Full-Release and Site-Based Mentoring of New Elementary Grade Teachers: An Analysis of Changes in Student Achievement

STEPHEN H. FLETCHER and MICHAEL A. STRONG

University of California, Santa Cruz, USA

Induction support for new teachers is widespread, particularly in the form of mentoring, but research evidence of effectiveness is limited. The majority of existing research has focused on the impact of induction on teacher retention. Of greater interest is the potential impact on student achievement, and on which forms of support are the most effective. One frequently encountered option is between full-release or site-based mentors. This study examines these two mentoring options employed in one large urban district. While mentors received the same training, they differed in case load and release time. A comparison of student achievement gains for classes taught by fourth and fifth grade new teachers, some of whom were supported by full release mentors and some by site based mentors, showed greater gains for classes of teachers in the full-release group, even though the demographic characteristics of the students would have led to the opposite prediction.

INTRODUCTION

Starting with the school reform movement of the 1980s, state agencies, university education departments, and school districts have been developing increasingly multifaceted programs to support new teachers. Notwithstanding the recent IES

report (Glazerman et al., 2008) of a randomized controlled study that shows no significant advantages of a comprehensive induction treatment for new teachers over a district's standard program, numerous other studies have shown that new teachers do appear to benefit from comprehensive induction support, in that the more intensive the program the less likely they will quit teaching (Ingersoll & Kralik, 2004; Kapadia et al., 2007; Smith & Ingersoll, 2004; Strong, 2005).

Variation across new teacher induction programs is likely to be found not only in the degree of their comprehensiveness, but also the nature of the program elements. Even the most frequently encountered support component, mentoring (Smith & Ingersoll, 2004), may differ according to the degree of formality of the mentoring relationship, the selection and training of mentors, the amount of release time a mentor is given, the amount and nature of the support mentors give teachers, and the appropriateness of match between mentor and mentee. The outcomes of interest regarding the potential impact of induction programs include teacher retention, student achievement, teaching practice, and participant satisfaction.

In the early 1990s before mentoring programs were widespread, Feiman-Nemser and Parker (1992) compared two different models, or contexts, by interviewing and shadowing mentors in Los Angeles and Albuquerque. The Los Angeles mentors were trained and given a stipend but continued to teach full time, and so mentored "around the edges". In Albuquerque, however, each mentor was trained and released from teaching full time, to support a caseload of 10–12 novice teachers. In this study, the outcome of interest was the kind of mentoring that happened in each context and its implied effect on teaching practice. Feiman-Nemser and Parker refer to the Los Angeles mentors as "local guides," who helped novices to feel comfortable and have a successful first year, but, because of time constraints, could not represent the broader image of mentoring reflected in their training and help their new teachers learn new skills and strategies. On the other hand, the Albuquerque mentors were described as "educational companions," who had the time to help teachers develop and study their practice. The implication is that the educative mentoring found in the Albuquerque context is likely to be the more effective approach to improving teaching practice.

In a second, slightly earlier, comparative study, Klug and Salzman (1991) studied differences between formal induction and informal mentoring. Another way of characterizing their study is that they, like Feiman-Nemser and Parker, examined the effects of release time for mentors. They followed teachers for two years, randomly assigning them to either a buddy mentor or formal induction system. They collected data through questionnaires, interviews, and videotapes of teaching practice, with a view to measuring teacher attitudes (as an indicator of professional growth), teaching performance, and implementation of the two induction program models. They concluded that, even though they were unable to see any improvements in practice for either group, structured models of teacher

induction for new teachers are preferable to buddy systems, both in terms of the attitudes of the participants and of the school administrators.

Smith and Ingersoll (2004) used data from the national Schools and Staffing Survey (SASS) to consider the cumulative effects of various induction components on one-year retention. They identify "basic induction" as having a mentor in supportive communication with administrators; one level up is "basic induction plus collaboration," which adds new teacher seminars and either common planning time or collaboration with other teachers; next comes "basic induction plus collaboration plus teacher network plus extra resources," where participation in an external teacher network, a reduced number of preparations, and a teacher's aide are added to the other types of support. Only 1% of the beginning teachers in their dataset had the full package, 26% had the next level of support, 56% had basic induction, and 3% had no induction at all. They demonstrated that the greater the number of induction components, the lower the turnover rates for both movers (those who changed schools) and leavers (those who quit the profession) after one year of teaching. However, the attrition rates for those teachers who had received basic induction (18% leavers and 21% movers) were barely different from those who had had no induction (20% leavers and 21% movers). At the next level of induction support the numbers went down to 12% and 15%. Only 9% of the few teachers who received the full package quit the profession, and another 9% transferred to other schools.

Kapadia et al. (2007) surveyed 1,737 novice teachers in Chicago Public Schools in order to look at induction support, teaching experiences, and retention. They determined levels of induction and mentoring support, dividing them into three groups: weak, average, and strong. About one fifth of the teachers reported that they were not involved in any induction program, even though it was a requirement. The researchers did not look at actual retention or turnover data, but measured the influence of participation in induction programs on three outcomes: the novice teachers' teaching experiences, their intentions to stay in teaching, and their intentions to stay in the same school. They found that, when adjusting for school context, participation in an induction program, by itself, had little effect on the three outcomes. However, the teachers in the strong induction group showed higher levels on all three outcomes. Mentoring was an important component, especially at the elementary level, but the comprehensive induction supports from all sources had the most effect on intentions to remain in the same school. Kapadia and colleagues conclude that programs should focus on selection and training of mentors to ensure they give high levels of support, but that teacher collaboration and principal assistance have greater influence on novices than mentoring.

Fewer studies have attempted to examine the influence of mentoring on student achievement. One exception is Rockoff (2008), who studied the impact of mentoring for beginning teachers in New York City on both retention and student achievement. His most consistent finding was that retention within a particular school was higher when a mentor had previous experience working in that school,

suggesting that an important part of mentoring may be the provision of school specific knowledge. He also found evidence that student achievement in both reading and math was higher among teachers receiving more hours of mentoring, lending credence to the assumption that more time with a mentor improves teaching skills. It leaves us wondering, however, if there is some kind of tipping point where the amount of mentoring time makes a difference, or what shape is formed by the curve representing the relationship between mentoring time and student achievement.

Fletcher, Strong, and Villar (2008) studied three models of teacher induction in different school districts. At one site, mentors worked full time for two years with a caseload of 15 new teachers. In the other two districts, mentors worked full time for the first year, but in the second year either caseloads were increased to 35 or the teachers received the services of an on-site mentor with no release time. Using regression analysis on the class-level value-added test score data, the authors found that classes taught by teachers who had the services of a full-release mentor over two years showed higher gains than classes of teachers in the other groups, suggesting, but by no means definitively, that mentoring can have an effect on student achievement if mentors have concentrated contact time over two years.

In the IES study referred to above (Glazerman et al., 2008) researchers from Mathematica Policy Associates used a randomized control trial (RCT) design to compare comprehensive induction treatment with the default standard support that is available to new teachers in many districts as a control. The advantage of RCT methodology is that one is able to make causal connections to the outcome variables. The disadvantages are that such studies are expensive and difficult to implement in education settings. Unfortunately there are too many opportunities to compromise the potential advantages of an RCT design, and thus lead one to have reservations about the findings. A reviewer of this study is likely to have many such reservations. The study concluded that, after one year of treatment, new teachers receiving comprehensive induction support showed no apparent differences from teachers receiving standard district support with regard to student academic achievement, teaching practice, retention, or attitudes. The researchers appeared to focus their attention more on the randomization strategies than on the control of the treatment. While comprehensive induction was made available to the treatment teachers, many of them appeared not to take advantage of it. Seven percent reported that they did not even have an assigned mentor, a fundamental component of the comprehensive induction treatment. At the same time, three quarters of the teachers in the control group had an assigned mentor and, on average, spent only 12.5 fewer hours with their mentors (assigned or not) over the entire school year compared to the teachers in the treatment group. The treatment consisted of hybridized versions of existing programs, modified especially for the study. Thus, not only were two-year programs compressed into a single year, but also all the mentors were new to the model. On the other hand, the mentors in the control group were likely to have been experienced providing service in their settings. These, and other problems,

Table 1. Number of teachers and students participating in the study

2006–2007 Grade Level	Type of Mentor	Teachers	Students
4	Site Based	5	69 (86)
	Full Release	11	142 (163)
5	Site Based	7	93 (97)
	Full Release	5	48 (58)

demonstrate why studies such as this are a challenge to researchers, as well as to those who try to make sense of the findings. Sometimes, a less rigorous, but also less expensive and less fragile, quasi-experimental design, such as that reported in this paper, provides a more desirable option that is less subject to flaws that might negate any findings a fully experimental study produces.

METHOD

A large urban school district wanted to improve the support of new teachers by using a mentoring model. The district, though, did not have sufficient resources to have all mentors released from full time classroom duty. The district chose to have some teachers work as mentors full time (full release) and others work as mentors within their own schools in addition to their own teaching schedule (site based). The caseload for full-release mentors was 12–15 new teachers and one or two teachers for site-based mentors. The variation in mentor assignment provided the district with an opportunity to look at how release time and caseload differences may be related to changes in class level student achievement.

One factor that needed to be controlled for in the study was the type of professional development received by the mentors. Past research has shown that new teachers supported by mentors who receive training in how to mentor are more likely to make changes in instructional practice (Evertson and Smithey, 2000). As instructional practice may be related to changes in student achievement, the district made sure that full-release and site-based mentors received the same professional development.

Table 1 shows the number of teachers and students involved in the study. As indicated, the study focused on teachers who taught fourth and fifth grades in 2006–2007. The district provided spring 2006 and spring 2007 achievement data on all students taught by the new teachers. The information in the Students column of Table 1 shows the number of students included in the analysis and the number of students, in parentheses, in the database. The difference between the numbers reflects the fact that some students did not have test data for both years. Comparison of the possible and actual count indicates that the majority of students had test data for both years and could be included in the analysis.

Table 2. Summary of 2006–2007 students' demographic characteristics by type of mentor.

		Grade 4		Grade 5	
		Site Based	Full Release	Site Based	Full Release
English Language Arts	LEP	+			+
	Low Income	+		+	
	Prior Achievement	+		+	
Mathematics	LEP	+			+
	Low Income	+		+	
	Prior Achievement	+		+	

Test Information

The study used the results of the state testing program. The assessments were developed to monitor students' learning with respect to the state's curriculum standards. Following a test administration, the state sends results to students, schools, and districts. For spring 2006 and spring 2007, English Language Arts and mathematics assessments were administered in grades 3–8 and 10, science and technology were administered in grades 5 and 8, and history and social science was administered in grades 5, 7 and high school. As we were interested in the change of student achievement across consecutive years, we chose to focus on English Language Arts and mathematics.

Test results are reported in terms of raw score (number of correct answers), scale score, and proficiency level. At the elementary level, raw scores for the English Language Arts test range from a maximum of 48 in grade 3 to 52 in grade 6. For the same grade level, the maximum raw score for mathematics ranges from 40 to 54. Scale scores are a transformation of raw scores and range from 200 to 280. The scale scores define proficiency levels, with 200–218 defined as Warning, 220–238 as Needs Improvement, 240–258 as Proficient, and 260–280 as Advanced. Because of the state's descriptive definition of the advanced proficiency level, there is no scale scores reported for Grade 3 Reading.

The advantage of using scale scores is that cutoff points for Advanced, Proficient, Needs Improvement, and Warning are the same across grade level and content. Comparison of scale scores across grade levels, then, is a simple subtraction. The grade 3 problem required us to convert raw scores for the different grade levels and assessments to z-scores using state-level means and standard deviations. By converting to z-scores, the results should be interpreted in terms of a student's performance with respect to grade level peers within the state. Thus, a negative gain score indicates a student performed worse than the average student, a zero gain score represents growth similar to an average student, and a positive gain indicates greater growth than the average student.
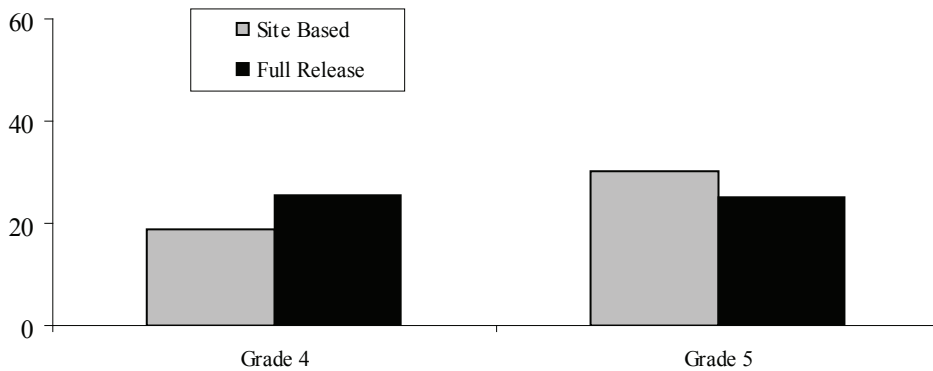
Figure 1. Percentage of former and limited English proficient students taught by new teachers in 2006–2007 by type of mentor.

Student Background

As student characteristics can account for differences in learning, it is important to look at similarities and differences in students taught by the new teachers supported by the two types of mentors. The results of our analysis, summarized in Table 2, indicate that, from student characteristics alone, achievement gains of students taught by site-based mentors should exceed the gains of students taught by full-release mentors.

LIMITED ENGLISH PROFICIENCY

The National Assessment of Education Progress results for 2007 indicate fourth grade Limited English Proficient (LEP) students score significantly lower than other students (Lee, Grigg, & Donahue, 2007). Mercado (2001) has argued that the lower achievement may be related to teachers not being prepared to teach LEP students. Whatever the reason, LEP status may be related to changes in student achievement.

Figure 1 shows the percentage of former and current limited English proficient students taught by new teachers in 2006–2007. As indicated, new teachers, regardless of the type of mentor supporting them, worked with classes consisting of 20–30% of students whose first language was not English. With the exception of Grade 4 Full-Release classes, most of the second language students were classified as Former Limited English Proficient. For Grade 4 Full-Release classes, half the students were classified as Former Limited English Proficient.

STUDENT POVERTY

Peng and Lee (1994) and Anyon (1981) have found that socio-economic status may be related to student achievement, in that low SES students tend to have lower
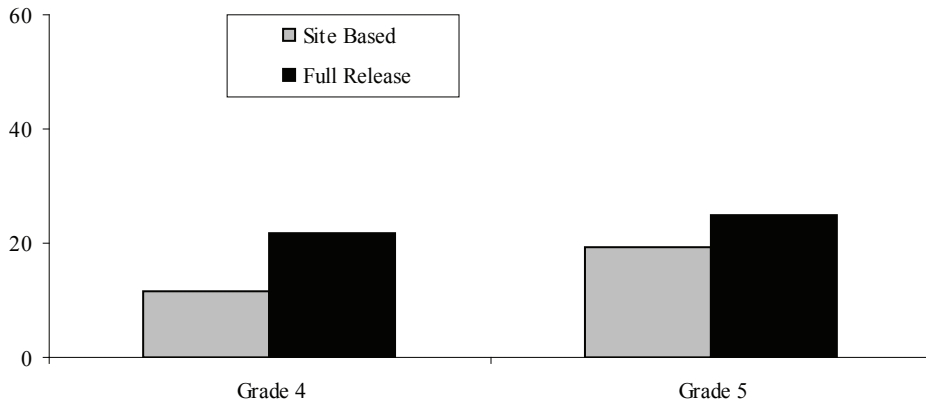
Figure 2. Percentage of students qualified for the free/reduced cost lunch program in 2006–2007 by grade level and type of mentor.

achievement than high SES students. According to the district's state department of education website, students are classified as low income if they qualify for participation in the free/reduced cost lunch program. As indicated in Figure 2, the proportion of low SES students was greater in classes taught by new teachers supported by full-release mentors than classes taught by new teachers in the site-based mentor condition.

PRIOR ACHIEVEMENT

Bower and Hilgard (1981) and Meyer (2000) have argued that prior achievement may be positively related to student learning. To estimate prior achievement, we looked at the mean z-scores for the spring 2006 administration of the English Language Arts and mathematics tests. Figures 3 and 4 indicate that, with respect to English Language Arts and mathematics, students in the Site-Based condition started the 2006–2007 school year with greater achievement, on average, than students in the Full-Release condition.

ANALYSIS AND RESULTS

The change in students' achievement from spring 2006 to spring 2007, disaggregated by grade level and type of mentor supporting new teachers, is illustrated in Figures 5 and 6. Although the hypotheses in Table 2 suggest students in the Site-Based program should have greater gains, the analysis indicates the opposite. In fourth grade English Language Arts and fifth grade mathematics, Full-Release students had greater gains between spring 2006 to spring 2007 than Site-Based students ($F_{1,209} = 24.427$, $p = 0.000$, for the change from grade three to four, $F_{1,140} = 9.279$, $p = 0.003$, for the change from grade 4 to 5). For fifth grade English
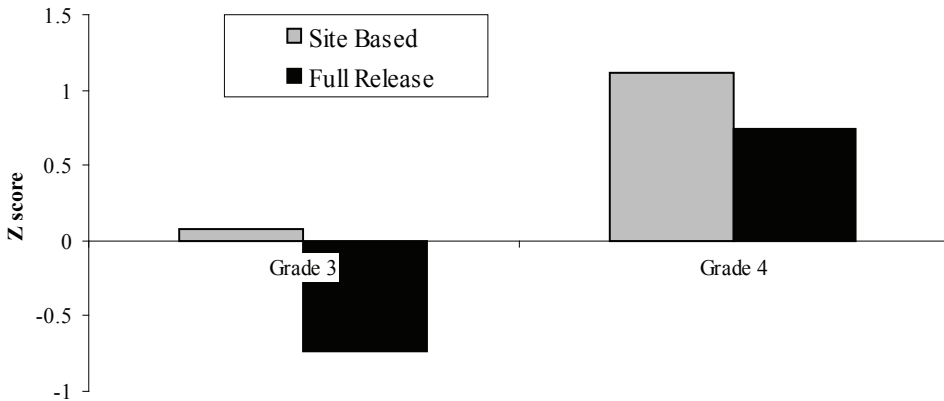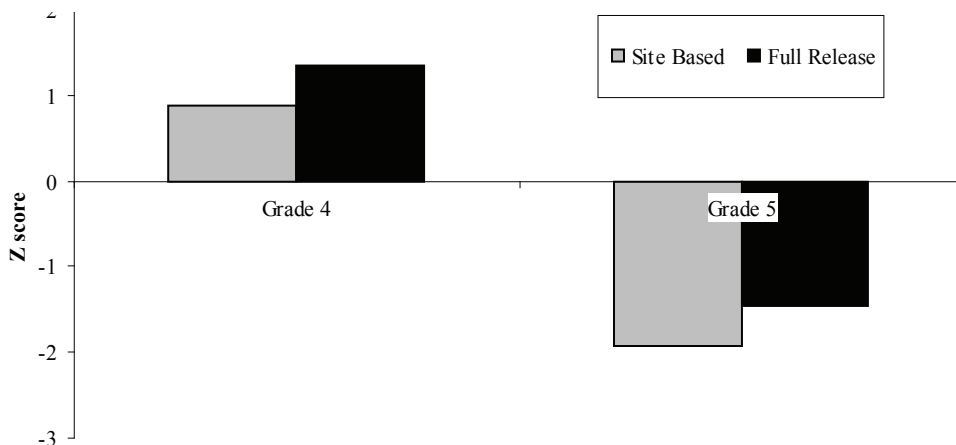
Figure 3. Spring 2006 MCAS English Language Arts test means by grade level and type of mentor.
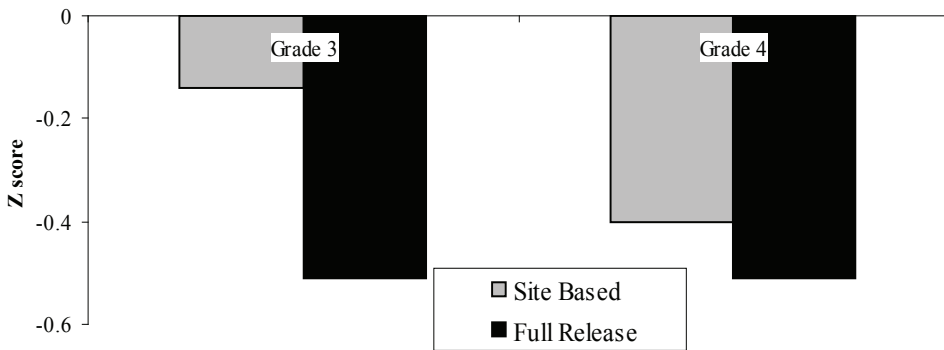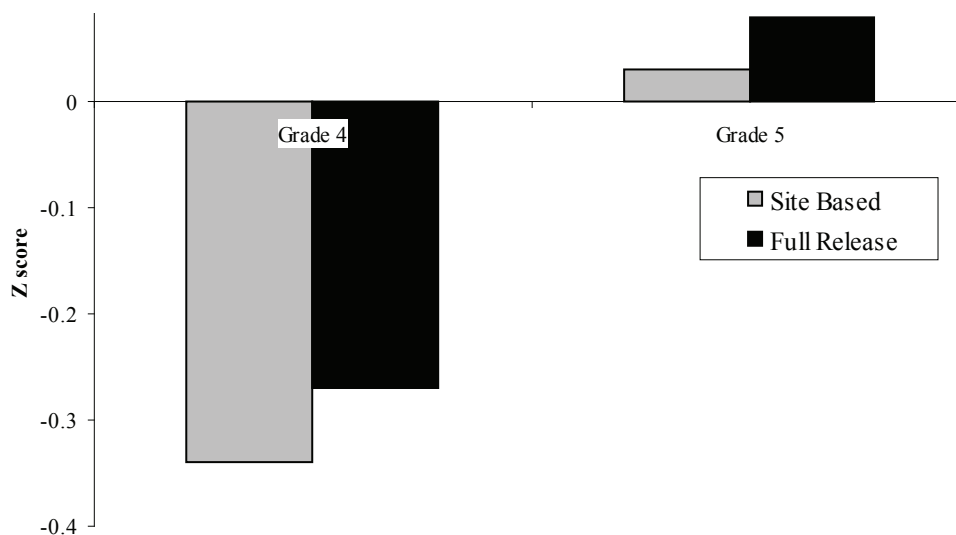
Figure 4. Spring 2006 MCAS mathematics test score means by grade level and type of mentor.

Language Arts and fourth grade mathematics, the negative gains of Full-Release students were less than those of Site-Based students ($F_{1,209}$ = 0.770, p = 0.381, for the change from grade three to four, $F_{1,140}$ = 0.157, p = 0.692 for the change from grade 4 to 5).

While Figures 5 and 6 illustrate the change for all students taught by new teachers, the goal of the No Child Left Behind Act is that all students reach the Proficient level on the assessments by 2014. Because of this goal, we wanted to determine the change in z-scores for students scoring at the Warning and Needs Improvement levels in spring 2006. The results for 2006–2007 fourth grade students are shown in Figure 7 and fifth grade students in Figure 8. For fourth grade students, low achieving students had greater gains in English Language Arts and mathematics if they were in the Full-Release program. The same pattern can be seen in Figure 8, though the gains for both programs were negative.

Figure 5. Change in Z-scores on the MCAS English Language Arts from Spring 2006 to Spring 2007.


Figure 6. Change in Z-scores on the MCAS mathematics from Spring 2006 to Spring 2007.

DISCUSSION

In the present climate of education in this country, where some kind of induction support is widely considered to be necessary for new teachers, there is a pressing need to learn which forms of support are the most effective. In particular, educators and policymakers are interested in programs that may have an impact on student learning. Much of the existing research on mentoring and induction focuses on possible connections with teacher retention, less on any relationship to student achievement. Existing research presents us with mixed findings, even regarding the effects of differing amounts of time spent with a mentor. The purpose of the present
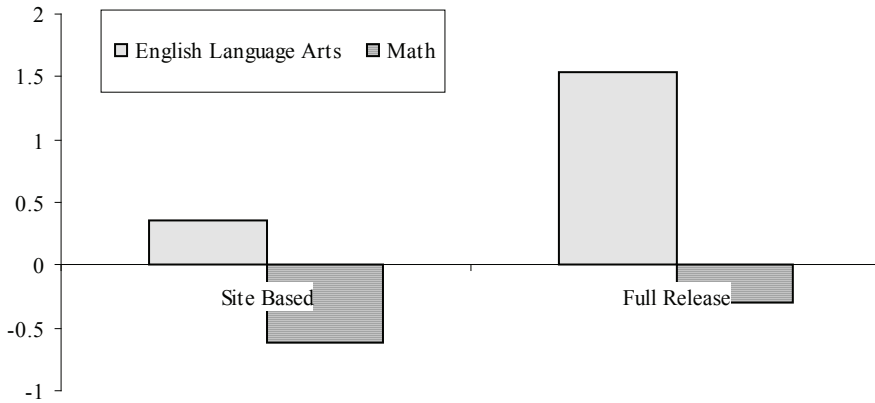
Figure 7. Change in MCAS Z-scores for low performing grade 4 students taught by new teachers in 2006–2007.
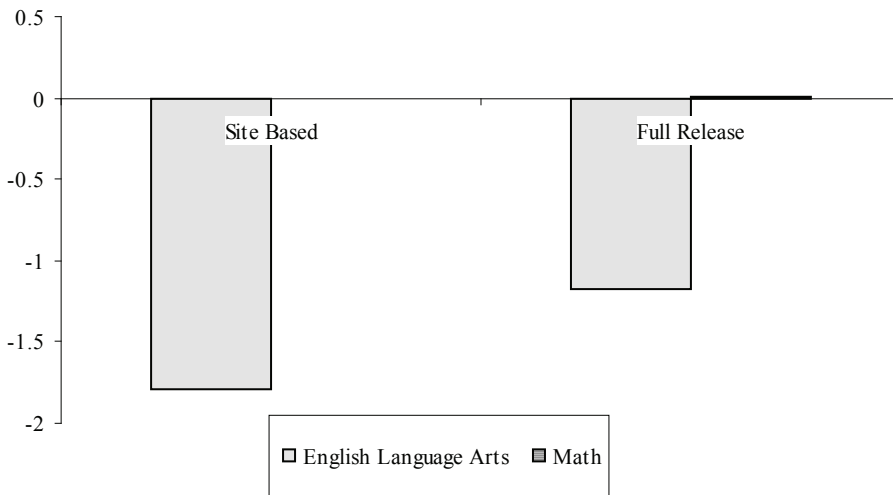
Figure 8. Change in MCAS Z-scores for low performing grade 5 students taught by new teachers in 2006–2007.

study was to look at whether different forms of mentoring (as defined by whether the mentors were fully released from teaching or worked on-site while retaining a full teaching load) may be related to changes in student achievement. We found that whether we focus on fourth or fifth grade, or English language arts or mathematics, students associated with full-release mentors had better achievement gains than students associated with site-based mentors.

The results of this study are interesting because the changes we observed in student achievement do not follow predictions indicated by the extant research literature, given the characteristics of the students and with all other things being equal. While we may reasonably hypothesize that the results are due to the different

levels of intensity of mentoring, it is also possible that they may be accounted for by cross-school differences, or some other unknown factors. We are also sensitive to the fact that the number of classes in the study is smaller than we would have liked. This is where an RCT design has an advantage. However, although the present results should be interpreted with caution, we look for further work to be done, which, if the findings are similar, will add robustness to the findings presented here.

The study also illustrates the value of gradually implementing a program. The district in this study chose to try site-based and full-release mentor models in order to maximize their financial resources. The result was that district leaders learned how to support full-release mentors as well as site-based mentors. Therefore, if the district's financial situation changed, either model could be expanded or contracted. In this way, staged implementation allows policy makers to collect data on program effectiveness. Staged implementation also gives district leaders a way to learn how to incorporate a new program into existing operations. Thus, this study is an illustration of an alternative method of implementing new programs which may be useful to school districts.

REFERENCES

Anyon, J. (1981). "Social class and school knowledge." *Curriculum Inquiry*, *11*(1): 3–42.

Bower, G. H., & Hilgard, E. R. (1981). *Theories of learning*. Englewood Cliffs, NJ: Prentice-Hall Inc.

Esch, C. E., Chang-Ross, C. M., Guha, R., Humphrey, D. C., Shields, P. M., Tiffany-Morales, J. D., Wechsler, M. E., & Woodworth, K. R. (2005). *The status of the teaching profession 2005*. Santa Cruz, CA: The Center for the Future of Teaching and Learning.

Evertson, C. M., & Smithey, M. W. (2000). Mentoring effects on protégés' classroom practice: An experimental field study. *Journal of Educational Research*, *93*(5): 294–303.

Feiman-Nemser, S., & Parker, M. (1992). Mentoring in context: A comparison of two US programs for beginning teachers. NCRTL Special Report. Retrieved June 12, 2009, from http://ncrtl.msu.edu/http/sreports/spring92.pdf

Fletcher, S., Strong, M., & Villar, A. (2008). An investigation of the effects of variations in mentor-based induction on the performance of students in California. *Teachers College Record*, *110*(10), 2271–2289.

Glazerman, S., Dolfin, S., Bleeker, M., Johnson, A., Isenberg, E., Lugo-Gil, J., Grider, M., Britton, E., & Ali, M. (2008). Impacts of comprehensive teacher induction: Results from the first year of a randomized controlled study. NCEE 2009-4034. Washington, D.C.: US Department of Education.

Ingersoll, R., & Kralik, J. M. (2004). The impact of mentoring on teacher retention: What the research says. ECS Research Review, Denver, CO: Educational Commission of the States. Retrieved June 30, 2004 from http://www.ecs.org/clearinghouse/50/36/5036.htm

Kapadia, K., Coca, V., & Easton, J. Q. (2007). Keeping new teachers: A first look at the influences of induction in the Chicago Public Schools. Research Report. Chicago: Consortium on Chicago School Research, University of Chicago.

Lee, J., Grigg, W., & Donahue, P. (2007). The nation's report card: Reading 2007, NCES 2007-496. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, US Department of Education.

Mercado, C. I. (2001). The learner: "Race", "ethnicity", and linguistic difference. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp 668–694) Washington, DC: American Educational Research Association.

Meyer, R. H. (2000). *Value-added indicators: A powerful tool for evaluating science and mathematics programs and policies*. Madison, WI: National Center for Improving Science Education.

Peng, S. S., & Lee, R. M. (1994). Educational experiences and needs of middle school students in poverty. In K. K. Wong & M. C. Wang (Eds.), *Rethinking policy for at-risk students* (pp. 49–64). Berkeley, CA: McCutchan Publishing Corp.

Rockoff, J. E. (2008). Does mentoring reduce turnover and improve skills of employees? Evidence from teachers in New York City, NBER Working Paper no. W13868. Cambridge, MA: National Bureau of Economic Research. Retrieved June 11, 2009 from http://papers.nber.org/papers/w13868

Smith, T., & Ingersoll, R. (2004). What are the effects of induction and mentoring on beginning teacher turnover? *American Educational Research Journal*, *41*(3), 681–714.

Strong, M. (2005). Teacher induction, mentoring, and retention: A summary of the research. *The New Educator*, *1*(3), 181–198.