

Charter Schools as a Vehicle for Education Reform: Implementation and Outcomes at Three Inner-City Sites

Aaron J. McDonald and Steven M. Ross

Center for Research in Educational Policy, The University of Memphis

Linda Bol

*Department of Educational Curriculum and Instruction, Old Dominion
University*

Brenda McSparrin-Gallagher

Office of Research and Evaluation, Memphis City Schools

This study examined the impact of 3 2nd-year charter schools (1 elementary, 1 middle, and 1 high school) on student achievement, school climate, and pedagogy. All schools served predominantly African American students in an inner-city district. Using a matched treatment-control student analytical design, charter school enrollees were individually matched to highly comparable control students of the same ethnicity, poverty level, gender, and ability. Qualitative and descriptive analyses showed reasonable to good progress in program implementation, very strong school climate, positive teacher and parent perceptions, largely traditional but academically-focused teaching, and positive student achievement on state-mandated tests ($p < .05$ on 12 out of 18 school \times cohort \times subtest comparisons). Interpretations of results stress the likely implications of teacher and family choice for effective implementation of the charter schools' academic and organizational programs.

Efforts to enact school reforms, particularly in urban districts, have been highly prevalent during the past 2 decades (Snipes, Doolittle, & Herlihy, 2002; Stringfield & Yakimowski-Sreblick, 2005). To historians and followers of school reform efforts, several national initiatives have established important groundwork for sys-

temic school improvement in these settings. Two key events were the passage of Goals "2000" and the reauthorization of Title I in 1994 to support both schoolwide and targeted programs (e.g., Fullan, 2000; Tyack & Tobin, 1994). Another important event was the enactment by Congress in 1999 of the Comprehensive School Reform Demonstration (CSR) program (U.S. Department of Education, 1999). By 2002, over 380 school reform models were adopted with CSR support (Desimone, 2002). Despite some documented success of the Comprehensive School Reform (CSR) approach in raising student achievement (American Institutes for Research, 2005; Borman, Hewes, Overman, & Brown, 2003; Rowan, Camburn, & Barnes, 2004), individual schools' adoption of externally developed restructuring models has been diminishing in favor of local integrations of internal, school-based, and district-wide programs and reforms (Ross & Gil, 2004).

An important lesson learned from the CSR movement is that struggling schools do not automatically improve by overlaying new curricula or research-based instructional programs on shaky structural and cultural foundations (Borman, Carter, Aladjem, & LeFloch, 2004). Strong leadership, along with effective teachers and involved parents, are critical factors for success, regardless of the model adopted (see reviews by Desimone, 2002; Stringfield & Yakimowski-Srebniak, 2005). Consistent with this philosophy is the recent surge of national interest in educational choice options as a foundation for educational reform (Tough, 2006). Such options are directly promoted by the allowance in the 2001 No Child Left Behind (NCLB) legislation (U.S. Congress, 2001) for students in failing schools to transfer to higher performing district schools. The underlying assumption is that if schools are forced to compete for students, they will be more likely to improve teaching and learning (Frankenberg & Lee, 2003). Charter schools, the focus of this study, are a central part of this growing national movement.

Charter schools can be broadly characterized as publicly funded schools that students can choose to attend (Bulkley & Fisher, 2002). Historically, the passage of the first charter school law and establishment of the initial charter school occurred in Minnesota in 1991. As of September 2006, approximately 4,000 charter schools were in operation in the United States (Center for Education Reform, 2006). Under NCLB legislation, children who attend schools identified as needing improvement have the opportunity to enroll in charter schools located within their district (U.S. Department of Education, 2001). Charter schools are intended to have maximum flexibility to achieve alternative ways for public schools to educate school children (see Wong & Shen, 2006).

Despite this factor and the potential appeal of flexibility and choice as reform elements, previous research on charter schools has indicated mixed results in their success (e.g., Anderson et al., 2002; Berends, Watral, Teasley, & Nicotera, 2006; Carnoy, Jacobson, Mischel, & Rothstein, 2005; Miron & Nelson, 2004). For example, a recent study by the American Federation of Teachers (Nelson, Rosenberg, &

Van Meter, 2004) concluded that charter schools scored lower on average than regular schools on the National Assessment of Educational Progress (NAEP). However, the failure of that research design to control for individual student and school variables represented a potentially serious limitation, particularly in view of the expectancy for most charter schools to give priority to enrolling low-performing and socioeconomically disadvantaged students (Buckley & Schneider, 2005; Paige & Huckabee, 2005).

Similarly, in a national study using hierarchical linear modeling (HLM) analyses to compare 150 charter schools to 6,764 traditional public schools, Braun, Jenkins, Grigg, and Tirre (2006) found that mean reading and mathematics scores on the NAEP were lower on average for the charter schools after adjusting for student characteristics. Noteworthy limitations of that study, however, include: (a) failure to control for students' prior achievement, given that charter school students tend to be lower performing as an enrollment requirement, (b) the NAEP not being directly aligned with or not being part of schools' NCLB accountability requirements, and (c) the short time the charter schools were in operation.

Solomon and Goldschmidt (2004) also used HLM in comparing longitudinal achievement growth trajectories in 873 charter and traditional schools in Arizona, controlling for both student and school characteristics. In contrast to the Braun et al. (2006) study, findings indicated that although charter school students began with lower test scores than the traditional school students, they demonstrated higher annual achievement growth in elementary school. No differences were obtained in the middle grades, whereas higher growth was indicated for the traditional students in high schools. The authors speculated that the elementary charters were more likely to focus on academics, and the middle and high school charters on vocational education and special needs students.

In contrast to the previously mentioned study, Bifulco and Ladd (2006) found negative effects, using a fixed effects model, for charter school students in Grades 3 to 8. These types of contrasting and varying achievement findings for charter school students are not uncommon. In a review of charter school studies by Hill, Angel, and Christensen (2006), for example, the authors concluded that the results were mixed.

Because charter schools are, by definition, chosen by students, it is rarely feasible to conduct randomized experimental studies on their effectiveness.¹ Where such studies are precluded, alternative designs can offer moderately high to high

¹A special circumstance would be where a sufficiently large number of students apply to different charter schools, but due to space restrictions at each school, selections are made on a random basis. Despite the advantages of such randomized designs for internal validity, the generalizability of results might be restricted to only those charter schools that are relatively popular in the community (perhaps due to already producing high achievement), thus threatening external validity for making inferences about the broader charter school population.

internal validity (Betts & Hill, 2006). A possible option if sufficient numbers of students transfer in and out of charter schools is a “fixed effects” analysis. By comparing growth trajectories in both charter and regular schools (e.g., Ballou, Teasley, & Zeidner, 2006), the fixed effects model offers the advantage of controlling for individual student variables. However, such models are weakened by essentially excluding the many students who continually attend charter schools or who never enter charters.

Another possible option, and the one best describing our mixed-methods research design, is the matched-samples design, in which charter school students are matched to students in regular public schools who are very similar in key characteristics (Betts & Hill, 2006; Hill, 2005). One advantage is establishing highly similar comparison groups with regard to student ethnicity, gender, poverty, prior achievement, and traditional school previously attended. Further statistical controls can then be applied via covariate analyses. Another advantage is ease of interpretability for policymakers and practitioners who can more readily comprehend comparison outcomes for the two groups.

This research consists of mixed-methods (Chatterji, 2005; Johnson & Onwuegbuzie, 2004) case studies of three charter schools—an elementary school, a middle school, and a high school—each in its second year of operation in a large urban district. One underlying assumption of our research design is that *charter schools* are not a unitary treatment but separate, potentially unique experiments in choice-based reform (Buckley & Schneider, 2005; Center for Education Reform, 2006). A second assumption is in agreement with both the tenets of mixed-methods research in general and many charter school researchers in particular, supporting the importance of “looking inside the black box of schools to better understand the conditions under which [individual] schools of choice have, or do not have, effects on achievement” (Goldring & Cravens, 2006, p. 3). As a foundation for this study, our preliminary research (Ross, McDonald, & McSparrin-Gallagher, 2005; Ross, McDonald, & Bol, 2004) revealed varied but mostly successful first-year implementation of curricular, instructional, and organizational goals, and directionally (though nonsignificant) positive achievement patterns when charter students were compared to matched traditional school counterparts.

Research questions addressed in this study concerned the impacts of three second-year inner-city charter schools on student achievement, school climate, stakeholder perceptions, and pedagogy. To address the student achievement question in the most powerful way possible, given the inability to randomly assign students to schools, we individually matched each charter school student to a demographically similar counterpart who attended the same grade in a traditional school serving the same geographic area. A slightly modified method was used for the elementary school, due to lack of pretest scores, which will be explained in the following. Along with student achievement, we also recognize the importance to

educational change of successful program implementation, positive school climate, improved pedagogy, and teacher support (see Rowan et al., 2004). Thus, additional research interests concerned the impacts of these factors as potential moderators and products of the charter school implementation.

METHOD

Participants and School Contexts

Charter school participants were faculty, administrators, parents, and students at the three urban charter schools, which were established in 2003–2004. The schools were all new, as opposed to being conversion schools. For clarity and anonymity purposes, we refer henceforth to the three schools using the pseudonyms ELEM, MID, and HIGH. All three schools were part of a large urban school district serving over 100,000 students, of which 71% are disadvantaged and over 80% are African American.

ELEM specifically served 72 African American students (92% poverty) in Grades K–3. The school will eventually serve students in Grades K–5. ELEM’s staff consisted of four teachers, with a teacher–student ratio of 1:18. Other staff members included a principal, a curriculum facilitator, a Title I coordinator, an administrative assistant, and three teacher assistants. The general focus at ELEM is literacy development, which its schoolwide program attempts to integrate across subject areas. A key school goal is to use a variety of teaching strategies to meet students’ needs. Organizational structures include the systematic monitoring of grade-level planning, classroom teaching practices, and student progress. Support structures include the incorporation of feedback from the principal and the curriculum facilitator, modeling of effective practices, and the provision of regular professional development opportunities for all staff.

MID served 180 students (100% African American, 75% poverty) in Grades 6 and 7. The school will add an eighth grade for the 2005–2006 academic year. The school’s faculty and staff consisted of 12 teachers (including one special education teacher), one principal, and one secretary. The teacher–student ratio was 1:20. The curriculum is a standards-based, interdisciplinary program that aims to incorporate projects and experiential learning centered on a health science theme. The goals established by the school for its second year of operation primarily focused on establishing structures that enable effective implementation and evaluation of the instructional program, productive communication among stakeholders, and a positive environment for teaching and learning.

HIGH served 251 seventh- and eighth-grade students, 98% of whom were African American (68% poverty). The school will eventually serve students in Grades 7–12. The school staff in 2004–2005 included 12 teachers, two support individu-

als, three administrators, and one employee having both administrative and support duties. The administration consisted of a principal, a vice principal, and an educational specialist. The school's goals include high student technology use, continuous monitoring of student progress, and an intense focus on core curricula (Math, Science, Language Arts) integrated across subjects. The curriculum emphasized preparation for higher education and careers in science and engineering.

Design

Multiple data sources for this study, to be described in more detail in a later section, included (a) the School Observation Measure (SOM; Ross, Smith, & Alberg, 1998); (b) the School Climate Inventory (SCI; Butler & Alberg, 1991); (c) the Charter School Teacher Questionnaire (CSTQ; McDonald & Ross, 2003); (d) teacher focus groups; (e) a principal interview; (f) an implementation benchmark review (Ross, McDonald, & Alberg, 2002); (g) a parent survey; (h) a student focus group; and (i) student-level Normal Curve Equivalent (NCE) scores on the Tennessee Comprehensive Assessment Program: Achievement Test (TCAP).

Given the different grade levels served and curricular objectives emphasized by the three charter schools, the results for each school were analyzed separately. Each analysis, excepting that for ELEM (see Results section), used a matched program-control design at the student level, whereby charter school students were individually matched with noncharter school students on all or most of seven criteria: (a) enrollment at the same school before attending the charter school, (b) grade level, (c) race, (d) gender, (e) free or reduced-price lunch status, and prior achievement (M 's within ± 5.0 NCE points) on TCAP for (f) Reading/Language Arts (R/LA) and (g) Mathematics.

A small percentage of charter students were not able to be matched on all criteria. In those instances, matches were made on prior achievement, grade level, and race. Consequently, lunch status or gender differed for a small number of pairs. Similarly, when an appropriate match for a charter school student could not be found from his or her former school (2002–2003 or 2003–2004), the closest match from one of the schools formerly attended by his or her present classmates was selected using all other criteria. Group equivalence on achievement variables at baseline was confirmed with one-way ANOVAs for each subject area within each school sample. Effect sizes were calculated on these preprogram achievement scores to confirm the similarities between charter and control student groups. Further details of the matching process for each school are included in the Results section.

Instrumentation and Measures

A summary of the instrumentation employed to address each research question is provided in Table 1. Descriptions of each follow.

TABLE 1
Evaluation Questions by Instrument

<i>Evaluation Questions</i>	<i>Instruments/Data Sources</i>
1. What are the immediate and long-term impacts of the charter school implementation on student achievement?	TCAP scores in R/LA and Math
2. What is the frequency of usage of various traditional and alternative instructional strategies?	School Observation Measurement (SOM) Rubric for Student-Centered Activity (RSCA)
3. What is the school climate at the charter school? How do charter school climate outcomes compare to those reflected in national norms?	School Climate Inventory (SCI) Teacher Focus Group Student Focus Group Principal Interview
4. To what degree are charter schools implementing their primary curriculum, instructional, and administrative programs?	Implementation Benchmarks Teacher Focus Group Principal Interview Charter School Teacher Questionnaire (CSTQ) Student Focus Group SOM/RSCA
5. What are teacher reactions to and experiences in the charter school?	CSTQ Teacher Focus Group Principal Interview
6. What are parent (caregiver) reactions to and experiences with the charter school?	Parent Survey Principal Interview Teacher Focus Group Student Focus Group SOM

TCAP. TCAP is a criterion-referenced multiple-choice test administered each spring. The test is mandated by the State of Tennessee for assessing adequate yearly progress in the subjects of R/LA and Math in Grades 3–8, in compliance with NCLB policies. In the participating school district, TCAP is also administered in Grade 2 in all schools by district requirement, and in Grade 1 at the discretion of individual schools.

SCI. The SCI consists of seven dimensions logically and empirically linked with factors associated with effective school organizational climates (Butler & Alberg, 1991). Each scale contains seven items, with 49 statements comprising the inventory. Responses are scored through use of Likert-type ratings ranging 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, 1991). 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 ($\alpha = .74$).

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 standard, or whole-school, SOM 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 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 the degree of student attention and interest.

The SOM strategies include traditional practices (e.g., direct instruction and independent seatwork) and alternative, predominately student-centered methods associated with educational reforms (e.g., cooperative learning, project-based learning, inquiry, discussion, technology use as a learning tool). The strategies were identified through surveys and discussions involving policy makers, 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 training, a manual providing definitions of terms, examples, and explanations of the strategies, and a description of procedures for completing the instrument. After receiving the manual and instruction in a group session, each observer participates in sufficient practice exercises to ensure that his/her data are comparable with those of experienced observers. 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. Further results establishing the reliability and validity of SOM are provided in the Lewis et al. (1999) report. In a reliability study using Generalizability Theory, Sterbinsky and Ross (2003) found reliability at the .74 level for 5 SOMs conducted at a school. Reliability increased to .82 with 8 SOMs and to .85 with 10 SOMs conducted at a school.

Rubric for Student-Centered Activities (RSCA). The RSCA was developed (Lowther, Ross, & Plants, 2000) as an extension to SOM, using comparable observer training preparation and certification. The RSCA is used by observers to more closely evaluate the degree of learner engagement in seven selected areas considered fundamental to the goals of increasing student-centered learning activities (cooperative learning, project-based learning, higher-level questioning, experiential/hands-on learning, student independent inquiry/research, student discussion, and students as producers of knowledge using technology).

Each item on the RSCA includes a two-part rating scale. The first is a four-point scale, with 1 indicating a very low level of application, and 4 representing a high level of application. The second is a *yes-no* option to the question: "Was technology used?" with space provided to write a brief description of the technology use. The RSCA was completed as part of SOM observation periods. In a reliability analysis, pairs of trained observers selected the identical overall response for 87% of the items. Slightly lower than 10% (8.0%) had a difference score of 1, 3.0% had a difference score of 2, 2.0% had a difference score of 3, and less than 1% (.3%) had a difference score of 4 (Sterbinsky, Ross, & Burke, 2004).

CSTQ. This questionnaire, adapted from the Comprehensive Reform Teacher Questionnaire (Ross & Alberg, 1999), was designed to assess teacher perceptions about the school in the areas of professional development, support, pedagogical change, and outcomes. Included on the questionnaire are 20 closed-ended items using a five-point Likert-type scale—with scores ranging from 1 (*strongly disagree*) to 5 (*strongly agree*)—and the following four open-ended questions:

- What do you see as positive or most successful aspects of your charter school?
- What do you consider to be negative aspects or areas in need of improvement at your school?
- In your opinion, what makes a charter school (like this one) different from a regular public school?
- Any other comments you would like to make regarding your experiences as a charter school teacher?

Charter School Parent Questionnaire. This instrument was designed to obtain parent perceptions of the school in areas such as instruction, curriculum, communication, and opportunities for involvement. Included on the questionnaire are 17 closed-ended items using a five-point Likert-type scale—with scores ranging from 1 (*strongly disagree*) to 5 (*strongly agree*)—and the following four open-ended questions:

- What are the strengths of this school?
- What would you like to see improved at this school?
- In your opinion, what makes a charter school (like this one) different from a regular public school?
- Any other comments you would like to make about this school or charter schools in general?

Focus groups and interviews. To supplement the survey data, site researchers at each charter school conducted a principal interview, student focus groups, and a teacher focus group. The time period was approximately 30–60 min for each interview. Teacher and 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 regular (noncharter) schools, reactions to major school components (e.g., teaching methods, curriculum, parent involvement), perceived strengths and weaknesses, and recommendations for improvement.

Procedure

A site researcher was assigned data collection responsibility for each charter school. This individual and support research staff visited the assigned school several times during the year to conduct six SOM visits, administer the questionnaires, and conduct the interview and focus groups. In addition, the site researcher worked with the school leadership team to develop implementation benchmarks describing beginning, intermediate, and full implementation phases (labeled Phases I–III, respectively) and associated evidence for major school components in the areas of curriculum, instruction, and organization. Each school developed customized implementation goals that were tailored to their specific programmatic objectives. At the end of the year, the site researcher met again with the leadership team to determine and identify the phase (implementation progress) that had been achieved for each benchmark goal.

RESULTS

Because achievement tests differed by grade level, we conducted separate analyses for each grade, using Analysis of Covariance (ANCOVA) or Multivariate Analysis of Covariance (MANCOVA) where prior achievement (pretest) scores were available, or ANOVA or MANOVA where there were no pretest scores. Effect sizes were calculated for both unadjusted and adjusted mean differences within each subject area within each school sample using Cohen's *d*. Qualitative analyses, guided by Miles and Huberman's (1994) analysis model, were performed on open-ended survey and interview responses. The procedure consisted of transcribing the responses, deriving codes, identifying themes, and revision and refinement based on member checking and interrater review. Triangulation across data sources and methods was used to validate the major findings. In the interests of brevity for this article, main findings from the quantitative and qualitative analyses of survey, interview, and observation data are summarized in the following by school.² Charter school outcomes on the SOM and SCI are summarized in Tables 2 and 3, respectively.

ELEM SCHOOL

Program Implementation Goals

Substantial progress had been made in the attainment of implementation benchmark goals at the conclusion of the school's second year of operation. Specifically, ELEM continued to extend its implementation to all grades of the Success for All Program in reading and math. The implementation of science and fine arts programs were in preliminary stages of development. Although progress in meeting instructional goals was noted, observational data indicated that diversity in instructional strategies could be improved. In the categories of organization, support, and evaluation, notable progress was apparent and all data sources indicated more advanced stages of goal implementation than in the first year of operation year (see Ross, McDonald, et al., 2004).

Pedagogy

As shown in Table 2, direct instruction was the most frequently observed instructional strategy at ELEM. In fact, the percentage of visits in which direct instruction

²A detailed reporting of qualitative and descriptive results from surveys, observations, and interviews is provided in Ross, McDonald, et al. (2005).

TABLE 2
Means for Observation of Instructional Strategies
on the School Observation Measure (SOM)

Observation Component	ELEM		Elementary Norm		MID		HIGH		Secondary Norm	
	M	SD	M	SD	M	SD	M	SD	M	SD
	Direct instruction	3.33	.82	2.75	1.07	3.17	.98	2.83	.98	2.85
Team teaching	.33	.52	.74	.90	0.00	0.00	0.00	0.00	.42	.61
Cooperative learning	2.17	1.17	.91	.94	1.00	.63	.33	.51	.97	.89
Individual tutoring	.50	.84	.89	1.00	0.00	0.00	0.00	0.00	.34	.74
Ability groups	.50	.84	1.43	1.41	.83	1.33	1.00	1.67	1.16	1.29
Multi-age grouping	.33	.52	.54	.98	.17	.41	0.00	0.00	1.19	1.47
Work centers	1.17	1.33	1.40	1.14	1.33	1.03	.50	1.22	.38	.63
Higher level feedback	2.17	1.17	1.20	1.00	.67	.82	1.83	.41	1.55	1.22
Integration of subject areas	.67	.52	.56	.79	.17	.41	.17	.41	.51	.86
Project-based learning	.33	.82	.37	.75	.67	1.21	1.00	.89	.67	.82
Higher level questioning	2.00	1.41	1.47	1.05	1.33	.82	1.33	.82	1.53	1.04
Teacher as coach or facilitator	2.50	1.22	2.19	.31	1.50	1.38	1.00	1.10	2.52	1.13
Parent/community involvement	.17	.41	.31	.60	0.00	0.00	0.00	0.00	.08	.31

was frequently to extensively observed increased from 67% in the first year (Ross, McDonald, et al., 2004) to 83% in the second year. These percentages were higher than the national elementary school norms (Center for Research in Educational Policy [CREP], 2005). In contrast, there was very limited usage of independent seatwork, another traditional strategy. With regard to student-centered teaching strategies, the percentage of time allotted to both cooperative learning and teacher coaching remained constant across years and higher than national norms across years. Based on RSCA outcomes, cooperative learning and higher level questioning were implemented with greater quality and intensity than in the first year. However, findings indicated very limited use of technology. In general, ELEM demonstrated moderate progress toward achieving its goal of varying instruction via greater usage of student-centered strategies.

School Climate

SCI means (see Table 3), along with focus group and open-ended survey responses, indicated that school climate was very positive at ELEM. The overall SCI mean rating was 4.59, with means for each dimension ranging from a low of 4.37

TABLE 3
Charter School Means, Standard Deviations, and Effect Sizes on School
Climate Dimensions

Climate Dimension	Charter Schools								
	ELEM			MID			HIGH		
	M	SD	ES	M	SD	ES	M	SD	ES
Collaboration	4.63	0.52	+1.23	4.09	0.81	+0.58	4.20	1.32	+0.71
Environment	4.37	0.90	+0.61	3.77	0.96	+0.06	4.06	1.29	+0.46
Expectations	4.71	0.49	+1.12	4.29	0.86	+0.73	4.57	0.74	+1.17
Instruction	4.66	0.48	+0.98	4.34	0.72	+0.54	4.43	0.91	+0.73
Involvement	4.64	0.57	+1.25	3.86	1.01	.37	4.06	1.06	+0.70
Leadership	4.75	0.47	+0.96	4.08	0.87	+0.19	4.24	1.16	+0.42
Order	4.39	0.47	+1.11	3.47	0.87	+0.24	3.68	1.16	+0.48
Total Climate	4.59	0.64	+1.20	3.99	0.93	+0.44	4.18	1.18	+0.76

Note. Effects sizes are derived by subtracting the national norm mean for elementary schools or middle/high schools from the school dimension mean and dividing by the national norm standard deviation.

for environment to a high of 4.75 for leadership. Compared to the initial year, the mean ratings increased on six of the seven climate dimensions and all exceeded national elementary norms (CREP, 2005), often substantially (median $ES = +1.12$).

The qualitative data support the SCI results. Student responses specifically reflected a great deal of respect and caring among teachers and students, which appeared to facilitate classroom management. Some students described their prior schools as very punitive and appreciated the more positive approach to discipline in their new school. According to one student respondent, "I used to get paddled, but teachers here give us a second chance or give us the bad grade we deserve." Another student said, "Teachers at [ELEM] don't accuse us. They say, 'If you fall short, you can lift back up.'" Similarly, teachers in the focus group characterized the relationship among faculty and students as familial, describing a "great sense of family and belonging" realized in a small school setting. Teachers were especially laudatory about school leadership, collaboration, and the creativity associated with a flexible teaching and learning environment. For example, one teacher commented, "You have the opportunity to be creative. The teacher is really the educator. She has more control of determining when students are ready to move forward."

Teacher Attitudes

The responses from the CSTQ and teacher focus group were organized into four themes: program/mission, professional development, resources, and external support. ELEM teachers were unanimously positive about the school's program and

mission and their professional development. The only suggestion was that new teacher orientation could be improved. One teacher noted, "Improvement in orientation could be enhanced. Since policies and procedure are different than the regular school setting, it is important to be sure that teachers know exactly what these differences are." Perceptions about the adequacy of resources were largely favorable. The only resource-related area perceived to be inadequate was the ability to serve children with special needs. Responses were mixed in reference to support received from the state and school district. Teachers unanimously agreed on the CSTQ, however, that parents were active partners with the school. In open-ended comments, teachers cited this involvement as "the most positive and successful aspect of our charter school." One teacher described their open-door policy with parents that promoted a "warm, family-centered environment for our students, parents, and faculty." Of note, several commented that most parents fulfilled and even exceeded the requirement of volunteering 20 hr of time to the school.

Parent Reactions

Parent questionnaire data indicated continued satisfaction on nearly all indicators. The only item that received a low rating concerned the adequacy of transportation services. Open-ended comments were overwhelmingly positive, particularly regarding the principal's vision and leadership, as well as the teachers' high expectations and caring for their children as individuals. Parents further appreciated the low student-to-teacher ratio and the positive, nurturing school climate. The following quote from one parent illustrates the intensity of response that characterized many comments.

[ELEM] is a nurturing place; they make each and every family feel special. They are led by a woman with knowledge, wisdom, devotion, sincerity, and a genuine love for education, the parents, and the children. The staff is a reflection of her and what a lovely picture they make. They care and allow every child to shine no matter what.

Suggestions for improvement pertained primarily to resources such as physical education equipment, musical instruments, uniforms, technology, and more nutritious meals.

Student Achievement

ELEM students with test score data included 19 first graders, 16 second graders, and 16 third graders. All students were African American and qualified for free or reduced-price lunch. Because TCAP testing begins in Grade 2 in the school district (optional in Grade 1), it was not possible to establish a true matched-pair control group, as done for the other charter schools. Instead, we created a comparison

group by randomly selecting students of like ethnicity and gender who qualified for free or reduced-price lunch from the same traditional schools that the third-grade ELEM students had attended (in 2002–2003) prior to their charter school enrollment. Due to the small sample sizes, we conducted separate ANOVAs within grades on the R/LA and Math posttest scores, using a conservative alpha of .025 (.05 divided by the 2 subject areas). Posttest means, standard deviations, and effect sizes are summarized in Table 4.

Grade 1. The univariate ANOVAs on posttest scores revealed significant ($\alpha = .025$) results favoring ELEM students in R/LA, $F(1,37) = 10.41, p = .003$, and Math, $F(1,37) = 6.96, p = .012$. Effect sizes reflected strong advantages in both subjects (respective ES 's = +1.21 and +0.84).

Grade 2. The ANOVAs yielded nonsignificant outcomes reflecting fairly comparable means ($ES = +0.03$) for both groups in R/LA, $F(1,34) = 0.01, p = .946$, but a directional advantage ($ES = +0.59$) for ELEM in Math, $F(1,34) = 3.71, p = .06$.

Grade 3. The univariate ANOVAs revealed nonsignificant results in both R/LA, $F(1,33) = .177, p = .68$, and Math, $F(1,33) = 0.02, p = .90$. The associated effect sizes of +0.13 and +0.04 were indicative of small program impacts.

MID SCHOOL

Program Implementation Goals

Notable progress toward full implementation of major program goals was recorded relative to the start-up year. Implementation benchmarks and indicators related to evaluation and assessment were in Phase III of development (full implementation). Across sources, the data indicated Phase II (intermediate) attainment of curriculum and instruction benchmarks. According to the responses from teachers and the principal, planning was still underway for adopting technology-based programs in math and science, as well as the increased utilization of interdisciplinary, project-based learning. Observed student-centered instruction was limited but still higher than the prior year. Implementation of goals directed at organization and support were either Phase II or approaching Phase III. Nearly all teachers indicated that they understood the school mission and were adequately involved in school decision making. Inhibiting factors were identified as lack of resources (staff, equipment) and low parent involvement.

TABLE 4
Year 2 Posttest Normal Curve Equivalent Achievement Scores in Reading/Language Arts and Mathematics for Charter and Control Students by School and Grade on the Tennessee Comprehensive Assessment Program

ELEM ^a	Group	Reading/Language Arts					Mathematics				
		n	M	SD	Adj M	ES	n	M	SD	Adj M	ES ^b
1st Grade	Charter Students	19	69.53	20.90	NA	+1.213**	19	54.05	18.06	NA	+0.841*
	Control Students	20	50.50	15.69	NA		20	38.70	18.26	NA	
2nd Grade	Charter Students	16	45.63	18.73	NA	+0.027	16	52.94	15.16	NA	+0.594
	Control Students	20	45.25	14.29	NA		20	41.75	18.85	NA	
3rd Grade	Charter Students	16	28.13	6.67	NA	+0.129	16	31.25	6.62	NA	+0.039
	Control Students	19	27.00	8.75	NA		19	30.95	7.67	NA	
MID											
6th Grade (1st Year)	Charter Students	78	32.64	7.76	32.64	+0.795***	78	34.71	8.37	34.69	+0.582***
	Control Students	78	26.27	8.01	26.27		78	29.85	8.30	29.86	
7th Grade (1st Year)	Charter Students	30	31.17	9.06	31.18	+0.501*	30	33.23	8.89	33.27	+0.604**
	Control Students	30	26.30	9.66	26.28		30	28.50	7.96	28.46	
7th Grade (2nd Year)	Charter Students	42	34.38	8.93	34.38	+0.777**	42	32.76	8.51	32.77	+0.376**
	Control Students	42	27.64	8.67	27.64		42	29.50	8.71	29.49	
HIGH											
7th Grade (1st Year)	Charter Students	91	35.68	7.39	35.66	+0.377***	91	38.24	7.08	38.31	+0.1444
	Control Students	91	31.92	9.88	31.94		91	35.92	9.19	35.85	
8th Grade (1st Year)	Charter Students	25	36.16	6.96	36.22	+0.518	25	38.24	7.08	38.31	+0.268
	Control Students	25	32.32	7.64	32.26		25	35.92	9.19	35.85	
8th Grade (1st Year)	Charter Students	99	35.68	7.26	35.71	+0.240	99	38.13	8.51	38.25	+0.431***
	Control Students	99	33.64	8.78	33.60		99	34.40	9.21	34.28	

^aBecause pretest scores were not available, no adjusted means or adjusted effect sizes were computed.

^bEffect sizes reported are unadjusted for pretest for ELEM, but adjusted for MID and HIGH.

p* < .05. *p* < .01. ****p* < .001.

Pedagogy

The percentage of time devoted to direct instruction remained high in the second year and was somewhat higher than the secondary school norm (see Table 2). Student engagement in independent seatwork was also frequently observed but at a normative level. Although student-centered strategies (e.g., cooperative learning, higher-order learning, and technology as a learning tool) were used more frequently than in the first year, their application was still limited. In contrast, computer use for instructional delivery was observed at least occasionally during 50% of the observations, a level that exceeded national norms. Overall, the results suggest reliance on traditional pedagogy with slight movement toward supplementing direct instruction with alternative teaching strategies.

School Climate

The mean climate scores for MID were largely positive (see Table 3) and higher on each dimension than national norms. The highest mean rating was obtained for *instruction* and the lowest was for *order*; effect sizes relative to national norms were also lowest for *order* (+0.24) but highest for *collaboration* (+0.58). Specific areas of concern were student tardiness and absences, and perceived lack of support from parents. Several climate strengths emerged from the interview and open-ended responses to questionnaire items. One strength was the principal's efforts to create ownership of the school by staff and students. The principal noted that, "Creating ownership is very important. I try to get my teachers what they want and need to do their jobs." The theme was echoed by the teachers themselves who talked about "taking ownership of the school" as well as the "willingness/commitment of teachers to implement the [educational] program." A second strength was a commitment to student success and improvement over time. One teacher wrote, "The students are given every avenue to attain and retain the material and information for standardized assessments and life skills." A third area of strength was the family atmosphere within the school, which was described by teachers as welcoming, warm, and inviting. Students were mostly positive about the school, noting the high expectations set by their teachers and principal. An exemplary comment was, "The teachers and principal are strict. They push us to do what we're supposed to do."

Teacher Attitudes

Results reflected overwhelming teacher support for the school's mission, educational programs, and perceived positive student outcomes. Professional development was viewed as adequate but less so than in the first year of implementation. According to the teachers, the principal was the primary source of professional de-

velopment this year, and more was needed from different sources. Teacher ratings and open-ended comments further indicated, more strongly than in the first year of operation, the need for increased resources for planning time, educational materials, technology, and serving special education students. Perceptions of support received from local and state educational agencies, external partners, and parents were noticeably less favorable when compared to last year. During focus groups, the teachers noted a lack of support from the local district, which appeared to be waiting to see if MID would sink or swim.

Parent Reactions

Consistently high ratings on nearly all questionnaire items indicated strong parent satisfaction at MID. However, only half of the parents agreed that transportation services were adequate. Open-ended comments about the school's strengths cited the academically challenging curriculum, the small student-to-teacher ratio, high expectations for student learning, principal leadership, and teacher quality. The following parent quote captures many of these strengths.

The curriculum is diverse. The teaching style is not mechanical. If a child is having trouble in a particular area, that child is not left behind to figure the subject out for himself. Every effort is made to ensure that each child has an opportunity to master a subject. Preventative measures are taken initially when a student is having problems, not at TCAP time nor at the end of the year.

The most frequently expressed dissatisfaction was the need for a new school building that would provide more space for physical education, sports, science labs, and performing arts. Parents also thought that communication, discipline, and attitudes of some teachers could be improved.

Despite the generally high parent satisfaction, questionnaire results revealed a sharp decline in agreement from the first year (from 86% to 36%) that parents are active participants in the school. Corroborating this perception, parent involvement was never observed in or within the vicinity of the visited classrooms during SOM observations. In the open-ended questionnaire items, a teacher wrote that "the participation of parents assisting with the academics, discipline, and growth of every child" was needed.

Student Achievement

In MID's first year of operation, 70 matched, sixth-grade charter-control student pairs were established. In the second year, 6 charter students moved out of the school district, 10 transferred to another school, 5 had missing achievement data for 2004–2005, and 7 lost their control school matches as a result of the latter fail-

ing a grade. The remaining longitudinal sample for the Year 2 seventh-grade analysis thus consisted of 42 matched pairs from Year 1. In addition, 30 new seventh graders enrolled in MID in Year 2 and were matched to control school counterparts. Of these, 15 pairs differed only on precharter school location. For the sixth-grade analysis, all 78 new enrollees were matched to comparison school counterparts. Of these, 40 pairs differed on precharter school location only. For the analysis of program effects on posttest scores, a MANCOVA was performed, separately by cohort, on achievement scores (see Table 4). In all three grade-level posttest analyses, as described in the following, both the R/LA and Math pretest covariates were highly significant (all $ps < .001$).

Pretest outcomes. To ensure comparability of MID and control group cohorts, pretest NCE scores in R/LA and Math analyzed for each grade-level cohort. The two groups performed nearly identically on all tests, with all outcomes nonsignificant and effect sizes approximating zero. Both groups scored below the national norm of 50, with the mean NCEs ranging from the high 30s to low 40s.

Posttest outcomes for 6th-grade new students. The multivariate effect of program was significant, $F(2,151) = 31.32, p < .001$. Follow-up univariate ANCOVAs revealed significant and strong program effects for both R/LA, $F(1,152) = 52.31, p < .001, ES_{adj} = +0.80$, and Math, $F(1,152) = 30.64, p < .001, ES_{adj} = +0.58$.

Posttest outcomes for 7th-grade new students. The multivariate effect of program was significant, $F(2,55) = 7.20, p < .001$. Univariate ANCOVAs revealed significant and strong effects for both R/LA, $F(1,56) = 7.16, p = .01, ES_{adj} = +0.50$, and Math, $F(1,56) = 13.20, p = .001, ES_{adj} = +0.60$.

Posttest outcomes for 7th-grade 2nd-year students. The multivariate effect of program was significant, $F(2,79) = 16.98, p < .001$. Univariate ANCOVAs were significant for both R/LA, $F(1,80) = 33.97, p < .001, ES_{adj} = +0.78$, and Math, $F(1,80) = 9.16, p = .003, ES_{adj} = +0.38$, favoring MID students.

TCAP proficiency levels. A supplementary analysis was performed to examine the percentages of MID and control students who scored at *Below Proficient*, *Proficient*, and *Advanced* levels on TCAP, as computed by the state for NCLB accountability. Two-way chi square (program \times proficiency level) analyses, all favoring MID, were significant for the 7th grade 2nd-year students in R/LA only, and 6th-grade new students in both R/LA and Math. Higher percentages of MID than control students tended to score at the Proficient and Advanced levels.

HIGH SCHOOL

Program Implementation Goals

The school leadership team judged most benchmark goals to be in Phase II (intermediate level of implementation). Although observations supported progress toward the goal of increasing student-centered activities, implementation was intermittent and occurred at rates lower than national norms. In contrast, technology use did increase to higher than average levels. Implementation of curriculum goals remained in early (Phase I) stages with regard to integration of subject areas, the elimination of remedial classes, and augmentation of textbooks in classroom instruction. Also at a beginning stage was ensuring adequacy of resources available for faculty and staff, professional development, and in particular, resources for serving special education students.

Pedagogy

Use of the traditional practices of direct instruction and seatwork remained high but consistent with norms for secondary schools. In accord with school goals, technology usage and higher order instructional feedback and questioning were all higher compared to the initial year of operation and national norms. Student-centered instruction, as evidenced by teacher coaching, project-based learning, experiential learning, and independent inquiry, although more frequent than in the first year of operation, were observed only rarely or occasionally and at lower rates than national norms. In general, the school maintained its reliance on traditional lecture and seatwork, while showing progress toward the goal of integrating technology usage with classroom learning.

School Climate

Suggestive of positive school climate, the overall SCI mean rating was a 4.15 on the 5-point scale compared to the national overall norm of 3.72 (see Table 3). In fact, the mean on only one climate dimension, order, fell below 4.00. All comparisons to national secondary school norms were positive, with ES ranging from +0.42 (leadership) to +1.17 (expectations). The latter outcome strongly reflected the school's core philosophy that it could prepare students to succeed in science and engineering careers, even though 2/3 of the students were economically disadvantaged and at-risk. Nonetheless, the dimension means were consistently lower than those obtained in the previous year. This decline was particularly apparent for the order dimension (from 4.70 to 3.68). The expansion of the school in terms of student enrollment (from 147 to 252) and grade levels served may have influenced

these results. Student discipline and enforcement of rules for misbehavior appeared to comprise a primary problem.

The principal described the slight decline in climate as due to "realistic optimism" replacing last year's "total optimism." He explained that although teachers see "the bumps in road," they remain strongly committed to the school's program and potential. High expectations were a dominant theme that emerged from the teacher interview data. As one teacher put it, "Here everyone is accountable. Everyone gets continual feedback, is expected to get results, and works collaboratively to address weaknesses." The teachers expressed respect for the principal, viewing him as a mentor who is supportive and understands their needs. Discipline was described as difficult and students were said to be "quite smart at breaking the rules."

High expectations were also a major theme in student responses. The students talked about the hard work and long hours but tempered their comments with recognition that teachers cared and wanted them to succeed. For example, one said, "Teachers are different here. Here they care and you must be respectful."

Teacher Attitudes

Teachers clearly supported the mission and educational program of HIGH. Satisfaction with the adequacy of professional development, however, showed some decline from the first year. More specifically, one teacher identified "a lack of ongoing professional development necessary to insure that student tasks are standards-based and adhere to the curriculum." Other open-ended comments identified lack of personal time as a result of the long school hours as a primary drawback. The following quote underscores the time commitment required: "You can talk about long hours and responsibility but until you are here and experience it, you really do not know what it is like. It is not for everyone. You have to be very dedicated ... you have to love the students."

Positive perceptions were indicated in the areas of planning time, educational materials, and technology. In contrast, low levels of agreement were expressed for having a sufficient number of faculty and staff and adequately addressing requirements of special needs students. Open-ended responses revealed that there was no specialist to work with special needs children and teacher turnover had been problematic because the number of staff was so small. According to the teachers, "those who left really did not buy into the mission and vision of [HIGH]." Support from parents and external partners was considered by the teachers and principal to be stronger than the support received from state and local education agencies.

Parent Reactions

Parents were highly satisfied with the school's educational program, but also complimentary of the school leadership, safety, discipline, and the respectful and

welcoming way in which they were treated. Due perhaps to the increase in school size, there was some decline in the percentages of parents who agreed that they were regularly informed about their child's progress and that teachers were readily available. For example, one parent wrote that "the communication between home and school needs to be greatly improved. That area has worsened this year." In other open-ended responses, most parents cited the challenging curriculum, high expectations, leadership, and teachers as strengths of the school. The following quote incorporates these themes: "The strengths of [HIGH] are quality instructors and instruction. The challenging curriculum. The tireless efforts of the principal. The desire of the staff to see real student achievement."

Evidence from other sources suggests moderate rates of parental involvement that exceeded normal levels for high schools. The relatively frequent parental involvement likely is attributable in part to parents' contract agreement to donate hours to HIGH. Interestingly, although nearly 3/4 (70%) of teachers agreed that parents were active partners with the school; this figure represents a decline from almost 90% agreement in the previous year.

Student Achievement

In its second year of operation, HIGH enrolled 124 students in Grade 7 and 91 students in Grade 8. In Year 1, 134 matched seventh-grade charter-control student pairs were established. Of these, 11 moved out of the school district, 17 transferred to another location, 5 were eliminated as a result of their matched control counterparts being retained, and 2 were eliminated as a result of their being retained. The remaining eighth-grade longitudinal sample for Year 2 thus consisted of 99 matched pairs from Year 1. In addition, 25 new eighth graders enrolled and were matched to control-school counterparts. Only one pair differed on pretest school enrollment. For the seventh-grade analysis, all 91 new enrollees were matched to control-school counterparts. Of these, 32 pairs differed on pretest school location. Posttest means, standard deviations, and effect sizes are summarized in Table 4. In all posttest MANCOVA analyses (see summary in Table 4), both the R/LA and Math pretest covariates were highly significant (all $ps < .001$).

Pretest outcomes. In all pretest analyses, HIGH and control students performed nearly identically on both tests, with all effect sizes approximating zero and all ANOVA outcomes nonsignificant. Both eighth-grade cohorts, but particularly the second-year enrollees, scored above the national norm of 50 in R/LA and Math, with the mean NCEs ranging from 51.92 to 60.31. The seventh-grade cohorts, however, scored slightly below the national norm.

Posttest outcomes for 7th-grade new students. The multivariate program effect was significant, $F(2,177) = 10.92, p < .001$. Follow-up univariate ANCOVAs

revealed results significantly favoring HIGH students in R/LA, $F(1,178) = 21.95, p < .001, ES_{adj} = +0.38$, and directionally favoring them in Math, $F(1,178) = 3.20, p < .08, ES_{adj} = +0.14$.

Posttest outcomes for 8th-grade new students. The multivariate effect of program was significant, $F(2,45) = 5.07, p = .01$. Similar to the seventh-grade results, the univariate ANCOVAs revealed significant results favoring HIGH students in R/LA, $F(1,46) = 8.30, p = .01, ES_{adj} = +0.52$, but only a nonsignificant directional advantage in Mathematics, $F(1,46) = 3.63, p = .06, ES_{adj} = +0.27$.

Posttest outcomes for 8th-grade 2nd-year students. The multivariate program effect was significant, $F(2,193) = 13.20, p < .001$. Significant univariate effects favoring HIGH students were obtained in both R/LA, $F(1,194) = 7.67, p = .01, ES_{adj} = +0.24$, and Math, $F(1,194) = 26.17, p < .001, ES_{adj} = +0.43$.

TCAP proficiency levels. Chi-square analyses of the percentages of HIGH and control students who scored at *below proficient*, *proficient*, and *advanced* levels on the TCAP subtests were significant ($p < .05$) for the eighth-grade 2nd-year students in Math only, and for the seventh-grade new students in R/LA only. Higher percentages of the HIGH than control students scored *proficient* or *advanced* on these tests.

DISCUSSION

In interpreting the second-year educational outcomes for the charter schools examined, it is important to recognize that all three demonstrated tangible progress but also individual challenges in implementing their curricular, instructional, and administrative programs. ELEM, in adopting the Success for All model as its major schoolwide framework, appeared to demonstrate the greatest advancement in most areas, particularly curriculum and instruction. Obtaining and allocating resources to meet programmatic needs was less successful at MID and HIGH, but also emerged as an intrinsic problem of the charter schools' separation from a larger district. Also challenging were efforts to shift pedagogy as desired from teacher- to student-centered approaches. Notably, teacher attitudes and school climate were positive at all three schools, with overall climate means exceeding national norms by 0.44 (MID) to 1.20 (ELEM) standard deviations. Relative to the start-up year, slight decreases in climate, seemingly most associated with discipline and behavior problems, occurred at both MID and HIGH, whereas noticeable increases in positive climate occurred at ELEM. The most salient teacher concerns at the three schools regarded needs for more resources and staffing, and perceived lack of dis-

trict support. Parent satisfaction appeared quite high at all three schools, although actual involvement varied between schools.

Design Considerations

Compared to prior research (Braun et al., 2006; Bulkley & Fisher, 2002; Greene, Forster, & Winters, 2003; Nelson et al., 2004), this quasiexperimental design reduced the potential for sampling bias by individually matching each charter student to a highly comparable control student. Based on the results obtained, students attending these charter schools were generally performing higher than their peers who remained in traditional schools. Why these schools are demonstrating early success in raising student achievement is a matter of considerable interest to both educational researchers and policymakers. A possible explanation, which cannot be proven or refuted from our data, is that the charter schools attracted more effective teachers and more involved families than did the traditional district schools (see Berends et al., 2006; Betts & Hill, 2006; Goldring & Cravens, 2006). Such factors are intrinsic to choice contexts and arguably a core attribute of the charter school treatment. That is, even if it were possible to randomly assign teachers to charter schools from some larger district pool, one might question to what degree an essential element of charterness (to coin a word) would be missing from that approach. In each of the three schools, teachers applied for position openings and were selected largely based on their interest in the particular school's academic theme, leadership, and administrative structure.

Similarly, when a charter school attracts so many student applicants as to require a lottery, possibilities for their random assignment to charter and traditional schools, and thus usage of a randomized experimental design, become possible (Goldring & Cravens, 2006). In this context, as for many other charter schools nationally, student applications met or barely exceeded available space, thus requiring active recruitment efforts on the part of school leaders and community members. As a result, a randomized design was not possible. Still, in the absence of random or clearly representative sampling, teacher and student selectivity must be considered as potential contributing factors to overall charter school effects.

Counter to the intrinsic advantages afforded by these factors, other conditions operated to limit the three charter schools' potential to demonstrate positive effects on student achievement. One obvious factor is the insensitivity of high-stakes state accountability systems, such as TCAP, for evaluating school effectiveness (Linn, *in press*; Raudenbush, 2004). Had supplementary, performance-based measures of achievement been available, greater insight into school impacts on varied and more complex types of learning (e.g., problem-solving, higher-order reasoning, writing skill) could have been obtained.

Second, regardless of how achievement is assessed, school change typically takes several years to manifest itself in observable effects (Fullan, 2000; Sizer,

1992; Tyack & Cuban, 1995). Notably, conditions in all three charter schools fostered rapid establishment of positive school climate and receptivity to implementing new programs. This pattern contrasts with that frequently depicted by studies of whole-school reform, in which lack of teacher support impeded or, worse, created active resistance to change (Datnow, Hubbard, & Mehan, 2002; Desimone, 2002; Vernez, Karam, Mariano, & DeMartini, 2006).

Third, research on school transitions performed specifically with the TCAP system (Sanders & Horn, 1995a, 1995b) and in other K-12 contexts (see Larson, Moneta, Richards, & Wilson, 2002), has shown decreases in student achievement associated with changing schools. In this regard, the entire charter school sample (and the MID matched-control students) changed schools from 2002-2003 to 2003-2004, whereas only 14% of the ELEM and 54% of the HIGH control students changed schools.

Impacts on School Climate, Pedagogy, and Teachers

With regard to variables related to school interventions, student achievement outcomes seem likely to be linked most directly to the quality of teaching and learning (Darling-Hammond & Bransford, 2005; Wright, Horn, & Sanders, 1997). Unfortunately, as cogently demonstrated in the extensive literature on CSR, teaching practices tend to be highly resistant to change (Borman et al., 2004; Datnow et al., 2002; Desimone, 2002) and do not become more effective simply because a new program or reform model has been adopted (Vernez et al., 2006). However, where school climate is positive, as reflected in collaboration between faculty, high expectations for student success, effective order and discipline, and enthusiasm for positive change, the implementation of reforms occurs more quickly and effectively (Bobbett, Ellett, Teddlie, Olivier, & Ruggett, 2002; Bryk & Schneider, 2002). In the process, professional development experiences and new curricula are more likely to impact teaching quality and, in turn, improve achievement (Desimone, 2002; Ross, McDonald, Alberg, & McSparrin-Gallagher, *in press*).

Although positive school climate was established rather quickly at the charter schools, seemingly as a function of strong teacher buy-in and principal leadership, the schools' implementation of their academic, organizational, and professional development programs were comparably slower to progress but still positive overall. Most resistant to change, predictably, were teaching methods, which at all three schools emphasized traditional lecture and teacher-directed seatwork. Held to the same NCLB accountability standards as their regular-school counterparts, the charter school teachers likely felt pressure to seek control and efficiency in covering the broad state curricula required for TCAP. Even so, their observed levels of academically focused instructional time and student engagement typically exceeded national norms.

Berends et al. (2006) presented a parallel framework that interprets charter school success as depending on the individual school's degree of "capacity-building" for promoting effective teaching and learning. Drawing from the work of Goldring and Cravens (2006), they identified six organizational properties (pp. 14–15; identified in the following section by italics) as indicators of such potential. As described here, these properties also receive direct support from our findings.

- *Shared mission and goals that establish educational priorities and academic activities.* In this study, intermediate to advanced implementation progress was demonstrated by ELEM on all implementation benchmark categories, and by MID and HIGH on most categories except external support. At all three schools, teacher reactions to the school program and mission were highly positive in focus groups and on the CSTQ.

- *Principal leadership.* Focus group and survey data from teachers and parents indicated clearly positive perceptions of principals at all three schools. In particular, means on the leadership dimension of the SCI approached the ceiling level at ELEM and HIGH, and slightly exceeded national norms at MID (see Table 3).

- *Expectations for instruction and focus on achievement.* SCI results revealed positive teacher reactions on the expectations dimension, with strong effect sizes relative to national norms for ELEM ($ES = +1.12$), MID ($ES = +0.73$), and HIGH ($ES = +1.17$). Similarly, on the instruction dimension, the respective effect sizes were $+0.98$, $+0.54$, and $+0.73$. Increasing instructional time via extending the school day or yearly calendar was also a core program feature at all three schools. However, the charter schools, particularly MID, were less successful in achieving stated goals of increasing active and meaningful learning through student-centered teaching strategies.

- *Expert teachers supported by coherent, consistent professional development.* Although judging teaching expertise was beyond the scope of our study, SOM observations across the three schools identified instruction that, although mostly traditional and teacher-centered, tended to be above-average in usage of academically-focused time and promoting student engagement. Synthesis of teacher responses on CSTQ Likert-type and open-ended responses suggested that professional development support was perceived to be strong at ELEM and at least adequate, though less positive, at MID and HIGH.

- *Professional community of teachers.* Without question, collegiality and professional community were evident at the schools. As described, key factors appear to have been strong principal leadership and hiring of teachers who specifically wanted to work at those schools. Revealingly, school means on the collaboration dimension of the SCI exceeded national norms by strong effect sizes of $+1.23$ (ELEM), $+0.58$ (MID), and $+0.71$ (HIGH).

CONCLUSIONS

The success of the three schools examined and other charter schools nationally must be judged on educational outcomes evidenced over time rather than during a restricted, mostly developmental period. For making evaluation judgments, we strongly agree with Berends et al. (2006) that the key question is not whether or not charter schools work (i.e., some do and some don't), but rather "Under what conditions do they work?" All factors considered, although limited to three schools, these findings are suggestive of the schools' establishment of promising foundations, largely due to progress in the areas identified in the preceding section. Regarding student achievement, it is noteworthy that out of 18 comparisons between mean performances of charter school and comparison students, 12 were statistically significant, with median effect sizes of $+0.38$ in R/LA and $+0.43$ in Math. By comparison, in a recent meta-analytic study of 29 CSR models, Borman et al. (2003) found overall effect sizes between $+0.10$ to $+0.14$, with the range for the *most successful* category being $+0.17$ to $+0.21$. Simply put, raising an entire school population by the present charter school median approximating .4 *SD* is certainly suggestive of substantive educational and economic benefits.

But will such dividends occur year after year? The recent RAND report on comprehensive school reform implementation (Vernez et al., 2006) and an extensive body of earlier literature (Datnow et al., 2002; Desimone, 2002; Fullan, 2000) on school restructuring describe the challenges associated with maintaining successful reforms as founders, leaders, and faculty retire or relocate, and the novelty of innovative programs wears off. Thus, it will be critical to examine to what degree and how the key elements of charter schools impact longitudinally the three schools' ability to sustain their initial enthusiasm and successes.

REFERENCES

- American Institutes for Research. (2005). *CSRQ Center report on elementary comprehensive reform models*. Washington, DC: Comprehensive School Reform Quality Center.
- Anderson, L., Adelman, N., Finnigan, L., Cotton, M., Donnally, B., & Price, T. (2002). *A decade of public charter schools: Evaluation of the public charter schools program: 2000–2001 Evaluation Report*. Washington, DC: U.S. Department of Education.
- Ballou, D., Teasley, B., & Zeidner, T. (2006). *Charter schools in Idaho*. Nashville, TN: National Center on School Choice at Vanderbilt University.
- Berends, M., Watral, C., Teasley, B., & Nicotera, A. (2006). *Charter school effects on achievement: Where are we going?* Nashville, TN: National Center on School Choice at Vanderbilt University.
- Betts, J., & Hill, P. T. (2006). *Key issues in studying charter schools and achievement: A review and suggestions for national guidelines* (NCSPP White Paper Series, No. 2). Seattle, WA: Center on Re-inventing Public Education.

- 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.
- 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*. 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*(2), 125–230.
- Braun, H., Jenkins, F., Grigg, W., & Tirre, W. (2006). *A closer look at charter schools using hierarchical linear modeling*. Washington, DC: National Center for Educational Statistics.
- Bryk, A., & Schneider, B. (2002). *Trust in school: A core resource for improvement*. New York: Russell Sage Foundation.
- Buckley, J., & Schneider, M. (2005). Are charter school students harder to educate? *Education Evaluation and Policy Analysis, 27*(4), 365–380.
- Bulkley, K., & Fisher, J. (2002). *A decade of charter schools: From theory to practice* (CPRE Research Report Series RB–35). Philadelphia, PA: Consortium for Policy Research in Education.
- Butler, E. D., & Alberg, M. J. (1991). *The Tennessee School Climate Inventory: Resource manual*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Carnoy, M., Jacobsen, R., Mischel, L., & Rothstein, R. (2005). *The charter school dust-up: Examining the evidence on enrollment and achievement*. Washington, DC: Economic Policy Institute and Teachers College Press.
- Center for Education Reform. (2006). *Charter school facts*. Retrieved November 10, 2006, from http://www.edreform.com/_upload/CER_charter_numbers.pdf
- Center for Research in Educational Policy. (2005). *National norms for the School Observation Measure*. Memphis, TN: University of Memphis.
- Chatterji, M. (2005). Evidence on “what works”: An argument for extended-term mixed-method (ETMM) evaluation designs. *Educational Researcher, 34*(5), 14–24.
- Darling-Hammond, L., & Bransford, J. (2005). *Preparing teachers for a changing world*. San Francisco: John Wiley & Sons.
- Datnow, A., Hubbard, L., & Mehan, H. (2002). *Extending educational reform: From one school to many*. New York: Routledge-Farmer.
- Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research, 72*, 433–480.
- Frankenberg, E., & Lee, C. (2003). *Charter schools and race: A lost opportunity for integrated education*. Cambridge, MA: Civil Rights Project.
- Fullan, M. (2000). The return of large scale reform. *Journal of Educational Change, 1*, 5–28.
- Golding, E., & Cravens, X. (2006). *Teachers' academic focus on charter and non-charter schools*. Nashville, TN: National Center on School Choice at Vanderbilt University.
- Greene, J. P., Forster, G., & Winters, M. A. (2003). *Apples to apples: An evaluation of charter schools serving general student populations*. Manhattan, NY: Center for Civic Innovation at the Manhattan Institute.
- Hill, P. T. (2005). *Assessing student performance in charter schools: Why studies often clash and answers remain elusive*. Retrieved on July 5, 2007, from <http://www.edweek.org/ew/articles/2005/01/12/18hillh24.html>
- Hill, P., Angel, L., & Christensen, J. (2006). Charter school achievement studies. *Education Finance and Policy, 1*, 139–150.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed-methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26.
- Larson, R., Moneta, G., Richards, M., & Wilson, S. (2002). Continuity, stability, and change in daily emotional experience across adolescence. *Child Development, 73*, 1151–1165.
- Lewis, E. M., Ross, S. M., & Alberg, M. J. (1999). *School Observation Measure: Reliability analysis*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Linn, R. L. (in press). Validity of inferences from test-based educational accountability systems. *Journal of Personnel Evaluation*.
- Lowther, D. L., Ross, S. M., & Plants, R. (2000). *Expanded rubric (ER)*. Memphis, TN: Center for Research in Educational Policy, University of Memphis.
- McDonald, A. J., & Ross, S. M. (2003). *Charter School Teacher Questionnaire*. 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.
- Miron, G., & Nelson, C. (2004). Student achievement in charter schools: What we know and why we know so little. In K. Bulkley & P. Wohlstetter (Eds.), *Taking account of charter schools* (pp. 161–175). New York: Teachers College Press.
- Nelson, F. H., Rosenberg, B., & Van Meter, N. (2004). *Charter school achievement on the 2003 National Assessment of Educational Progress*. Washington, DC: American Federation of Teachers.
- Paige, R., & Huckabee, M. (2005, January 26). Putting arts education front and center. *Education Week, 24*(20), 40, 52.
- Raudenbush, S. W. (2004). What are value-added models estimating and what does this imply for statistical practice? *Journal of Educational and Behavioral Statistics, 29*, 120–129.
- Ross, S. M., & Alberg, M. (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, M. (2002). *Program Benchmarking for Raising Student Achievement*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Ross, S. M., McDonald, A. J., Alberg, M., & McSparrin-Gallagher, B. (2007). Achievement and climate outcomes for the Knowledge is Power Program in an inner-city middle school. *Journal of Education for Students Placed At Risk, 12*, 137–165.
- Ross, S. M., McDonald, A. J., & Bol, L. (2004). *First year evaluation of TN charter schools*. Memphis, TN: The University of Memphis, Center for Research in Educational Policy.
- Ross, S. M., McDonald, A. J., & Bol, L. (2005). *Second-year evaluation of TN charter schools*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Ross, S. M., McDonald, A. J., & McSparrin-Gallagher, B. (2005). *Student-level analysis of year 1 (2003–2004) achievement outcomes for Tennessee charter schools*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Ross, S. M., Smith, L. J., & Alberg M. J. (1998). *School Observation Measure*. Memphis, TN: Center for Research in Educational Policy, University of Memphis.
- 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.), *Observation research in U.S. classrooms: New approaches for understanding cultural and linguistic diversity* (pp. 144–173). Cambridge, MA: Cambridge University Press.
- 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.

- Sanders, W. L., & Horn, S. P. (1995a). Educational assessment reassessed: The usefulness of standardized and alternative measures of student achievement as indicators for the assessment of educational outcomes. *Educational Policy Analysis Archives*, 3(6). Retrieved January 15, 2007 from <http://epaa.asu.edu/epaa/v3n6.html>
- Sanders, W. L., & Horn, S. P. (1995b). The Tennessee Value-Added Assessment System (TVAAS): Mixed model methodology in educational assessment. In A. J. Shinkfield & D. Stufflebeam (Eds.), *Teacher evaluation: Guide to effective practice* (pp. 337–350). Boston: Kluwer.
- Sizer, T. R. (1992). *Horace's school: Redesigning the American high school*. New York: Houghton Mifflin.
- Snipes, J., Doolittle, F., & Herlihy, C. (2002). *Foundations for success: Case studies of how urban school systems improve student achievement*. Washington, DC: Manpower Demonstration Research Corporation for the Council of Great City Schools.
- Solomon, L. C., & Goldschmidt, P. (2004). *Comparison of traditional public schools and charter schools on retention, school switching, and achievement growth* (Policy Report No. 192). Phoenix, AZ: Goldwater Institute.
- Sterbinsky, A., & Ross, S. M. (2003). *School Observation Measure reliability study*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Sterbinsky, A., Ross, S. M., & Burke, D. (2004). *TEAM Tennessee reliability study*. Memphis, TN: Center for Research in Educational Policy, University of Memphis.
- Stringfield, S. C., & Yakimowski-Srebniak, M. E. (2005). Promise, progress, problems, and paradoxes of three phases of accountability: A longitudinal case study of the Baltimore City Public Schools. *American Educational Research Journal*, 42, 43–75.
- Tough, P. (2006, November 26). Can teaching poor children to act more like middle class help to close the education gap? What it takes to make a student. *The New York Times Magazine*, pp. 44–51, 69–72, 77.
- Tyack, D., & Cuban, L. (1995). *Tinkering toward utopia*. Cambridge, MA: Harvard University Press.
- Tyack, D., & Tobin, W. (1994). The “grammar” of schooling: Why has it been so hard to change? *American Educational Research Journal*, 31, 453–479.
- U.S. Congress. (2001). *No Child Left Behind Act of 2001*. Washington, DC: Author.
- U.S. Department of Education. (2001). *The facts about supporting charter schools*. Retrieved April 20, 2004, from <http://www.ed.gov/nclb/choice/charter/charter.html>
- U.S. Department of Education. (1999). *Guidance on the Comprehensive School Reform Program*. Washington, DC: Author.
- Vernez, G., Karam, L. T., Mariano, L. T., & DeMartini, C. (2006). *Evaluating comprehensive school reform models at scale: Focus on implementation*. Retrieved January 4, 2007, from http://www.rand.org/pubs/monographs/2006/RAND_MG546.pdf.
- Wong, K. K., & Shen, F. X. (2006). *Charter law and charter outcomes: Re-examining the school marketplace*. Nashville, TN: National Center on School Choice at Vanderbilt University.
- Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11, 57–67.