

Achievement and Climate Outcomes for the Knowledge is Power Program in an Inner-City Middle School

Steven M. Ross, Aaron J. McDonald, and Marty Alberg
*Center for Research in Educational Policy,
The University of Memphis*

Brenda McSparrin-Gallagher
*Office of Research and Evaluation,
Memphis City Schools*

This study was designed to examine the effects of a whole school reform, the Knowledge is Power Program (KIPP), specifically designed to raise academic achievement of at-risk urban middle school students by establishing an extended school day and year, a rigorous curriculum, after-school access to teachers, and increased family-school connections. In our mixed-methods design, qualitative (interview and observation) and quantitative (survey and achievement test) measures are employed to determine first-year program implementation and student outcomes for the KIPP:DIAMOND (Daring Individual Achievers Making Outstanding New Dreams) Academy (KIPP:DA), an inner-city school in a large, high-poverty urban district. For the achievement analyses, 49 KIPP:DA students were individually matched to highly comparable control students of the same ethnicity, socioeconomic status, gender, and ability, who attended different district schools in the same neighborhood. Statistically significant and educationally meaningful advantages for KIPP:DA students were found on 4 out of 6 standardized tests. The achievement results are interpreted in relation to the key program elements, the positive school climate established, and the quality of the 1st-year implementation.

During the past two decades, both federal and state educational reform programs have concentrated on improving achievement of at-risk students (No Child Left

Behind, 2001). A primary focus, spurred by the enactment in 1997 of the Comprehensive School Reform (CSR) Demonstration program (U.S. Department of Education, 1999), has been attempting to realize positive changes across the entire school rather than implementing isolated reform programs (U.S. Department of Education, 1999). The target contexts for CSR have specifically been Title I schools, especially those in inner cities. As Snipes and Casserly (2004) pointed out, of the nearly 14,000 public school districts in the United States, the largest 100 of these districts contain 17% of all public schools, employ 21% of all public school teachers, and serve 30% of the nation's disadvantaged students and 40% of its minority students.

Unfortunately, recent studies of systemic urban CSR have indicated uneven progress in demonstrating success, due to factors such as low teacher buy-in, inadequate resources, insufficient professional development, conflicting district policies, and failure to demonstrably raise student achievement (Berends, Kirby, Naftel, & McKelvey, 2002; Ross, 2001; Ross & Gil, 2004; Rowan, Camburn, & Barnes, 2004; Snipes & Casserly, 2004). Ross (2003) recently analyzed the failure of urban CSR efforts in Memphis and in Toledo (Ross, Nunnery, et al., 2004), and found only weak to moderate congruence of schools' observed reform programs with the Correlates of Effective Schools (Edmonds, 1979, 1982) or with the 11 elements of the federal CSR program (U.S. Department of Education, 1999). For example, many of the schools attempting to enact reforms appeared to lack a clearly stated mission, a safe environment, high expectations, instructional leadership, opportunity to learn, monitoring of progress, formative evaluation activity, external partners, and effective communications.

For a variety of reasons, implementation of school reforms has been demonstrated to take place more slowly in high schools and middle schools than in elementary schools (Bodilly, 1998; Bodilly & Berends, 1999). In the higher grades, curricula tend to be more departmentalized and the teachers more resistant to experimenting with innovative programs and instructional strategies (Smith et al., 1997). The middle grades further bring unique challenges associated with the major psychological, social, and biological changes that impact students during these years (Simmons & Blyth, 1987). Ability to cope with stress and emotional stability in general tend to be lowest in Grades 5-7 (Larson, Moneta, Richards, & Wilson, 2002). Not surprisingly, the need for supportive middle school environments has been strongly emphasized in research studies (e.g., Roeser & Eccles, 1998; Way & Robinson, 2003).

In an effort to create a whole school reform program geared more closely to the needs of at-risk, urban adolescents, two student participants in Teach for America developed the Knowledge Is Power Program (KIPP) in 1994. KIPP is grounded on five foundational pillars: (a) high expectations, (b) choice and commitment by families, (c) more time to learn, (d) power to lead, and (e) focus on results. Although designed prior to the formal 1997 CSR legislation, KIPP addresses the

major CSR elements and encompasses virtually all aspects of school operations, including instruction, assessment, classroom management, professional development, parental involvement, school management, and curriculum (see reviews by Borman, Carter, Aladjem, & LeFloch, 2004; Ross & Gil, 2004; and Rowan et al., 2004). Operationally, KIPP's most salient elements include focusing its at-risk student enrollees on (a) graduating from both high school and college; (b) spending more time in the classroom through extended day, extended year, and Saturday classes; (c) assigning from 2–3 hr of homework each night; (d) providing after-school access to teachers via cell phones; and (e) offering extensive professional development for school leaders. Currently, KIPP is being implemented in 31 schools in 13 states (Manzo, 2004). Despite the long history and recent proliferation of whole school designs, there is surprisingly limited rigorous, scientific, or independent evidence on their effectiveness in either implementation quality or raising student achievement (see Borman, Hewes, Overman, & Brown, 2003; Herman, 1999; Northwest Regional Educational Laboratory, 2000; Slavin & Fashola, 1998; Wang, Haertel, & Walberg, 1997). Accordingly, although KIPP has gained a reputation as a promising educational system (Manzo, 2004), evidence for its effectiveness exists from descriptive or preexperimental studies only (see Doran & Drury, 2002; KIPP, 1994). For example, in the most recent study to date, the Educational Policy Institute (2005) examined growth scores on the Stanford Achievement Test for fifth-grade cohorts from 24 KIPP schools located in different states. Findings indicated pretest–posttest gains (from Fall to Spring, or Fall to Fall) that exceeded those for normal growth. However, as the authors themselves acknowledged, the ability to gauge program effectiveness was reduced by several design limitations, including the absence of comparison groups, usage of school-level rather than student-level data, and the pretest and posttest samples being unmatched (i.e., consisting of overlapping but not identical student cohorts). In their conclusions, the authors recommended that “comparison/control” groups be used in further research, stating: “Being able to analyze data from schools serving students most like KIPP would greatly enhance the findings” (p. 13).

Our study was designed to incorporate a rigorous control group comparison in examining first-year outcomes from implementing KIPP at an inner-city school, named the KIPP:DIAMOND (Daring Individual Achievers Making Outstanding New Dreams) Academy (KIPP:DA) in Memphis, Tennessee. According to the school district plan, the school consisted of fifth grade only in this first year, with the intention being to add one additional grade (sixth, seventh, and eighth) in each of the subsequent three years. Throughout the school year, KIPP:DA was in session from 7:30 a.m. to 5:00 p.m. during the week, 4 hr on Saturday, and a month during the summer. Teachers were provided with cellular phones and were available to students and their families outside normal school hours for assistance with homework or in case of emergency. There was no intellectual or documented achievement requirement for admission to KIPP:DA. However, all students and

their parents were required to sign commitment forms indicating their agreement with the educational mission of the school and their willingness to support the school's rigorous requirements for academic engagement and exemplary conduct. Teachers were selected through an extensive application and interview process that suspended traditional considerations such as length of service and allowed the new principal, Mr. Carter (a pseudonym), greater than usual autonomy in staffing. Teachers in this school received higher salaries than their peers in other Memphis City Schools because of greater than usual expectations at KIPP:DA regarding time in the school and after-hours accessibility to students and their parents.

A mixed-method research approach (Johnson & Onwuegbuzie, 2004; Onwuegbuzie & Teddlie, 2003), combining quantitative analyses with qualitative inquiry to describe influential processes and contextual factors, was employed. The primary research question was whether KIPP:DA students would achieve at higher levels than would matched control students in literacy and mathematics on the state-mandated standardized assessment. To address this question in the most powerful way possible, given the inability to randomly assign students to schools, we individually matched each KIPP:DA student to a demographically similar counterpart who attended the same grade in one of the five schools serving the same geographic area as KIPP:DA. However, in recognizing the importance to educational change of successful program implementation, positive school climate, improved pedagogy, and teacher support (see Rowan et al., 2004), additional research interests concerned the impacts of these factors as potential modulators and products of the KIPP:DA implementation.

METHOD

Design and Participants

Both descriptive and inferential analyses were conducted, the latter involving a quasi-experimental comparison of KIPP:DA and matched control groups on student achievement. Multiple data sources for this study, described in greater detail in the following sections, included (a) the School Observation Measure (SOM; Ross, Smith, & Alberg, 1998); (b) the School Climate Inventory (SCI; Butler & Alberg, 1989); the Comprehensive School Reform Teacher Questionnaire (CSRTQ; Ross & Alberg, 1999); (d) teacher focus group; (e) principal interview; (f) implementation benchmark review (Ross, McDonald, & Alberg, 2002); (g) parent survey (McDonald & Ross, 2004); (h) student focus group; and (i) student-level test scores (Language Arts, Reading, Writing, and Math) on the Tennessee Comprehensive Assessment Program: Achievement Test (TCAP:AT; CTB/MacMillan/McGraw Hill, 1997).

KIPP:DA student participants were the complete population of Year 1 enrollees ($n = 49$) from the three fifth-grade classes. All (100%) were African American,

60% were girls, and 92% were eligible for free or reduced-price lunch. There were three regular teachers, one special education teacher, and an administrative assistant. All except the administrative assistant ($n = 4$) completed the teacher surveys (SCI and CSRTQ), but all five participated in the focus group. Of the 53 parents who were mailed surveys, 29 (54.7%) returned completed forms.

For the achievement study, the potential control group sampling pool was fifth-grade classes in the five elementary schools that fed into KIPP:DA. All were located in the same geographic area and were highly comparable to KIPP:DA and to each other in both student and school demographics. Eligible control group selections were all fifth-grade students enrolled in these schools for whom both 2001–2002 and 2002–2003 scores in Reading and Mathematics on the TCAP Norm-Referenced Test (NRT) were available ($N = 317$). We then individually determined the closest individual match for each KIPP student based on (a) gender, (b) free- or reduced-price lunch status, (c) ethnicity, and (d) 2001–2002 NRT-Reading and Mathematics subtest scores. On both pretests, the KIPP:DA and control group means were nearly identical with associated effect sizes (ES) close to zero (see Table 1).

Despite the aforementioned similarities between KIPP:DA students and their control counterparts, a variable that could not be controlled experimentally or quantitatively was student and parent interests in education. In the principal interview, Mr. Carter explained that about half of his students enrolled in KIPP:DA as a result of their families learning about the school through word of mouth or media information and submitting an application. With the desired enrollment quota not nearly filled a few months before school opened, Mr. Carter went door to door and visited local community organizations to publicize the school. Through this recruitment effort, the remaining slots were filled. KIPP:DA students were therefore those for whom a parent or guardian, either on his or her own or through explicit invitation or encouragement, decided to choose KIPP:DA over the regularly assigned school. One might infer, but to an unknown degree, that such families would tend to be more involved in their children's education than those who did not exercise such choice. This factor constitutes a limitation of this nonrandomized design.

Another limitation is that observations, interviews, and surveys were conducted at KIPP:DA only, not at the control sites. The reason was that control students were distributed among five neighborhood schools that were not formally participating in the study and that were identified in midyear. Fortunately, confidence in interpreting KIPP:DA results is bolstered by two factors. One is that, for both of the primary surveys (SCI and CSRTQ) and the classroom observations (SOM), national norms from over 400 schools were available from The Center for Research in Educational Policy (CREP) at The University of Memphis. Although the majority of the schools in that database are Title I and urban, their overall socioeconomic status and achievement scores were higher than for our control sites. Thus, comparison of KIPP:DA outcomes to the various norms would tend to be conservative with regard to showing advantages for KIPP:DA.

TABLE 1
Pretest (2001–2002) and Posttest Means and Standard Deviations
for KIPP:DA and Control Students on the Tennessee Comprehensive
Assessment: Achievement Test (TCAP:AT) Subtests

Subtest	KIPP		Control		Effect Size
	M	SD	M	SD	
Pretests					
NRT-Language Arts	41.16	18.36	43.16	18.61	-0.11
NRT-Reading	40.25	17.72	39.92	16.45	+0.02
NRT-Math	41.65	16.87	41.37	16.10	+0.02
Writing	3.68	0.98	3.51	0.76	-0.03
Posttests					
NRT-Language Arts	42.80	16.01	39.96	16.28	+0.24
Adjusted	42.98	—	38.81	—	+0.26
NRT-Reading	43.08	18.47	38.18	15.40	+0.29
Adjusted	43.38*	—	37.89*	—	+0.31
NRT-Math	42.84	16.42	37.80	11.73	+0.35
Adjusted	42.87*	—	37.72*	—	+0.35
CRT-Reading/Language Arts	641.92	31.59	633.00	30.91	+0.28
Adjusted	642.38*	—	632.53*	—	+0.31
CRT-Math	633.82	32.94	622.15	22.59	+0.41
Adjusted	634.28**	—	616.29**	—	+0.63
Writing	3.88	0.86	3.91	0.92	-0.03
Adjusted	3.89	—	3.90	—	-0.01

Notes. * $p < .05$ for KIPP versus control means in ANCOVA. ** $p < .01$. KIPP:DA = Knowledge Is Power Program: Diamond Academy; NRT = Norm-Referenced Test portion of the (TCAP:AT); CRT = Criterion-Referenced Test portion of the TCAP:AT.

Second, although no climate data were available for the control schools during the year in which this study was conducted, the school district administered its own climate inventory to all schools starting in the following year (2003–2004). On all five scales, the control school median score approximated district norms, with the control schools scoring slightly higher on two scales (Teacher Performance/Collegiality and Respect for Teachers by Students), and the district slightly higher on three (Safety, Principal Performance, Parent/Student Investment). Although year-to-year changes might have occurred, these results suggest that the climate of the control schools was very typical of those in the district and therefore not indicative of the control schools experiencing unusual conditions.

Instrumentation

SOM. The SOM was developed to determine the extent to which different common and alternative teaching practices are used throughout an entire school (Ross, Smith, & Alberg, 1998). The procedure involves observers' visiting 10–12

randomly selected classrooms, for 15 min each, during a 3-hr visitation period. The observer examines classroom events and activities descriptively, not judgmentally. Notes are taken relative to the use or nonuse of 24 target strategies. At the conclusion of the 3-hr visit, the observer summarizes the frequency with which each of the strategies was observed across all classes in general on a data summary form. The frequency is recorded via a 5-point rubric that ranges from 0 (*not observed*) to 4 (*extensively*). Two global items are used to rate, respectively, the level of academically focused instructional time and degree of student attention and interest.

The SOM strategies include traditional practices (e.g., direct instruction and independent seatwork) and alternative, predominantly student-centered methods associated with educational reforms (e.g., cooperative learning, project-based learning, inquiry, discussion, using technology as a learning tool). The strategies were originally identified through surveys and discussions involving policymakers, researchers, administrators, and teachers, as those most useful in providing indicators of schools' instructional philosophies and implementations of commonly used reform designs (Ross, Smith, Alberg, & Lowther, 2004).

To ensure the reliability of data, observers receive 1 full day of training, a manual providing definitions of terms, examples, and explanations of the strategies, and a description of procedures for completing the instrument. In a reliability study (Lewis, Ross, & Alberg, 1999), pairs of trained observers selected the identical overall response on the five-category rubric on 67% of the items and were within one category on 95% of the items. In a second reliability study using generalizability theory, Sterbinsky and Ross (2003) found reliability at the .74 level for five SOMs conducted at individual schools. Reliability increased to .82 with eight SOMs and to .85 with 10 SOMs conducted at a school.

SCI. The SCI consists of seven dimensions logically and empirically linked with factors associated with effective school organizational climates (Butler & Alberg, 1989). Each scale contains seven items, with 49 statements comprising the inventory. Responses are scored through use of Likert-type ratings from 1 (*strong disagreement*) to 5 (*strong agreement*). Scale means can range from 1 to 5, with higher scores being more positive. Additional items solicit demographic information.

Face validity of the school climate items and logical ordering of the items by scales were established by the research team during the development of the inventory (Butler & Alberg, 1989). Subsequent analysis of responses collected through administration of the inventory in a variety of school sites substantiates validity of the items. Dimension descriptions and current internal reliability coefficients on the seven dimensions of the inventory, obtained using Cronbach's alpha, are as follows: *order*, the extent to which the environment is ordered and appropriate student behaviors are present ($\alpha = .84$); *leadership*, the extent to which the administration provides instructional leadership ($\alpha = .83$); *environment*, the extent to which

positive learning environments exist ($\alpha = .81$); *involvement*, the extent to which parents and the community are involved in the school ($\alpha = .76$); *instruction*, the extent to which the instructional program is well developed and implemented ($\alpha = .75$); *expectations*, the extent to which students are expected to learn and be responsible ($\alpha = .73$); and *collaboration*, the extent to which the administration, faculty, and students cooperate and participate in problem solving ($\alpha = .74$).

CSRTQ. The CSRTQ was designed to assess teachers' perceptions of school conditions that affect capacity to enact reforms (e.g., see Ross & Alberg, 1999; Ross et al., 1997). The CSRTQ contains 28 items to which teachers respond using a five-point Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). In a second section, respondents report their perceived progress toward implementation benchmark goals. In a third section, they provide open-ended comments regarding the positive and negative aspects of their school's reform program. Face validity and logical ordering of the items by scales were established by the research team during the development of the inventory (Ross et al., 1997). Principal component analyses of responses collected through administration of the inventory in multiple restructuring and comparison sites substantiated the instrument's scales (Bol et al., 1998). All items loaded unambiguously on the factors, which explained 63.2% of the variance in item scores. The factors (and concomitant reliability coefficients) included Resources (.64), Professional Development (.76), Pedagogical Change (.67), and Student Outcomes (.91). More recently, Nunnery, Ross, and Sterbinsky (2003) conducted a construct validation study that expanded the CSRTQ to include five factors. This study showed the Support and Capacity dimensions to account for 84% of the variance in Focus, which in turn had strong direct effects on both Pedagogy and Outcomes. It is important that 84% of the variance in Outcomes was explained by Focus and Pedagogy combined.

Parent questionnaire. This instrument was designed to obtain parent perceptions of the school in areas such as instruction, curriculum, communication, and opportunities for involvement (McDonald & Ross, 2004). Included were 17 closed-ended items using a five-point Likert-type scale (from *strongly disagree* to *strongly agree*), and 4 open-ended questions regarding satisfaction and involvement with the school.

The instrument was evaluated for content and face validity based on reactions by a diverse review panel consisting of principals, teachers, and researchers, and was approved by the Tennessee Department of Education for, and employed in, a 2003–2004 study of urban charter schools (Ross, McDonald, Bol, et al., 2004).

Focus groups and interviews. To supplement the survey data, we conducted a principal interview, a teacher focus group, and three student focus groups that included 14 different students. The time period for each was approximately 1

hr. Student participants were randomly selected to participate. A semistructured protocol, involving standard questions with flexibility for follow-up on selected responses, was used. In all three protocols, the basic questions concerned experiences during the year, differences from other schools, reactions to major school components (e.g., teaching methods, extended day, curriculum, parent involvement), perceived strengths and weaknesses, and recommendations for improvement.

Program Implementation Benchmarking. As part of the implementation benchmark development process, the school staff developed statements or goals for each major program component. Each statement was then accompanied by a specific indicator, and evidence for implementation in Phases I (Beginning), II (Intermediate), and III (Full). The draft benchmarks were then shared with the entire faculty for review. In the spring, the teachers evaluated progress and specified program goals for the next year.

Procedure and Analyses

During the school year, three researchers visited KIPP:DA a total of six times to conduct half-day classroom observations. In the spring, they administered surveys or interviews to the faculty, students, and parents. Quantitative data were analyzed via descriptive summaries of frequency counts and means, and in the case of student achievement, via multivariate analysis of variance (MANOVA) comparing KIPP:DA and control group scores on TCAP:AT measures. Qualitative analyses, guided by Miles and Huberman's (1994) model, were performed on open-ended survey and interview responses. The procedure consisted of transcribing the responses, deriving patterns and concepts, identifying themes, and revising and refining based on member checking and inter-rater review.

RESULTS

Results are presented in three sections, respectively addressing (a) perceptual data from the principal, teachers, parents, and students regarding school climate, experiences, and program implementation; (b) observations of classroom teaching; and (c) student achievement.

Participant Perceptions

Qualitative analysis of the interview and survey data yielded several major categories of respondent perceptions. These results, along with survey results on

TABLE 2
 Knowledge Is Power Program: Diamond Academy KIPP:DA School Climate
 Dimension Means Compared to National Middle/High School Norms

	<i>KIPP:DA</i>		<i>National Norms</i>		<i>ES</i>
	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	
Collaboration	4.37	.41	3.64	.61	+1.60
Environment	4.51	.54	3.74	.67	+1.14
Expectations	4.66	.54	3.71	.60	+1.58
Instruction	4.71	.20	3.94	.50	+1.54
Involvement	4.54	.33	3.65	.58	+1.53
Leadership	4.63	.52	3.85	.70	+1.11
Order	4.11	.81	3.27	.80	+1.05
Overall	4.51	.45	3.69	.55	+1.49

Note. Scale ranges from 1 = *strongly disagree* to 5 = *strongly agree*. *ES* = effect size.

Likert-type items, are presented by category. Responses to the two teacher surveys, the SCI and CSRTQ are summarized in Tables 2 and 3, respectively.¹

High expectations. The KIPP:DA school plan describes “clearly defined, measurable high expectations for academic achievement and conduct that make no excuses based on the background of students.” The principal, Mr. Carter, described the school’s academic expectations for students as “(preparing) students to be competitive for the schools that they will attend after 8th grade.” Progress toward achieving these goals was reflected by the noticeably high mean of 4.66 out of 5.00 ($ES = +1.58$ compared to national norm) on the Expectations dimension of the SCI (see Table 2). For example, all respondents (100%) strongly agreed with the item, “Teachers have high expectations for all students.” The SCI results further revealed that behavioral expectations were generally being met for this at-risk preadolescent student population, as reflected by the relatively high order dimension mean of 4.11 ($ES = +1.05$). All respondents agreed or strongly agreed with SCI Item 46: “Student behavior is generally positive in this school.”

Corroborating these SCI outcomes was teachers’ unanimous agreement on Item 18 on the CSRTQ, “Students have higher standards for their own work because of our school’s program” (see Table 3). Teacher focus group responses to the

¹For the purposes of brevity, a shortened version of the qualitative results is presented here. For a full reporting of quotations and supplemental impressions from respondents, see Alberg (2003); Ross, McDonald, and McSparrin-Gallagher (2004). 3. The first author began the research with these stakeholder interviews, as described in the full technical report (Ross, McDonald, & McSparrin-Gallagher, 2004).

TABLE 3
 Percentages of Teachers Indicating Varied Levels of Agreement
 on the CSRTQ

CSRTQ Items	Strongly Agree and Agree (%)
1. I have a thorough understanding of this school's comprehensive school reform (CSR) program.	100.0
2. I have received adequate initial and ongoing professional development/training for CSR program implementation.	100.0
3. Professional development provided by external trainers, model developers, and/or designers has been valuable.	100.0
4. Guidance and support provided by our school's external facilitator, support team, or other state-identified resource personnel have helped our school implement its program.	100.0
5. Teachers are given sufficient planning time to implement our program.	75.0
6. Materials (books and other resources) needed to implement our comprehensive school reform program are readily available.	100.0
7. Our school has sufficient faculty and staff to fully implement this program.	75.0
8. Because of our CSR program, technological resources have become more available.	100.0
9. Because of our CSR program, I use textbooks, workbooks, and worksheets less than I used to for basic skills or content area instruction.	75.0
10. Our comprehensive school reform program has changed classroom learning activities a great deal.	100.0
11. Students in my class spend at least 2 hr per school day in interdisciplinary or project-based work.	75.0
12. Students in my class spend much of their time working in cooperative learning teams.	50.0

TABLE 3 (Continued)

<i>CSRTQ Items</i>	<i>Strongly Agree And Agree (%)</i>
13. Students are using technology more effectively because of our CSR program.	75.0
14. Student achievement has been positively impacted by CSR.	100.0
15. Students in this school are more enthusiastic about learning than they were before we became a CSR school.	100.0
16. Because of CSR, parents are more involved in the educational program of this school.	100.0
17. Community support for our school has increased since comprehensive school reform has been implemented.	100.0
18. Students have higher standards for their own work because of our school's program.	100.0
19. Teachers are more involved in decision making at this school than they were before we implemented comprehensive school reform.	100.0
20. Our program adequately addresses the requirements of children with special needs.	100.0
21. Because of our school's program, teachers in this school spend more time working together to develop curriculum and plan instruction.	100.0
22. Teachers in this school are generally supportive of our CSR program.	100.0
23. Because of CSR, interactions between teachers and students are more positive.	100.0
24. The elements of our CSR program are effectively integrated to help us meet school improvement goals.	100.0
25. As a school staff, we regularly review implementation and outcome benchmarks to evaluate our progress.	100.0
26. Our school has a plan for evaluating all components of our comprehensive school reform program.	100.0
27. My school receives effective assistance from external partners (e.g., university, businesses, agencies, etc.).	100.0
28. I am satisfied that Federal, State, local, and private resources are being coordinated to support our CSR program.	75.0

question, “To what degree do you feel that the (school) goals are being accomplished,” offer further insight into the culture of high expectations at KIPP:DA:

- “The difference in the children is that they now believe college is a possibility.”
- “We are trying to instill in the students a sense of achievement that they can achieve. It’s just a given. That expectation is there.”
- “I think there has been a paradigm shift from feeling that ‘I am a failure’ to ‘I celebrate me, I am unique in this world, no one can take my place.’ Many of the children really do believe that they are college-bound. We talk about it daily.”
- “The idea of ‘no excuses’ has been a part of the paradigm shift as well. If you, for example, do not take the responsibility to finish projects or homework, there is a consequence.”

Of the 29 parents returning questionnaires, 96.5% agreed or strongly agreed with the statement, “The KIPP Academy expects students to perform at a high level.” Open-ended comments were also positive, as reflected, for example, by: “The best thing about the Academy is that they work very hard to help the kids be the best that they can be and reach their highest goals in life,” and “The teachers show a genuine concern about whether my child learns what is being taught.”

Perhaps the most compelling evidence for high expectations came from the students themselves. Illustrative responses were:

- “[They make sure that I do] my homework every day so that I get good grades and probably get to go to more schools that are better.”
- “By bringing your grades up and then when you go to high school, you can go to college and find a job and get more money.”
- “Raising the bar means getting better TCAP scores, to bring our grades up.”
- “My grades. Last year I would get low grades, like all Cs. This year I am getting As and Bs instead of Cs and Ds.”
- “It helps me to achieve high because I know that all of my teachers put their best foot forward to help me do my work and learn.”

Choice and commitment. In the review by Mr. Carter and faculty of the school’s Implementation Benchmarks, there was consensus that KIPP:DA reached full implementation of: “The students, faculty, staff, and administration sharing a common understanding of the goals and purposes of the school, supporting these goals, and communicating ideas and concerns about the school.” Other benchmark assessment showed progress in formally involving parents and community members with school activities.

Several items on the Environment and Involvement scales of the SCI addressed issues of commitment:

- 100% of the faculty agreed that “they make important contributions to the school,” “community businesses are active in the school,” and “parents actively support school activities.”
- 80% agreed that “people in the school really care about each other.”
- 80% agreed that “teachers are proud of the school and its students.”

Further on the CSRTQ, 100% of the faculty (Item 22, see Table 3) agreed that “Teachers are generally supportive of the school’s program.” Teachers’ open-ended comments directly related to choice and commitment were all positive. Exemplary reactions were:

- “There [is] a freedom to be able to do and implement what you love. And at the same time, you are surrounded by people who support that.”
- “We have a wonderful group of teachers who are committed and dedicated to the cause. Our parents want what is best for their children. And with the parents and the teachers and the principal all working together, I really think that we can succeed.”
- “If you are here for any reason other than to help our students to succeed, you are in the wrong place.”
- “Parents are happy. Teachers are happy. Everyone is here because of their choice and it shows.”

Increased learning time. KIPP:DA teachers were highly supportive of the school’s extended hours. As one teacher stated in the focus group, “I felt a long time ago that children got in trouble a lot because parents got off work at 5:00 and children get out of school at 2:15. Children need more school.” Focus group and observational data suggested that teachers primarily use the longer hours to extend lessons and provide tutorials for students who need them. Teacher awareness of the increased potential to bore students with “more of the same” motivated them to try to vary instructional strategies as much as possible. As one student stated, “The teacher doesn’t just talk all day. She will teach you a song or something so you will understand better.” These perceptions were partially supported by the SOM data on teaching methods to be described in the following sections.

When asked if it was easy or hard to adjust to the differences between KIPP and their previous school, several students mentioned getting out of school at 5:00 pm as a difficult adjustment. Nonetheless, based on responses during the focus groups, most students appeared to view the extended hours, at worst, as just a standard part of KIPP:DA or, at best, as a beneficial increased opportunity to learn.

Power to lead. Principal Carter fully participated in professional development offered by both the national organization (i.e., The School Leadership Program) and the local district. He considered administrative preparation and expertise to be critical aspects of KIPP. Specifically, the school's benchmark goal for leadership states, "In a fully implemented KIPP school, the principal has complete knowledge of the program, can articulate the components to colleagues and the community, and functions as the instructional leader of the school." The consensus of teachers in the surveys and interviews was that the principal attained this goal within the first year. When asked to describe his role in program implementation, Mr. Carter listed the following:

- "I still consider myself a teacher. I take an active role. I model what I am expecting from my teachers and students."
- "I am a facilitator in presenting ideas and showing proven ways of teaching and communicating. I basically take my role as principal to be the principal teacher."
- "I am available for answering questions, modeling, showing, providing needed resources, and keeping communication lines open."

Consistent with the qualitative data, the school mean on the Leadership dimension of the SCI (see Table 2) was an unusually high 4.63 out of 5.00 ($ES = +1.11$). Teachers were unanimous in strongly agreeing that the principal (a) communicates that all students can learn, (b) encourages teacher creativity and trying new methods, and (c) protects instructional time. Other mostly favorable responses concerned receiving feedback from the principal, and his visibility and instructional leadership. Illustrative open-ended comments from the different participant groups include:

- Teacher: "We have a principal who really believes in the concept of KIPP. He is an excellent communicator. He has had to really fight for the things that KIPP was supposed to have, some of which we do have now and some we are still waiting on."
- Parent: "The principal communicates information directly to me."
- Student: "He [Mr. Carter] had always dreamed about opening a school like this. I was surprised that they opened it up in a neighborhood like this because it is bad. Not the school, the neighborhood."
- Student: "When we were at our old school, Mr. Carter said we need to raise the bar up to here [gesture]. Raising the bar means getting better TCAP scores; to bring our grades up when we are being bad, he will say we are acting like a size 2 pair of shoes but we really aren't. We are supposed to be a normal size shoe, size 100."

Focus on results. The consensus of the faculty and the principal was that explicit benchmarks for establishing measurable goals for student achievement outcomes were developed and are understood by stakeholders. Additionally, instructional plans and professional development activities were adjusted based on review of progress toward goals. Attesting to KIPP:DA's perceived progress in establishing strong academic focus, all (100%) teachers agreed on CSRTQ Item 14 (see Table 3) that student achievement was positively impacted by the KIPP program. They all also agreed that the school has a plan for evaluating all components of the program and that, as a school staff, they regularly review implementation and outcome benchmarks to evaluate progress toward goals. Open-ended responses further elaborated on how the school's results orientation was manifested. For example:

- Teacher: "They [the students] want to pass all their courses. They are concerned about their grades. They want to achieve."
- Teacher: "When the testing situation came up, all they [students] could do is say, 'We're going to do it. We're going to do it.' They were self-motivated."
- Parent: "My child has changed a lot. He is more eager to learn and he comes right home to do his homework so that he can get a good grade."
- Parent: "When she first started the Academy, her grades were low. Now her grades are two grades up."
- Student: "At my old school, I made *F*s every single day but at this school I am making *A*s."
- Student: "KIPP is a wonderful school and you should go there. They will teach you all kinds of things. You can learn a lot of stuff. Sometimes they will help you change your attitudes."

Curriculum and instruction. KIPP:DA incorporated the school district's curriculum and state content standards, while emphasizing interdisciplinary studies, especially writing across subject areas. Teachers agreed that by the end of this first school year, they had attained the second (intermediate) phase of their benchmarks for curriculum: (a) "Core curriculum areas have been aligned with state performance standards and with the most current assessment results," and (b) "Based on pilot data, integrated units have been evaluated and revised, and lesson plans have been extended to contain KIPP:DA components."

As a fundamental part of KIPP:DA's goals, teachers were expected to use a wide variety of strategies, such as musical mnemonics, writing across the curriculum, project-based learning, cooperative learning, and peer and individual tutoring. Teacher responses on the instruction dimension of the SCI conveyed 100% agreement that: (a) a variety of teaching strategies were being used; (b) learning activities were designed to support student needs, the curriculum, and higher order skills; and (c) student achievement was appropriately evaluated. The scale mean

for the instruction dimension was 4.71 (see Table 2), close to ceiling and well above the national norm ($ES = +1.54$). On the CSRTQ, three-fourths (75%) of the teachers indicated that students in their classes spend at least 2 hr per day in interdisciplinary or project-based work (Item 11; Table 3), and half indicated that students spend much of their time working in cooperative learning groups (Item 12). Open-ended statements showed much consistency of perceptions between respondent groups:

- Principal: “We incorporate strategies from the Multiple Intelligences model such as mnemonics, chants, body movements, or whatever it takes to tap into how our students are learning and what works most effectively. Also used are traditional reading, writing, direct instruction, lecture, cooperative learning.”
- Teacher: “We incorporate music in each classroom. We learn chants and songs to go along with our program to help enhance learning.”
- Teacher: “We integrated (the curriculum) as much as possible. We used various teaching strategies: cooperative groups, individual learning, peer tutoring, and tutors that came from outside the school.”
- Teacher: “We wrote across the curriculum. In every subject, the students were doing writing and that was important to us.”
- Student: “They give you chants to go with math, reading, and social studies. They have better teachers.”

Organization and resources. There was consensus among the teachers and the principal that the organizational structures and resources at KIPP:DA supported and sustained the school’s mission and goals. The single benchmark in which lack of satisfactory progress was perceived was “staffing to promote student attendance and good health.” Included in the school’s staffing plan, but not yet represented, were a full-time nurse, a guidance counselor, a parent advocate, and a librarian. Parents provided volunteer support in the form of tutoring, fund raising, coaching, and other services as requested.

In responding to the CSRTQ (Table 3), all teachers agreed that the materials (Item 6) and technological resources (Item 8) needed to implement the KIPP program were available. When asked to describe negative aspects of the program, however, teachers noted the lack of support staff and some other resource needs such as fully operable administrative software, materials for Accelerated Reader, and educational manipulatives. Sample comments were:

- “[The district’s] lack of cooperation to get appropriate personnel to KIPP:DA in a timely manner [was a concern]. The school psychologist was assigned late and a guidance counselor has not been assigned at all and the school year is just about over.”
- “We need a science lab and library.”

- “A guidance counselor is much needed because of the problems children face and services that need to be rendered to students.”

Overall satisfaction. Teachers were asked in the focus group to rate the school’s progress on a scale of 1 to 10. The consensus rating was level 9, based primarily on the changes they observed in the students over the course of the year. On the parent survey, nearly all respondents (96.6%) indicated satisfaction with what their children were learning. Many positive aspects, but only a few very specific concerns (e.g., PTO, dismissal time, and school location), were identified in their open-ended comments. Included among the positives were sense of family and togetherness, teacher quality, raising children’s self-esteem, emphasis on performance, increased time to learn, concern for children, and the teachers’ availability after hours.

Even students had a short list of “worst things about the Academy.” Almost exclusively listed were disciplinary components such as “the Dugout” (time-out area). Positive aspects noted included the supportive learning atmosphere, encouragement for learning, increased success in learning, and the inclusion of fun activities to break up the longer day.

Observations of Teaching

Using the SOM, we conducted six 3-hr observations, encompassing approximately 60 classroom visits (15-min each) at different times during the school year. Results indicating the frequency with which the various strategies were observed are summarized in Table 4. As indicated, the most prevalently observed strategy was direct (teacher-centered) instruction, which was rated as frequent to extensive in 50% of the SOM visits. At the other extreme, strategies never or very rarely observed were team teaching, multiage grouping, systematic individualized instruction, individual tutoring, parent/community involvement (in the classroom), sustained reading, independent inquiry, computer for instructional delivery, performance assessment, and student self-assessment. Of note, high academic focus (instructional time used predominantly for curriculum and instruction, few interruptions) and high student engagement (interest, on-task behavior, active learning) were both frequently or extensively observed in all (100%) of the SOM visits.

Given that certain teaching strategies are intrinsically more difficult or less appropriate to implement regularly than others, we compared the percentage of times each strategy was observed frequently or extensively at KIPP:DA relative to the SOM national norms (CREP, 2004). Strategies associated with absolute differences exceeding the arbitrary cut-off of 10% are identified on Table 4. As shown, KIPP:DA classes were less likely than other schools to feature direct instruction, seatwork, higher order feedback, ability grouping, and multiage grouping. However, KIPP:DA classes more frequently employed project-based learning and

TABLE 4
 Percentages of Observations Reflecting Varied Levels of Frequency
 on the School Observation Measure

<i>The extent to which each of the following was used or present in the school</i>	<i>None(%)</i>	<i>Rarely(%)</i>	<i>Occasionally(</i>
Instructional Orientation	—	—	—
Direct Instruction (lecture)*	0.0	16.7	33.3
Team Teaching	100.0	0.0	0.0
Cooperative/Collaborative Learning	0.0	50.0	33.3
Individual Tutoring (teacher, peer, aide, adult volunteer)	83.3	0.0	16.7
Classroom Organization	—	—	—
Ability Groups*	83.3	16.7	0.0
Multi-age Grouping*	100.0	0.0	0.0
Work Centers (for individuals or groups)	66.7	16.7	16.7
Instructional Strategies	—	—	—
Higher-level Instructional Feedback (written or verbal) to Enhance Student Learning*	50.0	0.0	33.3
Integration of Subject Areas (interdisciplinary/thematic units)	66.7	16.7	16.7
Project-based Learning+	50.0	0.0	33.3
Use of Higher-level Questioning Strategies	16.7	33.3	33.3
Teacher Acting as a Coach/Facilitator+	0.0	33.3	16.7
Parent/Community Involvement in Learning Activities	83.3	16.7	0.0
Student Activities	—	—	—

TABLE 4 (Continued)

<i>The extent to which each of the following was used or present in the school</i>	<i>None(%)</i>	<i>Rarely(%)</i>	<i>Occasionally</i>
Independent Seatwork (self-paced worksheets, individual assignments)*	0.0	50.0	16.7
Experiential, Hands-on Learning	66.7	33.3	0.0
Systematic Individual Instruction (differential assignments geared to individual needs)	100.0	0.0	0.0
Sustained Writing/Composition (self-selected or teacher-generated topics)	33.3	33.3	33.3
Sustained Reading	83.3	16.7	0.0
Independent Inquiry/Research on the Part of Students	83.3	16.7	0.0
Student Discussion	66.7	33.3	0.0
Technology Use	—	—	—
Computer for Instructional Delivery (e.g., CAI, drill & practice) Assessment	83.3	0.0	16.7
Performance Assessment Strategies	—	—	—
Student Self-assessment (portfolios, individual record books)	83.3	16.7	0.0
Summary Items	83.3	16.7	0.0
High Academically Focused Class Time	—	—	—
High Level of Student Attention/Interest/Engagement	0.0	0.0	0.0
	0.0	0.0	0.0

Note. The 10% criteria were subjectively selected as a means of highlighting the approximately 25% of the SOI ences.

* KIPP:DA% frequently/extensively was 10% less frequent than national norms.

+ KIPP:DA% frequently/extensively was 10% more frequent than national norms.

teacher coaching. In general, the observation data support the impression of a slightly nontraditional orientation where teachers act more than usual as coaches and facilitators than as presenters of information.

Student Achievement

Descriptive statistics for pre- and post- implementation measures are presented in Table 1. Although the two groups scored very similarly on the pretests, KIPP:DA students performed directionally higher than control students on all CRT and NRT posttests. Moderate to strong effect sizes, ranging from +0.24 to +0.63, are indicated. Despite the described emphasis on writing across the curriculum, nearly identical KIPP:DA and control group means were obtained on the Writing assessment. Inferential results are described in the following sections.

Posttest NRT Language Arts, Reading, and Mathematics. A multivariate analysis of covariance (MANCOVA), using the three 2001–2002 (4th grade) pretest scores as covariates, was conducted on the 2002–2003 NRT Language Arts, Reading, and Mathematics subtests of NRT. Although all covariates were highly significant (all $ps < .02$), the multivariate Program effect did not reach significance, $F(3,91) = 2.52, p = .063, \eta^2 = .077$.

Given the relatively small sample sizes, a priori hypotheses, and the approximation to $\alpha = .05$ in the MANCOVA (see Wainer & Robinson, 2003), we conducted univariate tests (ANCOVA) on each of the dependent measures. The univariate results were significant for Reading, $F(1,93) = 5.55, p = .021, \eta^2 = .056$; and Math, $F = 5.74, p = .019, \eta^2 = .056$; but nonsignificant for Language Arts, $F = 2.77, p = .099, \eta^2 = .029$. Adjusted means and effect sizes are summarized in Table 1.

Posttest CRT Reading/Language Arts and Mathematics.

As shown in Table 5, data reflecting the percentages of students who scored at below proficient, proficient, and advanced levels on the CRT subtests directionally favored KIPP:DA. For example, 10% and 16% of KIPP:DA students scored at the advanced level in Reading/Language Arts and Math, respectively, compared to only 2% and 0% of the control students. Two-way chi square (Program \times Proficiency Level) analyses were significant for Math, $\chi^2(2) = 8.62, p = .013$, but nonsignificant for Reading/Language Arts, $\chi^2(2) = 4.17, p = .124$.

A MANCOVA, using the three 2001–2002 (4th grade) NRT pretests as covariates, was performed on the 2002–2003 CRT Reading and Mathematics subtest scale scores. The Reading and Math covariates were significant (both $ps < .001$), as was the Program main effect, $F(2,91) = 5.70, p = .005, \eta^2 = .111$. Univariate ANCOVAs were significant for both Reading, $F(1,92) = 4.76, p = .032, \eta^2 = .049$; and Math, $F = 10.82, p = .001, \eta^2 = .105$. Effect sizes associated with the

TABLE 5
The Percentages of KIPP:DA and Control Students Scoring
at Different Proficiency Levels on the CRT-Reading/Language Arts
and CRT-Mathematics in Fifth Grade

<i>Group and Subject</i>	<i>Proficiency Levels</i>		
	<i>Below Proficient</i>	<i>Proficient</i>	<i>Advanced</i>
Reading/Language Arts			
KIPP:DA	35	55	10
Control	50	48	2
Mathematics*			
KIPP:DA	41	43	16
Control	46	54	0

* $p < .05$ in chi-square test.

adjusted means (see Table 1) were +0.31 in Reading/Language Arts and +0.63 in Math.

Posttest Writing Assessment. An analysis of covariance, using the 2001–2002 (4th grade) Writing scores as a covariate, was performed on the 2002–2003 Writing assessment. Although the covariate was highly significant ($p < .01$), the Program effect was close to zero, $F(1,82) = .004$, $p = .953$, $\eta^2 = .000$.

DISCUSSION

These findings reflect positive first-year attainments for KIPP:DA with regard to the implementation of curricula, instructional, and organizational programs, and particularly buy-in and support by faculty, parents, and students. These process outcomes, in turn, were associated with some advantages in achievement by KIPP:DA students relative to their matched control counterparts. Specifically, although KIPP:DA and control fifth graders had virtually identical means on all fourth-grade pretests, the KIPP:DA students demonstrated significantly higher achievement on 4 out of the 6 fifth-grade tests, with ES ranging from +0.31 to +0.63. Across all six achievement measures, the median adjusted ES was +0.31, indicating a moderate to strong effect. By comparison, in a recent meta-analytic study of 29 CSR models, Borman et al. (2003) found an overall ES from +0.10 to +0.14, with the range for the “most successful” category being +0.17 to +0.21. Only 3 out of the 29 models (none specific to middle schools) achieved this high status: Direct Instruction, School Development Program, and Success For All.

The achievement outcomes are noteworthy in addition given the consensus by scholars and researchers of school reform (Desimone, 2002; Fullan, 2000; Levin, 1993) that school change takes multiple years to produce implementation success and subsequent measurable effects on student achievement. Further, KIPP:DA also had to overcome multiple challenges in its first year, including sharing a building with another school, enrolling a greater than expected number of special needs students, and adopting new curricula, schedules, and administrative structures.

On the other hand, choice schools such as magnets, charters, and KIPP:DA inherit the potential advantage of parents and students being more committed and informed about educational opportunities relative to their counterparts attending regular neighborhood schools (Bulkley & Fisler, 2002; Collins, 1999; RAND, 2001). Also, similar to most charter schools and some magnet schools in different districts, KIPP:DA had the benefit of greater autonomy relative to regular district schools in hiring teachers. Whether or not KIPP:DA teachers were superior to other middle school staffs in pedagogical or management skills, it certainly seems likely that they possessed high levels of commitment and collegiality. The significant impacts of teacher buy-in on the success of planned educational change are continually emphasized in the literature on school reform (Borman et al., 2004; Desimone, 2002; Ross & Gil, 2004; Rowan et al., 2004).

By conducting rigorous quasi- or true-experimental studies of school programs, researchers can obtain valid evidence to support broad theories and generalizations about effective practices (Creswell, 2002; Feuer, Towne, & Shavelson, 2002). However, in field-based contexts, especially involving schoolwide reforms (i.e., CSR), the simultaneous occurrence and interaction of multiple social, academic, administrative, and cultural events significantly complicates research efforts to isolate the effects of specific program components (Berliner, 2002; Ross, 2003). Accordingly, it is possible only to speculate from our data as to which elements of KIPP:DA had substantive impacts. Clearly, for today's schools, the culminating criterion for judging program success is student achievement (U.S. Congress, 2001). It is therefore relevant to consider that one of the school variables most consistently and strongly linked to student achievement gains is increasing allocated instructional time (Bloom, 1980) and, even more directly, students' time on task or engaged time (Good & Brophy, 1987). In this regard, KIPP:DA's extended school day and year acquires obvious importance as a primary program element. Furthermore, according to the SOM results, KIPP:DA demonstrated higher than average instructional focus, student engagement, and diversity in instructional methods.

Research on CSR strongly supports the idea that teaching improvements and strong program implementations do not occur in a vacuum, merely because some new reform model has been selected (Borman et al., 2004; Datnow, Hubbard, & Mehan, 2002). To improve their effectiveness, teachers need quality professional development (Rowan et al., 2004) and the motivation to learn and practice new in-

structional strategies. For fostering such changes, the critical role of positive school climate has long been established (Bobbett, Ellett, Teddlie, Olivier, & Ruggett, 2002; Bryk, & Schneider, 2002; Desimone, 2002). Based on the SCI and CSRTQ results (Tables 1 and 2), it can be inferred the KIPP:DA teachers and staff were highly supportive of the program, and willing to meet the increased demands (time and effort) required. Simply put, positive climate facilitates program implementation, which, in turn, can improve teaching effectiveness and curricula, and ultimately, student achievement.

These results can also be interpreted relative to broader theoretical frameworks posited to explain effective school reform. For purposes of brevity, we have selected two—one classic, the *Correlates of Effective Schools* (e.g., Edmonds, 1979, 1982), and the other contemporary—the five-component model of Desimone (2002; as adapted from Porter, 1994). The main elements of each framework and our subjective ratings, based on the present evidence, of KIPP:DA's status on each are presented in Table 6.

Considering first the *Correlates of Effective Schools*, we rated all seven Correlates as having been fairly strongly addressed at KIPP:DA. The extended learning time schedule, focus on academic rigor, and explicit promotion of college as the attainable goal for every student reflected the uncompromising commitment of the teachers and principal to raising academic achievement. Every classroom was named after a historically African American college, conveying the goal and expectancy that all KIPP:DA students would attend college when they completed high school. The principal participated in extensive professional development activities and emerged as an effective and popular leader who involved teachers and the community in decision-making processes. Parent involvement in their child's education in general and day-to-day schoolwork far exceeded what has been seen in similar urban schools serving many at-risk students (e.g., Ross, Nunnery, et al., 2004). KIPP:DA also devoted substantial time and resources in monitoring and

TABLE 6
Rating of KIPP:DA Relative to Two Frameworks of Effective School Reform

<i>Correlates of Effective Schools</i>		<i>Desimone-Porter Model</i>	
<i>Elements</i>	<i>Rating</i>	<i>Elements</i>	<i>Rating</i>
Clearly stated mission	Strong	Specificity	Moderate
Safe and orderly environment	Strong	Consistency	Strong
High expectations	Strong	Authority	Strong
Instructional leadership	Strong	Power	Moderate
Opportunity to learn	Moderate	Stability	Weak
Monitoring of progress	Moderate		Weak
Enhanced communication	Strong		Weak

assessment, regularly using, for example, the STAR tests from Accelerated Reader to identify individual student reading levels, adapt instruction accordingly, and assess progress. However, the school's strong academic focus and goal of bringing its disadvantaged student population to proficiency levels appeared to limit student participation in free choice activities. This concern was noted by the school staff as an improvement need for the second year.

In Desimone's (2002) model of CSR implementation (see Table 6), *specificity* incorporates clarity and concreteness in strategies, materials, information, and monitoring. Although KIPP is not a highly prescriptive program, such as Success For All or Direct Instruction (see Herman, 1999), its basic principles, clear mission statement, extensive and structured professional development for teachers and leaders, and specific core components (e.g., extended school day, teacher availability) appear to give it at least moderate status in specificity. *Consistency* denotes smooth integration with other reform efforts at the school, district, and state levels. Although there were some communication and resource problems with the district in the first year, the KIPP:DA implementation was generally smooth, fully aligned with district and state curricula and standards, and compatible with district philosophies for improving achievement of at-risk students. In fact, in preliminary interviews with the superintendent and board members, high interest was expressed in disseminating effective KIPP components in other district schools.

The *authority* dimension entails the interactive roles of teachers, the principal, and the district in the model selection and implementation. A related dimension, *power*, concerns the degree to which these and other stakeholders exercise control over decision making and policies. In the case of KIPP:DA, all data sources indicated extensive teacher participation in decision making, site-based autonomy, principal support, and parent involvement (relative to comparable district schools). District accommodation of operational needs (e.g., space, equipment, staffing), however, was perceived by teachers and the principal to be limited or inadequate.

Stability indicates that there has been sufficient time and consistency in the school and policy environment to enact sustainable reforms. Although there was generally high stability for KIPP:DA during the school year (no teacher and minimal staff or student mobility), obviously, there has not been adequate time for full program implementation to take hold, especially given the planned addition of one grade level (6–8) each year.

Given the mobility and burn-out of many urban teachers (Haycock, 1998), an important question for future research concerns the degree to which the increased work and time demands of KIPP:DA impact the sustainability and scale-up of the program. Even when schools make early progress with CSR, implementation quality may diminish as teacher enthusiasm for innovations decreases over time (see Berends et al., 2002; Muncey & McQuillan, 1996; Ross, 2003). Still, evidence associating systematic whole-school change with immediate achievement

gains by at-risk students raises hope for the benefits of KIPP:DA and comparable approaches for middle school reform.

References

- Alberg, M. (2003). *KIPP D.I.A.M.O.N.D Academy: The first year*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Berends, M., Kirby, S. N., Naftel, S., & McKelvey, C. (2002). *Implementation and performance in New American Schools: Three years into scale-up*. Santa Monica, CA: RAND.
- Berliner, D. C. (2002). Educational research: The hardest science of all. *Educational Researcher*, 31(8), 18–20.
- Bloom, B. (1980). *All our children learning*. Highstown, NJ: McGraw-Hill.
- Bobbett, J. L., Ellett, C. D., Teddlie, C., Olivier, D., & Ruggett, J. (2002, April). *School culture and school effectiveness in demonstrably effective and ineffective schools*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Bodilly, S. J. (1998). *Lessons from New American Schools' scale-up phase: Prospects for bringing designs to multiple schools*. Santa Monica, CA: RAND.
- Bodilly, S. J., & Berends, M. (1999). Necessary district support for comprehensive school reform. In G. Orfield & E. H. DeBray (Eds.), *Hard work for good schools: Facts, not fads, in Title I reform* (pp. 111–119). Cambridge, MA: Harvard University, Civil Rights Project.
- Bol, L., Nunnery, J., Lowther, D., Dietrich, A., Pace, J., Anderson, R., Bassoppo-Moyo, T., et al. (1998). Inside-in and outside-in support for restructuring: The effects of internal and external support on changes in the New American Schools. *Education and Urban Society*, 30, 358–384.
- Borman, K. M., Carter, K., Aladjem, D. K., & LeFloch, K. C. (2004). Challenges for the future of comprehensive school reform. In C. T. Cross (Ed.), *Putting the pieces together: Lessons from comprehensive school reform research* (pp. XX–150). Washington, DC: National Clearinghouse for Comprehensive School Reform.
- Borman, G. D., Hewes, G., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, 73, 125–230.
- Bryk, A., & Schneider, B. (2002). *Trust in school: A core resource for improvement*. New York: Russell Sage Foundation.
- Bulkley, K., & Fisler, J. (2002). *A review of the research on charter schools* (CPRE WP-01). Pittsburgh, PA: Consortium for Policy Research in Education.
- Butler, E. D., & Alberg, M. J. (1989). *The Tennessee School Climate Inventory: Resource manual*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Collins, T. (1999). *Charter Schools: An approach for rural education?* Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools. (ERIC Document Reproduction Service No. ED425896).
- Center for Research in Educational Policy. (2004). *SOM national norms*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Pearson Education.
- CTB/MacMillan/McGraw-Hill. (1997). *TerraNova Technical Bulletin I*. Monterey, CA: Author.
- Datnow, A., Hubbard, L., & Mehan, H. (2002). *Extending educational reform: From one school to many*. New York: Routledge-Falmer.
- Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research*, 72, 433–480.

- Doran, H. C., & Drury, D. W. (2002). *Evaluating success: KIPP educational program evaluation*. Alexandria, VA: New American Schools.
- Edmonds, R. R. (1979, October). Effective schools for the urban poor. *Educational Leadership*, 37, 15-18, 20-24.
- Edmonds, R. R. (1982, December). Programs of school improvement: An overview. *Educational Leadership*, 40, 4-11.
- Education Policy Institute. (2005). *Focus on results: An academic impact analysis of the Knowledge Is Power Program (KIPP)*. Virginia Beach, VA: Educational Policy Institute.
- Feuer, M. J., Towne, L., & Shavelson, R. J. (2002). Scientific culture and educational research. *Educational Researcher*, 31(8), 4-14.
- Fullan, M. (2000). The return of large scale reform. *Journal of Educational Change*, 1, 5-28.
- Good, T. L., & Brophy, J. E. (1987). *Looking into classrooms*. New York: Harper & Row, Publishers, Inc.
- Haycock, K. (1998). Good teaching matters: How well-qualified teachers can close the gap. *Thinking K-16*, 3(2), 3-14.
- Herman, R. (1999). *An educators' guide to schoolwide reform*. Arlington, VA: Educational Research Service.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed-methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Knowledge is Power Program. (1994). *About KIPP*. Retrieved on September 1, 2006 from <http://www.kipp.org/01/>
- Larson, R., Moneta, G., Richards, M., & Wilson, S. (2002). Continuity, stability, and change in daily emotional experience across adolescence. *Child Development*, 73, 1151-1165.
- Levin, H. (1993). *Learning from accelerated schools*. Unpublished paper, Stanford University, Palo Alto, CA.
- Lewis, E. M., Ross, S. M., & Alberg, M. (1999). *School Observation Measure: Reliability analysis*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Manzo, K. K. (2004). Scholar advocates 'brand name' charter school networks. *Education Week*, 23(37), 11.
- McDonald, A. J., & Ross, S. M. (2004). *Parent survey*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Miles, M. B., & Huberman, A. M. (Eds.). (1994). *An expanded sourcebook: Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage.
- Muncey, D. E., & McQuillan, P. J. (1996). *Reform and resistance in schools and classrooms: An ethnographic view of the Coalition of Essential Schools*. New Haven, CT: Yale University Press.
- No Child Left Behind Act of 2001, XX U.S.C.A. § *et seq.* (West 2001).
- Northwest Regional Educational Laboratory. (2000). *Catalog of school reform models* (2nd ed.). Portland, OR: Author.
- Nunnery, J. A., Ross, S. M., & Sterbinsky, A. (2003). *Constructs underlying teacher perceptions of change in schools implementing comprehensive school reform models*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Onwuegbuzie, A., & Teddlie, C. (2003). A framework for analyzing data in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 351-383). Thousand Oaks, CA: Sage.
- Porter, A. (1994). National standards and school improvement in the 1990s: Issues and promise. *American Journal of Education*, 102, 421-449.
- RAND. (2001). *Rhetoric versus reality: What we know and what we need to know about vouchers and charter schools*. Washington, DC: Author.
- Roeser, R., & Eccles, J. (1998). Adolescents' perceptions of middle school: Relation to longitudinal changes in academic and psychological adjustment. *Journal of Research on Adolescence*, 8, 123-158.

- Ross, S. M. (2001, December). *Creating critical mass for restructuring: What we can learn from Memphis* (AEL policy brief). Charleston, WV: AEL Organization
- Ross, S. M. (2003). Effective schools correlates as indicators of educational improvement: An examination of three urban reform initiatives. *Journal for Effective Schools*, 2(2), 67–81.
- Ross, S. M., & Alberg, A. (1999). *Comprehensive School Reform Teacher Questionnaire*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Ross, S. M., & Gil, L. (2004). The past and future of comprehensive school reform: Perspectives from a researcher and practitioner. In C. T. Cross (Ed.), *Putting the pieces together: Lessons from comprehensive school reform research* (pp. 151–174). Washington, DC: National Clearinghouse for Comprehensive School Reform.
- Ross, S. M., McDonald, A. J., & Alberg, A. (2002). *Program benchmarking for raising student achievement*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Ross, S. M., McDonald, A. J., Bol, L., Dietrich, A., Scott, J., & Siegel-Robertson, J. (2004). *First Year Evaluation of Tennessee Charter Schools 2003–2004*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Ross, S. M., McDonald, A. J., & McSparrin-Gallagher, B. M. (2004). *Year 1 evaluation of the KIPP DIAMOND Academy: Analysis of TCAP scores for matched program-control group students*. Center for Research in Educational Policy, The University of Memphis.
- Ross, S. M., Nunnery, J. A., Goldfeder, E., McDonald, A., Rachor, R., Hornbeck, M., & Fleischman, S. (2004). Using school reform models to improve reading achievement: A longitudinal study of Direct Instruction and Success For All in an urban district. *Journal of Education for Students Placed at Risk*, 9, 357–388.
- Ross, S. M., Smith, L. J., & Alberg, M. (1998). *School Observation Measure*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Ross, S. M., Smith, L. J., Alberg, M., & Lowther, D. L. (2004). Using classroom observation as a research and formative evaluation tool in educational reform: The School Observation Measure. In H. Waxman (Ed.), *Observational research in U.S. classrooms: New approaches for understanding cultural and linguistic diversity* (pp. 144–173). Cambridge, MA: Cambridge University Press.
- Ross, S. M., Troutman, A., Horgan, D., Maxwell, S., Laitinen, R., & Lowther, D. (1997). The success of schools in implementing eight restructuring designs: A synthesis of first-year evaluation outcomes. *School Effectiveness and School Improvement* 8, 95–124.
- Rowan, B., Camburn, E., & Barnes, C. (2004). Benefiting for comprehensive school reform: A review of research on CSR implementation. In C. T. Cross (Ed.), *Putting the pieces together: Lessons from comprehensive school reform research* (pp. 1–52). Washington, DC: National Clearinghouse for Comprehensive School Reform.
- Simmons, R., & Blyth, D. (1987). *Moving into adolescence*. New York: Aldine de Gruyter.
- Slavin, R. E., & Fashola, O. S. (1998). *Show me the evidence?* Thousand Oaks, CA: Corwin Press.
- Smith, L., Maxwell, S., Lowther, D., Hacker, D., Bol, L., & Nunnery, J. (1997). Activities in schools and programs experiencing the most, and least, early implementation success. *School Effectiveness and School Improvement*, 8, 125–150.
- Snipes, J. C., & Casserly, M. D. (2004). Urban school systems and education reform: Key lessons from a case study of large urban school systems. *Journal of Education for Students Placed at Risk*, 9, 127–142.
- Sterbinsky, A., & Ross, S. M. (2003). *School Observation Measure Reliability Study*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- U.S. Department of Education. (1999). *Guidance on the comprehensive school reform program*. Washington, DC: Author.
- Wainer, H., & Robinson, D. H. (2003). Shaping up the practice of null hypothesis significance testing. *Educational Researcher*, 32, 22–30.

- Wang, M. C., Haertel, G. D., & Walberg, H. (1997). *What do we know? Widely implemented school improvement programs*. Philadelphia: Temple University Center for Research in Human Development and Education.
- Way, N., & Robinson, M. (2003). A longitudinal study of the effects of family, friends, and school experiences on the psychological adjustment of ethnic minority, low-SES adolescents. *Journal of Adolescent Research, 18*, 324–346.