ABSTRACT

A thorough search of the literature on school evaluation produced less than a dozen resources for parents that would (1) enable them to decide what types of information they needed, (2) suggest how they might obtain the information, and (3) indicate how the information might be interpreted to make a judgment on school quality. The characteristics, strengths, weaknesses, similarities, and differences of these resource materials are summarized. Further steps that might be taken are recommended. (RC)
INFORMATION FOR PARENTS ON SCHOOL EVALUATION

Robert A. Feldmesser and Esther Ann McCready

A family seeking to relocate in a new area usually has a number of communities to choose from, and the quality of the schools in each may be one of its primary considerations. A family which is otherwise content with the community in which it lives may wonder how good its children’s schools are, perhaps because the parents are being called upon to vote on a school budget or to choose among candidates for a school board or because they are interested in helping improve the schools if improvement is needed or because they may want to consider relocating if that will improve the education their children will receive. What kinds of guidance are available to families in situations such as these? That is the question this study was designed to answer.

Through the facilities of the ERIC system and a variety of other sources, we searched for suggestions to parents that would (1) enable them to decide what types of information they needed, (2) suggest how they might obtain the information, and (3) indicate how the information might be interpreted to make an evaluative judgment of school quality. We were looking for guides that could be used by adults who had no technical sophistication in educational measurement, that did not require large-scale organization or massive data-gathering, and that could lead to a systematic comparison among several school systems or among the schools of a particular community.

There is, of course, no dearth of literature on educational evaluation. Most of it, however, does not meet the criteria set for this study. Much of it is highly technical, directed to the professional evaluation researcher (for a recent review of this type of literature, see Walberg [1974]). Some of it is intended for accreditation or similar purposes and involves procedures that would be far too elaborate for the lay citizen to use and that do not readily yield comparative judgments among several schools or school systems (for example, Bradford and others, 1972: National Education Association, 1966). Real estate agencies and chambers of commerce may publish brochures on the local schools, but they conspicuously lack objectivity and are neither systematic nor comprehensive. Many school systems and some schools issue booklets—in some cases, even books—about themselves, but while some of these contain a wide range of valuable information presented in an orderly way (e.g., Council Rock, 1974), they, too, are likely to lack objectivity and (probably deliberately) do not help the citizen compare the system with other systems.

When these types of materials had been eliminated, we were left with less than a dozen items that met the stated criteria to a reasonable degree. They ranged in size from articles of a few pages to whole books; in specificity, from vague generalities to minute details; in underlying attitude toward the school system, from smug confidence that it is basically sound to belligerent insistence upon drastic change. The following report attempts to summarize their characteristics, identify their similarities and differences, and suggest further steps that might be taken.

Types of Information Needed

Presumably out of a desire to give a reassuring sense of definiteness and orderly inquiry to the bewildered parent, several of the authors of these school publications present their advice in the form of a list of questions to be asked or categories of information to be obtained. There is very little agreement on the number of such questions or categories. At one end, Silberman (1971) boldly states that “[t]here are only three requirements for a good elementary school.” At the other end, Postman and Weingartner (1973) pose “…slightly more than one hundred questions…” to be answered. In between, Knox (1971) puts his evaluative procedure into nine questions; Area Consultants (Harrison, 1972) apparently believe that there are 16 important items of information; Smith (1965) suggests 36 questions; and Blitzer and Ross (1951) offer a guide consisting of 75 questions.

Beneath this diversity of numbers, however, there is considerable consensus on what the relevant dimensions of evaluation are. Probably the one that most frequently recurs is the extent to which instruction is adapted to the needs and interests of the individual child. Others that are often mentioned are the physical attributes of the building, the pupil-teacher ratio or the class size, the “atmosphere,”
the standards of performance that are set for students, and
the levels of student achievements as shown in test scores.

This set of dimensions is hardly surprising. Yet it is some-
what unexpected to find so much agreement about them
even among authors who approach evaluation from very
different viewpoints. This is most vividly illustrated by the
comparison between Blitzer and Ross (1951) and Lurie
(1970), who are separated by 20 years and much more
besides. The text of the former, leading up to their guide, is
written as a series of amiable conversations between two
fictional couples of obviously middle-class status who
evidently believe that school systems are run by well-
meaning people like themselves—indeed, one of the
husbands is cast in the role of an elementary school
principal, which allows him to be the spokesman for "what
research shows"—and that the differences among systems
are chiefly matters of community support, as reflected, for
example, in the size of the school budget, and (implicitly)
of "traditional" versus "liberal" and "scientific" tastes in
education. Lurie, a veteran of the battles over education in
New York City, speaks on behalf of the poor and the
minorities; she bases her intensely passionate book—
significantly subtitled A Parents' Action Handbook on How
to Fight the System—on the conviction that the great
majority of schools are staffed by self-serving bureaucrats
from whom children must be protected at all costs. They
agree, nevertheless, on most of what to look for in evalu-
ating a school; their differences are manifested in phraseol-
ogy and tone. Both believe, for instance, that test scores are
important information. But Blitzer and Ross suggest (p. 79)
that parents ask "Do students in this school system
compare favorably with national norms on standardized
subject-matter achievement tests?", and they assume the
answer will be forthcoming. Lurie is, to put it mildly, more
skeptical:*

Find out how many children are not reading. This is not
as easy as it sounds. Many educators prefer to cover up
the facts.... You must insist upon precise and
complete information.... If your principal tells you
that "norm" means half the children should score
below grade level, tell him he is being ridiculous (pp.
45-56).

There is also a good deal of agreement on which ends of
the dimensions are indicative of high and low quality. A
school where there is more individualized instruction is
better than one where there is less; spacious, clean, and
well-cared-for schools are to be preferred over schools that
are crowded and littered and have broken windows and
stopped-up toilets; low pupil-teacher ratios are better than
high ones, and small classes are better than large ones (with
an exception to be noted); and schools where teachers and
students are friendly, happy, and mutually respectful,

* Lurie gives no single list of questions but scatters hundreds of them
throughout her book and offers suggestions through anecdotes and
in other forms as well.

where uniformly high standards are set for all students, and
where achievement levels are high, are of better quality
than schools with the opposite characteristics. One should
not be misled by the fact that a suggested evaluative pro-
cedure is couched in the form of questions, for the wording
almost always makes quite clear the answer which the
author is looking for. When Knox asks "Is accomplishment
genuinely expected of the students?", or when Blitzer and
Ross ask (p. 79) "Is there evidence . . . that there is concern
for individual differences of pupils?", there is little doubt
about the answer that is supposed to reveal high quality.

Only Lurie, despite her brash beginning and a general
assertiveness, displays some uncertainty about what
"quality" is, and then only on some points. On the matter
of pupil-teacher ratio and class size, for example, she says
(p. 90):

It sometimes happens that your school will have a
pupil-teacher ratio of twenty children for every teacher,
but an average class size of twenty-eight. This is because
many teachers are not assigned to a classroom. . . . Your
school may . . . have a number of specialists who do not
necessarily have special training but who do not have
responsibility for a specific class. Some parents believe
that, in the elementary schools, it would be better not
to have so many specialists, but instead to have smaller
classes. However, in very overcrowded schools this is
not always possible. What do you think would be best
for your school?

With respect to preservice training, e. e. asks (p. 91) "Do
you feel that teachers who have attended traditional
teachers colleges are better prepared than the others? Or do
you feel that the liberal-arts 'generalists' are more respon-
sive and less conservative?" Compare Smith (p. 58): "Does
the majority of the teachers have a broad training in the
liberal arts or are most of them products of institutions
which lay more stress on matters of technique than of
content?" Blitzer and Ross phrase the issue in different
terms but are equally sure of the high-quality answer: "Are
there teachers from a number of different teacher-training
institutions rather than most of the teachers from only one
place?" (p. 78).

Alone among these sources, Postman and Weingartner
explicitly acknowledge that there may be systematic and
legitimate philosophical differences among parents in their
views of what a "good" school is. Consequently, they point
out (p. 94) "the process of evaluating a school must begin
with your making some judgments about yourself." To help
in doing this, they present, in parallel columns, "two sets of
assumptions about learning and schooling, one constituting
a 'traditionalist's creed' and the other a 'progressive's
creed.'" Here it is recognized that some parents may want
organized regularity rather than individualization, or order-
liness and quiet rather than free expression. The authors
admit, however, that they themselves incline strongly
toward the "progressive" view, and this "bias" markedly
affects their evaluative criteria.
Smith, on the other hand and again, alone among these sources—is a fairly straightforward representative of "traditionalism." Long a leader in the Council on Basic Education, his bute naive is what he scornfully refers to as the "social-adjustment" philosophy of education. His evaluative criteria are sternly focused on "the values inherent in the academic program" (p. 61), "the heart" of which "consists of the sciences, English, mathematics, foreign languages, and history" (p. 58). Differences among students are perceived in terms of greater or lesser aptitudes and appetites for learning rather than of variant personal interests or needs, and he exhibits little patience with those who do not measure up. Thus, he suggests asking (p. 61) "Does the school make clear to parents of a recalcitrant student that it will not retain him over the legal leaving-age if over a long period of time the student resists all efforts of the school to teach him?" Lurie, who would be defined as a "progressive" by Postman and Weingartner, asks instead (p. 56) "Does [the principal] want to change the children or the teachers?"

Yet the divergence even between these two should not be exaggerated. Once more, it is sometimes a matter of tone rather than of substance. On the issue of the standards set for students, for example, Smith's question (p. 60) is: "Is the marking system one that seeks to measure a child only "against himself" and in terms of his own capacity, or is it a system which attempts to indicate clearly to a parent where his child stands in relation to his class, to others at the same grade level, and to a publicly known standard of achievement?", while Lurie asks (p. 26) "Do you think high standards have been set for the children? Some teachers tell us that the children should be encouraged to set their own standards... is this a professional-sounding excuse for no standards?"

Another issue which is often the subject of intense public debate, the racial and ethnic makeup of the school, is alluded to in only two of the sources. Postman and Weingartner admit (p. 103) that this is one of the "most important questions that many parents tend to ask when evaluating a school," yet, as they point out, they have not included anything about it among their questions: they merely plead with parents to accept integration as a necessity. For Lurie, it is a dominant theme, and her position—reiterated in many ways without ever being explicitly stated—is that a good school is fully integrated in every sense of the word, at every grade and in every program. She takes it for granted that her readers agree.

On the whole, the dominant impression that emerges from the review of these materials is that there is a wide area of agreement about the criteria by which schools should be evaluated. Indeed, if one were to pare away the polemical terms and state the respective evaluative criteria more neutrally, one might well be left wondering whether there is much genuine disagreement, at the non-technical level, about what a good school is. Perhaps the contro-

versies that do, after all, occur are about the priorities of the moment, or perhaps these commentators find themselves unwilling to put into cold print the kinds of positions they might feel compelled to take in the heat of a confrontation. Another possibility, however, is that it is precisely the polemical terms that mask the underlying agreement; if so, there is hope for developing a set of evaluative criteria that would command broad public assent.

How the Information Is To Be Obtained

Most of the sources offer suggestions for ways in which the citizen evaluator can gather the information he needs. The suggestions are of three types, which are usually (though not always) to be used in combination. One is to obtain published material from the school or school system, from a parents' organization, or from a local chamber of commerce or similar group. This is useful primarily for statistical data such as budgets and enrollments and for descriptions of programs and facilities. It is recognized that material of this sort is limited in scope, if not tendentious, and so cannot serve as the sole basis of evaluation.

A second method is to request a school administrator—usually the building principal—to fill out a questionnaire especially prepared for purposes of evaluation. Its advantages are that it can be used to obtain comparable information from several schools; that it may elicit statistical data that are often not published, such as test scores and racial and ethnic composition of the student body; and that it can ask for information about school policies and administrators' attitudes and judgments as well. Lurie presents (pp. 59-65) four sample questionnaires (one for all schools and one each for elementary schools, intermediate and junior high schools, and high schools, with supplementary questions for academic and vocational high schools). They concentrate on the testing program, test scores and other achievement indicators, special programs for slow and fast learners, and the distribution among these programs of students of different racial and ethnic backgrounds; but they also include such questions as "Are you satisfied with these [test] results?" Area Consultants (Harrison, 1972) uses a questionnaire, developed by the Metropolitan School Study Council, containing 16 items of a quantitative type; all of them are concerned only with high schools, since data about them, the firm contends, reflect the same policies that govern the elementary schools and are more easily obtained. The firm relies exclusively on this source of information, which is probably a necessity if it is to maintain, as it claims it does, a file of "up-to-date information on 450 school systems serving 230 communities in 13 suburban counties within commuting distance of Manhattan." Gayfer (1971) also reports on an evaluative procedure based entirely on a questionnaire, this one prepared by the Ontario Institute for Studies in Educa-
tion. It asks for the principal's judgment, on a scale of 0 to 3, of "the degree of success of your school in meeting each of" 39 criteria.

By far the most commonly recommended procedure is a personal visit to the school. Most of the authors—with the exceptions mentioned—believe it is indispensable, whatever other information may be used, and someone believe it is sufficient in itself. Knox, for example, says (p. 26) that "...you can size the school up correctly after just a couple of hours there." Others believe repeated visits are necessary, or even visits to several schools, in order to develop a frame of reference for judging one. *

When it comes to giving directions for the kinds of observations to be made on these visits, however, the sources are not very helpful.† They suggest looking at the outside of the building, walking through the hallways (especially while students are changing classes), observing some classrooms in session, talking with a few teachers and the principal, but they offer no systematic observational procedures. Rather, they tend to describe procedures through a series of anecdotes, real or invented, focusing particularly on the kinds of things that would not be seen in a good school.

Look around the room. The little towhead, squirming in his seat, has zipped through six pages in his math workbook while the boy sitting next to him is still struggling with the first page. And the pale child biting her nails is frantically trying to catch up because she was absent when her class learned long division (Silberman, p. 40)

Often when observations indicative of high quality are mentioned, they are contrasted with extremes that are unlikely to be observed in any school, leaving the potential observer at a loss to know what the two ends of the dimension might really look like. Lurie (p. 23) asks "Do the children seem relaxed and friendly as they move through the halls, or does it remind you of a prison march?"

Bulletin board and wall displays are favorite targets of observation: "Are any of the bulletin board exhibits examples of recent work by the children themselves? Do the displayed works of children reflect a freedom from stress or uniformity as a goal in itself?" (Blitzer and Ross, p. 79; emphasis in the original).

"What objects decorate the walls? Is there any evidence of student creativity or even effort? Artwork? Photography? Sculpture? If there are mottoes, flags, trophies, etc. in evidence, what do they tell you about the kinds of people and values the school honors?" (Postman and Weingartner, p. 100).

"Bare walls suggest a school that is barren in other ways, too" (Silberman, p. 38).

This last may be contrasted with a comment by Lurie (pp. 22-23):

...if the bulletin boards are empty, find out why. I once visited a school where the teacher was determined to leave a blank bulletin board until the children decided what should be placed there; the principal, however, was much less democratic, and had ordered the teacher to put up a display... if parents had barged in demanding that the bulletin board be filled, they would have done a great disservice to that very able young teacher. (Emphasis in the original.)

Yet Lurie herself devotes two pages (pp. 22, 26) to commenting on what can be learned from bulletin boards, never raising a general question about the validity of the observations. The same question could be raised with respect to any of the observations made during a visit, but it is not in these sources. On the contrary, they show a remarkable degree of trust in a visitor's intuition, "Most visitors can tell instantly if a classroom is 'happy,' even when no students or teachers are around," writes Knox (p. 26)—and the clue, interestingly enough, is on the walls, which in a "happy" classroom are brightly decorated and "abound" with student artwork and the results of class projects. (See also Lurie, pp. 17, 107.)

It must be said, too, that the expectations of what can be discovered in a school visit, or even in several visits, seem overly optimistic in many instances. "Are the classrooms varied enough in their furniture, arrangements, decoration, and so forth that the teachers' freedom and ingenuity are encouraged?" (Blitzer and Ross, p. 80), "What attitudes toward authority is the school training students to accept?" (Postman and Weingartner, p. 101). "Children sitting in a row with stiffly folded hands may look orderly, but are they really listening to what is going on?" (Lurie, p. 25). "Are the 'assembly' programs uniformly of such unmistakable educational value that they have a legitimate claim on precious schoolroom hours, or are they mainly entertainment?" (Smith, p. 62).†

*The problem of the reliability of observations is sometimes referred to, indirectly, by pointing out that the presence of a visitor may alter the situations being observed.

†Smith has less to say than any of the others about data-collection methods. At the same time, he proposes questions that appear difficult to answer with any feasible and reliable method—e.g. (p. 57). "Are the members of the school board... men and women of sound personal education and broad cultural concerns whose primary interest is in fostering an academic program of solid content?"
Making the Evaluation

If the citizen evaluator has gathered his information, what does he do with it? How does he interpret it and arrive at some final judgment as to whether the school is a "good" one or better than others with which it is compared or the most suitable one for his child or whether one aspect or another of it ought to be changed? A close examination reveals much ambiguity.

An illustration of this is the use of test-score information. The sources that recommend including such information in the evaluation usually show an awareness that it should be interpreted in the light of student backgrounds. Silverman points out (p. 39) that "children who come from homes full of books, magazines and dinner-table conversation should score higher," but does not suggest any techniques for taking that into account. Knox goes a little further: "Probably the best way to discover how your child's school stacks up is to measure it against others that serve similar neighborhoods" (p. 28) but that is an injunction the unaided parent may find difficult to follow. Lurie, however, rejects this kind of procedure altogether. Speaking from the standpoint of low-income and minority groups, she seems strongly inclined to suspect that any reference to students' backgrounds is an inadmissible excuse for low scores (pp. 56-57). It may be noted that none of the guides goes beyond a discussion of mean scores, neglecting the valuable evaluative information that might be derived from other ways of analyzing score data (Klitgaard, 1974).

The sources do not, of course, assert that the measured achievements of students are the only basis of evaluation but neither do they indicate how achievements are to be balanced against other considerations. In fact, it is a general characteristic of these guides that they offer no help in assigning weights or priorities to the various evaluative dimensions, in estimating trade-offs among them, or in combining the dimensions into a single score for a school. Yet no school is likely to be found excelling on all dimensions.

Even within a single dimension, the sources provide little guidance for placing a value on something other than an extreme position. "One teacher," writes Lurie (p. 26), "will put long, detailed social studies reports on display. Another has only a poster which she brought in from a travel agency." Few classrooms, however, will fall into either of these polar types; in the typical one, there will most likely be some moderately detailed reports, some that are longer and some shorter, and perhaps several posters—which may or may not have been "brought in from a travel agency"—as well as other materials. Or, using an example from Smith that was previously cited, most assembly programs probably combine educational value and entertainment, in varying—and usually unknown—proportions. How is an evaluator to "rate" such findings?

A related deficiency is that the guides sometimes suggest that a school of high quality should be in between the extremes, but fail to say just where in between the point of "maximum quality" is located. "An atmosphere of chaos should not be tolerated. But the opposite—grim silence and strict regimentation—is equally inappropriate" (Knox, p. 26). Similarly, words like "some" and "often" and "wide" are used which no precise numerical referent, or terms like "evidence" or "activity" which have no clearly defined units of n. So that it would be hard to determine the relationship of two schools on the dimension at issue. Perhaps it seems as if excessive expectations are being made of guides which are, after all, intended for use by persons without technical training, but unless they enable the citizen evaluator to reach a reasonable degree of precision, it may be doubted that they offer any improvement over what he could accomplish without them.

One device reported in the sources may help resolve some of the problems that have been discussed above. As has already been mentioned, the questionnaire developed by the Ontario Institute for Studies in Education (Gayfer, 1971) calls for a principal's judgments, on a scale from 0 to 3, of the "degree of success" of his school in meeting each of 39 criteria. This technique would permit the calculation of a single "evaluation score" for each of a number of schools of interest. If the evaluator wished, he could even weight the responses differentially. The questionnaire has been administered to the principals of 65 Canadian schools nominated by a panel of experts as being the "best" in that country, thus providing a frame of references for interpreting the scores. Of course, questions could be raised about the reliability and validity of principals' judgments (and about whether judgments by principals alone are adequate), and the usefulness of the questionnaire to a particular evaluator would depend on the extent to which he accepted the criteria. Perhaps he could devise items to fit his own criteria, but that would complicate the reliability and validity problems.

Further Steps

It seems clear that none of the guides that have been examined is sufficient to enable a citizen to evaluate a school or school system, or even particular aspects of it. There are crucial shortcomings in the procedures they suggest for collecting and interpreting information. But the need to which they are addressed is an important one, and more ought to be done to meet it.

School systems themselves are becoming more responsive to the public's demand for better information, either on their own initiative or by state mandate. An increasing number of them publish school-by-school mean scores on student achievement tests, usually in some sort of normative framework (if only in the form of grade equivalents), and sometimes with additional interpretive information...
(Badal and Larsen, 1970). Some provide data on community and staff characteristics and on selected items from attitude surveys (see, for example, Cincinnati Public Schools, 1973). Much improvement could be made in the ways this information is reported, but reporting it should be a routine practice for every school system. Not only would it assist parents in selecting a school or school system and raise the level of sophistication of citizens in educational matters, but it could also have salutary effects in educational reform (Wynne, 1971).

Proposals have been made for the establishment of a public agency to perform this function (Gooier, 1973; Evaluation Research Center, 1974). Such an agency would collect and make available, from a central source, comparable and "nonpartisan" information on a large number of schools and would work with schools and the community in interpreting, using, and improving the data. The idea sounds like a promising one. This review has shown that agreement on a set of criteria of quality may not be as hard to reach as has been imagined; but even if it were, the use of alternative (almost surely overlapping) criteria might not be impracticable. Selection of valid, reliable, and efficient indicators of the criteria would be a formidable but probably not an insuperable problem. The public's need for a yardstick of school quality is great enough to warrant the effort to solve it.
REFERENCES


Klitgaard, R.E. *Achievement scores and educational objectives*. Santa Monica, Cal.: Rand, 1974. (R-1217-NIE)


