This paper provides a cross-program view of the results of three successful school improvement programs that regional educational laboratories developed and implemented. The programs and their developers are: (1) Achieving Excellence from the Mid-Continental Regional Educational Laboratory; (2) Onward to Excellence (OTE) from the Northwest Regional Educational Laboratory; and (3) the School Improvement Partnership Process (SIPP) from the Southwest Educational Development Laboratory. All three programs have been field tested and have demonstrated their efficacy for improving schools. Their intellectual roots lie in research on effective schools and school improvement processes. All three programs attend closely to involving school-based practitioners in decisions regarding improvements, maintain an information component so participants share knowledge, foster data-based decision making, provide personnel with a perspective that change takes time, and protect schools from unreasonable expectations. In each case, outside assistance is a feature of the program. Limitations include: (1) disaggregation of data beyond student achievement; (2) a limited research base; and (3) inadequate methods of assessment. These programs represent solid efforts to provide research-based guidance to teachers as they work to improve schools. There is a 21-item list of references. (SLD)
A Cross-program View of School Improvement Results

Naida C. Tushnet
Southwest Regional Laboratory
AERA Annual Meeting, San Francisco, CA

April 1992

BEST COPY AVAILABLE
Introduction

This paper provides a cross-program view of the results of three successful school improvement programs that regional educational laboratories developed and implemented. The programs and their developers are: A+chieving Excellence from the Mid-Continent Regional Educational Laboratory (McREL); Onward to Excellence (OTE), Northwest Regional Educational Laboratory (NWREL); and the School Improvement Partnership Process (SIPP), Southwest Educational Development Laboratory (SEDL). All three programs have been field-tested and have demonstrated their efficacy for improving schools. Their intellectual roots lie in two areas of research—the effective schools research and research on school improvement processes—thereby combining process and substance.

Along with an analysis of the strengths the programs share, the paper notes some of their limitations. It concludes with tentative recommendations for improving these research-based school improvement programs.

Strengths

All three programs are built on research on effective schools, effective instruction, and school improvement. Consequently, the programs attend closely to involving school-based practitioners in decisions regarding improvements. Further, all three programs contain an information component so participants share knowledge of and a language about research findings related to school and classroom effectiveness. In addition, the programs foster databased decisionmaking, providing staff members with a framework for analyzing the current status of the school, identifying gaps between what is and what they want the school to be, and reviewing current practice to align curricula and implement research findings concerning effectiveness. All three programs also provide district and school-site personnel with a perspective that change of any value takes time, and, in this way, protects the school from unrealistic expectations about improvement. Finally, the programs offer outside assistance.
Involving School Staff

The most common method for involving school staff is through the development of leadership teams that include teachers and administrators. These teams may also include other school-site professionals, parents, central office administrators, and, occasionally, students.

Using representative teams serves three functions. First, it reduces the number of people the outside assisters have to work with, thus making the work manageable. This is important because the process of change involves building new understandings. That is, the team comes to view what is happening in the school and what could be through a common set of lenses. They "construct reality" (Berger & Luckman, 1967) in a way that can guide their action. The outside assisters need, then, to assess constructions of reality at the start of the process and provide the information and tools in ways meaningful to the participants. Clearly, it is easier to do that with smaller, rather than larger, numbers. Working with smaller numbers also facilitates the assisters' understanding of the local context, the culture, mores, and ways of interacting within the school, district, and community. This "local knowledge" (Goldenberg & Gallimore, 1991) has important implications for how the improvement process proceeds. The solutions to local problems must come through the development of shared meanings (Fullan, 1991), and the teams facilitate such development.

Leadership teams also provide a basis for "bottom-up" involvement. Although there are facile claims for the importance of staff involvement (see, for example, Onward to Excellence: "Research shows that involving the school principal, other school staff members, and district office personnel jointly in a school improvement effort increases success."), the issue is more complex (Fullan, 1991). However, whatever the particular role is of staff involvement in planning and implementing the school improvement program, the leadership team at least serves as a touch point within the school for the outside assisters. In addition, it is likely the team members are better able to "sell" the improvement process to their colleagues than outsiders could.

Finally, the leadership team develops a fairly sophisticated understanding of research, social problem-solving processes, and school improvement. As a result, the outside agents leave behind both structure and capacity for continuing improvement. Developing collegial relationships and database decision-making can have an impact on the school long after the specific improvement program is "complete."
Research Knowledge

All three programs contain a "knowledge-transfer" component. In all, the school-site professionals, particularly the leadership team, are introduced to concepts drawn from research on effective schools and effective instruction. In addition, either directly (as in the case of A+chieving Excellence and SIPP) or as part of the program itself, site practitioners become knowledgeable about research on school improvement.

The research base that is drawn upon covers curricula, instruction, and testing. For example, within OTE, staff members engage in a process of analyzing current instructional practice and compare it to effective instruction. The A+chieving Excellence program also includes information about effective instruction, as well as the relationship of the curriculum to the testing program (curriculum alignment).

Staff members at participating schools develop action plans to guide the implementation process. The external agents provide a framework for planning, based on what is known about school improvement processes and effective staff development. That is, the plans include descriptions of how progress will be monitored, new skills taught and practiced, outcomes judged, and procedures for adjusting the plans. The plans also focus the attention of staff to a limited number of issues or provide phased attention to greater numbers. The focus is important for success because trying too many things at once can be overwhelming.

The programs' limitations stem from the limitations in the research bases being used. That is, the "effective-instruction" research that is drawn upon highlights generic instructional approaches and does not deal with content-specific issues. This matter will be elaborated in the discussion of program limitations.

Databased Decisionmaking

Perhaps the most important aspect of the programs is their attention to databased decisionmaking. School personnel do not tend to use hard data to evaluate their activities and make decisions. Rather, they operate on immediate feedback and more "intuitive" views of their classrooms and schools (Huberman, 1983; Lortie, 1975; Cuban, 1988). Although the programs use different mechanisms, they carefully assess student outcomes, including achievement and attendance. As the leadership team and others analyze the data, they can determine priorities for improvement.
The use of data is facilitated by the attention the programs give regarding how to provide the information to the school staff. OTE, for example, prepares a one-page display for each type of data accompanied by a narrative description. The profile data are nonevaluative so that staff members can exchange understanding of the meaning of the data and judge how well they believe they are doing with regard to each group of data. Somewhat differently, A+chieving Excellence provides data about the relationship of the curriculum to the testing program to assist teachers in adjusting to one or the other. This type of information helps teachers understand student achievement patterns.

Importantly, the programs facilitate attention to students from different backgrounds. The SIPP program, for example, focuses much attention on disaggregating data to help teachers see how well educationally disadvantaged students are performing. The data disaggregation draws attention to particular students, and, in the case of SIPP, clearly has payoffs in their achievement. That is, when compared to similar schools, the schools in which principals received SIPP training show statistically significant greater closing of the gap between the achievement of educationally disadvantaged and higher-income students. The importance of disaggregation cannot be overemphasized and should, I will argue below, be applied beyond student achievement issues.

Finally, data also are used to determine how well instructional practices align with research on effective instruction. This allows staff members to identify strengths and weaknesses and indicate areas for staff development activities. The picture of current practice compared to “best” practice becomes the basis for the plan. The process implemented has two advantages over other potential approaches. First, it works with schoolwide data, and thus avoids dealing with any particular teacher as “deficient.” It harnesses the professional ethos of all teachers and helps build collegial investment in the improvement process. Second, it anchors instructional improvement in research, rather than in opinions (of which there are many) of what constitutes good teaching. Consensus building activities can shift focus on goals rather than on procedures to reach the goals. Those procedures are empirically derived. (As noted below, there is a downside to this approach.)

**Time Perspective**

The programs take as given that “change is a process, not an event” (Hall & Hord, 1987). All have multiyear time perspectives to introduce the process, implement the procedures, and evaluate outcomes. In some sense, this may be the most significant assistance to the schools. The program staff members make it clear to school and district personnel that improvement involves dedicating time and effort and results will not be immediate.
The fact that the regional laboratories are willing to adhere to what we know about school improvement is a credit to them, and an argument for continuing support of them. Public education has long searched for the magic bullet and quick fix. Indeed, given where we currently are, one can see the six “national educational goals” as ignoring the findings of research on educational improvement in the call for their achievement “by the year 2000.” The time perspective brought to these programs emphasizes again that improvement is a difficult process, requiring people to rethink current practice and learn new skills. And, to the extent that the improvements face the basic problems confronting educators, the task is more difficult and time consuming.

By providing the perspective that change takes time, the programs protect the school participants from demands for immediate payoff. In schools with major problems, this protection may also serve as a morale booster for teachers. That is, they have the space to attend to problems without being criticized for their failures. The outlook of the programs is more positive than negative and recognizes that feelings of failure lead as often to paralysis as to action.

The time perspective, coupled with using leadership teams, also helps build a culture of collegiality and continuous improvement, which Bird and Little (1985) found to be associated with improving schools. It takes time to build such cultures, in part because they rely on mutual trust among school-based professionals, and such trust is not automatic. The time also is important because schools as currently structured tend to isolate teachers from one another (Rosenholtz, 1989). Collegial action requires attention to the logistics of getting teachers together, and, in fact, the time available is more limited than anyone would like.

Outside Assistance

Assistance from a nonschool organization such as a laboratory serves several functions in implementing school improvement. First, lab staff members have synthesized and translated research into practical applications. Second, they view the school with fresh eyes and provide interpretations of actions, processes, and culture that add to or modify school personnel perspectives. Finally, outside assistance provides a continuing pressure to attend to the improvement plan and protection from being distracted by the next fad that comes along.

The research on school and classroom effectiveness is distilled in each of the programs. School-based practitioners are clearly capable of reading and understanding the research base, and many of them do. However, each of these programs has developed techniques that ease
application. For example, the school profiles in OTE facilitate data display and understanding. Other techniques could be used, and school staff could develop their own, but efficiency is served by the labs’ activities.

Perhaps even more important, the external agents bring a different set of perspectives to the school from those that already exist. The school improvement process requires that teachers view their work in new ways. That is, few teachers wish to teach badly or irrelevant curricula. Because they judge their own efforts by how well classes go (Lortie, 1975), alternate ways of assessing their efforts help them think about improvements to undertake. It is not that the external agents know “better”; rather, they know “different.” They bring interpretations of the meaning of policies and practices to the group. Their understandings will, in turn, be modified by the group’s reactions and feedback. Through this process, new understandings are developed, and plans for action can be developed. This outside assistance is particularly helpful when the change is a “multilevel, complex system-oriented innovation where what is being changed is the organizational culture itself” (Fullan, 1991, p. 73). The three school improvement programs are that type of innovation.

Finally, the external assistance both prods and protects the improvement effort. As Crandall et al. (1982) found, the combination of pressure and assistance led to improvement. Once the district or school has “signed on” to the lab program, scheduled dates for meetings and staff development are taken seriously. The pressure may be subtle, because after all, the major role of the lab is to provide assistance, but it is there. No school staff wants to fail to live up to its side of a bargain.

Along with the pressure that comes from the involvement with the laboratory program comes protection. School and district staff buy time from their critics by pointing to the improvement program. Indeed, improvement program staff members have to take care that their involvement goes beyond the need to say to critics, “Look how hard I’m trying.” They need to gain real, not just political, commitment to offer the protection. In addition, if the superintendent comes back from a meeting with a new plan, school staff members can be saved from immediate action on it by their involvement with the lab program. Obviously, this does not always work, but it does provide a barrier against a constant barrage of new programs.
Summary

The three programs have much to recommend them. They are built on research and provide extended assistance to participating schools and districts. Further, they draw attention to the achievement of disadvantaged students. Finally, they provide a framework for developing locally meaningful improvement plans. However, the programs also share limitations, which will be discussed in the following section.

Limitations

I am a bit hesitant in discussing the limitations of the programs. First, there is no doubt the programs are effective in their own terms. Consequently, pointing out limitations may be asking the programs to change their goals. Second, I fear that changing the programs may cause them to lose some of their value, and, indeed, impose upon them untested ideas. Nonetheless, I will raise two limitations that could be addressed within the context of the current programs and one that requires more long-term consideration.

Disaggregation of Data Beyond Student Achievement

Currently, the programs assist staff in disaggregating student-related data. This clearly has helped participating schools direct improvement efforts toward disadvantaged students. But focusing only on student outcomes limits the information that could be useful in developing the school improvement plan. It is important to find out the courses and instruction students receive as well as student outcomes.

Within American schools, there are patterns of course taking and types of instruction received related to educational disadvantage. As Oakes (1990) states:

Assessments of academic ability, placement in different tracks or ability-grouped classes, and the reduced educational opportunities that characterize low-track classes often parallel race and social class differences...to the extent that placement in classes at different ability levels affects students’ opportunities to learn—and the evidence from our study suggests that the effects are quite profound—minority students disproportionately suffer whatever disadvantages accrue to students in low-track classes. (pp. vi-vii)

Further, the effects of race, social class, and locale exist on the elementary, as well as the secondary, level. A school engaged in an improvement process, then, needs to look at patterns in
the courses students are taking and the opportunities available to them. Data disaggregation of this type will strengthen the diagnosis and prescription for improvement.

Another type of data disaggregation concerns the type of instruction students receive. Again, research has revealed different patterns of instruction offered to disadvantaged students from those who are from backgrounds of educational privilege. Metz (1990), for example, found that teachers in schools serving largely lower- and working-class students were likely to emphasize rote learning and obedience to authority in contrast to teachers of the privileged, who were likely to emphasize critical thinking. Data displays of surveys or observations of instructional practice should be disaggregated so such discrepancies can be seen.

Finally, an analysis of the National Educational Longitudinal Study of 1988 (NELS: 88) revealed major differences in the quality of instruction educationally disadvantaged and privileged students receive (Hafner, 1991). For example, high socioeconomic-status (SES) students are twice as likely as low-SES students to report conducting experiments in science classes daily and weekly. Further:

There are indications that more advantaged, higher-achieving students are more likely than less-advantaged youth to have teachers who majored or minored in the subject they teach (e.g., math or science) and that poorer, limited-English proficient and lower-achieving students are more likely than advantaged students to have math and science teachers who majored in education. (Hafner, 1991, p. 16)

The programs presented here demonstrate success in narrowing the gap between the achievement of educationally privileged and educationally disadvantaged students. Further attention to the educational program that students actually receive would provide an even stronger base for changes in practice. Disaggregation of information about teachers and instructional practice will serve that purpose. Such information also helps focus staff development and other improvement activities.

Broadening the Research Base

The programs build on research from a set of identifiable research traditions—the effective schools research, research on effective instruction, research on school improvement processes, and staff development research. There is some use, as well, of the research on school-site leadership. These research bodies focus primarily on generic characteristics and are neither subject-matter nor context-specific. Including subject-matter- and context-specific research would, indeed, change
the nature of the school improvement approaches these programs offer. However, there is an argument to be made for doing so.

The procedures for diagnosing problems and developing prescriptions, to use the language of OTE, employed in all three programs focus on generic instructional and curriculum issues. They ask: Are teachers using the type of instruction found to be effective by research? Is the curriculum aligned with the test? But important questions are left out by these foci.

In terms of instruction, there are differences in effective instructional techniques designed to accomplish different objectives. For example, there is a “consistent picture of teaching that yield[s] achievement gains: teaching that was highly structured and directed, involved explicit explanation and modeling by the teacher, and kept students highly engaged with academic content” (Putnam et al., 1991, p. 124). However, “achievement” is defined here as scores on standardized tests, a limitation to be discussed later. In mathematics, for example, this picture of successful teaching works well when math is conceived of as involving computation and routine problem solving. But if math is conceived to involve a broader range of tasks “involving various kinds of nonroutine problem solving, the mathematization of situations and judgments about the appropriateness of mathematical models for various purposes, and the use of mathematical argument and justification” (Putnam et al., 1991, p. 126), then a different sort of instruction is warranted. A similar case can be made with regard to other subjects.

Shulman (1987) argues that “pedagogical knowledge” is rooted in knowledge of the particular discipline being taught. The three school improvement models presented here do not provide a framework for looking at content areas, and therefore, do not guide an analysis of instruction appropriate to the goals and objectives of a particular subject. The SIPP program did address specific content areas (for example, writing), but it is not clear how the process led to the focus. One reason that this is important is that current reform efforts in the subject areas include profound changes in the ways knowing and learning are viewed, and school improvement efforts should address those changes.

The core of the curriculum issue for schools and teachers is determining what knowledge is worth knowing and, therefore, teaching. Arguably, the essence of the “reflective practitioner” (Schon, 1987) is attention to “the role that schools actually play within a race-, class-, and gender-divided society” (McLaren, 1987, p. 40). Consideration of such issues raises key issues of the knowledge validated by inclusion in the curriculum. The three school improvement programs do not provide the space for such questions.
One reason for including content-related concerns in a school improvement effort is to strengthen reform efforts. That is, what the three programs embody about process, capacity building, and staff development could be used well within the context of other efforts. For example, in studying the implementation of the California mathematics framework, Cohen and his colleagues (1990) found uneven use. The kind of assistance provided through the leadership teams, for example, could help Mark, who says, “I can’t teach what I don’t know” (Wilson, 1990), if the leadership team were empowered with content, as well as process, knowledge.

Including context-specific issues could address both the variations in school structures associated with their level and variations in treatment of students associated with socioeconomic status. Given the data disaggregation, the improvement models would be stretched, but not transformed, to attend to issues of differential treatment of students both within schools and between schools. These contextual concerns are not currently part of the programs, but providing teachers with tools to deal with them would be empowering.

Assessment

The final limitation of the programs that I will discuss has to do with assessment. Currently, standardized achievement tests are used as the base for evaluating student achievement. Most of such tests measure student achievement on a narrow band of skills and knowledge. The problem is, as Romberg (1992) says, “a growing consensus on the shift in goals for students... these goals cannot be achieved by current instructional practices in most schools, nor can school efforts to meet them be assessed using current tests” (p. 23). Within the school improvement programs, reliance on standardized tests to judge the current status and progress in the school is a limitation. The limitation operates in two ways. First, school-based practitioners can only know how well students are doing in areas the tests measure. To the extent the tests fail to measure important things, the leadership team cannot direct improvement efforts at them. Second, the tests build on assumptions concerning “the separability of ends and means, and the moral neutrality of technique” (Berlak, 1992, p. 15). But there is no moral neutrality, and the tests reify and strengthen the very forces that influence stratification by class, race, and gender. Both these issues are addressed below.

Standardized tests tend to use “short-answer, closed-ended format[s] [that] preclude the assessment of higher-order thinking and mastery of complex material” (Berlak, 1992, p. 8). Even when they involve problem solving, the focus is on the answer rather than “the process of
achieving it or the nature of the task" (Romberg, 1992, p. 24). Consequently, when such tests are used as the "baseline" on which faculty members judge the quality of the school and plan improvements, there is much missing information. School personnel can know, for example, how well students add and subtract, but they are unable to tell whether the students can address open-ended problem situations. Students may do well on the existing standardized measures, but fail to meet the standards embraced by the National Council of Teachers of Mathematics (1989). The point is that using such tests does not provide important information.

Perhaps more important, given the emphasis in these programs on improving schools for all students, the tests reinforce existing divisions by class, race, and gender. They rank-order individuals on their "knowledge" of what is included in the test. The tests measure "constructs" that inevitably include value judgments and subjective choices and, as a result, are the "products of power" (Berlak, 1992, p. 185). Moves toward "performance assessment," then, are likely to be more fair to students from poor families, ethnic-minority students, and girls than are current tests so long as the performance deemed important does not also reflect issues of power and control. That is, there is a need for debate and discussion within the local context, of what knowledge is worth knowing and how we will know if someone else knows it. The programs discussed here do not raise those questions. In the end, teachers not involved in discussions about such matters are not fully empowered, no matter what the rhetoric, because they are not involved in decisions about the central issues of education.

Conclusion and Recommendations

The three programs discussed here, A+chieving Excellence, OTE, and SIPP, are excellent programs. They represent solid efforts to provide research-based guidance to teachers as they work to improve schools, particularly how schools serve low-performing students. School staffs have used the programs to improve educational opportunities for students. Further, the processes used help build the capacity of school-based professionals to address new problems and concerns. The attention to process and support is an important attribute to all the programs. It separates these programs from those that hope for improvement through fiat or by providing short-term workshops. Consequently, the success rate for the programs is high.

Despite the strengths, the programs share limitations. Perhaps most important, whatever the rhetoric used about local empowerment, teachers rarely engage in thinking through the core issues of education. Second, the data disaggregation, a strength of the programs, is confined to
disaggregating data about student achievement. From that information, the local teams develop plans for improvement. However, the plans could be strengthened by disaggregation of teacher and program information as well. Then, the teams could focus more clearly on relationships between school and classroom processes and outcomes. Third, the programs draw upon research that addresses generic issues and does not relate to specific content areas or contexts. This limits their potential because key problems in schools may not be recognized. Also, current state reform efforts focus largely on curriculum issues and could benefit from the kind of assistance the school improvement programs provide. Finally, the reliance on standardized tests is limiting.

The program developers must be careful in trying to make their programs better, even along the lines suggested here. They could graft onto carefully developed and tested programs some ideas that do not fit the assumptions that frame the existing processes and approaches. However, I think it is important for the developers to work on two issues: First, they should develop some ways of involving school staff in basic curriculum questions. Staff should not only look at whether the curriculum is aligned with the test, but whether the curriculum focuses on important matters. The discussion of what constitutes an important matter requires guidance and methods for addressing the fact that school staff may hold different (and sometimes competing) goals for educating students. The second point is related to the first: Staff need to attend to the issue of testing and whether current tests focus on what matters. Exploring alternative approaches to assessment will strengthen discussions of what the curriculum should be. This combination will empower teachers far beyond most current site-based management programs.
References


