

SCHOOL DISTRICT EVALUATION EFFORTS:
CONTRADICTIONS AND IMPLICATIONS

David W. O'Shea

CSE Report No. 141

June 1980

Evaluation Design Project
Center for the Study of Evaluation
Graduate School of Education, UCLA
Los Angeles, California 90024

Table of Contents

	<u>Page</u>
Introduction.....	1
Origin of Evaluation Requirements.....	4
National Evaluations.....	7
Conflicting Logics.....	8
..... Technical Versus Institutional Organizations.....	11
Program Evaluations.....	13
Local Use of Standardized Tests.....	15
Alternative Approaches.....	16
Qualitative Evaluation.....	17
Conclusion.....	18
Footnotes.....	20
References.....	22

INTRODUCTION

Widespread emergence of school district evaluation units in recent years reflects a local response to the Elementary and Secondary Education Act (ESEA), and to similar legislation in individual states, requiring participating states to secure, from each school district, data on student academic achievement by selected grades. The legislation also requires each district receiving federal funds

"... to assess the effect of its program and to identify weaknesses as well as strengths of the project, thus serving as a tool for program revision and improvement (USCAN, 1974, p. 411, cited by David, 1978: 4).

Eighty-five percent of school district evaluation units, which now exist in 43 percent of the nation's 320 districts enrolling 10,000 pupils or more, have been established since 1965, the first year of ESEA (see Lyon, et al, 1978). Curiously, however, while these evaluation units expend great effort on gathering achievement data, little attention is given to the analysis of these data with a view to exploring the "weaknesses as well as strengths of the project," as legislators hoped. In practice, district evaluation units are more likely to engage in achievement monitoring rather than in analytic evaluation (see Kerins, 1973: 76; Ebel, 1980: 281-292). This outcome provokes two questions, providing the focus for this paper. First, what inhibits analytic evaluation within public school districts? Second, of what utility is achievement monitoring as typically practiced?

In principle, achievement monitoring is an appropriate alternative to formal analytic evaluation. While the latter is directed toward the

identification of what works, and at what cost, monitoring focuses upon whether a program works, in the sense of achieving its intended purpose.

Ubiquity of the monitoring approach reflects an adaptive response on the part of school district officials to the reality of schools and their instructional processes. The analytic approach to evaluation is designed for use in the context of technically rational systems whereas schools, and their instructional processes, operate in accord with practical rationality.

However, in adopting achievement monitoring as a way of resolving one problem, evaluation units are likely to create a second. Commonly, norm-referenced tests are used for monitoring. Such tests are not always related to the actual instructional content of the program being monitored.

In exploring factors inhibiting formal evaluation at the school district level, and limiting the utility of conventional achievement monitoring, we begin with a brief review of the actual origin of ESEA's evaluation requirements and the attempts to implement national evaluation of Title I programs, the largest component within ESEA. This review indicates that a major factor inhibiting evaluation is simply the contradiction between the assumptions underlying prevailing approaches to evaluation and the actual logic of school-based instruction.

The problem of the logic of instruction centers in the fact that schools, Meyet (1977) argues, are not technically rational, but institutional organizations. While technical organizations conform to the imperatives of their technology, institutional organizations, which do not originate from a developed technology, conform to the demands of their social environment. Before pursuing the implications for evaluation of the

institutional rather than the technological character of schools, it is helpful to consider the origin of ESEA's evaluation requirement, and the problems confronting implementation of this requirement.

ORIGIN OF EVALUATION REQUIREMENTS

Prior to passage of ESEA in 1965, most evaluations within local school districts focused upon students, not upon instruction. The instructional program was taken as given, leaving in question how well individual students were performing in their effort to master the program content. Beginning in the 1950s, the civil rights movement drew public attention to the fact that not only did individual students differ in their academic achievement, but whole groups of students were performing at lower levels than other groups. Specifically, by the 1960s the phenomenon of the "achievement gap" was well documented (see especially Coleman, et al., 1966). Students graduating from high schools in low income, predominantly minority communities, have average levels of achievement three grades below the average for students from schools serving the majority white population.

Until the findings of Coleman, et al (1966), emphasizing the dominance of out-of-school factors upon in-school achievement, civil rights leaders attributed the achievement gap to factors such as the discriminatory allocation of financial resources between schools and the absence of staff accountability to parents, or both. Congress responded with legislation addressing both possibilities. ESEA makes available over three billion dollars annually for local schools. Title I of the Act, oriented to poverty area programs, receives the largest proportion of the funds allocated. The Act also mandates state-level collection of student achievement data for each school, and annual national evaluations of the Title I program.

Requirements regarding evaluation were written into the Act in response to demands from Senator Robert Kennedy, then representing New York (see Baily & Mosher, 1968; McLaughlin, 1975). Robert Kennedy supported ESEA on condition that the legislation included a reporting requirement and "good faith administration efforts to hold educators responsive to their constituencies and to make educational achievement the touchstone of success in judging ESEA" (McLaughlin, 1975: 3). In Kennedy's opinion, if achievement data were available, parents would be in a position to hold educators accountable for learning outcomes, thus creating an essentially political pressure for program improvement.

Kennedy's insistence upon the collection of achievement data was welcomed by officials in the Department of Health, Education, and Welfare (HEW), though not because they shared his interest in facilitating parental pressure upon the schools. It so happened that coincident with the passage of ESEA, President Lyndon Johnson, impressed by Secretary of Defense McNamara's techniques for controlling Pentagon budgets, announced that the Defense Department's system of fiscal management, the Planning, Programming, Budgeting System (PPBS), was to be installed throughout all sections of the executive branch (McLaughlin, 1975: 6).

To implement PPBS in HEW, the department with overall responsibility for ESEA, William Gorham was brought in as Assistant Secretary for Program Evaluation (ASPE). From the outset the ASPE staff, unlike Senator Kennedy, took an analytic approach to evaluation. As Gorham (1967) wrote later (see McLaughlin, 1975: 7):

"Title I of the Elementary and Secondary Education Act... is essentially a vast experiment designed to find effective ways of reaching disadvantaged children. We know that most school systems have not been doing a very good job with these children, but there is no consensus among educators about how to do better."

Obviously, Gorham and Kennedy had quite different approaches to program assessment in mind.

"Kennedy thought of evaluation as a means to ensure that schools focused on the needs of the poor, to make sure that Title I 'worked.' Gorham's ASPE staff thought of evaluation as a way to find out 'what works,' to identify the most effective compensatory strategies." (McLaughlin, 1975: 7.)

Kennedy's approach really involves achievement monitoring, rather than program evaluation. Evaluation is analytic, involving what Thompson (1967: 86) calls "efficiency tests," focusing upon the worth, or effectiveness, of program processes in relation to program outcomes. Achievement monitoring, as noted earlier, when linked to what Ebel (1980: 288) calls "judicious deliberation" on the part of program staff, actually may be better adapted than evaluation to serving the task of program improvement, for the reason that it

"... avoids the implication that the program is the principal determiner of achievement regardless of pupil effort, teacher skills, or the general educational context. It avoids the almost impossible task of abstracting the program from its context and studying it in isolation." (Ebel, 1980: 290.)

Inhibiting the development of evaluation, as opposed to monitoring, at the district level are contradictions, implicit in the above comments by Ebel, between the assumptions underlying analytic evaluation and the nature of instructional processes in school settings. These contradictory assumptions become more apparent if one looks at past attempts at national evaluation of Title I programs.

NATIONAL EVALUATIONS

In enacting ESEA, Congress called for both the development of school level achievement data and an annual evaluation of the Title I program at the national level, without specifying the form that such an evaluation might take. However, as noted earlier, the approach to evaluation has been influenced by the government's desire to implement PPBS as the system for management of all federal programs.¹ Such an objective serves to focus evaluation upon cost-benefit analyses, orienting evaluation studies toward identification of the most cost-effective, or efficient, means of improving student learning. To facilitate cost-benefit studies it is necessary to discover which program elements actually are effective in promoting learning gains in classrooms. For this purpose the typical, though not exclusive, approach has been to base evaluation upon experimental design, comparing the progress of students in funded programs with the progress of others in selected control groups. As Apsler (1977: 14) confirms, the experimental paradigm, so fruitful in the natural sciences, is "advocated by nearly every methodological discussion of evaluation research (See, for example, Campbell, 1975; Houston, 1972; Suchman, 1967; Riecken & Boruch, 1974)."

In practice, this paradigm has proven extremely difficult to apply to national evaluations of Title I programs, a factor inhibiting downward extension of evaluation activity to the local district level. Nationally, an early attempt was the study of compensatory education conducted for the ASPE section of HEW by the TEMPO division of General Electric. Data were obtained from schools in eleven districts considered to have "exemplary" Title I programs. Hopefully, analysis of these data would determine whether

"different amounts of money spent in different ways would be significantly and differentially successful in effecting academic achievement" (McLaughlin, 1975: 37).

Unfortunately, the TEMPO study foundered. The main problem was that identified by Ebel's (1980: 290) observation cited earlier, that in the case of schools it is almost impossible to abstract a program from its context and study it in isolation. As McLaughlin (1975:37) concludes:

TEMPO analysts were able to identify neither a Title I population, nor Title I program, nor significant achievement gains that could be attributed to Title I funds. The incomplete, confusing, or non-existent records at the local level made it difficult for TEMPO analysts to determine which services were being purchased with Title I dollars, or to specify the group receiving special attention through Title I. In practice, there seemed to be no real Title I program to evaluate.²

Basically, confronting the staff of TEMPO, and of subsequent Title I evaluations, has been the fact that the logic of evaluation using the experimental paradigm is at odds with the logic underlying schools.³ This same conflict of logic also frustrates attempts at the school district level to develop formal program evaluation.

CONFLICTING LOGICS

Insight into the basis for these conflicting logics is provided by Aspler's (1977: 15) comments:

"Although the needs of action programs might be very specific, such as to raise children's test scores or to reduce the incidence of crime, the procedures by which these ends will be attained are generally quite unspecific. Unfortunately, little is known yet about how social action programs can raise children's test scores, reduce crime, and so on. As a result, social action programs often become operational without clear approaches or means of attaining their goals.

Not surprisingly, the actual operations of a social action program may then be primarily determined by intuitions and exigencies. The confusing situation that develops - in which approaches are poorly thought through and in which approaches frequently vary in response to recurrent crises and changing intuitions - present enormous difficulties to the evaluators."

To the evaluator, one might add, who operates within the constraints of the experimental or even the quasi-experimental, paradigm. Evaluation according to the experimental paradigm, with its assumptions about treatments causally related to outcomes, follows the logic of instrumental, or technical, rationality. Schools cannot fit this mode, lacking a theory that can specify which instructional interventions are likely to produce specific learning outcomes. Were such a theory available it would guide specification of the interventions necessary to reduce the substantial gap in average levels of academic achievement between schools serving low and middle-income communities. As Averch et al. (1972) determined, educational research has not yet identified "what works." Instruction remains more an intuitive and practical art than a science, following the normatively grounded logic of practical rationality, like the family or the community, rather than the factory.

However, as Bernstein (1975: 64) reminds us, in our technological society we are dominated by the values of an instrumental culture that is grounded in theoretical knowledge, and committed to technical rationality. In this context, it is not surprising that most evaluators subscribe to the notion that all institutions should operate in a technically rational mode, and be evaluated in terms of the experimental paradigm. From this technocratic perspective, schools appear somewhat diffuse. If evaluators remain

unaware of the basis of this diffuseness, with its accommodation of multiple objectives, there is a danger, as House (1978: 392) has noted, that while evaluators are called in to help make educational programs effective, they could end up trying to impose instrumentally, or technically, rational programs on the schools in order to make evaluation effective.

Goodlad (1975) argues, as does House (1978), that once you begin thinking about schools in the framework of the experimental paradigm, with its dependent and independent variables, there is a danger of absorbing the reality of schooling into the model of a technically rational system epitomized by the modern factory.⁴ Factories embody technology based upon tested knowledge of means-ends relationships. In designing an automobile plant, for example, planners draw upon a body of knowledge that is sufficiently developed to serve as a basis for specifying in advance all the functions that must be performed in order to transform an array of raw materials into a finished automobile. Systems embodying technical rationality are greatly extending man's domination over nature, facilitating the exploration of space, and providing the material basis for contemporary industrial civilization.

Schools, however, are not of the same organizational genre as factories and other technically rational systems. Given the enormous prestige of such systems there is, of course, continuous pressure for schools to move toward the technically rational mode. Inhibiting such movement in the field of schooling is the absence of the ingredient most necessary to technically rational systems: tested knowledge of the relationships between instructional means and educational ends. (See Averch et al., 1972). In effect, as

Dreeben (1970) reminds us, we lack a technology of instruction, a condition leading Metz (1978: 20) to conclude:

Technology, then, is a major problem for the public schools. They are faced with the task of creating changes in diverse raw material through processes that are poorly understood, in the absence of any universally effective means, and without any trustworthy way of measuring the success or failure of whatever methods they finally apply.

Schools' weak technology, coupled with their highly variable material, in the form of students, constrains a decentralized type of organizational structure. Coordination is achieved not by staff implementing a specific technology, but by the collective enactment of a moral order, or institutionalized social system. (See Metz, 1978: 30.) As Meyer (1977) argues, schools are institutional rather than technical organizations. Schools, as organizations, articulate a particular set of rules derived from societal ideals regarding the way in which emerging generations should be inducted into adult society. Evaluation of institutional organizations and their component programs calls for a different approach than that which is appropriate for technical organizations.

TECHNICAL VERSUS INSTITUTIONAL ORGANIZATIONS

The major determinant of organizational type, in Meyer's (1977) opinion, is the nature of the social environment within which the organization evolves. Technical organizations, such as factories, evolve in environments with complex technologies. By contrast, institutional organizations, such as schools, evolve in environments with very limited technologies but with elaborated institutional rules. The task of a technical organization is to coordinate and control technical work. The task of an institutional organization is to

create structures that conform with institutional rules generated over time within society to order the maintenance of the social system.

Technical organizations, in their pure form, are cybernetic systems, oriented to goal attainment, and regulated by feedback from their own results (see Habermas, 1970: 87-93). By contrast, institutional organizations are social systems oriented to the implementation of their constitutive norms or institutionalized rules.

In evaluating technical organizations, priority is given to assessing the cost-effectiveness of alternative means to achieve specific outcomes. When institutional organizations are being evaluated, priorities shift. The dominant concern becomes the degree to which organization structure and processes conform to the values to which the organization is expected to give expression. The effectiveness of a school in promoting the academic achievement of its students plays a minor role in school accreditation. The priorities are facilities, staff qualification, teacher-pupil ratios, and the scope and content of the educational program.

Not only do technical and institutional organizations differ in terms of the premises upon which they are structured, but their structural characteristics take on radically different forms. Knowledge of means-ends relationships allows the administrators of technical organizations to direct all operations from the top down. In the case of institutional organizations, the relative absence of technical knowledge constrains administrators to leave operations to the discretion of staff at the operational, or technical, level. One consequence in the case of schools, as Metz (1978: 21) concludes, is that:

The persons who perform the actual work of the organization need to be given relatively large and diffuse tasks with the right to make important decisions independently as they use their intuition to adjust their methods to the requirements of each specific instance.

While technical organizations are highly centralized, with the administrative and technical or production levels tightly coupled, schools as institutional organizations, are decentralized, with administration and the technical or teaching level loosely coupled, at least with regard to the formal work of the schools, the instruction of students. (See Weick, 1976; Meyer, 1977.)⁵ School principals, for example, though including instructional leadership among their responsibilities, seldom visit classrooms or involve themselves in the details of instructional processes. The latter are delegated almost entirely to teachers.

In summary, therefore, schools as institutional organizations combine a relative absence of technical rationality with loosely coupled relationships between administrative and instructional levels, two conditions that constrain the nature of school and program evaluation.

PROGRAM EVALUATION

Given that institutional and technical organizations each operate on the basis of different premises, approaches to evaluation that are appropriate for the one type prove inappropriate for the other. In exploring the theory of organizations, Thompson (1967: 83-101) points out that the appropriate mode of organizational assessment is determined by two variable factors: (1) the degree to which there is clarity regarding the effects, or outcomes, desired; (2) the degree of knowledge available regarding the means that actually can produce the desired outcome.

The same principles apply to program evaluation. Where it is clear just what outcome is desired, and there is full knowledge of the means needed to attain this outcome, administrators can both prescribe and control the relevant means. In this context, evaluation can focus upon the cost-effectiveness, or efficiency, of the program.

Where knowledge of means/ends, or cause/effect, relations are incomplete, as with schools and their programs,

"...the efficiency test is inappropriate, for there is no way of assessing the net effect of causal action. In this case, the appropriate test is not the economic one, but the instrumental one - whether a desired state of affairs is achieved." (Thompson, 1967: 86.)

In the case of schools, knowledge of the causes of student academic achievement are notably incomplete, a condition that leads Ebel (1980: 288) to conclude:

"In seeking to evaluate an educational program, we may be led astray if we try to follow too closely the model of the scientific research study. An educational program is not a stable, natural phenomenon with built-in operating characteristics that people may discover and put to use, but not alter in any fundamental way. An educational program is a human artifact, highly complex, infinitely variable, and subject to incessant change. It defies precise definition or accurate measurement. The outcome of a precisely controlled scientific study of an educational program is almost certain to be either inconclusive or misleading."

Logically, therefore, from Thompson's (1967) perspective, evaluation of educational programs should be conducted in the instrumental mode. In effect, this corresponds to Ebel's notion of achievement monitoring. Rather than struggle with the methodological hazards associated with attempts to identify "what works," the focus can shift to whether a program works.

Given that districts already gather data on student achievement, both for Title I reports and for statewide assessment, the adoption of local-level achievement monitoring would seem relatively nonproblematic. Unfortunately, this is not so, as revealed by Jane David's (1978) survey of local uses of Title I evaluations in 30 "exemplary" districts across six states.

LOCAL USE OF STANDARDIZED TESTS

Ironically, David (1978) found that few local district evaluation units view the collection of standardized achievement test data as being relevant to the evaluation of local school programs. Rather, such data are viewed as serving informational needs of the state and federal levels of school governance. Locally, the rationale for using standardized tests is that they are the simplest means of responding to state and federal reporting requirements. (David, 1978: v.)

Of course, when these tests reveal student gains, school officials view them as providing confirmatory evidence of program effectiveness. Negative results, by contrast, are ignored. Behind this pattern of response lies the fact that school staff, and parents, generally are happy with their programs, an attitude that positive results tend to reinforce, "but negative results are ignored and, if necessary, explained away as inappropriate." (David, 1978: 22.)

Inappropriateness of standardized tests, David found, was likely to be based on one or other of four arguments: (1) The tests are biased; (2) the tests were administered under bad conditions; (3) test scores are

not meaningful by themselves; they do not take account of student IQ and family status; (4) the tests do not measure what is actually being taught in the local schools. (See David, 1978: 22.)

The low credibility of standardized tests at the local level inhibits their utility for achievement monitoring or instrumental evaluation. However, there is no reason why schooling outcomes cannot be measured by alternative types of tests, thus facilitating achievement monitoring.

ALTERNATIVE APPROACHES

Given that little authentic evaluation activity actually takes place in local school districts, David (1978) asked respondents in the 30 districts she studied how they would demonstrate the success of their programs. Preferred approaches, for both staff and parents, emphasized utilization of criterion-referenced, or "curriculum-embedded and other skill tests." (David, 1978: vi.) Such tests incorporate items projected forward from the actual curriculum content. Test results thus bear directly upon the extent to which local instructional objectives are being attained, allowing teachers, the ultimate instructional decision-makers, to determine which curriculum components need more, or less, emphasis.

Devolution of responsibility for testing to the teachers at the instructional level of the school certainly respects the logic of schools as institutional organizations. To make this alternative succeed, Ebel argues, at least three things are needed:

"First, periodic evaluations of educational accomplishments need to be required by some authority external to the local school. Second, the evaluation reports submitted by the local school need to be audited periodically. Third, local

school personnel need to receive training in evaluation procedures to the point where they can do the job as adequately as an external evaluator can be expected to do it. Practical evaluation techniques are not so highly technical or so remote from typical school operations as to require a separate profession of evaluation. People competent to design programs and to teach should also be competent to assess the results of teaching." (Ebel, 1980: 284.)

Achievement monitoring, based upon what is actually taught combined, Ebel (1980: 288) suggests, with "judicious deliberation" by professional staff appears to be the most credible alternative to the utilization of norm-referenced tests for monitoring purposes. Certainly this would be true in the case of attempts to conduct evaluation relevant to program improvement at the local level.

A limitation of this approach, of course, is that the focus is upon what is taught, not how it is taught. To pursue understanding of how required adoption of qualitative rather than quantitative studies, drawing upon the hermeneutic, rather than the experimental, tradition of research design.

QUALITATIVE EVALUATION

The alternative to what Patton (1978: 204) calls "the dominant hypothetic-deductive paradigm" is being applied to evaluation by persons such as Parlett and Hamilton (1976), Robert Stake (1975), and Kenneth Strike (1972). While "the natural science paradigm aims at prediction of social phenomena, the holistic-inductive, anthropological paradigm aims at understanding of social phenomena" (Patton, 1978: 204). This paradigm draws upon the methods and perspectives of phenomenology, symbolic interactionism, ethnomethodology and anthropology.

Specific objectives associated with qualitative evaluation are outlined by Parlett and Hamilton (1976: 144) as cited by Patton (1978: 209). These objectives, in relation to innovative programs such as those funded in schools under Title I of ESEA, are likely to include:

... how it operates; how it is influenced by the various school situations in which it is applied; what those directly concerned regard as its advantages and disadvantages; and how students' intellectual tasks and academic experiences are most affected. It aims to discover and document what it is like to be participating in the scheme, whether as teacher or pupil, and, in addition, to discern and discuss the innovations' most significant features, recurring concomitants, and critical processes. In short, it seeks to address and to illuminate a complex array of questions.

In practice, implementation of such "illuminative" evaluation requires specialized personnel, thus restricting its use to evaluations sponsored by state or federal agencies, rather than local districts, though the latter may well opt to encourage universities in their area to explore this qualitative approach.

CONCLUSION

In summary, school evaluation is presently at a critical juncture. Contradictions that become increasingly apparent with the passage of time are creating growing pressures for change. Among the more obvious contradictions are:

- (1) conflict between the promise and the performance of national evaluation, typically based upon the experimental research paradigm;
- (2) attempts to use the experimental research paradigm for evaluating instructional programs in schools, despite the fact that schools are not technically rational but institutional organizations;

- (3) utilization of standardized norm-referenced tests for reporting the outcomes of externally funded programs, despite the low credibility of these tests as a means of local program assessment.

To date, response to the above contradictions takes a different, though related, form at the local and national levels. Locally, school district evaluation units tend to remove themselves from the cross-pressures inherent in the several contradictions by opting out of any sustained commitment to program evaluation, focusing instead upon monitoring program outcomes in the form of student achievement, administering tests for funded program participants, and reporting test results to funding agencies. The utility of such monitoring could be enhanced greatly by linking it directly to the actual content of instructional programs in each school.

At the national level, a combination of high costs and null findings is leading to pressure for an end to national evaluations of Title I. It is to be hoped that such pressure can be re-directed toward substituting more appropriate modes of assessment, emphasizing qualitative approaches that advance understanding of instructional processes.

The danger to avoid is that identified by House (1978), and discussed in more general terms by Habermas (1971). This danger is that rather than adopt evaluation approaches appropriate to the nature of schools and school processes, governmental pressure will be exerted in the direction of making schools into technically rational organizations, in the belief that the null findings reflect the inadequacies of schools, rather than the inappropriateness of current methods of evaluation.

FOOTNOTES

¹The objective at the federal level to rationalize program administration did not terminate with changes in the White House. Though the move to adopt PPBS began under Johnson, officials in OE were arguing for the same objective as recently at 1977, as evidenced by the following comments of John W. Evans and Janice K. Anderson (Evaluation, 1977: 160), senior officials in the Office of Planning, Budgeting, and Evaluation at the U.S. Office of Education.

Since the late sixties, Congress has become increasingly disillusioned with the policy relevance of most general or basic research produced through the open-ended grant process, and useful research/evaluation results. Mission-based federal agencies have responded by moving more in the applied direction, and in the process have assembled technically trained evaluation staffs to design and closely monitor the conduct of evaluations and other analyses... They are part of a larger trend, with other elements being the emphasis on accountability, management by objectives, program planning and budgeting, and the creation of the new budget committee's procedures and ceilings within the Congress. The basic thrust of this larger trend is a belated effort in the social program area to rationalize federal policies, and the allocation of scarce resources.

²As McLaughlin (1975: 40) emphasizes:

A central requirement of an impact, cost-benefit study - the ability to tie inputs to outputs - does not coincide with the operational reality of Title I. An attempt to trace the flow of Title I dollars to specific programs and outcomes is beset with problems. In what Michael Krist calls the "Byzantine world of school accounting," it is difficult if not impossible to trace the course of Title I dollars through the school system. Some cities, especially larger cities, have over 100 sources of income. As the number of revenue sources increases, the ability of evaluators to identify the impact of any single source diminishes.

³In discussing McLaughlin's (1975) study of ESEA, House (1978) draws conclusions similar to those offered in this paper. Commenting upon the approach to Title I evaluation, House (1978: 388) concludes that:

The evaluation policy developed by the federal government reflected a particular ideology. It reflected the belief of systems analysts and economists that evaluation should be used primarily to detect the most efficient programs. This concern for efficiency led to evaluations that could find no differences between the new programs and those already existing in the public schools.

The result has been that educational funds have been constrained at the federal level for lack of visible results and education has been discredited for its inability. Actually the lack of results reflects more the type of evaluation employed than the quality of the educational programs.

⁴One then becomes entrapped in the phenomenon Wise (1977: 44) labels "hyper-rationality." In essence, hyper-rationality represents an assumption that the logic of technical rationality is at work when, in fact, it is not.

⁵Others concerned with the nature of schools as organizations also have noted their loosely coupled structure. See, for example, Bidwell (1968); Dreeben (1970); Metz (1978); Weick (1976); Wolcott (1977).

REFERENCES

- Apsler, R. "In Defense of the Experimental Paradigm as a Tool for Evaluation Research." Evaluation, 4:14-18, 1977.
- Averch, H.A., et al. How effective is Schooling? A Critical Review and Synthesis of Research Findings. Santa Monica, CA: Rand Corp., 1972.
- Bailey, S.K. & Mosher, E.K. ESEA: The Office of Education Administrators a Law. Syracuse: Syracuse University Press, 1968.
- Bernstein, B. Class, Codes and Control, Volume 3. Towards a Theory of Educational Transmissions. Boston, Mass.: Routledge & Kegan Paul, 1975.
- Bidwell, C. "The School as a Formal Organization." In James F. March (Ed.) The Handbook of Organization. Chicago: Rand McNally, 1965.
- Campbell, D.T. "Reforms as Experiments" in E.L. Struening and M. Guttentag (Eds.) Handbook of Evaluation Research: Volume 1. Beverly Hills: Sage Publications, 1975.
- Coleman, J., et al. Equality of Educational Opportunity. U.S. Department of Health, Education, and Welfare. Washington, D.C., U.S. Government Printing Office, 1966.
- David, J.L. "Local Uses of Title I Evaluations." Research Report EPRC 21. Educational Policy Research Center. Stanford Research Institute, 1978.
- Dreeben, R. The Nature of Teaching. Glenview, Ill.: Scott, Foresman, 1970.
- Ebel, R.L. Practical Problems in Educational Measurement. Lexington, Mass.: D.C. Heath and Company, 1980.
- Evans, J.W. & Anderson, J.K. "Comments on an RFP for Evaluation of Title I Programs: Will OE Learn What it Wants to Know?" Evaluation, 4: 157-60, 1970.
- Goodlad, J.I. The Dynamics of Educational Change. New York: McGraw-Hill, 1975.
- Gorham, W. "Notes of a Practitioner." The Public Interest, Summer, 4-8, 1967.
- Habermas, J. Toward a Rational Society. Boston: Beacon Press, 1970.
- Habermas, J. Knowledge and Human Interest. Boston: Beacon Press, 1971.

- House, E.R. "Evaluation as Scientific Management in U.S. School Reform." Comparative Education Review, 22:3, 388-401, 1978.
- Houston, Jr., T.R. "The Behavioral Sciences Impact-Effectiveness Model." In P.H. Rossi and W. Williams (Eds.) Evaluating Social Programs: Theory Practice, and Politics. New York: Seminar Press, 1972.
- Kerins, C.T. "Accountability and the State." Chapter 7 in Ernest R. House (Ed.) School Evaluation: The Politics and Process. Berkeley, CA: McCutchan Publishing Corp., 1973, pp. 74-79.
- Lyon, C.D., Doscher, L., McGranahan, P., and Williams, R. Evaluation and School Districts. Los Angeles; Center for the Study of Evaluation, UCLA, 1978.
- McLaughlin, M.W. Evaluation and Reform: The Elementary and Secondary Education Act of 1965/Title I. Cambridge, Mass.: Ballinger Publishing Company, 1975.
- Metz, M.H. Classrooms and Corridors. Berkeley, CA: University of California Press, 1978.
- Meyer, J.W. "Research on School and District Organization." Paper presented at annual Sociology of Education Association Conference, San Diego, 1977.
- Parlett, M. & Hamilton, D. "Evaluation as Illumination: A New Approach to the Study of Innovative Programs." In Gene V Glass (Ed.) Evaluation Studies: Review Annual, Vol. I. Beverly Hills, CA: Sage Publications, 1976.
- Patton, M.Q. Utilization-Focused Evaluation. Beverly Hills, Sage Publications, 1978.
- Riecken, H.W. & Boruch, R.F. (Eds.) Social Experimentation: A Method for Planning and Evaluating Social Intervention. New York: Academic Press, 1974.
- Sarason, S.B. The Culture of the School and the Problem of Change. Boston: Allyn & Bacon, 1971.
- Stake, R.W. Evaluating the Arts in Education: A Responsive Approach. Columbus, Ohio: Charles E. Merrill, 1975.
- Strike, K. "Explaining and Understanding: The Impact of Science on our Concept of Man." In Lawrence J. Thomas (Ed.) Philosophical Redirection of Educational Research: The Seventy-First Yearbook of the National Study for the Study of Education. Chicago: University of Chicago Press, 1972.

- Suchman, E.A. Evaluative Research: Principles and Practice in Public Service and Social Action Programs. New York: Russell Sage Foundation, 1967.
- Thompson, J.D. — Organizations in Action. New York: McGraw Hill, 1967.
- USCAN, "Legislative History of P.L. 92-380: Education Amendments of 1974," U.S. Congressional and Administrative News (1974).
- Weick, K.E. "Educational Organizations as Loosely Coupled Systems." Administrative Science Quarterly, 21: 1-19, 1976.
- Wise, A.E. "Why Educational Policies Often Fail: The Hyperrationalization Hypothesis." Curriculum Studies, 9:1, 43-57, 1977.
- Wolcott, H. Teachers versus Technocrats. Eugene, Oregon: University of Oregon, CASEA, 1977.