1978, its leaders have consistently supported external and internal evaluations to discern and improve the efficacy of the Institute's work. The distinctive approach of the Institute does pose challenges for rigorous empirical evaluation. Most scholars agree that causality can best be determined via randomized controlled trials, in which subjects are randomly assigned to two groups. One, the "treatment" group, is subjected to the causal variable under investigation. The other, the "control" group, is left unaffected by it. Researchers hope that, apart from this treatment, all other characteristics will be randomly distributed among the two groups, so that differences in their subsequent performance can confidently be attributed to the study's causal variable.

But random assignment of teachers to treatment groups is not appropriate for evaluating the Teachers Institute approach. That approach relies heavily on teacher leadership, including not only the decisions of teachers concerning their own participation, but also the decisions of Teacher Representatives on applicants' eligibility for the program and on the seminars to which teachers are assigned. If these features were replaced by random assignment of teachers to seminars, the result would be an evaluation of a program quite different from a Teachers Institute.²⁶

The second-best research option for determining causality is generally considered to be quasi-experimental designs, in which subjects are not randomly assigned to treatment and control groups. Instead, researchers construct or identify a control group that matches the treatment group as closely as possible on all measurable characteristics. To obtain a close match, researchers generally need a great deal of information on a large number of potential group members. It is often difficult to meet these needs in real world settings. That is why one recent review of more than 1300 studies of teacher professional development concluded that only 9 were sufficiently rigorous to support confident attributions of causality.²⁷

In light of these difficulties, Teachers Institutes have from their inception relied most heavily on annual surveys of teachers, to determine whether they evaluate Institute seminars and curriculum units as preferable to available alternatives; on observational studies of Institute seminars and the teaching of curriculum units in school classrooms; and on the assessments of nationally recognized educational experts. Beginning in 1981, distinguished experts have been invited periodically to conduct site visits to the Yale-New Haven Teachers Institute and to provide evaluations. Their assessments have been strongly favorable, including those of Ernest L. Boyer, former U.S. Commissioner of Education, who wrote in 1981 that the Institute was a "dramatic exception" to a pattern of unsuccessful school-college collaborations; Theodore R.Sizer, Founding Director of the Annenberg Institute for School Reform, who termed the program "remarkable" for its "clear and useful focus" and "constructive collegiality"; and Norman C. Francis, President of Xavier University, who stressed the "absolute

need for" and "inestimable value" of the Institute's example. In addition, Gordon M. Ambach, Executive Director of the Council of Chief State Officers, has termed the Institute "one of our nation's most effective school-university partnerships to improve student achievement" in 1995; and U.S. Secretary of Education Rod Paige applauded the Institute in 2001 "for supplying models for what universities should do."²⁸

The Institute has also pursued more continuing and systematic assessment. In 1981 it began requiring all seminar participants to complete detailed questionnaires on their experience, and in 1981 and 1985 it undertook surveys both of teachers who had and teachers who had not been seminar participants to determine their experiences with using curriculum units. The results showed that the use of Institute-developed units was growing rapidly, with units taught in more than 1500 school classes by 1984. Over 97% of the teachers using the units praised them highly for their quality and the positive responses they invoked in students. In 1986, Gita Z. Wilder of the Education Policy Research Division of the Educational Testing Service assisted the Institute in refining the questionnaires for seminar participants, crafting the format that New Haven and the other Institutes continue to use, with only minor modifications. She wrote at that time that the surveys showed New Haven Fellows to be representative of the larger population of New Haven teachers in terms of subjects and grades taught, and that all teachers rated their seminar experiences highly with remarkable consistency. In 1991, Wilder helped develop a "Progress Report" that analyzed the accumulated data from the annual questionnaires and the surveys on unit use. She found that teachers from all grade levels and diverse subject matters rated the seminar experiences and the impact of the units on students in "consistently positive" ways, as subsequent studies have also shown.²⁹

The National Demonstration Project and Institute Evaluations

From 1999 to 2002, the Yale-New Haven Teachers Institute undertook a National Demonstration Project, seeding the creation of four new Teachers Institutes in different urban environments: Pittsburgh, Pennsylvania; Houston, Texas; Santa Ana, California; and Albuquerque, New Mexico. The Project included evaluations at each site and evaluation of the Project as a whole. Every Institute administered similar annual questionnaires to all participants on their seminar experiences, and the Project also sponsored a one-time survey on unit use by participating and non-participating teachers at each site. These surveys were analyzed in a 2004 study sponsored by the Institute, "To Motivate My Students," the precursor to this report. The study found that teachers at all sites rated the Institute programs higher than other professional development programs in developing the knowledge, skills, enthusiasm, high expectations of students, and capacities to motivate students that most studies indicate to be central to successful teaching. Roughly 60% of teachers surveyed rated student motivation and attention as higher when using these units than when using other curriculum, with only 1% rating them less.³⁰

The external evaluator of the National Demonstration Project, Policy Studies Associates, also found "clear evidence of important accomplishments" and concluded that the Project had "succeeded in reaching its goal" of showing that the Institute model could be replicated in different contexts, especially districts that respect "teachers as leaders" and value "teachers' professional learning."³¹ Using qualitative techniques including seminar and classroom observations, in-depth interviews, and focus groups, sociologists at the University of Houston and educators at Carnegie Mellon University and Chatham College, Pittsburgh conducted their own evaluations of their Teachers Institutes during the National Demonstration Project, with similarly positive results. The external evaluator's report included suggestions for improved continuing evaluation that have been incorporated into the ensuing Yale National Initiative. These recommendations included the importance of conformity to the Institute model. The Santa Ana Teachers Institute relied solely on state funding instead of the recommended diversified funding base, while the Albuquerque Teachers Institute reported to a dean, not a President or Provost. After the completion of the National Demonstration Project, these departures from the Institute approach appear to have contributed to financial difficulties that forced the suspension of operations in Santa Ana, and to substantial changes in the forms of professional development offered by the Albuquerque Teachers Institute.

The Yale National Initiative

The Houston and Pittsburgh Teachers Institutes continued in conformity with the Institute model along with the original Yale-New Haven Institute, each engaging in annual surveys of teacher participants and in other forms of continuing evaluation. In New Haven, Institute leaders planned and then undertook the Yale National Initiative, seeking to create new Institutes in other cities by adopting guiding principles and procedures for establishing Institutes; by forming a League of Teachers Institutes that includes a National University Advisory Council (composed of university faculty) and a National Steering Committee (composed of public school teachers); and by sponsoring an annual series of National Seminars and National Conferences in which representatives of both existing and prospective Institutes participate. Through the National Initiative, the Teachers Institute of Philadelphia was established, joining the League of Teachers Institutes in 2007. At this writing similar efforts are underway in seven other cities.

IV. RECENT RESULTS

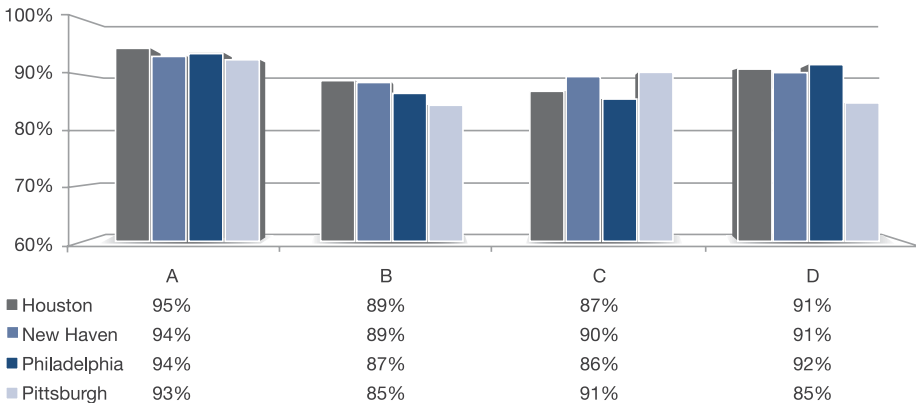
The administration by each Institute of similarly structured annual surveys of teacher seminar participants, which all must complete in order to receive their stipends, makes it possible to update the survey findings of the 2004 National Demonstration Project evaluation, determining whether teachers at the various Institutes continue to rate their seminar experiences favorably. In brief, the results from each individual Institute, from the National Seminars, and in the aggregate

national data all show that teachers throughout the country have consistently evaluated their Institute experiences much as the teachers did in the National Demonstration Project. Along the dimensions most associated with teacher quality, they rate the Institute approach very highly, markedly better than the low evaluations scholars report for most other existing forms of professional development.

First, the Fellows' surveys from over 1200 Fellows at all four existing sites from 2003-2008 show that teachers *chose to participate* in Institute seminars out of desires to improve themselves in precisely the areas that research indicates to be vital to teacher quality. When asked to choose among 12 reasons for participation in the seminars, the teachers at every site listed the "opportunity to develop materials to motivate my students" as their leading reason. As in the previous study, the sites displayed minor variations, but teachers everywhere also rated three other motives highly.³²

- A. *Desire for materials to motivate students*: cited by **93.2%-94.7%** of Fellows at the four sites.
- B. *Desire for curriculum fitted to teachers' needs*: cited by **84.6%-89.3%** of Fellows.
- C. *Desire to increase teachers' mastery of their subjects*: cited by **85.5%-90%** of Fellows.
- D. *Desire to exercise intellectual independence*: cited by **85%- 91.6%** of Fellows.

GRAPH 1



Pittsburgh teachers placed slightly greater relative stress on improving their mastery of their subject matter than teachers at the other three sites, while Philadelphia teachers cited this motive somewhat less often than teachers in the other cities. In New Haven and Philadelphia, the opportunity to work with university faculty members was another strong motivation, cited by 89% of the teachers in both those cities, as compared with 82.1% in Houston and 68.7% in Pittsburgh. A similar pattern is visible in the survey results of teachers who have been "National Fellows" in

recent years. Since 2005, the National Initiative has conducted National Seminars in New Haven for teachers from cities contemplating starting Institutes, as well as for Fellows from existing sites. These seminars provide teachers with experiences comparable to those in local Institutes in a more compressed time frame. Teachers are accepted into seminars and begin research in April; meet twice in May and then ten times in July; and they complete their units in August. The National Fellows listed the opportunity to work with Yale faculty most often as their motivation for spending substantial time in New Haven in May and July (96.4%), followed closely by their desires for curriculum that could motivate their students (95.5%), for increased mastery of their subject matter (94.6%), and for opportunities to exercise intellectual independence (93.7%).

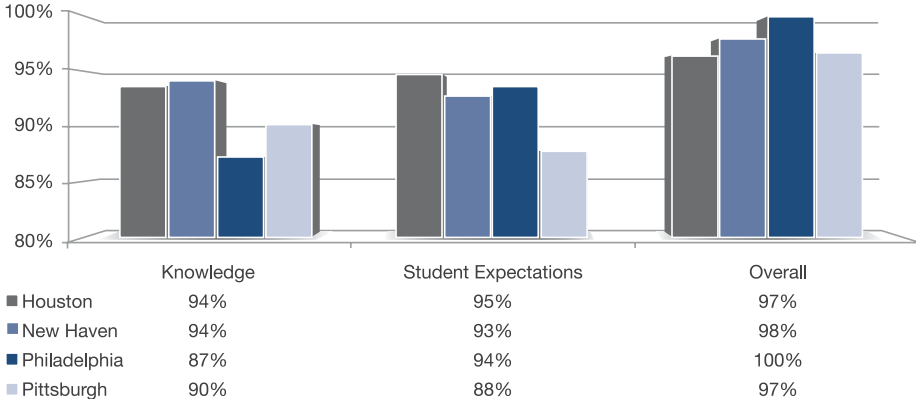
These responses from the local Institutes in New Haven and Philadelphia and from the National Initiative seminars may indicate the comparative appeal for teachers of working with faculty in Ivy League institutions. Nonetheless, teachers at all the sites rated their overall experiences at their Institutes very similarly. In all four local Institutes and in the national seminars, high percentages of teachers chose Teachers Institute seminars because they wanted to increase their content knowledge; to create curriculum that fit their professional needs and exercised their minds; and again, above all, to develop curriculum that could motivate their students.

There were also interesting variations in the responses among categories of teachers, though none were consistent across all four local sites and the National Initiative seminars, and they cannot be considered statistically significant.³³ At three of the four sites middle-school teachers were somewhat more likely than high-school or elementary-school teachers to stress the goal of motivating their students (rather than high-school teachers, as in the National Demonstration Project). Elementary-school teachers most stressed the goal of developing curriculum that fit their needs. Older teachers and those with post-B.A. degrees were often particularly likely to emphasize their desire to increase their control over the curriculum they teach, as were teachers in the humanities.

After completing their seminars, teachers at all four local sites overwhelmingly reported that they had benefited along dimensions that the professional literature indicates to be vital to teacher quality. In contrast to most professional development programs, the surveys revealed no widely shared criticisms of the seminars. Instead, *over 96% of participating teachers praised the overall program, rating it "moderately" or "greatly" useful*, in a remarkably consistent range running from 96.5% in Houston to 100% in Philadelphia. Only 3% of the respondents said the program was useful only to a small extent, with the remainder non-responses. Fellows also overwhelmingly "agreed" or "strongly agreed" that the seminars provided them with professionally useful new *knowledge* and that the seminars raised their *expectations of their students*:

- *Agreed or Strongly Agreed seminars provided useful knowledge: 87.3%-94.2% of Fellows.*
- *Agreed or Strongly Agreed seminars raised their expectations of students: 87.8%-94.75% of Fellows.*
- *Program overall rated Moderately or Greatly Useful: 96.5%-100% of Fellows.*

GRAPH 2: MAIN DIMENSIONS OF TEACHER QUALITY



The patterns among National Seminar Fellows were even more strongly favorable: 98.2% said those seminars increased their professionally useful knowledge, 96.9% said the seminars raised their expectations for their students, and 99.6% rated their experiences overall as "moderately" or "greatly" useful. These data strongly support the conclusion that virtually all teachers who complete Institute seminars feel substantially strengthened in their mastery of content knowledge and their professional skills more generally, while they also develop higher standards for what their students can achieve.

As in the National Demonstration Project, the surveys show other variations by categories of teachers, though again these patterns are not statistically significant. At most sites, elementary-school teachers were most likely to say that the seminars had increased their knowledge. And perhaps surprisingly, teachers with more than 10 years of teaching experience most often indicated that the seminars had raised their expectations for their students — a pattern also discerned in the National Demonstration Project. But whereas during that Project science teachers were slightly more likely than other teachers to list their seminar experiences as "moderately" rather than "greatly" useful, in these surveys teachers currently teaching or planning to teach in the humanities, social sciences, or the sciences showed no consistent variations across the sites in their program evaluations.³⁴ All rated their seminar experiences very highly.

The Curriculum Units in the Classroom

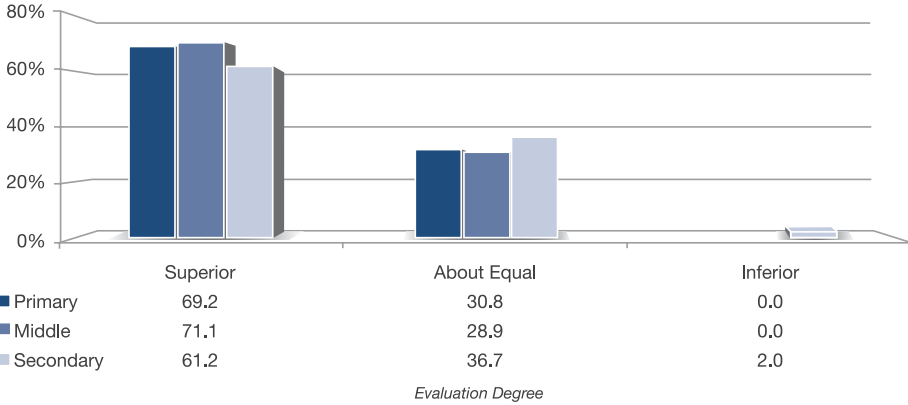
The National Demonstration Project included a one-time survey on unit use; and in New Haven, the annual survey of teachers provides an opportunity for open-ended comments on teachers' previous use of Institute curriculum. Though such comments are not available from the other Institutes and so they do not constitute a systematic national data set, they indicate that in this regard, too, the patterns discerned in the National Demonstration Project appear to persist.

The results of the National Demonstration Project survey showed that, except for those who are unable to do so due to shifts in assignments or health reasons, virtually all Fellows at each site went on to teach the units they prepared, in whole or in part; and many also reported use of other Teachers Institute units. About 87% taught their units in from 2 to 5 classes, usually over a full academic year, sometimes half a year. Most teachers chose to present their units via teacher-led discussion rather than extensive lecturing. They also stressed writing exercises and activities designed to strengthen speaking, listening, vocabulary and reasoning skills, much more than test taking. About a fifth of the teachers used units to develop math skills, largely but not exclusively teachers in the physical sciences. Much recent research indicates that these teaching methods, employed by teachers with good content knowledge, are especially effective in enhancing student knowledge, critical thinking skills, and problem-solving capabilities.³⁵

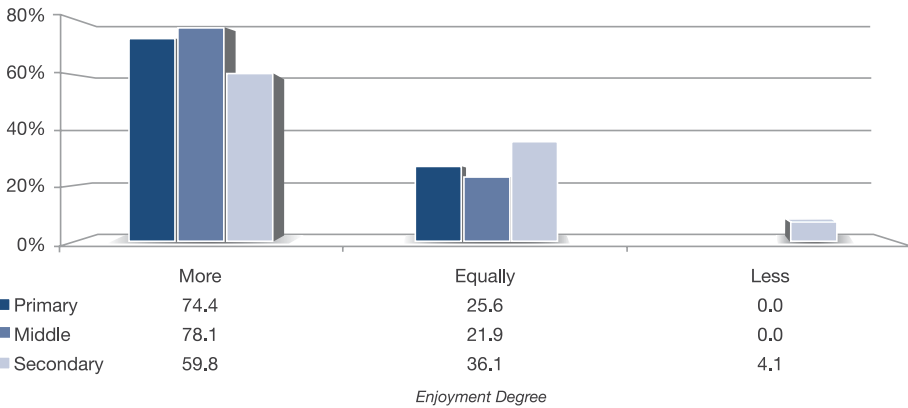
The Fellows in the National Demonstration Project expressed very strong satisfaction with the units. Over 98% rated them either "very useful" (57.3%) or "somewhat useful" (40.9%). Only Fellows teaching in the physical sciences were more likely to rate them "somewhat" instead of "very useful." All teachers particularly valued the sample lessons and activities the units provided, along with the resource lists for teachers and students. Roughly 75% of teachers also compared the units favorably to commercially available curriculum in terms of sample lessons and activities, teaching strategies, resources and other features. In these comparisons, teachers in the physical sciences rated the units as high as or higher than their counterparts in other subject areas.

Overall, an impressive 65% of all Fellows rated the units written by themselves or other Institute Fellows as superior to all other types of curriculum they had used, with almost 34% rating them the same. Roughly the same percentages rated the units as more enjoyable or equally enjoyable to teach than other curricula. There were no statistically significant variations according to the grade levels or subject areas in which Fellows taught. In these regards, physical science teachers again resembled other teachers. Due to their very low numbers, the responses from Fellows using other Fellows' units, as well as from non-Fellows using units, must be interpreted with caution. Still, these responses were essentially indistinguishable from those of Fellows teaching their own units.

GRAPH 3: OVERALL EVALUATION OF UNITS BY FELLOWS BY GRADE LEVELS



GRAPH 4: ENJOYMENT OF UNITS BY FELLOWS BY GRADE LEVELS



It should be noted that during the National Demonstration Project, no Institute made a concerted effort to persuade teachers to use Institute-prepared units instead of other curriculum. Instead, teachers learned about units most often from knowing their authors, other Institute Fellows, or from the Institute Teacher Representatives at their schools — essentially "word-of-mouth" forms of dissemination largely limited to the schools participating in the Institutes — and from the Institute Web sites. Even so, at least 90 teachers used units that they did not write during the National Project, and about half of those had not participated in an Institute.³⁶

Student Responses to Curriculum Units

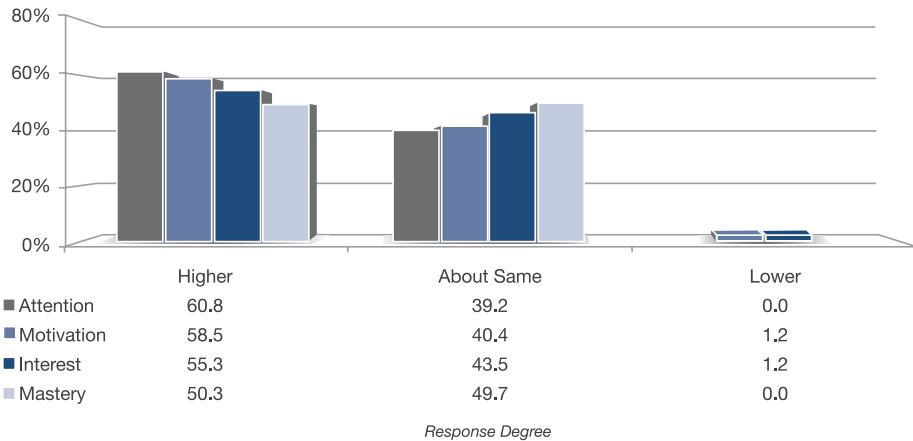
Because the DeWitt Wallace-Reader's Digest Fund, sponsor of the National Demonstration Project, wished evaluators to focus on the Project's impact on teachers, it did not authorize funds for the complex task of directly evaluating the

impact on students of having had teachers who had been Fellows. Teachers using units did, however, provide a great deal of data concerning student responses to those units.

Again, these data were highly positive. Teachers believed these units were especially challenging for students: **55.4%** of responding Fellows rated those challenges as *high*, 43.0% as moderate, and only 1.6% as minimal. Yet they found students responded well to these challenges. **60.4%** rated student interest during these units as *high*, 34.5% rated it as moderate, 5.2% as minimal. In comparison with other sorts of curriculum they had used, Fellows found the Institute-prepared units received strikingly superior student responses. The different grade levels and subject matters in which Fellows taught produced only minor variations, although again science teachers rated units somewhat less highly. Overall:

- **60.8%** of Fellows rated *student attention as higher* during these units than during other work, 39.2% about the same, 0% lower.
- **58.4%** of Fellows rated *student motivation as higher* during these units, 40.4% about the same, 1.2% lower.
- **55.3%** of Fellows rated *student interest as higher* during these units, 43.5% about the same, 1.2% lower.
- **50.3%** of Fellows rated *student content mastery as higher* during these units, 49.7% about the same, 0% lower.

GRAPH 5: STUDENT RESPONSE TO UNITS BY FELLOWS



In commenting on their experiences teaching units in the classroom during the five-year (2000–2005) period of the New Haven quantitative study discussed in the next section, teachers spoke often of how student attention, enjoyment, comprehension and retention improved whenever they used Institute units, both ones they had personally written or ones written by other Fellows.³⁷ Many also reported that their units were written to be interdisciplinary and served as bases for successful team-teaching, sometimes in pairs, sometimes in larger groups. These collaborative

teaching efforts often involved teachers who had not been Fellows, expanding the influence of the Institute's work.³⁸ And teachers frequently stated that Institute curriculum units enabled them to meet "state and district standards" with materials that were more exciting for their students and for themselves, in part because these were units "tailor made to suit" the "personalities and interests" of the students to whom they were taught. In a number of cases, teachers indicated that their units had become part of their school's "overall Comprehensive School Plan."³⁹

Teachers also saw the Institute as "influential in retaining existing teachers" because the teachers find the seminars stimulating and feel "respected and acknowledged as creative, caring, educated colleagues." The Institute, one teacher reported, "fires me back up to go back to the trenches."⁴⁰ These comments suggest that in New Haven, and presumably at the other similarly structured Institutes, Institute curriculum units are helping both Fellows and non-Fellows to meet state and district curricular goals, often through collective teaching efforts, using materials that improve student interest and learning. For many teachers, Institute experiences reinforce their desires to continue working in the nation's most challenging educational environments.

Site Variations

As in the earlier National Demonstration Project, the populations of Fellows participating at the four current sites from 2003-2008 display some demographic differences. In comparison with the other local sites and the National Seminar participants, New Haven Fellows were roughly twice as likely to be under 30. Although all sites attracted similar ratios of more experienced and less experienced teachers, Philadelphia and, particularly, National Fellows were somewhat more likely to have less than 10 years teaching experience. Both Philadelphia and National Fellows were also more likely to be high-school teachers, though in Philadelphia this pattern arose as a condition of one of the Institute's grants. At all Institutes, teachers in the humanities were the most frequent participants, though the varying, often interdisciplinary responsibilities of teachers make these categorizations difficult, and for Pittsburgh they are not fully available. New Haven, Philadelphia, and National Fellows were more likely to have advanced degrees, Master's or PhD's, than those in Houston and Pittsburgh.

DEMOGRAPHIC AND EDUCATIONAL CHARACTERISTICS OF PARTICIPANTS

	<i>Houston</i>	<i>New Haven</i>	<i>Philadelphia</i>	<i>Pittsburgh</i>	<i>National</i>
Total Fellows	489	290	117	385	223
Age Group	[478]	[284]	[117]	[379]	[221]
21-30	90 (18.8%)	101 (35.6%)	17 (14.5%)	47 (12.4%)	34 (15.4%)
31+	388 (81.2%)	183 (64.4%)	100 (85.5%)	332 (87.6%)	187 (84.6%)
Total Years Teaching	[477]	[273]	[117]	[364]	[218]
1-9	116 (24.3%)	69 (25.3%)	39 (33.3%)	82 (22.5%)	83 (38%)
10+	361 (75.7%)	204 (74.7%)	78 (66.7%)	282 (77.5%)	135 (61.9%)
Grade	[489]	[287]	[117]	[385]	[223]
Secondary 9-12	165 (33.7%)	125 (43.5%)	63 (53.8%)	149 (38.7%)	112 (50.2%)
Middle 6-8	157 (32.1%)	45 (15.7%)	18 (15.4%)	51 (13.2%)	61 (27.3%)
Primary K-5	146 (29.8%)	87 (30.3%)	26 (22.2%)	86 (22.3%)	37 (16.6%)
Multiple Grades	21 (4.3%)	30 (10.4%)	10 (8.5%)	99 (25.7%)	13 (5.8%)
Field	[489]	[290]	[117]	[385]	[223]
Humanities	256 (52.3%)	105 (36.2%)	48 (41%)	(34%)*	105 (47.1%)
Soc. Sci. and History	78 (15.9%)	64 (22%)	27 (23.1%)	(17%)*	46 (20.6%)
Science and Math	77 (15.7%)	72 (24.8%)	33 (28.2%)	(32%)*	66 (29.6%)
Special Ed. and Other	78 (15.9%)	49 (16.9%)	9 (7.7%)	(17%)*	6 (2.7%)
Highest Degree	[489]	[290]	[117]	[385]	[223]
Bachelor's	255 (52.1%)	72 (24.8%)	23 (19.7%)	261 (67.8%)	64 (28.7%)
Master's or PhD	234 (47.8%)	218 (75.2%)	94 (80.3%)	124 (32.2%)	159 (71.3%)

Numbers in brackets indicate numbers of Fellows who completed surveys in that category.

* indicates percentages estimated due to coding omissions.

Corollary Benefits

Experiences at the four sites during the three years of the Demonstration Project, in New Haven for three decades, and in Pittsburgh, Houston, and Philadelphia since a Demonstration Project, all indicate that the Institute approach generates significant corollary benefits that are not easily grasped through survey responses and not always visible in a relatively short time period. Perhaps the most important of these include:

- The development of *teacher leadership* capabilities, as many teachers serve as Teacher Representatives or Seminar Coordinators;
- The development of *teacher collaborations and teacher networks*, as teachers gain knowledge of who their fellow teachers in other subjects, at other grade levels, and in other schools are, and what they are doing in their classrooms;
- The development of *university faculty who see themselves as partners in improving public education*, rather than passive, often critical recipients of its graduates;
- The development of *university-public school institutional partnerships* in ways that promote respect and strengthen education in both settings.

Qualitative studies in Houston, Pittsburgh, and New Haven provide many strong testimonials by teachers, university faculty members, and university and public school administrators of the Institutes' benefits in all these regards.⁴¹

V. THE NEW HAVEN QUANTITATIVE EVALUATION

To augment those findings, the Yale National Initiative sponsored a retrospective quantitative study of the impact of the Teachers Institute on New Haven teachers and students from the 2000-2001 academic year to the 2004-2005 academic year. The study was based on a state of the art multilevel design by Dr. Ellen E. Kisker of Twin Peaks Partners and executed with the aid of graduate student research assistants from the Yale-New Haven Teachers Institute and the University of Pennsylvania Graduate School of Education.⁴² It used data compiled by the Yale-New Haven Teachers Institute and by the Department of Research, Assessment, and Student Information in the New Haven Public Schools. The data allowed researchers to identify all teachers teaching in the New Haven schools in these years, including all who had been Institute seminar participants before or during the study period; the schools in which all teachers taught; and the students of all teachers during the study period. The research also included demographic and performance data for those teachers, schools, and students. Because randomized assignment of teachers to seminars was impossible and inappropriate, the study relied on the second approach noted above: comparing the outcomes for teachers participating in Institute seminars (Fellows), their schools, and their students with comparison groups of nonparticipating teachers, schools, and students.⁴³

Not all of the data sought proved available, particularly in regard to teacher and student outcomes. The study design called for obtaining information on teacher retention, promotion to leadership positions, awards, and attendance, and data on student test scores, grades, awards and recognition, attendance, and promotion to higher grades. In practice, researchers were only able to document teacher retention and attendance and student performances on standardized tests and course grades.⁴⁴ These results were examined both at the level of individual teachers and students and at the school level.

Hypotheses

In regard to the demographics of teachers who participate in Institute seminars, the chief hypothesis was that they would be broadly representative of all New Haven teachers, not drawn primarily from particular demographic subsets, with the partial exception of elementary-school teachers.⁴⁵ In regard to teacher outcomes for which data were available, the chief hypothesis was that participation in Teachers Institute seminars would correlate with heightened teacher retention, as the seminars would increase teacher enthusiasm and satisfaction with their teaching. The fact that teachers self-selected to apply for seminars means, however, that the analysis can

only identify correlations, not support claims for causality. The study cannot rule out the possibility that teachers who choose to participate in the Institute are more likely to continue teaching in New Haven for reasons other than Institute participation. It was also hypothesized that Institute participation might be positively but weakly associated with teacher attendance (an outcome affected by many factors, including release time for professional development and service).

In regard to student outcomes for which data were available, the central hypothesis was that student exposure to teachers who had been Institute Fellows would be positively but weakly associated with standardized test scores and grades. Because Fellows' curricular units are designed to fulfill general state and district content standards and goals, not to prepare the specific knowledge and skills examined in state standardized achievement tests, no strong impact on test scores was anticipated, even when the math test scores of students of math and science teachers and the reading test scores of students of English and history and social science teachers were examined. Because more effective teachers may also grade more strictly, it was also not clear that even enhanced student interest and performance in courses led by Institute Fellows would be associated with higher grades in relation to students of non-Fellows. Still, in the aggregate, improved student interest and learning should be associated with better test scores and better grades.

Results

A multilevel model was used to account for variation explained at the school, teacher, and student levels. The results at the *school* level did not produce statistically significant results varying from those found at the teacher and student levels, leaving unsettled the question of whether collective teacher participation by school produces benefits beyond individual teacher participation. In regard to *teachers*, there was no statistically significant relationship between participation as a Fellow and teacher attendance. But the hypothesis that Fellows are roughly demographically representative of New Haven teachers as a whole was supported by the data. Women were found to be slightly more likely than men to become Fellows, but the difference was not statistically significant. Participation by teachers with and without master's degrees was virtually identical. Though as a group, teachers who at some point had been Fellows had a mean of 17.02 years of teaching experience while those who had never been Fellows had a mean of 10.56 years of teaching, a difference statistically significant at the 95% confidence level, that difference derives from the fact that teachers who had taught longer had more opportunities for seminars. The teachers who were Fellows in 2005 had a mean of 7.38 years of teaching experience, while non-Fellows in that year had a mean of 12.45 years, an equally statistically significant difference that shows that many less-experienced teachers were taking seminars. On average, teachers who were Fellows during the study period had 1.87 fewer years of experience than non-Fellows, a difference that was *not* statistically significant. In short, the seminars attracted both new and experienced

teachers, so that over time, experienced teachers were more likely to have been Fellows in one year or another. As anticipated, holding all else equal, teachers in higher grades were also more likely to be Fellows than teachers in lower grades.

During the study period, African American teachers were 83% more likely than white teachers to be Fellows when holding years of teaching experience, sex, subject, grade level, and degrees constant. This result is statistically significant at the 95% confidence level; but the Yale-New Haven Teachers Institute actively sought to encourage participation by African American teachers during these years. The data show that these efforts succeeded and that the Institute was at least as attractive to these racial minority teachers as to whites.

In regard to retention, of those teachers who had been Fellows by the end of the first year of the study, 2000-2001, 63.25% were still teaching in New Haven at the end of the study period, versus 42.71% for those who were non-Fellows at the end of 2000-2001. Because these descriptive differences in teacher retention could be traceable to other teacher characteristics, the researchers ran an inferential statistical test to illuminate the Fellow correlation. Fellows were almost twice (1.93 times) as likely to remain teaching in the district five years later, controlling for race, sex, and years of teaching experience, a statistically significant result ($p < .001$). Overall, over 50% of those teaching in New Haven in 2000-2001 had stopped being New Haven teachers by 2004-2005, a turnover rate consistent with national averages, especially for urban districts.⁴⁵ That turnover percentage was substantially less for those who had been Fellows. Again, the study design does not permit a claim of causality. But in light of the high percentages of New Haven teachers who become Fellows, it is reasonable to view this correlation between participation as a Fellow and teacher retention in the district as substantively, as well as statistically, significant. Because research suggests that experience within a district is more strongly associated with teacher effectiveness than earlier experiences elsewhere, this finding is especially notable.⁴⁷

In regard to *students*, no statistically significant relationships were found between having a teacher who had been a Fellow and student performances on standardized tests and in course grades. Though unsatisfying, this result is not surprising. It underlines the need for future evaluations to use data more closely tied to the goals of the curriculum units written in Institute seminars, and the need for the accumulation of much more extensive data on student outcomes. If the impact on standardized test scores and grades is, as hypothesized, positive but highly indirect, those relationships will only be discernible when a great deal of reliable data can be analyzed.

VI. CONCLUSIONS AND RECOMMENDATIONS FOR CONTINUING EVALUATIONS

The continuing positive results of annual surveys of teachers at each Institute site and of National Seminar participants leave little doubt that teachers consistently

rate their Institute experiences and the curriculum units that result favorably along the five dimensions agreed to be key ingredients of teacher quality: teacher knowledge, teacher skills, teacher enthusiasm, teacher expectations of their students and teacher capacity to motivate students. Though we have less data on teachers' experiences in using Institute curriculum units, those data are also positive. The New Haven quantitative study indicates that Institute seminars attract a broad range of teachers from every observable demographic category and that those who choose to be Fellows are much more likely to continue teaching in the district than those who do not. These results are all the more credible in light of the ways the Institute approach embodies the different elements that researchers have found to contribute to successful professional development programs: a focus on content and pedagogy linked to content; active teacher learning; teacher leadership; duration; alignment with state and local standards; and, somewhat less extensively, collective participation and continuing evaluation.

At the same time, these results underscore the importance of continuing and improved program evaluation. The annual surveys of Fellows at each site and in National Seminars, with high rates of participation guaranteed by the requirement to complete the surveys to obtain Institute stipends, provide invaluable data on teacher responses that need to be sustained. In addition, a number of further steps should be considered. Dr. Ellen Kisker has designed a technical research guide that can assist Institutes and school districts in preparing programs for continuing evaluation.⁴⁸ It describes alternative means that may be appropriate in different settings.

To summarize the main options: Insofar as possible, each site should seek to institutionalize annual data collection on the use of curriculum units and on their results for students. These data could include annual surveys of former Fellows on their use of the units they have written, on their use of units written by other Fellows, and surveys of non-Fellows on their unit use — though past experience suggests that achieving high response rates will be difficult even with modest financial incentives. If resources permit, Institutes might also survey *students* on their experiences of units used in their classrooms; and they might seek to use pre-unit and post-unit student tests to determine the extent of student learning. Institutes should also seek to work with school districts to gather more data on student outcomes that are more closely tied than standardized tests to the content and goals of units written for Institute seminars and used in classrooms. Pertinent data could include student performance on measures of achievement in regard to district and state standards that the curriculum units address; student attendance during the teaching of the unit; and evaluations of student written work fulfilling unit assignments. Less direct but still pertinent evidence of the impact of having had Fellows as teachers might also come in students' relative success in winning academic awards and in graduating. Where possible, these indicators should be supplemented by observational research on the teaching of units or material derived from units, both by Fellows and non-Fellows, in actual classrooms. Analyses of in-depth inter-

views with the Fellows and students in those classes would also be desirable. Ideally, these should be accompanied by comparable observations and interviews with teachers and students not directly affected by Institute experiences.

Institutes should also continue to collect and analyze descriptive data on teachers and on teacher outcomes. It is vital to monitor whether Institutes continue to attract teachers from a broad range of demographics. It would also be useful to assess systematically at every site whether experiences as Fellows are associated with heightened teacher retention and with the promotion of teachers to positions of academic leadership within their districts and beyond.

If circumstances permit, Institutes might also seek to obtain and analyze more systematic data that can help answer the question of whether collective participation by teachers from the same school, grade, or department has benefits exceeding those obtained when individual teachers are mixed without any such clustering. Because teachers' voluntary choices do sometimes produce collectivities of teachers from the same school or grade in particular seminars, and because the number of Fellows teaching in a school often grows over time, there are opportunities to study the patterns of student and teacher performance associated with these clusters; although the impossibility of constructing randomized controlled tests will still prevent claims for causality.

Both individual Institutes and the National Initiative should also seek to identify and employ researchable forms of randomization that are consistent with the Institute approach. One possibility in large school districts may be to assign schools randomly to be targeted for recruitment into Institute seminars, though the availability of units written by Fellows to all teachers in the district would arguably limit the "controlled" character of such a study. Even so, if students of teachers at schools randomly assigned for recruitment appeared over time to perform better on appropriate measures than those from the same pool of schools whose teachers were not eligible for Institute participation, the results would provide additional reasons for confidence in the value of the distinctive Teachers Institute form of professional development.

There are, then, many routes for continuing and improved evaluation that merit consideration, and there are strong reasons for undertaking such steps. The consistently positive results of the great variety of evaluations conducted at multiple sites over the past three decades suggest that, while further research can surely suggest improvements, the fundamentals of the Teachers Institute approach will continue to be shown to be a distinctive and valuable contribution to both teacher and student learning and achievement throughout the nation.

1. This report was prepared for the Yale-New Haven Teachers Institute by Rogers M. Smith, Christopher H. Browne Distinguished Professor of Political Science at the University of Pennsylvania, with the assistance of Dr. Ellen Kisker, Twin Peaks Partners LLC; Jing Zhang, Emily Vince, Cris Malino, Jaime Waters, and You-Liang Deng of the Yale-New Haven Teachers Institute; Catherine Carter of the Teachers Institute of Philadelphia; Christopher P. Dean and David Seidenfeld of the University of Pennsylvania Graduate School of Education; Carla Medalia and Allison Sullivan of the Penn Sociology Department; and Timothy Weaver of the Penn Political Science Department. Special thanks go to Dr. Catherine McCaslin, Director of Research, Assessment, and Student Information of the New Haven Public Schools, and her staff for generous, skilled, and invaluable assistance.
2. Few dismiss the older view, elaborated in the works of James S. Coleman and Christopher Jencks, that socioeconomic backgrounds strongly affect student achievement (Coleman, J. S. et al. *Equality of Educational Opportunity*. Washington, D.C.: Government Printing Office [1966]; Jencks, C. et al. *Inequality: A Reassessment of the Effects of Family and Schooling in America*. New York: Harper & Row [1972]). But evidence suggests quality teachers can succeed with students of all backgrounds. The most widely cited study on the importance of teacher quality is Sanders, W. L. and Rivers, J. C., "Cumulative and Residual Effects of Teachers on Future Student Academic Achievement," University of Tennessee Value-Added Research and Assessment Center, Knoxville, Tennessee, 1996; see also Rivers, J.C. and Sanders, W.L., "Teacher Quality and Equity in Educational Opportunity: Findings and Policy Implications," in L.T. Izumi and W.M. Evers, eds., *Teacher Quality*. Stanford, CA: Hoover Institution Press Publication #505 (2002). Similar conclusions are reached in H. Wenglinsky, "How Schools Matter: The Link Between Teacher Classroom Practices and Student Academic Performance," *Educational Policy Analysis Archives* 10, no. 12 (2002); Hanushek, E. A. and Rivkin, S. G., "How to Improve the Supply of High Quality Teachers," in D. Ravitch, ed., *Brookings Papers on Education*, Washington, D.C.: Brookings Institution Press (2004); Bill and Melinda Gates Foundation Staff, "High-Performing School Districts: Challenge, Support, Alignment, and Choice," Gates Foundation, June 2005, 17, at <http://www.heartland.org/pdf/18328.pdf>, accessed June 23, 2008; Kane, T. J., Rockoff, J.E., and Staiger, D.O., "What Does Certification Teach Us About Teacher Effectiveness? Evidence from New York City," March 2006, 6, at <http://gseweb.harvard.edu/news/features/kane/nycfellowsmarch2006.pdf>, accessed December 29, 2008.
3. Desimone, L.M., Porter, A.C., Garet, M.S., Yoon, K.S., and Birman, B. F., "Effects of Professional Development on Teachers' Instruction: Results from a Three-year Longitudinal Study," *Educational Evaluation and Policy Analysis* 24: 82 (2002).
4. The value of content-area knowledge is stressed in e.g. Monk, D. H., "Subject Matter Preparation of Secondary Mathematics and Science Teachers and Student Achievement," *Economics of Education Review* 13: 125-145 (1994). Researchers who stress the importance of pedagogy acknowledge the comparable importance of content knowledge. See Darling-Hammond, L. and Youngs, P., "Defining 'Highly Qualified

Teachers': What Does 'Scientifically-Based Research' Actually Tell Us?" *Educational Research* 31: 18-19 (2002). See also Frome, P., Lasater, B., and Cooney, S., "Well-qualified Teachers and High Quality Teaching: Are They the Same?" Atlanta, GA: Southern Regional Education Board, April 2005, 2; Odden, A. and Kelly, J.A., "Strategic Management of Human Capital in Public Education," Madison, WI: Consortium for Policy Research in Education, Wisconsin Center for Educational Research, June 2008, 11, 13.

5. The impact on student achievement of teachers with better basic skills, and the impact of teachers' expectations, is documented in Ferguson, R.F., "Teachers' Perceptions and Expectations and the Black-White Test Score Gap" and "Can Schools Narrow the Black-White Test-Score Gap?" in C. Jencks and M. Phillips, eds., *The Black-White Test Score Gap*. Washington, D.C.: Brookings Institution Press (1998), 273-374; see also Wayne, A. J. and Youngs, P., "Teacher Characteristics and Student Achievement Gains: A Review," *Review of Educational Research* 73: 89-122 (2003).
6. Ferguson, "Teachers' Perceptions and Expectations and the Black-White Test Score Gap," 273-317; Ferguson, R. F. and Ladd, H.F., "How and Why Money Matters: An Analysis of Alabama Schools," in H. F. Ladd, ed., *Holding Schools Accountable: Performance Based Reform in Education*, Washington, D.C.: Brookings Institution Press (1996); Kannapel, P. J. and Clements, S. K. with Taylor, D. and Hibpshman, T., "Inside the Black Box of High-Performing High-Poverty Schools," Lexington, KY: Prichard Committee for Academic Excellence, February 2005, 3, 15; Frome et al., "Well-qualified Teachers and High Quality Teaching: Are They the Same?" 1, 4-8.
7. Kannapel et al., "Inside the Black Box of Higher-Performing High-Poverty Schools," 19; Raffini, J. P., *Winners Without Losers: Structures and Strategies for Increasing Student Motivation to Learn*. Boston: Allyn and Bacon (1993), esp. 245-247; Bill and Melinda Gates Foundation Staff, "Challenges to Teaching Capacity in New and Redesigned Schools," Gates Foundation, 2006, 4.
8. Raffini, *Winners Without Losers*, esp. 245-247. Stipek, D. J., *Motivation to Learn: Integrating Theory and Practice*, 4th ed., Boston: Allyn and Bacon (2002); Frome et al., "Well-qualified Teachers and High Quality Teaching: Are They the Same?" 5-8.
9. Goe, L. and Stickler, L. M., "Teacher Quality and Student Achievement: Making the Most of Recent Research," Washington, D.C.: National Comprehensive Center for Teacher Quality, March 2008; Kane, et al, "What Does Certification Teach Us About Teacher Effectiveness? Evidence from New York City," 22-24, 26.
10. Arzi, H. J. and White, Richard T., "Change in Teachers' Knowledge of Subject Matter: A 17-Year Longitudinal Study," *Science Education* 92: 246 (2008).
11. Desimone et al., "Effects of Professional Development on Teachers' Instruction: Results from a Three-year Longitudinal Study," 105; Shkolnik, J., Song, M., Mitchell, K., Uekawa, K., Knudson, J., and Murphy, R., "Changes in Rigor, Relevance, and Student Learning in Redesigned High Schools: An Evaluation for the Bill & Melinda Gates Foundation," Washington D.C.: American Institutes for Research, April 2007, 25.
12. Desimone et al., "Effects of Professional Development on Teachers' Instruction: Results from a Three-year Longitudinal Study," 83; Corcoran, T. B. "Helping Teachers Teach

Well: Transforming Professional Development," CPRE Policy Briefs, RB-16-June-1995. Philadelphia: Consortium for Policy Research in Education (1995). As one Houston teacher put it, "My experience with most professional development is that I dread it. They put me to sleep — they're mind numbing." Lorence, J. and Kotarba, J., *The Houston Teachers Institute: Goals Accomplishments, 1999-2002*. Houston: Houston Teachers Institute, 2003, 75. See also Odden and Kelly, "Strategic Management of Human Capital in Public Education," 5, 20.

13. Desimone et al., "Effects of Professional Development on Teachers' Instruction: Results from a Three-year Longitudinal Study," 82, 102; Yoon, K.S., Duncan, T., Lee, S. W-Y., Scarloss, B., and Shapley, K. L., "Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement," *Issues and Answers Report, REL 2007-No. 033*, Washington, D.C.: U. S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest, October 2007, 5; Blank, R. K., Alas, N., Smith, C., "Does Teacher Professional Development Have Effects on Teaching and Learning? Analysis of Evaluation Findings from Programs for Mathematics and Science Teachers in 14 States," Washington, D.C.: Council of Chief State School Officers, February 2008, 1, 3, 15-19, 26; Goe and Stickler, "Teacher Quality and Student Achievement: Making the Most of Recent Research," 6.
14. Desimone et al., "Effects of Professional Development on Teachers' Instruction: Results from a Three-year Longitudinal Study," 82, 86, 102; Blank, et al., "Does Teacher Professional Development Have Effects on Teaching and Learning?" 3, 5; Kannapel et al., "Inside the Black Box of High-Performing High-Poverty Schools," 3, 13; Gates Foundation Staff, "High-Performing School Districts: Challenge, Support, Alignment, and Choice," 17.
15. Desimone et al., "Effects of Professional Development on Teachers' Instruction: Results from a Three-year Longitudinal Study," 82, 85, 102; Blank, et al., "Does Teacher Professional Development Have Effects on Teaching and Learning?" 3, 5; Gates Foundation Staff, "High-Performing School Districts: Challenge, Support, Alignment, and Choice," 17-18.
16. Desimone et al., "Effects of Professional Development on Teachers' Instruction: Results from a Three-year Longitudinal Study," 102, express surprise at not finding support for the importance of duration. Most scholars do, including Blank et al., "Does Teacher Professional Development Have Effects on Teaching and Learning?" 3, 26; Gates Foundation Staff, "High-Performing School Districts: Challenge, Support, Alignment, and Choice," 18; Kannapel et al., "Inside the Black Box of High-Performing High-Poverty Schools," 13; Yoon et al., "Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement," 5; Goe and Stickler, "Teacher Quality and Student Achievement: Making the Most of Recent Research," 6.
17. Desimone et al., "Effects of Professional Development on Teachers' Instruction: Results from a Three-year Longitudinal Study," 82, 86, 102; Blank et al., "Does Teacher Professional Development Have Effects on Teaching and Learning?" 1, 5, 27; Gates Foundation Staff, "High-Performing School Districts: Challenge, Support, Alignment, and Choice," 18; Odden and Kelly, "Strategic Management of Human Capital in Public Education," 26.

18. Gates Foundation Staff, "High-Performing School Districts: Challenge, Support, Alignment, and Choice," 10, 13, 18; 20; Goe and Stickler, "Teacher Quality and Student Achievement: Making the Most of Recent Research," 6.
19. Blank et al., "Does Teacher Professional Development Have Effects on Teaching and Learning?" 1, 5, 26-27, Gates Foundation Staff, "High-Performing School Districts: Challenge, Support, Alignment, and Choice," 10; Kannapel et al., "Inside the Black Box of High-Performing High-Poverty Schools," 13; Yoon et al., "Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement."
20. Yoon et al., "Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement."
21. Desimone et al., "Effects of Professional Development on Teachers' Instruction: Results from a Three-year Longitudinal Study," 105.
22. Gates Foundation Staff, "High-Performing School Districts: Challenge, Support, Alignment, and Choice," 18.
23. Pittsburgh public schools have experienced severe and continuing declines in student enrollments in the decade since the Pittsburgh Teachers Institute was established.
24. Suggesting the value of collective participation, one Philadelphia teacher stated that initially she was the only teacher from her school enrolled in the program and that "I'm so happy now there are four. It makes the level of conversation in your building rise." Popp, T., "Teaching Those Who Teach," *The Pennsylvania Gazette* 106: 21 (2008).
25. By 2008 the total number of New Haven teachers who had been Fellows was 657.
26. To use a political science analogy, it would be similar to trying to evaluate the comparative success of "winner take all" and "proportional representation" presidential primary systems by randomly assigning winning candidates to the two different parties. Little could be learned about the topic in question.
27. Yoon et al., "Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement," 5.
28. "Principal Recognition Accorded to the Yale-New Haven Teachers Institute," New Haven: Yale-New Haven Teachers Institute, at <http://www.yale.edu/ynhti/brochures/A9.html>.
29. "A Progress Report on Surveys Administered to New Haven Teachers, 1982-1990," New Haven: Yale-New Haven Teachers Institute, 1991, at <http://teachers.yale.edu/publications/index.php?page=ReportsEvaluations>.
30. The study can be found at <http://teachers.yale.edu/publications>.
31. Haslam, B. M., Rouk, U., with Laguarda, K. G., "Professional Development with the Yale-New Haven Teachers Institute Model: Impact, Lessons, and Future Prospects," Policy Studies Associates, 2006, 45-46, 52, at <http://www.policystudies.com/studies/professional/ynhti.html>, accessed June 23, 2008.

32. All percentages reported here are percentages of the Fellows who actually responded to particular questions. This section reports data from the Teacher Surveys on Seminar Experiences conducted in Houston, New Haven, and Pittsburgh from 2003-2008 and Philadelphia in 2006-2008. Because the overall completion rate for these surveys was over 98%, percentages of all Fellows surveyed, whether or not they responded to a particular question, are only slightly smaller. Chart figures are rounded to whole numbers; graph bars are accurate to one decimal place. All the survey data reported here can be obtained through a link at the Yale National Initiative Web site, <http://teachers.yale.edu>.
33. Because the numbers of teachers in particular categories were sometimes low, and because some teachers were Fellows multiple times, standard tests for statistical significance are not appropriate. In any case, the variations were not large, generally less than 10 percentage points.
34. In the National Demonstration Project, more experienced teachers, presumably more familiar with alternative forms of professional development, rated the overall program especially highly. Teachers with Master's as well as Bachelor's degrees were more likely than those only with Bachelor's degrees to say the seminars raised their expectations of their students, especially in the humanities. Though teachers in the physical sciences rated their seminar experiences highly, their ratings were slightly less positive than those provided by teachers in the humanities, social sciences, and other areas. For these and other variations found in the evaluation of the National Demonstration Project, see Huang C., "Conclusion of Further Analysis," research memo available at <http://teachers.yale.edu>.
35. See especially Bransford, J. D., Brown, A. L., and Cocking, R.R., eds., *How People Learn: Brain, Mind, Experience, and School*. Washington, D.C.: National Research Council; executive summary available at <http://www.nap.edu/html/howpeople/es.html>; and Cotton, K., "Teaching Thinking Skills," NW Regional Educational Laboratory, available at <http://www.nwrel.org/scpd/sirs/6/cu11.html>.
36. Overall statistics are drawn from the National Demonstration Project Survey of Fellows' Unit Use, Table 111. Data on relative ratings by teachers in different subject areas are drawn from Tables 311 and 321. All tables available at the Yale National Initiative Web site, <http://teachers.yale.edu>.
37. A complete set of these comments is available at <http://teachers.yale.edu>. One teacher commented that whereas the previous year, only 4 out of 20 regularly completed all reading and writing assignments, with the Institute unit those numbers were reversed. Another reported that students showed detailed knowledge of a unit involving films a year and half after it was taught. An elementary teacher said that returning fourth graders immediately asked, "Are we going to do an Institute unit?"
38. These team efforts sometimes involved classes, sometimes projects for student groups. One teacher noted that "In my school, we had projects generated by the seminars of the Institute involving several teachers not all of them Fellows. Science is particularly prone for projects with broad participation of students and staff." Another reported that "we have formed a number of teams, often involving teachers who never were Fellows. These teams have allowed Institute material to benefit many, many children other than those in classes taught by Institute teachers." A third stated, "My school has become a

strong YNHTI center school with a Fellow involvement of at least 40% of the faculty. Not only do we write curricula but we developed a summer school program that utilized written units by other Fellows."

39. Referring to Connecticut state curriculum and standardized tests, various teachers stated that although their units are not designed to improve standardized test scores, the units did "incorporate CAPT activities"; satisfied "the CMT objectives and curriculum requirements"; became "a feature in the school's Science and English curriculum" for years; added "depth" to topics that were "already part of the district's curricula"; and integrated easily into "the math program that is used at my school" and the "mandated 'literacy block' time."
40. Drawn from Brown, J., "Memo on Retention Comments in Fellows' Questionnaires, 2000-2005," 2008, available at the Yale National Initiative Web site, <http://teachers.yale.edu>.
41. For qualitative evidence of these benefits, see the Pittsburgh Teachers Institute Transcripts of Focus Groups with Elementary/Middle School Teachers and Principals, and High School Teachers and Principals; and the publications on the Yale-New Haven Teachers Institute and the National Demonstration Project available at the Yale National Initiative Web site, <http://teachers.yale.edu>.
42. The multilevel design by Dr. Ellen Kisker is available as an online appendix to this report at <http://teachers.yale.edu>.
43. The data set and the data analysis memos by David Seidenfeld on which this report is based can be found at <http://teachers.yale.edu>.
44. The standardized tests employed in the study are the Connecticut Mastery Test (CMT), which was administered in the study period in grades 4, 6, and 8, and the Connecticut Academic Performance Test (CAPT), which was administered in grade 10. During the period of the study, the tests assessed student learning in Mathematics, Reading, and Writing (grades 4, 6, 8, and 10) and Science (grade 10).
45. Elementary teachers, who were not initially eligible for New Haven seminars, came to be more proportionately represented over the years. During the test period, a number of New Haven elementary schools were made part of K-8 schools. Many teachers were reclassified as K-8 school teachers, leading to an apparent decline in elementary school teacher participation.
46. Ingersoll, R., "Teacher Turnover and Teacher Shortages: An Organizational Analysis," *American Educational Research Journal* 38: 499-534 (2001).
47. Kane et al., "What Does Certification Teach Us About Teacher Effectiveness? Evidence from New York City," 30.
48. The technical research guide, designed by Dr. Ellen Kisker, is available as an online appendix to this report at <http://teachers.yale.edu>.

