Improving the Performance of an Educational System

WILLIAM W. COOLEY
Learning Research and Development Center
and
School of Education
University of Pittsburgh

Preparation for this address has certainly been an awesome experience. In an early draft, I outlined all the fun I’ve been having these past few years working directly with a school district. I recounted how rewarding it has been and the important lessons I think I’ve learned about doing research that seems directly relevant to the current crisis in public education. Then I shared this draft with a few valued colleagues. The typical reaction was: “It’s an interesting story, Cooley, but I had hoped you’d say something presidential.” So I got to worrying about what a Presidential Address was supposed to be. About then my sociologist colleague, Leslie Salmon-Cox, shared with me the late Erving Goffman’s 1982 Presidential Address to the American Sociological Association. In his introduction he says:

For an evening’s hour, it is given to each current president of the Association to hold captive the largest audience of colleagues that it can provide. Someone you have selected from a very short list takes to the center of this vast ballroom on a hobby horse of his own choosing. . . . Taking office, presidents find a podium attached, along with encouragement to demonstrate that they are indeed obsessed by what their election proved they were already known to be obsessed by. Preparing and then presenting their addresses, presidents come to feel that they are temporarily guardians of their discipline. However large or oddly shaped the hall, their self swells out to fill it.

Certainly inspiring stuff. So I briefly revisited what I’ve been trying to accomplish this past year as AERA’s President.

The main goal that I set for my term was to see what AERA might do to better achieve its broad mission of improving education through research. There are two important aspects to this mission. One, exemplified by the annual meeting, is to improve the quality of educational research by facilitating communication among researchers. The other aspect of the mission is to increase the educational relevance of our research by facilitating communication between researchers and the possible users of the results of that research. As an Association, much less was being done toward that end.

I felt we needed to find better ways of helping researchers become more knowledgeable about the problems confronting practitioners, as well as helping practitioners to see the value of what we are learning about education in our research. In my President’s column in the Educational Researcher this past year, I’ve summarized some of the fruits of Council deliberation as we identified ways in which AERA could be more effective in this regard. The two new awards for the contributions of educational research to the improvement of educational practice represent one outcome of those efforts.

Looking back on those deliberations, one regret I have is that in our brief, agenda-packed Council meetings, we never took time to have debates about an issue that Elliot Eisner raised early in the discussions. He wondered what model of how research affects practice was guiding these various AERA initiatives. Let me briefly sketch out a few possible models.

One model, which I like to call the eminent scientist model, shows how research influences practice through tracing the ways in which important theoretical notions find their way into the culture of the schools. The volume Suppes (1978) edited for the National Academy of Education illustrates this very nicely by showing how the notions of such dissimilar scientists as Freud, Piaget, Thorndike, and Skinner eventually...
influence what happens in classrooms in very significant ways.

Then there are those who are impressed with what Simon (1982) says in his *Science of the Artificial*. In this model of research impact, the science of design guides the development of artifacts (i.e., tools and techniques) that educators can use to improve their performance.

The bottom-up model, on the other hand, recognizes the fact that a successful innovation is often one that has been developed by, or is at least significantly altered by, the teachers who are expected to implement the innovation. Berman and McLaughlin’s (1974) concept of mutual adaptation is illustrative of this notion of how educational research and development influences practice.

Then there is the decision-oriented inquirer model. This model recognizes the importance of generating information that is immediately useful to the policy-shaping community (Cronbach and Associates, 1980). In this model, educational research affects practice through research that is conducted as it is needed, and for whom it is wanted.

I think it is important to recognize that these various models of research influencing practice are not really competing models. They seem to me to be describing the various ways in which research can affect practice. Sometimes in our enthusiasm to promote a particular kind of educational research, the irrelevance of other forms is implied, occasionally even claimed. Scarce resources will do that to people.

Of course even presidents can be wrong. These various models may not all be valid even though they are not necessarily contradictory. Also, there seem to be a few folks out there who doubt the validity of any model that tries to show the value of our enterprise. More effort needs to be devoted to the task of validation. In the process, we might even find some models more valid than others.

One way of validating this type of informal model is to document particular instances in which research results have influenced educational practice. Time does not permit me to do more than contribute to the documentation of one such model this afternoon. So I thought it would be useful to describe just how decision-oriented inquiry can have an impact as the inquirer seeks to improve the performance of an educational system. I hope this will stimulate those engaged in other forms of research to contribute to this much needed documentation.

“It is important to recognize that these various models of research influencing practice are not really competing models.”

And now I’ve got the rationale for the address I wanted to give in the first place! As I tell my tale, there are some specific themes that I would like to ask different audiences to be listening for.

*To University-based Researchers:* If you have been depressed because a dead silence followed the publication of your latest article in one of our journals, try taking your research skills to an educational setting and offering to work on their problems. It can be a very stimulating and educationally rewarding experience (as Larry Barber tried to tell me here at AERA in 1977).

*To District-based Evaluators:* If you’ve been putting all your energies into discrete studies that try to establish the effects of specific programs or policies, and are frustrated that your results are not influencing district decisions, try something like the monitoring and tailoring approach (which I will describe shortly).

*To People Interested in Educational Reform:* Before launching some new broad-brush solution to an educational problem, be sure that what is already in place is working as well as it can be. If it is clear that fine-tuning won’t do the job, be sure the context of the problem is adequately understood, or you may end up doing more harm than good.

---

**Working With A School District**

Let’s begin with a little background. At the Learning Research and Development Center, Bill Bickel and I have been trying to find ways to improve educational practice through the improvement of the evaluation processes that take place within school districts. Our approach assumes that a lot can be learned about how to do district evaluation by trying to do it ourselves...an old idea from John Dewey. This approach was possible because the Pittsburgh Public Schools (PPS), the school district in which LRDC is located, did not have an evaluation office. In 1978, we offered our services as district evaluators and have been working hard at the task ever since.

Our relationship with the PPS provides us the unusual opportunity to conduct district-level evaluations in close cooperation with school district personnel. The problems, the need for timeliness, and the political context that we experience are similar to those experienced by other district-based evaluators. However, this research has an important difference. Additional resources are available from NIE to study the evaluation process as it occurs and to communicate specific findings, methodological insights, and information about the process of applying results to our colleagues in educational research.

Over the 2 years prior to fall 1980, we had developed a good working relationship with the District, building up the trust and confidence that this kind of effort requires. In spring 1980, however, things became a little unsettled in the District. For example, the latest desegregation plan (in the District’s 14-year-old struggle to develop an acceptable one) was rejected by the Pennsylvania Human Relations Commission; the School Board chose not to renew the Superintendent’s contract; and the teacher’s union announced that it would go on strike in the fall if its demands were not met. (I was ready to retire to tidy little methodological studies. Fortunately Bill Bickel convinced me to hang in there.) Eventually the teacher’s contract was settled, a desegregation plan was tentatively
approved, and the School Board selected Richard Wallace as their new Superintendent.

**A Districtwide Needs Assessment**

We met with Superintendent Wallace during his first week in office (September 1980) and offered the continued assistance of LRDC’s Evaluation Unit. During this meeting, he indicated that it would be very important to conduct a Districtwide needs assessment that would provide a basis for establishing a set of priorities for the District. Several subsequent meetings with the Superintendent and others were held to define the general nature, purpose, and procedures to be used in this assessment.

The primary purpose of the assessment was to determine the degree to which the Pittsburgh Public Schools were meeting the educational needs of children in the District, and to do it in a way that would suggest priorities for improving their educational program. The general objective was to identify conditions in the District that could and should be improved.

Two approaches were used in this assessment. One was to survey the various stakeholder groups regarding their perceptions of current conditions in the District. The other approach was to examine such data as student achievement, failure and dropout rates, attendance, suspensions, and demographic descriptions of the students in each school, looking for clues that might further indicate the educational needs of students. Stufflebeam’s (1977) paper was very helpful in conceptualizing these two approaches to the needs assessment.

We spent a very intensive 3 months in collecting and analyzing data. Fortunately, we had begun building a District data base during the prior 2 years, and some of the data extended back 5 years. This made it possible to examine trends over time as well as relationships among the many different variables in any given year. The results were organized in terms of five broad areas that were perceived to be the District’s most pressing needs. Although some stakeholder groups ranked them differently than others, there was consensus regarding the District’s most pressing problems: (1) student achievement was not satisfactory; (2) no one was happy with available procedures for personnel evaluation; (3) almost half of Pittsburgh’s students were going elsewhere to school and one in three of those that did not graduate; (4) enrollment decline was not being well managed; and (5) there was no systematic effort in place for improving individual schools. The more specific survey results and the 5-year data base were used to further define the nature of each of these problems, and possible strategies for dealing with them.

Dissemination of the results began in January, using both interactive slide show sessions and written reports. The illustrated presentations proved to be especially useful because they stimulated dialogue with different audiences (School Board, administrators, teachers, press). The first indication that the assessment results were going to be used came in the form of a newsletter that followed a day long retreat we had with the Superintendent and School Board. It reported that Wallace and members of the Board of Education had just had the opportunity to review results of the Districtwide needs assessment and that action plans would be developed to deal with the results. At its next regular meeting, the Board voted to assign resources and concentrate its efforts to resolve the problems identified in the needs assessment. Then in September 1981, 1 year from the beginning of the needs assessment, the resulting plans were formally adopted and initiated.

In addition to the districtwide needs assessment, we have taken on a variety of other evaluation activities for the District. These experiences are being summarized in a series of case histories (e.g., Bickel & Cooley, 1981). Each case history describes what the School District asked us to do, what we did, why we did it, what was learned from what we did, and what the District did with the resulting new information. The case histories are being written to help us clarify our own thinking about our work in the schools, and in the hope that what we are learning might be useful to other district evaluators, to those who are training evaluation researchers, to clients of evaluators, and to scholars of knowledge use.

**Generalizing from our Evaluation Activities**

So far, I think there are two important generalizations that can be drawn from our various experiences. One is the importance and nature of a client orientation, the other is the need for a shift in emphasis from formal program evaluation to a systems approach to program improvement. Those of you who have been toiling in this field will recognize that these notions are not exactly novel. (For example, the recent books by Meyers, 1981, and Borich and Jemelka, 1982, and a chapter by Cook and Leviton, in press, consider the connection between evaluation and the systems approach.) But neither are they settled issues. First, let’s take a brief look at what I mean by client orientation; then we will turn to the second, more important generalization.

**Client Orientation**

We have set as our objective the generation of information that can be useful in a decision context. At the Board level, the important decisions have to do with priority setting and allocation of resources. Our client orientation involves working hard to try to understand the decision context in which the client operates, identifying information that is missing in this context, then working up data in ways that might inform the dialogue among those in the policy-shaping community. The case histories we are preparing detail how we go about this critically important task. The value of the new information was not that the data determined priorities or settled policy issues, but that “they permitted those issues to be argued more productively” (Duncan, 1969).

Our experiences have convinced us that it is important to identify a primary client, and to serve that client well. We have identified the Superintendent as our primary client. This is a little bit at variance with the stakeholder notion, but not as different as it seems. (A good analy-
esis of the stakeholder approach by Weiss, Stake, Farrar, House, Gold, and Cohen is found in the 1983 volume edited by Bryk. We do try to determine the perceptions of various people with a stake in the Pittsburgh Public Schools. It has proven to be very useful to contrast the perceptions of different groups, and compare perceptions with other kinds of data. The problem is, if the evaluation activity is guided by the information requirements of a variety of stakeholders with a broad range of interests, the evaluation effort tends to try to serve everyone's needs for information and ends up serving no one very effectively. The evaluation of the Cities in Schools program (Murray, Bourque, & Mileff, 1981) illustrates this point.

Client orientation does not mean doing the client's bidding. Rather, it takes the form of a mutual educational process. It requires a dialogue between client and evaluation researcher out of which the needs for information are identified and strategies for obtaining it are defined. It also requires the researchers to be methodologically eclectic. Evaluation has suffered from evaluators trying to force the evaluation task presented to them into their particular "model" for doing evaluations. The book that Brandt (1981) edited illustrates this point very nicely. He had several notable evaluators describe how each of them would evaluate a particular program, and the contrasts are quite revealing. (I must admit that I do remain hopelessly quantitative, but when the problem requires qualitative documentation, my social historian colleague Bickel takes the lead.)

Another important feature of our work is how we try to get new information into the working knowledge of the participants in the decision context (see Kennedy, 1982, for a good discussion of this notion). This is as important as trying to figure out what information was needed in the first place. Busy managers don't read big reports (Sproull & Larkey, 1979). Thus we have done our primary dissemination through interactive slide shows and informal conversations with the participants.

Now I must admit that we have a very special client. Not too many school superintendents have previously been a director of a Regional Laboratory or a director for planning and evaluation at a University R&D Center, as Dick Wallace has. Clearly his determination to apply educational research in his various efforts to improve the Pittsburgh schools has made possible our search for fruitful ways of helping. Although his research orientation is a critical factor in our explorations, the methodological approaches that are emerging from this work seem to be generalizable to other districts.

Notice that I am not claiming that generalizable knowledge about educational phenomena is an outcome of this work. What seems generalizable are the approaches we are developing for generating information needed by clients in a particular context at a particular time.

Partial basis for this claim of methodological generalizability are the findings of the several studies of district evaluation and testing offices, which suggest that others are coming to similar conclusions. (Alkin, Daillak, & White, 1979; Bank, Williams, & Burry, 1981; Chase, 1980; Hendrickson, 1982; Kennedy, 1982; King & Pechman, 1982; Lyon, Doscher, McGranahan, & Williams, 1978) as well as personal communications with people (e.g., Robert Hammond in Oregon and Thomas Saterfield in Mississippi) who are developing similar approaches in very dissimilar school districts.

Monitoring Indicators and Tailoring Practice

The main generalization emerging from our Pittsburgh work so far is that educational research that takes place within school districts (a research enterprise of considerable promise and which is currently larger than the one NIE sponsors) would profit greatly if the emphasis were shifted from discrete studies of particular programs or policies, which generally fall under the rubric of program evaluation, to a continuous activity of data collection and analysis, which I refer to as monitoring and tailoring.

Reform efforts in education have tended to assume that the best way to improve educational practice is to adopt a new program that seems to address a particular problem, implement that innovative program, and then evaluate the program to determine its effectiveness in dealing with the problem. Berman (1980) calls this the technological-experimental paradigm of educational change.

Our Pittsburgh experience has convinced me that this paradigm is inadequate as a guide for district-based evaluators. Formal, summative program evaluations that attempt to estimate the impact of a particular program or policy on student outcomes tend to produce results that should not be used, because of their invalidity, or cannot be used because valid impact studies, if they can be done at all, take too long to be timely. Worse than that, such studies represent a substantial opportunity cost. That is, they require so much time and effort that other, potentially more useful approaches, are not employed.

An alternative to the experimental paradigm, one which has not been widely used in education but which has considerable promise, is the cybernetic paradigm. It involves developing and monitoring a variety of performance indicators. Then whenever an indicator moves into an unacceptable range, an attempt is made to determine just where that condition is most severe. Focused corrective action is then taken which I call tailoring practice.

Although monitoring and tailoring resembles the cybernetic model used in controlling physical systems (as in the thermostat), there are some very important differences when applying it in an educational system, as Sproull and Zubrow (1981) point out.

In applying the cybernetic paradigm to educational systems, it must be recognized that, compared to physical systems, the available indicators are more fallible, it is usually less clear what an unacceptable range is, and it is not as obvious what the corrective action must be when an indicator moves into an unacceptable range. It certainly sounds hopeless. However, the monitoring and tailoring approach can be designed in ways that take these shortcomings into account.
and can serve a very useful function within an educational system.

Our Districtwide needs assessment is useful to illustrate the main features of a monitoring and tailoring approach. One purpose of monitoring indicators is to help districts establish priorities for improving the system. As I mentioned, we worked up a variety of districtwide data that indicated the state of the educational system in the district.

The indicators included observed variables, such as student attendance, very simply derived variables, such as student/teacher ratio, or more complex, indirectly measured latent variables, such as socioeconomic status. The original unit of observation may have been students, classrooms or schools, and the level of aggregation used depended on the point to be made. What is common to indicators is that they are a function of a construct that describes some aspect of educational phenomena that people care about. More on constructs in a moment.

Extensive dialogue with the Superintendent and Board of Education regarding these indicators led to the establishment of the priority areas for improving the educational program in the District. The dialogue concerned the degree to which the indicators seemed to be in an unacceptable range, as well as which indicators were getting at more fundamental problems which in turn might be affecting the performance of other indicators.

So one function of indicators in a monitoring and tailoring system is to contribute to districtwide priority setting. Examining a variety of indicators (in the form of districtwide aggregates) makes such priority discussions more productive. Since the objective is to improve the performance of the system, district-level aggregates are important for a dialogue about what aspects are in greatest need of improvement and, over time, for indicating whether progress is being made.

The other useful way to examine a performance indicator is to look at how it is distributed. Noticing where unusually low performance is occurring on a priority indicator provides a basis for guiding the action-system that is supposed to improve that performance. The unit of analysis for examining distributions, whether it is students, classrooms, or schools, must be consistent with the unit that is the focus of the action system. Let’s examine this important point a little more carefully.

As part of the districtwide needs assessment, we analyzed all the achievement data from the District that we could find—in terms of 5 year trends, contrasts across grades, and differences among the various subject areas. The results of these analyses suggested that a major problem within achievement was primary grades reading. For example, district level aggregates indicated (on a criterion-referenced basis) that approximately 25 percent of the students were leaving third grade with reading comprehension skills inadequate to deal with the fourth-grade curriculum. If those skills were not learned in the primary grades, subsequent remedial efforts became more and more costly and seemed less and less effective.

Further looks at the data (this time at the classroom level) revealed some second- and third-grade classrooms in which little or no reading growth was occurring in the course of the year for students placed in those teachers’ classrooms. (Note that in this particular example the indicator is not end-of-year achievement level but achievement growth, the units being monitored are classrooms, not individual students or schools, low growth was a trend for that teacher’s students, not a one time event, and initial student abilities were also taken into account in comparing growth.)

Now the question is, what is to be done when the growth in reading ability of students assigned to a particular teacher is discovered to be low year after year? Tailoring requires a deployable resource, an action system, that can respond to such indications. In its response, the action system must recognize that the indicator is fallible. A procedure is needed for confirming the indication. In this particular case, the person that responds to the indicator might be an instructional supervisor, trained in clinical supervision and capable of visiting that classroom, initially to confirm (or disconfirm) the indication that reading instruction is not going well in that classroom, and, if confirmed, to diagnose the situation and then take corrective action, generally employing intensive clinical supervision.

As another example, if district-level aggregates indicate that student absence has moved into an undesirable range, and if the action system available for correcting truancy consists of social workers working with individual students and their families, then the extreme cases in the distribution of student-level truancy rates would be the focus of that action system. If other levels of aggregation revealed a classroom or school with particularly low attendance, further investigation might reveal the need for another kind of action system that can work with particular teachers or schools to create a more attractive learning environment.

One assumption central to the monitoring and tailoring approach is that important, significant improvements can be made in the educational system through fine tuning the system. Some might call it a form of incrementalism (Lindblom, 1972). Recognizing that, I also must emphasize that there indeed may be fundamental changes that must be made in the system in order to adjust to fundamental changes that occur in society. But it does seem rational to make sure that current programs and policies are working as well as possible before trying some dramatic departure from current practice. I am not saying we should not innovate. There comes a time when the saber-
tooth curriculum has to go. But when a problem is detected in a school district, the tendency is to launch a districtwide solution, generally involving a new program, rather than to determine just where (i.e., in which schools or classrooms) things are not working well and to tailor their practice to improve performance. The districtwide innovation can frequently disrupt those schools or classrooms in which things had been working smoothly and seldom corrects situations where they were not. Without focused assistance, some principals or teachers will not implement the new solution any better than the previous practice. Too little has been done to get the programs that are in place working well. (What is impressive in the Follow-Through data, for example, is the vast variability in effectiveness among sites implementing the same instructional model, not differences between models. See, for example, Cooley & Lohnes, 1976.) A monitoring and tailoring system can help a district decide how to focus its available energies for staff development and remedial attention.

Now let us take a little closer look at two of the major components of such a system, the indicators and the action systems.

Major Ingredients of a Monitoring and Tailoring System

Indicators

It helps to think about possible indicators in terms of major constructs. One set of constructs relates to the efficacy of the system as it prepares students for adulthood. Here it is important to have indicators of such constructs as entering student abilities and interests, their educational programs and progress in them, and achievement and other personal outcomes that are indicative of their expected futures. This set represents the familiar systems constructs of inputs, processes and outcomes. Lohnes and I (Cooley & Lohnes, 1976) have suggested one way of organizing such data.

Another set of constructs is needed to describe the quality of the present experience. Schooling is a large part of everyone’s life. It is important to seek indicators of the richness of the present experience as well as to satisfy the clamor for indicators of how well the students are being prepared for adulthood. For example, Epstein’s (1981) measure of the quality of school life is a useful indicator of school climate.

A third set of constructs is necessary to satisfy the need to consider whether the system is fair. Questions concerning the equality of educational opportunity dominate policy discussions, and indicators that can reflect such inequities in the system can make those debates more productive. Disaggregations of these indicators can also reveal where more resources may be needed to correct currently unjust distributions.

Dialogue regarding inequities in an education system is admittedly not a straightforward affair (e.g., Green, 1971). But arguments in terms of the appropriateness of different indicators makes those discussions less abstract. Also, concern soon shifts from noticing the inevitable differences in student outcomes and justifiable differences in resource allocations, to unjust inequities in opportunity to learn.

Having indicators from all three domains (efficacy, quality and equality) also facilitates discussions about “balancing” the different demands on the schools (e.g., quality of the present experience versus preparation for the future; demand for excellence versus the need for equity; liberal education versus training marketable skills.)

For all constructs, it is important to develop multiple indicators and be able to display them at multiple levels (i.e., student, classroom, school and district levels). Because this requires considerable information processing, an essential ingredient is to build a computer-based information system that allows the development and display of the necessary indicators. Today such computer systems are quite feasible. One is being implemented in the Pittsburgh District under Jim Angevine’s leadership with terminals in all 90 school buildings.

One problem in using indicators is that it is frequently possible to corrupt them. That is, indicators are corruptible if it is possible to affect the indicator without affecting the underlying phenomena that the indicator is attempting to reflect. For example, if suspensions are being monitored as one indicator of school climate, and if having many of them reflects poorly on the principal, it’s very easy to see how principals could modify their behavior with respect to issuing suspensions (or reporting them!) and still have the same level of chaos in that building. The corruptibility of indicators is one reason why it is important to have multiple indicators of the same construct and to continuously refine them. This is an important task of the evaluation researcher in such a system. It should be pointed out that indicators will be corrupted more readily if rewards or punishments are associated with extreme values on that indicator, than if the indicator is used for guiding corrective feedback.

A lot can be learned about the construction and use of indicators in education from the research on social indicators that emerged as an active area among social scientists in the mid-1960s. That work was stimulated by two Secretaries of HEW, John Gardner and Wilbur Cohen, who felt that social indicators could provide a better basis for federal social policy (Land & Spilerman, 1975). Krathwohl (1975) organized an ETS Invitational Conference on this theme in 1975. There now are entire books and journals devoted to social indicators research. In education, the book by Johnstone (1981) is particularly relevant.

It must be noted, however, that there has been considerable debate about the utility of the social indicators movement. I sense that it has two major weaknesses as far as education is concerned. One is the emphasis in that literature upon federal level aggregates. Noticing that a national indicator has moved into an alarming range encourages the launching of federal solutions to an isolated problem. That can easily produce programs that end up being counterproductive, because of the counterintuitive ways in which large, complex systems react to change (Meyers, 1981, p. 20). Everyone has their favorite example of an unintended negative side
effect from a well-intentioned innovation. The message in Rachel Carson's "Silent Spring" is the classic example.

The other weakness in the social indicators movement is the tendency to justify indicators as a way of estimating the impact of social programs and policies. To do that, of course, requires rather well specified causal models. Economists have been at this for a long time, and it seems pretty clear they are not there yet. It is important to work on developing models that can describe, even predict, how indicators change over time, how they interrelate, even how they seem to be influencing one another. But adequately specified causal models may be a long way off, even at the district level. Meanwhile, indicators can play an important role in the type of monitoring and tailoring system I am advocating for school districts. Now a few more words about action systems.

Action Systems

If tailoring is to follow from monitoring, the information system must be somehow "connected" to an action system. It is very important how this is done because it is usually not clear why a student, or classroom, or school, is in an undesirable range on some valued indicator. Indicators are a function of many factors in the system. Indicators can only tell you where to look for possible problems. For that reason, the action system that is called into play must be first and foremost a diagnostic system. Corrective action is generally not clear from the indicators because our causal models for explaining their rise and fall are still not adequately specified.

The monitoring and tailoring approach requires the availability of services that can be deployed to correct the most serious cases that are found within the district. At the student level, those who deal with extreme cases on such indicators as attendance, suspensions, and achievement growth would be social workers, counselors, and remedial tutors. Specialists trained in clinical supervision would work with principals and teachers who are low on classroom level growth indicators, and a school improvement team would work in schools that are extreme on building level indicators.

Notice I am not calling for new staff to perform these functions. What the information system does is show how existing personnel might focus their energies. It also helps to justify that focusing. Guided by an unexamined sense of equality, those capable of this type of corrective action tend to be spread so thin they cannot possibly be effective.

So the necessary action systems are financially feasible because they represent a way of focusing the efforts of existing personnel. Of course it is one thing to be feasible, and another to be effective. But even here I am optimistic. One basis for the optimism is the documentation that Bickel and Arzt (1983) are conducting in the context of our school improvement effort. Focusing on seven elementary schools that had a history of low student achievement, the director of that program (Louis Venson) and his school improvement team are demonstrating how monitored student progress data, accompanied by focused, team supervision are capable of improving the quality of instruction going on in those buildings.

Another basis for optimism is all the good work that has been going on in our field the past decade on effective teaching and effective schooling. For example, by carefully examining the behavior of unusually competent teachers, Leinhart (1983) has found that these teachers buy considerable amounts of instructional time through the use of well-rehearsed, easily operationalized routines. The effective schools research has just been summarized in a special issue of Educational Researcher (Bickel, Ed., 1983). Such results contribute to the knowledge base needed for guiding the type of focused, clinical action required in monitoring and tailoring. But our Pittsburgh experience is also making it clear that Purkey and Smith (1982) are correct when they point out that school improvement is certainly possible, but don't expect it to be easy.

Also, I'm optimistic because of the way the indicators in the District are starting to turn around. Of course, the improvement we are seeing in student achievement and attendance, for example, is not necessarily the result of monitoring and tailoring. What is clear is that there is a new energy in the District. Its sources include a Superintendent who is willing and able to provide the kind of educational leadership he expects of others in the District, and a School Board that has recognized the importance of what he is trying to accomplish. What monitoring and tailoring can do is indicate potentially fruitful ways in which that energy can be focused, and provide a basis for new hope as indicators begin to improve.

"The Systems Approach and Its Enemies"

There is a question that very properly arises as one begins to build a computer-based monitoring system: Is it humane? Part of my answer to that question is a series of other questions. For example:

—Is it humane to expose children year after year to a teacher who is apparently unable to improve the reading performance of children in his or her charge?

—Is it humane to have a child absent 160 days in a school year without so much as a home visit because the social worker's energies are being diverted by a principal to the task of preparing the justification for a long-term suspension?

—Is it humane for a school district to solve the problems of an unsatisfactory principal's performance by an annual game of musical chairs among principals in the hopes that eventually a principal will end up in a school in which the parents don't complain?

The other part of my answer is to note that the monitoring and tailoring system I am proposing is designed in a way that recognizes the fallibility of the indicators and the inadequacies of our causal models in determining why an indicator has moved into an undesirable range. It is not designed, for example, to issue pink slips to teachers whose students turn in an unsatisfactory standardized achievement performance.

We did not set out to develop a monitoring and tailoring system for

Educational Researcher
the Pittsburgh Public Schools. We set out to provide evaluation services for the District. We began by doing program evaluation studies, trying to find the impact of particular programs. We moved from that to a districtwide needs assessment. It was in the context of that activity that we became impressed with the vast variance in any indicator we looked at—variance among students, among classrooms, among schools. Measures of central tendency for the district can be a cause for alarm, but noticing how the indicator distributes and where the outliers are can be the first step in an effective improvement strategy.

Superintendents too often become distracted with something that might be called “mess management.” What sometimes happens is that those responsible for managing a large school system see problems all over the place, and it all seems so unmanageable. The tendency then is to run from one problem to another. The most serious error is what Churchman (1979) calls the “environmental fallacy,” which is not taking into account the context in which a problem occurs.

Let me illustrate the environmental fallacy with some earlier history from our Pittsburgh experience. Several years ago, the School Board passed a resolution to abolish social promotions. People had become unhappy with illiterate high school graduates. Subsequently, people began to notice that a lot of students were failing. Staff members at the Board then designed a special program for students who had failed a grade. The failures were to be removed from the regular classroom for a year and placed in a class called Pass. Limited to 14 students, Pass classrooms were “self-contained,” with a teacher and an aide in each classroom. The program was designed as an intensive, individualized program of remedial instruction in the basic skills. The good intentions of the designers of the program were that the failed students would be taken off the track for a year and then reinserted with their original cohorts if they passed Pass. (Notice that “passing” would have required these students to master 2 years work in 1 year.)

We were asked to evaluate this innovative program. The details of our evaluation are one of the case histories that we are preparing; so I simply want to point out here that we uncovered a major problem with the Pass program as designed. It did not directly address truancy—one the main reasons children were failing.

About the same time the District passed the no-social-promotion policy, the District had negotiated an agreement with the teachers union that allowed teachers to fail students who were absent from class more than 9 days in a quarter. When we dug into individual attendance records for the Pass participants, we found that over one-fourth of the students were absent more than one-third of the time. Many of the students were absent more than 100 of the 180-day school year. Meanwhile, Project Pass was not directly connected to the “action system” that was supposed to deal with truancy. Many of the serious cases apparently never even had a home visit (Carter, 1982). The procedures for dealing with truancy in the District are to assign social workers to buildings. A relatively informal system within each building is used to bring attendance problems to the attention of the social worker.

We also learned how unconnected the remedial programs were to the mainstream program, and how extensive those remedial efforts had become in the school. So prior to Pass the schools already had massive efforts to attempt to help students who were not keeping up with their classmates. At some grade levels, over half the children were in one or more such programs. These programs were implemented in ways that operated independently of the child’s regular classroom experience. This rather typical phenomenon in urban districts was in part a result of the way the Federal Title I guidelines were written.

That little history reminds us that a school district is a complex educational system. The performance of particular subsystems cannot be improved in isolation. Board politics and policy, union contracts, bureaucratic turf, state code, federal guidelines, and court actions, all influence the behavior of the system. So do the value orientations of the thousands of participants in the system. As Green (1980) points out so well, the behavior of the system as a whole usually frustrates the intent of particular changes in policy.

Systems theory provides a framework for designing ways to improve a system. Its cybernetic model is one such approach. The environmental fallacy is its great enemy. You don’t improve systems by solving problems in isolation. You can improve systems by monitoring indicators and tailoring practices.

Many have concluded that our system of public education is beyond salvage, that a new system must be designed to replace it. My own view is that a greater chance for success lies in fine-tuning the system we have. In the United States there are 40 million students being served in 85,000 public schools by 2 million teachers. Most students are being served well but many are not. What I’m recommending are procedures for locating just where services are not being delivered effectively, how they are not, and what can be done about it.

I hesitate to use national statistics to describe the educational system. Although it helps to remember the magnitude of the enterprise that has evolved over the centuries, and which has served this country well, national statistics sometimes encourage very specific federal programs. My view is that big improvements can be made in education by monitoring and tailoring if it is applied within a more reasonably bounded system. The system that can be improved in this way is the school district.

**Feedforward**

I think it is important to make one final observation about monitoring and tailoring. A fundamental feature of a cybernetic model is the notion of feedback. The monitored indicators provide information that guides priority setting and provides a focus for the available action system. At the height of the feedback craze in the 1960s, I. A. Richards
(1968) wrote a thoughtful essay in the *Saturday Review*, about “feedback.” He pointed out the importance of setting goals for the system in the first place. The design of monitoring and tailoring systems and the development of indicators must be guided by a coherent set of ideas about the goals of education, what forms it should take, and what conditions should be considered intolerable. As Churchman (1979) points out so well, system approaches can not avoid the deeper problems of values. Implementation of a monitoring and tailoring approach within a school district can force a dialogue about fundamental goals of the system. A well-articulated set of goals is required to provide feedback for the system. They can and must emerge from that dialogue.

**Footnotes**

'I am indebted to Churchman (1979) for this subtitle.

“Ackoff (1973) defines a mess as ‘a system of external conditions that produces dissatisfaction. A mess can be conceptualized as a system of problems. Problems, even as abstract conceptualization, do not exist in isolation: they are elements of a system.’

**References**


Chase, F. Educational quandaries opportunities. Dallas, Tex.: Urban Education Studies, 1980.


Green, T. F. Predicting the behavior of the educational system. New York: Syracuse University Press, 1980.

Hendrickson, L. Personal communication, December 1, 1982.


