

The Sustainability of Inclusive School Reform

PAUL T. SINDELAR

DEIRDRE K. SHEARER

DIANE YENDOL-HOPPEY

TODD W. LIEBERT

University of Florida

ABSTRACT: *For over a decade, University of Florida researchers worked with middle schools in a large urban and suburban south Florida district, as they developed and then worked to sustain inclusive reform. One middle school, Socrates, was notably successful, having built its inclusion model on a foundation of previous reform and a school culture characterized by shared decision making, collaboration, and teaming. For 4 years, we studied Socrates and the sustainability of its program. Inclusion was not sustained; our analysis of teacher and administrator interviews revealed three primary factors that help explain why: leadership change, teacher turnover, and state and district assessment policy change. Reduced support for the program, a by-product of the primary factors, also contributed to the lack of sustainability.*

The idea of including students with disabilities in general education classrooms is nearly 30 years old. In 1975, the Individuals with Disabilities Education Act (IDEA) introduced the concept of instructing students in the least restrictive environment (LRE), and for many students with disabilities, the LRE is a general education classroom. A decade later, Will's (1986) call for shared responsibility in educating students with disabilities set schools and researchers on a quest for successful models of inclusion. During the 1990s, we were able to follow one school's transition from traditional special education to a teacher-developed inclusion program. This transformation took

place with the assistance of University of Florida researchers and a federal grant. Two years after the original project ended, we returned to Socrates Middle School (SMS) to study the sustainability of those reforms.

SUSTAINABILITY RESEARCH

Despite the growing body of knowledge about school reform and special education practices, researchers know little about the extent to which innovations are sustained over time and what factors influence their sustainability. In short, empirical research on sustainability factors is limited (Gersten, Chard, & Baker, 2000).

Research on the sustainability of innovation has focused on both classroom-specific (e.g., reading strategies) and schoolwide reforms (e.g., magnet programs); (Gersten et al., 2000). Schoolwide reforms affect the structure and day-to-day operation of a school, and research on them has been limited to general education reforms (Florian, 2000; Huberman & Miles, 1984). By contrast, research on the sustainability of classroom-specific reform includes several special education studies (Gersten et al.; Klingner, Arguelles, Hughes, & Vaughn, 2001; Vaughn, Klingner, & Hughes, 2000).

Despite the growing body of knowledge about school reform and special education practices, researchers know little about the extent to which innovations are sustained over time and what factors influence their sustainability.

In studies of classroom reforms, researchers have identified three main factors related to sustainability: district and state policy, leadership, and teaching/classroom factors. Districts that show strong commitment to a reform recognize schools for adopting new practices and take measures to ensure that principals follow through. Both actions have been linked to sustained use of reforms (Klingner et al., 2001; Vaughn et al., 2000). In contrast, innovations stand less chance of survival when districts are not committed to them (Klingner et al., 2001) or when districts invest heavily in high-stakes assessments (Furney, Hasazi, Clark/Keefe, & Hartnett, 2003) and improving student outcomes on them (Klingner et al., 2001). A second factor affecting the sustainability of classroom-specific innovation is school leadership (Klingner et al., 2001). Schools at which principals devote time to the development of an innovation are more likely to have teachers committed to its practice. Further, districts that procedurally rotate principals may have more difficulty sustaining a classroom-specific strategy than schools where principals are retained (Klingner et al., 2001). The third factor is teachers' acceptance of the practice. Successful adop-

tion of innovative practices occurred when it was consistent with teachers' beliefs or teaching style (Klingner et al., 2001; Vaughn et al., 2000), when the practice helped the most difficult-to-teach students, and when teachers received supportive training (Gersten et al., 2000). Because learning novel teaching practices places new demands on teachers, they will not sustain their use of innovative practices unless they see benefits for students (Gersten et al.; Klingner et al., 2001).

In research on the sustainability of schoolwide reform, policy and leadership also play a role. In addition, school culture and factors related to the innovation itself are likely to affect schoolwide reform. For the most part, district and state policy and principal leadership have the same impact on schoolwide reform as on classroom-specific reforms (Florian, 2000; Huberman & Miles, 1984), but Huberman and Miles also emphasized the role of teacher leaders. They noted that a group of teacher leaders—whom they call *enforcers*—often provide resources and encouragement to other teachers for adopting a new practice. Enforcers understand the innovation and are invested in its continued use. Huberman and Miles found that enforcers were often motivated by the opportunities for advancement that a new reform created in the district (Huberman & Miles).

A third factor in the sustainability of schoolwide reform is school culture: Schools with shared vision and cultures of communication and shared decision making, and schools that involve teachers in the design of an innovation, are more likely to sustain innovations (Florian, 2000; Huberman & Miles, 1984). Another important element of school culture is teacher mobility. Huberman and Miles described two waves, one when an innovative practice commences and one later. The second wave, which is likely to have the most impact on sustainability, results from teachers who demonstrate success with the innovation and advance to higher positions. Successful teachers were frequently enforcers in the Huberman and Miles sense; when they left the classroom, support for and communication about the innovation waned (Huberman & Miles).

The final factor derived from the nature of the innovation itself. In their study of various large-scale, longitudinal innovations, Huberman

and Miles (1984) found that innovations that were smaller in scope and that placed fewer demands on teachers were more likely to take root and be sustained. By contrast, innovations that created too many demands or were too complex to understand did not have the success of more manageable programs. Furthermore, innovations that required too many changes in the current functioning of the school were less successful than more proscribed innovations (Huberman & Miles).

The inclusion of students with disabilities in general education is one such complex and demanding reform. Inclusion is often misunderstood and sometimes resisted by teachers, and it is not always fully understood or supported by school administrators. The Individuals with Disabilities Education Act (1997) stipulates that students with disabilities be educated in the least restrictive environment but also requires that districts provide a continuum of placement options. Thus, states and districts have some latitude with regard to IDEA implementation, and as a consequence special education practice varies dramatically from district to district and state to state. Inclusion, in short, would seem to be a challenging schoolwide reform to establish and implement, and its sustainability would seem difficult to achieve. No previous research on the sustainability of inclusive reform is available in the literature.

PURPOSE OF THE STUDY

The purpose of the current study was to address this gap in knowledge about factors leading to the sustainability of a schoolwide special education reform. For a decade, faculty from the University of Florida (UF) collaborated with teachers and administrators at Socrates Middle School (SMS), first assisting with the design and implementation of an inclusion program, then focusing on the sustainability of the reform. Six years into the collaboration, the researchers removed themselves from the role of participant observers to assume an observational research stance that focused on the documentation of the sustainability effort. In this article, we report results of a qualitative analysis of teacher, administrator, and staff interviews concerning the factors that led to the rise and fall

of the school's inclusive reform. In doing so, we address two questions: Would inclusive reform be sustained beyond the life of the initial university-school collaboration? and What factors influenced sustainability?

METHOD

We addressed these questions of sustaining inclusive reform using a case study approach. Ethnographic case studies produce what Stake (1980) has called "naturalistic generalizations" (p. 69). Naturalistic generalizations are useful for evaluators because they identify hidden variables and produce hypotheses researchers can later verify using quantitative strategies (Ashton & Webb, 1986; Sherman & Webb, 1988; Smith, 1978). Case study research, as Fetterman (1983) has said, "with its close attention to details . . . can identify causal features and causal linkages that may be overlooked or misinterpreted on the basis of correlational analysis of survey data, or predetermined observational category systems" (p. 21). Naturalistic generalizations also are useful to practitioners who want practical guidance on program implementation in similar settings. Detailed descriptions of program activities, identification of barriers, clarification of significant variables, surfacing of taken-for-granted assumptions, and analyses of causal program processes are useful to those working in the setting being studied or others like it.

In our case studies, Bronfenbrenner's (1979) ecological framework served as a lens for our analysis. We used rewriting (Stake, 1980), coding (Pfaffenberger, 1988), and the constant comparative method of data analysis (Strauss, 1987) in our layered case study analysis of interview data. These activities appear linear, but the process was recursive and occurred throughout data collection (Spradley, 1980; Taylor & Bogdan, 1984; Webb & Glesne, 1992). Rewriting involved transcribing individual interviews (Spradley, 1979); coding involved attaching category names to basic units of field-generated data (Strauss & Corbin, 1990; Taylor & Bogdan).

Initially two members of the research team, who were not involved in the study design or data collection, independently coded the data, in the

process producing 47 codes. They discussed similarities among the first set of codes and collapsed them into 16 codes. After discarding 3 codes that did not generate adequate support from the data, two additional members joined the research team for the next stage of analysis. This group clustered the remaining 13 codes into 4 themes.

In an effort to enhance credibility, the authors engaged in multiple researcher triangulation. Researcher triangulation (Patton, 2002) involved multiple investigators in the analysis, three of whom were not involved as participant observers during the Project RISES (Restructuring for the Inclusion of Special Education Students) and had little stake in or knowledge of the development of the SMS program. Additionally, the lead author's statement of researcher bias is provided to make known any preconceived notions regarding the importance of sustainability:

When we began the sustainability study, I believed that teachers at Socrates Middle School had developed a first-rate inclusion program and that structures—supportive leadership, collaborative culture, high-quality professional development, and shared decision making—were in place to sustain it. I had come to admire the SMS teachers and their commitment to students with disabilities, but seldom agreed completely with the decisions they made about the inclusion program. As a result, entering into the sustainability study, I was curious about the shape the program had taken and whether SMS stakeholders retained their commitment to it. I felt less investment in the particulars of the program; I knew its flaws and understood that I would have little opportunity to contribute substantially to the process of improving it. Yet, in this regard, little had changed; it was the teachers who shaped and refined the original program, not RISES researchers.

The final four themes created a conceptual framework that organized and explained the variables associated with successful or unsuccessful school restructuring for inclusion (Pfaffenberger, 1988; Spradley, 1980; Strauss & Corbin, 1990). During these discussions, we paid special attention to the micro- and macro-politics of education (Ball, 1987) and variables at all levels of the educational system (Bronfenbrenner, 1979). Dur-

TABLE 1
Data Collection

<i>School Year</i>	<i>Interviews</i>	<i>Quantity</i>
1998–99	Teachers	45
	Key Players	7
1999–00	Administrators	5
	Key Players	5
2000–01	Teachers	30
	Administrators	11
	Key Players	8
2001–02	Teachers	30
2002–03	Key Players	11

ing this final analysis, we included key informants from every level by member checking themes.

DATA COLLECTION AND PROCEDURE

From 1998 to 2002, we conducted individual interviews with 95 teachers and 16 administrators. Individual interviews generally lasted 30 to 50 minutes. Table 1 describes the chronology of the interviews; the term “key players” refers to stakeholders who contributed substantially to the development of the inclusion program during RISES. All interviews were held in private, out of eyesight and earshot of other school personnel. All data were typed into laptop computers or taped and edited later. Because we did not type fast enough to capture every word, our contemporaneous transcriptions were not verbatim. We left some information out and compressed some ideas into shorter statements; however, we believe we fairly represented stakeholders' comments and opinions. Because teachers and administrators at SMS had taken part in many research projects, they were familiar with the interview process. In general, most stakeholders were friendly, open, and cooperative. Their observations typically were thoughtful and candid.

BACKGROUND

SOCRATES MIDDLE SCHOOL

Socrates Middle School is a large middle school in a large urban and suburban district in southeast

Florida. Like many schools in South Florida, it serves a fast changing, culturally diverse student population. From 1990 to 1996, the student population at SMS grew from 1,200 to over 2,000 students, and the faculty expanded from 60 to over 100 teachers. In the 1996-97 school year, a new middle school opened nearby causing SMS to lose 700 students and 30 faculty. However, because the district was growing so rapidly, enrollment at SMS recovered. By the 2001-2002 school year, 1,765 students attended SMS.

To meet the challenge posed by population changes, the faculty and staff at SMS initiated and sustained several reforms including shared decision making (SDM), a district-sponsored initiative, and the Coalition of Essential Schools (CES). In 1992, SMS collaborated with the University of Florida (UF) on Project RISES, a U.S. Department of Education grant with a focus on including students with disabilities in the process of systemic reform. In the following section, we briefly describe the evolution of inclusion during RISES; a more complete account can be found in Kilgore, Griffin, Sindelar, and Webb (2001, 2002).

INCLUSION AT SMS

SMS moved toward inclusion gradually. In the late 1980s, teachers first developed a model they called "ESE (Exceptional Student Education) with EASE," in which students with disabilities were taught by a single team of teachers at each grade level. ESE with EASE offered two placement options: general education classrooms with support as needed or self-contained ESE classes taught by teams of content area and ESE teachers. Students who made the transition from ESE to general education classes were taught by the same content teachers, making the move less difficult.

The ESE with EASE program was considered a success by the faculty. Although teachers saw growth in special education students and their classmates, teachers and principals were not satisfied with what they had created. The model still segregated special education students. Putting all students with disabilities on one team concentrated their numbers in inclusion classes, and such concentrations sometimes frustrated teachers, students, and parents.

Program development continued apace. Faculty and staff, parents, district and state administrators, and university researchers discussed limitations of the ESE with EASE model and ways to improve it. SMS made significant progress toward implementing a broader vision of inclusion during a series of retreats facilitated by the UF research team, in 1994 and 1995. As facilitators, we helped plan the retreats, worked to make sure all stakeholders were heard, raised questions, reported on relevant research findings when asked, kept minutes, and reported those minutes to the faculty. The group used several principles to guide program improvement, among them (a) natural proportions, (b) general class placement, (c) heterogeneous grouping, and (d) multidisciplinary intervention.

The movement toward inclusion was interrupted briefly when the principal was transferred during the 1994-95 school year. However, the new principal shared the faculty's commitment to democratic governance, academic excellence, school improvement, professional development, CES principles, and inclusion. She quickly entered the school's ongoing conversation about inclusion and supported the faculty's effort to solve the problems they identified in the ESE with EASE Program.

The 1995-96 school year began with the new teacher-developed inclusion model in place. All at-risk students were assigned proportionately to teams. Teams were assisted by co-teachers, and co-teachers helped all students who needed special attention, regardless of their classification. Six full-time co-teachers each worked with two or three teams, while two half-time co-teachers worked with one team each. The exact role of the co-teachers was not prescribed by the planning group and evolved differently on different teams. Most co-teachers worked in different ways with different teams. The program was implemented in this form through the 1996-97 school year.

SUCCESS AND LIMITATIONS DURING PROJECT RISES

The focus of our RISES research was on the culture of the school and the organizational processes in place that allowed teachers to participate meaningfully in decision making. The project was

funded in a competition designed to support studies of efforts to include students with disabilities in the process of systemic reform. As a result, the effectiveness of the SMS inclusion program was not the primary focus of our work, and our data relate primarily to how the school organized itself to develop and implement the inclusion program. We collected no student achievement data, but observed in classrooms and conducted focus groups with teachers and students.

In focus groups and interviews, we asked teachers to identify practices that contributed to their ability to include students with disabilities successfully, and, in response, they discussed thematic units, cooperative learning, peer tutoring, explicit instruction, and small group and one-to-one instructional arrangements (Kilgore et al., 2001, 2002). Teachers learned and taught one another such techniques as split-page note taking, mnemonic strategies, and organizational strategies. Furthermore, SMS teachers consistently demonstrated high-quality group instructional practices, which they supplemented with small group and individualized remediation, as needed.

Despite enthusiasm for the inclusion program and other indicators of its success, problems emerged (Kilgore et al., 2001, 2002). Some teachers felt they did not get enough help from their co-teacher, and, in fact, the quality of implementation varied from team to team (and sometimes from teacher to teacher within a team). Furthermore, not all teachers were confident about their ability to accommodate students with disabilities successfully. (Math teachers tended to be more discouraged with the performance of included students than other teachers.) Teachers also were concerned about grading and the fairness of using different criteria to evaluate different students. They were reluctant to adjust their evaluation methods but disliked giving included students poor grades. Finally, in many cases, teachers were comfortable including students with less significant disabilities—students with learning disabilities, for example—but were not prepared for students with mental or emotional disabilities.

In spite of these problems, Project RISES was considered a success when its implementation was completed in 1996. In 1998, UF researchers returned to SMS to study the extent to which RISES reforms had been sustained. This second

project was called SIR, an acronym for Sustaining Inclusive Reform. In the interim, much had changed, both at the school and in the policy context in which the school operated.

FINDINGS

STATE AND DISTRICT CHANGES

During the 1990s, Florida developed a statewide system of school accountability, initiated by the development of standards—the Sunshine State Standards—that provided benchmarks in reading, mathematics, and writing. To assess student progress, the state developed and validated the Florida Comprehensive Achievement Test, or FCAT, first administered in 1997. In the following year, Florida instituted the “A⁺ Accountability” system for measuring a school’s progress toward achieving the standards. School grades were first assigned in the summer of 1999 using FCAT scores, the percentage of students tested, and the percentage of students making gains on the FCAT. Students with disabilities (except those with speech impairments) were not included in the calculation of school grades.

The Florida School Recognition Program provided financial rewards and greater autonomy to schools that sustained high performance or demonstrated exemplary improvement due to innovation and effort. Schools receiving “A” grades and schools improving at least one grade were eligible for school recognition awards, which could be used for performance bonuses for faculty and staff or for equipment and materials. School advisory council members and school staff determined how these funds would be used. The district and state placed great emphasis on improving test scores and school grades. The district developed curricula (keyed to the standards) and scripted teaching methods for each subject and grade level. In addition, it announced it would support only “laser-focused” professional development that specifically addressed teachers’ needs vis-à-vis assigned subject areas and approved teaching strategies.

CHANGES AT SMS

Leadership. Before we returned to SMS, in the spring of 1997, the SMS principal learned

that she was to be transferred. Her replacement, the third SMS principal we worked with, had been principal of a nearby elementary school but had no middle school experience. She faced the difficult challenge of fitting into a school with strong traditions and a faculty with strong opinions. She differed from her predecessors in being less familiar with and committed to either CES or inclusion principles.

School Demographics. As the 1998-99 school year began, SMS's student population increased from 1,376 to 1,582 (see Table 2), which led to an increase in the number of teachers and staff. In a departure from previous practice, new teachers were not chosen because of their understanding of and commitment to inclusion (or CES), and few had experience in inclusive or CES schools. Moreover, SMS continued to grow throughout the project. As a result of redistricting, the population of the school grew to 1,725 in 2000-01, and the number of teachers increased commensurately.

At the same time, the inclusion support staff was reduced. In 2000-01 (see Table 2), for example, there was the same number of support staff as in 1996, when SMS served 500 fewer students. The number of full-time co-teachers had dropped from six to three from 1996 to 1999, and although there were still two part-time co-teachers, guidance counselors and assistant principals no longer served as co-teachers. Furthermore, only two of the eight co-teachers from the Project RISES era were still working at SMS.

Program Changes. Initially, SMS held to its inclusion ideal, but significant events during the 1998-1999 school year presaged change. For one thing, co-teachers' responsibilities were expanded. In addition to their usual duties (collaborating with colleagues, adapting assessments, modifying curriculum, offering small group and one-on-one instruction, and completing individualized education programs and other paperwork), co-teachers were asked to substitute teach and were assigned regular lunchroom duty.

SMS also began to provide services for students who performed poorly on the FCAT. Concern about meeting these students' academic needs led to the establishment of a basic skills instruction program, known as the Direct Instruction (DI) Lab. The DI Lab served roughly 140

students found eligible on the basis of performance on screening tests. They participated in the lab 2 days a week for reading and 4 days a week for math, depending on their needs.

To pay for the DI Lab instructor, SMS invested less in co-teaching. Figure 1 shows the declining number of co-teachers (by total FTE) from 1995 to 2001; the decline in co-teachers from the original cohort was even more precipitous. These data show that SMS's investment in the inclusion program decreased markedly over these years. It was reduced first in 1997-98, when total FTE dropped from 7 to 5.5. By 1998-99, with the school's investment down to 5.0 FTE, full-time co-teachers worked with three or more teams, and individual teams had assistance available to them for no more than a day and a half a week. In 2002, when the project ended, SMS employed only 3.5 FTE co-teachers.

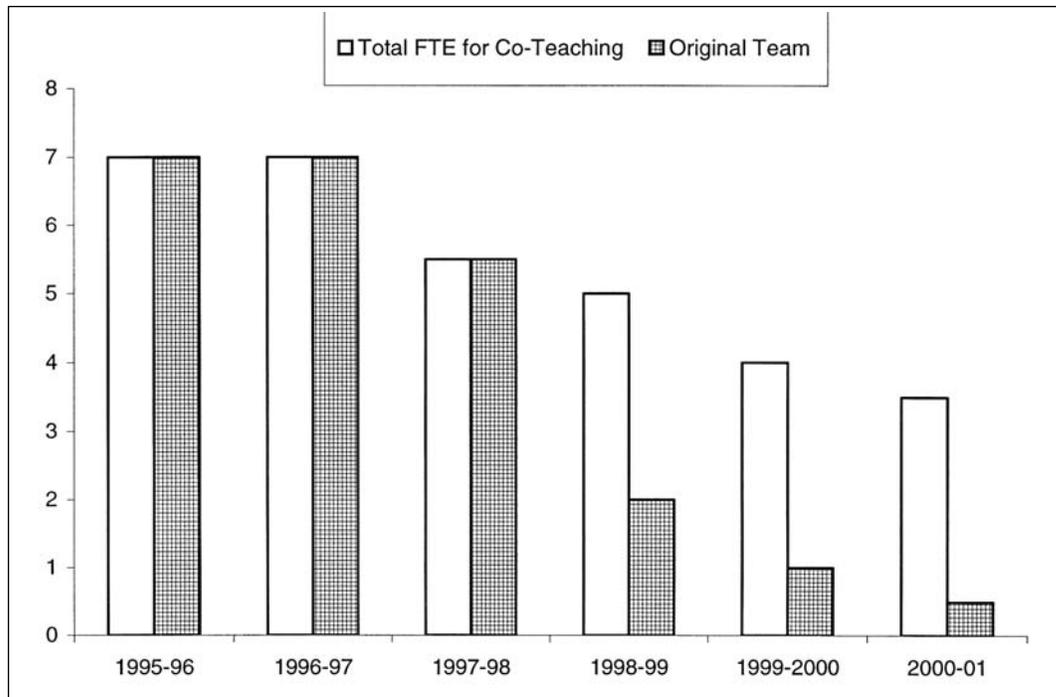
Between 1998 and 2002, SMS's A+ Accountability grade fluctuated between A and B (see Table 2). After earning an A in 1998-1999, SMS added a computer-based reading lab to its array of student pullout services. Most students attending the Read 180 Lab, as it was called, scored between the 20th and 40th percentiles on FCAT reading, and most were not special education students. This group was targeted for intervention because SMS stakeholders believed that they were most likely to improve their test performance substantially; students who scored below the 20th percentile apparently were deemed poor investments for remediation. Nonetheless, in large measure due to the stringent criteria for sustaining A grades, SMS fell to B in 1999-2000.

After a second B in 2001-02, SMS brought its grade back up to A. During these 2 school years, enrollment grew substantially, and the percentage of students on free and reduced lunch peaked (albeit at 16%). The percentage of students with disabilities also peaked in 2001-2002. Although the faculty grew by 10 positions from 1999-2000 to 2001-2002, the number of full-time co-teachers dropped to three, and the number of part-time co-teachers dropped to one. Most students with mild disabilities were included in general education classes, but a self-contained class and resource room were added (both serving primarily students with behavioral disorders), and

TABLE 2
SMS Population Changes

Year	School Enrollment	School Grade (began in 1998)	% of Disabilities	% Free and Reduced Lunch	Per Pupil Exp. Regular Students	Per Pupil Exp. Exceptional Students	Number of Teachers	Number of Support Staff	Number of FT Co-Teachers	Number of PT Co-Teachers
1996-97	1272		8.8	13.8	N/A	N/A	66	24	6	2
1997-98	1376		9.3	13.3	4,225	6,547	70	22	5	1
1998-99	1582	A	8.8	14.4	4,618	5,425	78	21	4	2
1999-00	1587	B	8.9	13.7	4,383	8,074	80	22	3	2
2000-01	1725	B	8.5	16.0	4,306	9,074	85	24	3	1
2001-02	1765	A	10.0	14.5	4,495	9,583	90	27	N/A	N/A

FIGURE 1
Co-Teacher Turnover



both the DI and Read 180 Labs continued to operate.

During this time period, our observations of teachers led us to conclude that the quality of group instruction remained strong, and experienced teachers continued to use strategies they had learned previously. On the other hand, there seemed to be less opportunity to share ideas and expertise, and professional development focused exclusively on FCAT preparation. Teachers seemed less collegial, and schoolwide activity gave way to concentrating on the work of teams. Thus, although the quality of instruction remained strong, there seemed to be less potential at SMS to sustain its high-quality instruction.

ANALYSIS OF CHANGES

The purpose of this longitudinal qualitative study was to identify factors affecting the sustainability of inclusion reform at SMS. Analyses indicated that the school underwent dramatic changes, moving from an inclusion program in the mid 1990s to a special education program characterized by a menu of pullout services and self-con-

tained placements. Our analysis suggests that changes in leadership, teacher turnover, and a shift in district and state priorities led to these changes. Reduced resources, a by-product of these primary factors, played a secondary role.

Shifting Leadership Priorities. Given keen public interest in school grades and the link between funding and FCAT performance, the district's continued focus on accountability and test scores came as a surprise to no one. SMS's first two principals were both effective at shielding teachers from district pressures; however, teachers felt the new principal at SMS was less effective:

Whenever the area office requests anything, the principal stops everything else to fulfill the request. (L. J.)

Our past principals have never said no to us. They always said "no" [to] downtown, and we would figure something out. Now it is everything comes from downtown. (M. S.)

In defense of the new principal, one teacher acknowledged the power of district pressure.

[The principal] has a pile of things to share with us from the district. They want more documentation. [Now] too much time spent in telling us how to teach and not enough in letting us do it. (P. B.)

The new principal also placed less emphasis on CES principles than the previous principals had. The teachers also began to sense less support for inclusion.

My first two principals carried the torch and inspired us to use the coalition principles to help kids. Now, there is no talk of purpose, no use of coalition principles and language to create a community educational agenda. There is no sense of teaching as a common good in the service of the community. (M. L.)

She [current principal] will keep our school out of the red. She cares very much for the bottom line, but she may end up holding the bottom line by herself. In the past, teachers were emotionally in the black, the budget was in the red; today, teachers are in the red emotionally, and the budget is in the black. (W. J.)

Shifting District and State Priorities. In contrast to CES evaluation principles, which emphasize “un anxious expectation,” state mandated reform seemed coercive. Teachers felt that the district was pressuring schools to spend their time and money on preparing students for the FCAT. Because SMS was historically a high-performing school, there was added pressure to produce high student achievement and good school grades. Teachers put it this way: “I think that accountability that the state puts on us . . . inhibits us from having time to do those kinds of things we used to do at SMS.” (K. S.)

Despite the added pressure, many teachers at SMS reported that they did not change how they taught and had no intention of doing so to meet the needs of the FCAT: “I was already teaching the thinking skills, the reading and writing in the context of the curriculum. I did not need to change anything.” (B. P.)

However, in practice, most teachers adopted state-developed practice materials. We observed most of these activities in math, where many teachers began lessons with FCAT practice prob-

lems. The activities themselves were organized and afforded students direct practice on skills on which they would be assessed. Most teachers carefully articulated FCAT practice activities to the lessons they planned independently. All were active and directive during these activities, sometimes more so than during the rest of the lesson.

In respect to content I don't teach to FCAT. However, I do spend 25 minutes warm-up period where I do FCAT content. (B. N.)

I did some more simulations and prompts; other than that, I did the same things I have always done. (T. J.)

“In the past, teachers were emotionally in the black, the budget was in the red; today, teachers are in the red emotionally, and the budget is in the black.”

Teachers on strong, longstanding teams—almost exclusively four-person, seventh- and eighth-grade teams—were least likely to report changes in their approach to teaching. The quality of instruction on these teams had been and continued to be outstanding, and the practices they commonly employed supported inclusion. In addition, strong teams occasionally shielded relatively inexperienced team members from FCAT pressures.

Our team does very little FCAT drill. We let the kids see what it looks like but we don't spend time practicing for the test. What we are doing is working on specific skills, like evidence practice in reading and social studies, that has affected what we teach. We are constantly asking ourselves, are we teaching higher order thinking? (B. R.)

On the other hand, teachers on weak or newly constituted teams seemed to be more heavily influenced by the FCAT pressure. Some of them told us:

Absolutely, no doubt, the strongest influence on my teaching is knowing that there will be an FCAT Test. It has become an all-encompassing goal. I was one of those people who became very stressed by it. I think I sacrificed some of the creativity and warmth that I would have used. I got very businesslike. I

may have made some of the kids stressed and tense. (B. M.)

There is no time for innovation; you have to teach to the FCAT test to some degree. Innovation takes trial and error, reflection, adjustment, and renewed effort. We don't have time for trying new things, and trying new things is not valued here. It is all about getting the kids to score better on the test, not getting them to learn things in depth or in a new way. (P. B.)

Thus, although FCAT pressures had a profound impact on SMS teachers' sense of common purpose and collegiality, it did not necessarily diminish the quality of teaching or hamper teachers' efforts to accommodate students with disabilities. Competent teachers, particularly those on strong, longstanding teams, were unlikely to change, and teachers who did change wound up adding an explicit teaching activity to their lessons—a change from which students with disabilities were likely to benefit.

Teacher Turnover. During Project RISES, SMS's teacher ranks were rife with enforcers. These key individuals worked closely with administrators to develop the inclusion program and were heavily invested in its success. However, between 1996 and 1998, as the number of staff increased, the number of staff with knowledge of the history of inclusion diminished. The departure of enforcers also impacted the inclusion program. Figure 1 shows that only one of the original seven SMS co-teachers remained in 2001-2002. Furthermore, the turnover rate at SMS during the 1990s averaged 21.3%; average turnover in suburban schools has been estimated at roughly 15% (Ingersoll, 2001).

Stakeholders were well aware of these changes. They told us: "It is real hard to sustain [reform] if the principal doesn't stay there, especially when she takes her faculty with her." (B. L.) "I think it is a level of general accepted maxim that it [inclusion program] was better and all the good people left." (K. D.)

The turnover also diluted SMS's commitment to the inclusion philosophy. With key players leaving and new teachers entering SMS, there was less understanding and less acceptance of the ideal.

I have been dealing with inclusion for a long time. When it doesn't work for some students there needs to be something in place to take care of them. I agree with the inclusion philosophy, but it doesn't work for everybody. (C. R.)

I think that students that are behind for whatever reason should get what they missed. I shouldn't be teaching seventh grade math to kids that don't know fourth grade math yet. They can't do algebra when they can't multiply. It is hard as a teacher to teach that kind of a class. I think that everyone loses. I feel really bad for the good students. (S. L.)

It also became clear that not all teachers understood what was meant by inclusion. We were told: "We have a total inclusion program; severely handicapped students are in a self-contained class." (B. P.) "We have full inclusion, so we have all kinds of students in the room. There are some pull out programs, and we have a self-contained class." (F. J.) "[We are a] full inclusion school. We don't have separate, self-contained classes. It is an excellent program. There are some separate classes." (N. K.) "We have full inclusion except for kids who get pulled out for needs. I guess that is not really full inclusion, but that is what teachers call it." (F. P.)

The failure to socialize new teachers to SMS ideals may have resulted from dilution of the school culture. For example, less emphasis was placed on collegiality and collaboration. In response to a question about the status of shared decision making, teachers told us:

No, it [SDM] is not as important. There does not seem to be as much consideration of teacher voices; not as much support for creative programs. There are some problems with communication after decisions are made; the results don't always get clearly communicated to the faculty. (F. J.)

I am not even sure what it is any more. I know what a decision is and what shared is. The joke around here is they will make a decision and share it with us. (G. K.)

Commitment to CES principles also waned. New teachers in particular did not understand the principles or how to apply them in their class-

rooms. Furthermore, the successful teaming model set up at SMS suffered as a result of turnover and lack of consistent team membership.

We were given the essential principles to put in our room, but we have never really discussed the coalition and the school's involvement with it. No one really oriented us to it. I have not been to any coalition activity this year. (C. T.)

We only talk about the coalition in the new teacher program monthly meeting, not at the school. I don't know anything about the way the coalition works at the school. I have not attended any coalition activity yet. (L. B.)

Again [Teaming] not as much as it used to be, and there has been a tremendous turnover of teaching staff. And if it is not your planning or your lunch you don't get to meet the new people. There are not as many social activities as there used to be, except by the same little cliques. So, the new people don't show up. (H. S.)

Reduced Support for the Inclusion Program.

The SMS inclusion program was never without flaws. RISES teachers bemoaned the lack of adequate support and expressed uncertainty about their ability to include students with disabilities successfully.

From 1999–2001, teachers grew even more frustrated with reductions in co-teacher support. The changing tenor of their comments—from “She is a Superwoman!” (L. M., 1999) to “She is spread too thin to be useful.” (M. L., 2000)—reflected this frustration.

The Special Education Department is not familiar with these kids. . . . Even at IEP meetings the Special Education person might not be able to pick a kid out of a lineup. She'd come in and ask me to point out kids she's supposed to be seeing. (B. M., 2001)

Our students with disabilities are pulled out of class and then put in special classes; there is no support facilitation. (B. R., 2001)

Only one teacher ever told us she received adequate support from her co-teacher. Resources for the inclusion program had been depleted, and it took a toll on the faculty.

It's ridiculous the number of teams they (co-teachers) are supposed to service. This subbing thing is ridiculous; there's a shortage of subs, and they are pulled away from their teams, and their credibility is shot. A lot of teams have not bought into inclusion because they don't get the help they need and can't depend on their co-teacher. (G. M.)

Furthermore, resources once invested in inclusion were redirected to programs for students who did not score well on statewide assessments.

We were in a parent conference yesterday and we were asked: “Do we have a guidance person?” Before we had guidance at each level. Now we don't have adequate support for our inclusion kids or our regular kids. (R. S.)

There are things that we need to be doing as co-teachers that we can no longer do. The SOS lab, for example, where kids could come for help. Programs have been dropped this year because there is not enough support. (R. S.)

One teacher summed things up this way:

Putting kids that were not emotionally or intellectually equipped to handle the work into the regular classes and not having the resources to help them be successful, later on down the road . . . it wasn't good for them. When you have a small number of students with disabilities combined with support teachers to help them be successful, inclusion works, and under those conditions I support it fully. It does not work for all students, in all settings. You have to honestly evaluate the conditions and the students. (V. J.)

DISCUSSION

In our research on the sustainability of reform at Socrates Middle School, three primary assertions helped us explain why inclusion was not sustained: leadership change, teacher turnover, and state and district policy change. An additional factor—reduced support for the program—also contributed to the lack of sustainability. In this section, we discuss these factors, how they interrelate, and how our understanding of these findings compares with principles established in previous research. We also consider the implications of our findings for school practice.

Three primary assertions helped us explain why inclusion was not sustained: leadership change, teacher turnover, and state and district policy change.

LEADERSHIP CHANGE

SMS changed principals twice, once during the initial project and a second time in 1997, a year before the second project began. Previous literature on special education classroom reform suggests that leadership change may undermine sustainability, particularly when a new principal does not devote time to the innovation (Klingner, Vaughn, Hughes, & Arguelles, 1999). Our findings about schoolwide reform were consistent with this assertion. Because the first new principal was committed to her predecessor's reform agenda and was successful at keeping teachers working towards the inclusive vision, that change had little impact on the program. By contrast, the second new principal seemed more committed to recent district initiatives related to student assessments and school grades. She expressed support for CES, inclusion, and shared decision making, the backbones of SMS reform, but she understood that her school—and her own performance as its leader—would be judged primarily on how well SMS students scored on the FCAT.

In their study of classroom specific reform, Klingner et al. (2001) determined that procedurally rotating principals added to the difficulty of sustaining reform. To this generalization, we would add that changes in school leadership may affect schoolwide reform in different ways, depending upon the principal's affinity for and commitment to an established schoolwide reform agenda.

TEACHER TURNOVER

Teachers transferred out of SMS at a surprisingly high rate, even before the second principal change. In a sense, all of the principals brought on new teachers whose beliefs were aligned with the prevailing reform philosophy. Thus, during the tenure of the third principal, teachers were hired with less attention paid to their knowledge of inclusion and their commitment to the co-teaching model than had been the case before we

returned to SMS. It is not surprising that this process diluted faculty commitment to the original reform or that, as a group, the SMS faculty grew less knowledgeable about including students successfully—and less enthusiastic about trying.

Teacher turnover had a second deleterious effect on the inclusion program. Many of the teachers who left were the enforcers (Huberman & Miles, 1984) who had helped to create and establish inclusion at SMS. Huberman and Miles noted that enforcers used reform expertise to demonstrate their leadership and move to administrative positions, and they considered such upward mobility a problem inherent in reform. Our findings at SMS were consistent with these assertions; several enforcers devoted to the inclusion program left with the first two principals and followed them to new schools, often in leadership roles. Others left with the arrival of the second new principal. If other teachers were being groomed for leadership roles—for the inclusion program or otherwise, it was never apparent to us or to the teachers we interviewed. Apparently, at SMS, teacher leadership was a victim of high teacher turnover and the loss of focus that came with new and competing reform.

CHANGE IN STATE AND DISTRICT POLICY

Long before the passage of the No Child Left Behind Act in 2001, the state of Florida initiated a program of high-stakes assessments and accountability in which schools were graded, largely on the basis of FCAT performance. Previous work on the sustainability of both classroom specific and schoolwide reforms (Huberman & Miles, 1984; Klingner et al., 2001) emphasized the importance of the match between a reform and the policy context in which it is implemented. Furney et al. (2003) concluded that high-stakes assessment was a poor policy context for classroom-specific reforms; our findings suggest that high-stakes assessment also proved to be a poor context for inclusion, a schoolwide reform. The third principal began at SMS a year before school grades were first awarded, as pressure to perform was being ratcheted up. She made clear to her teachers the importance of improving FCAT performance. All teachers felt pressure; they were being held to high accountability standards but lacked the re-

sources to meet all students' needs. They did not have adequate co-teacher support, social supports for students, or training to deal with students with diverse needs. Programs like the DI Lab were added that focused on low-performing—but not necessarily special education—students. There was less communication and teaming as teachers began to sacrifice creativity to the demands of test preparation.

Research on classroom reform has found that teachers are more likely to use new practices if they see improvements in students, especially hard-to-teach students (Klingner, et al., 1999; Vaughn, Klingner, & Hughes, 2004). At SMS, district accountability policy forced teachers to redefine student improvement as performance on standardized tests rather than performance on more sensitive academic assessments (e.g., curriculum-based measures) or measures of social growth. Redefining success on the basis of academic test performance obscured the benefits of inclusion, particularly for students with disabilities and other students with learning difficulties, and thus undermined the sustainability of the reform.

IMPLICATIONS FOR SCHOOL PRACTICE

Our findings make clear that change in school leadership can sustain reform—or drive new reform, depending on the new leader's commitment to a particular agenda. At SMS, we witnessed an example of both. The first principal change had little impact on the program because both the first and second principals were committed to inclusion and the broader reform agenda on which it was built. The second principal change coincided with the emergence of a new reform agenda—high-stakes assessment and school accountability. As a result, inclusion and high-stakes accountability were pitted against one another, with the new principal thrust into the role of championing a new agenda. Her circumstance was unenviable: Marching orders in hand, she entered an environment where most teachers remained committed to inclusive reform.

Given state and federal pressure to implement high-stakes assessments and judge school performance, district policy to improve FCAT scores overwhelmed its commitment to inclusive reform—and that priority was communicated in

no uncertain terms to principals. In fairness, there is no way to determine how the previous principals would have dealt with such pressure. Furthermore, it does not seem likely that the second change was intended to subvert inclusion at SMS. After all, the second principal was reassigned—to the largest high school in the district, a prestige position, and the new principal had administrative experience at a nearby feeder school. In a large district, principal change is a gamble; sometimes it works to sustain reform, and sometimes it does not. That such a change would render SMS vulnerable to the powerful pressure of high-stakes assessment and the need to improve test performance, in retrospect, seems obvious.

SUMMARY AND IMPLICATIONS

In sum, we found three major factors that contributed to the demise of the inclusion program at SMS: changes in leadership, shifting district/state policy, and teacher turnover. These factors also led to diminished philosophical and financial commitment to the reform. In turn, lack of resources stressed an already fragile program structure. This study provides further support for the importance of strong principal leadership, proper teaching training, and adequate resources in maintaining reform. It adds to the literature by demonstrating these same factors are important in schoolwide special education reform.

REFERENCES

- Ashton, P. T., & Webb, R. B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement*. New York: Longman.
- Ball, S. J. (1987). *The micro-politics of the school: Towards a theory of school organization*. London: Methuen.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Education of All Handicapped Children Act. (1975). Pub. L. No. 94-142.
- Fetterman, D. M. (1983). Ethnography in educational research: The dynamics of diffusion. *Educational Researcher*, 11(1), 17–22.
- Florian, J. (2000). *Sustaining education reform: Influential factors*. (Report No. EA 031023). Aurora, CO: Mid Continent Research for Education and Learning. (ERIC Document Reproduction Service No. ED 453 583)

- Furney, K. S., Hasazi S. B., Clark/Keefe, K., & Hartnett, J. (2003). A longitudinal analysis of shifting policy landscapes in special and general education reform. *Exceptional Children, 70*, 81–94.
- Gersten, R., Chard, D., Baker, S. (2000). Factors enhancing sustained use of research-based instructional practices. *Journal of Learning Disabilities, 33*, 445–457.
- Huberman, A. M., & Miles, A. B. (1984). *Innovation up close. How school improvement works*. New York: Plenum.
- Individuals with Disabilities Education Act of 1997. (1997). Pub. L. No. 105-17, 611 et seq.
- Ingersoll, R. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal, 38*, 499–534.
- Kilgore, K., Griffin, C. C., Sindelar, P. T., & Webb, R. B. (2001). Restructuring for inclusion: A Story of middle school inclusion (Part I). *The Middle School Journal, 33*(2), 44–51.
- Kilgore, K., Griffin, C. C., Sindelar, P. T., & Webb, R. B. (2002). Restructuring for inclusion: Changing teaching practices (Part II). *The Middle School Journal, 33*(3), 7–13.
- Klingner, J. K., Arguelles, M. E., Hughes, M. T., & Vaughn, S. (2001). Examining the school wide “spread” of research-based practices. *Learning Disability Quarterly, 24*, 221–234.
- Klingner, J. K., Vaughn, S., Hughes, M. T., & Arguelles, M. E. (1999). Sustaining research-based practices in reading. *Remedial and Special Education, 20*(5), 263–274.
- No Child Left Behind Act: Reauthorization of the Elementary and Secondary Education Act. (2001). Pub. L. No. 107-110, §2102(4).
- Patton, M. Q. (2002). *Qualitative evaluation and research methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Pfaffenberger, B. (1988). *Microcomputer applications in qualitative research*. Newbury Park, CA: Sage.
- Sherman, R. R., & Webb, R. B. (Eds.). (1988). *Qualitative research in education: Focus and methods*. London: Falmer.
- Smith, L. M. (1978). An evolving logic of participant participation in educational ethnography and other case studies. In L. Shulman (Ed.), *Review of Research in Education*, (Vol. 6, pp. 316–377). Ithaca, IL: Peacock Press.
- Spradley, J. P. (1979). *The ethnographic interview*. New York: Holt, Rinehart, & Winston.
- Spradley, J. P. (1980). *Participant observation*. New York: Holt, Rinehart, & Winston.
- Stake, R. (1980) The case study method in social inquiry. In H. Simons (Ed.), *Towards a science of the singular* (pp. 62–73). Norwich, England: Centre for Applied Research in Education.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research methods: The search for meaning*. Newbury Park, CA: Sage.
- Strauss, A. L. (1987). *Qualitative analysis for social scientists*. New York: Cambridge University Press.
- Taylor, S. J., & Bogdan, R. (1984). *Introduction to qualitative research methods*. New York: John Wiley & Sons.
- Vaughn, S., Klingner, J. & Hughes, M. (2000). Sustainability of research-based practices. *Exceptional Children, 66*, 163–171.
- Vaughn, S., Klingner, J. K., & Hughes, M. T. (2004). Sustainability of research-based practices: Implications for students with disabilities. In A. M. Sorrells, H. J. Rieth, & P. T. Sindelar (Eds.), *Critical issues in special education: Access, diversity, accountability* (pp. 135–153). Boston: Allyn & Bacon.
- Webb, R., & Glesne, C. (1992). Teaching qualitative research. In M. D. LeCompt, W. L. Millroy, & J. Preissle (Eds.), *The Handbook of qualitative research in education* (pp. 771–813). San Diego, CA: Academic Press.
- Will, M. C. (1986). Educating children with learning problems: A shared responsibility. *Exceptional Children, 52*, 411–415.

ABOUT THE AUTHORS

PAUL T. SINDELAR (CEC FL Federation #1024), Professor, Department of Special Education; **DEIRDRE K. SHEARER**, Doctoral Candidate, Department of Educational Psychology; **DIANE YENDOL-HOPPEY**, Assistant Professor, School of Teaching and Learning; and **TODD W. LIEBERT**, Research Assistant, Department of Counselor Education, University of Florida, Gainesville.

Address all correspondence to Paul Sindelar, Department of Special Education, G-315 Norman Hall, 605 SW 13th Street, Gainesville, FL 32611-7050 (e-mail: pts@coe.ufl.edu).

Manuscript received October 2004; accepted April 2005.