DOCUMENT RESUME

ED 102 209 95 TM 004 510

AUTHOR Stodolsky, Susan S.
TITLE Open Education: A Challenge to Evaluators. Report No. 29.
INSTITUTION ERIC Clearinghouse on Tests, Measurement, and Evaluation, Princeton, N.J.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C. Office of Dissemination and Resources.
REPORT NO ERIC-TM-29
PUB DATE Dec 74
CONTRACT OEC-0-70-3797(519)
NOTE 25p.

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Affective Objectives; Classroom Environment; Classroom Techniques; Educational Objectives; *Evaluation Needs; *Literature Reviews; *Open Education; *Program Evaluation; Student Evaluation; *Teacher Role; Testing Problems

ABSTRACT

A review of selected literature on evaluation of open education programs provides a background for the author's discussion of the conditions essential to open education and the evaluation of its effectiveness. Other factors mentioned are: (1) the importance of the types of measuring instruments used to evaluate these programs, (2) the application of the classic evaluation paradigm, (3) lack of procedures for assessing the kinds of behaviors of interest to open education programs, i.e. instruments in the affective domain and classroom observational techniques. (DEP)

# Open Education: A Challenge to Evaluators

by

Susan S. Stodolsky



ERIC/TM Report 29

December 1974

1

The report upon which this article is based was funded by the ERIC Clearinghouse on Tests, Measurement and Evaluation. The Clearinghouse operates under contract with the National Institute of Education, U.S. Department of Health, Education and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not therefore represent official National Institute of Education position or policy.

# Open Education: A Challenge to Evaluators

Susan S. Stodolsky
University of Chicago

The purpose of this paper is to review selected literature on the evaluation of open education programs. It seems important, as a first step, to clarify what class of educational environment is being discussed when people refer to open education. What is the nature of the beast?

It is not really easy to delimit the meaning of open education, but there have been a number of systematic attempts to define the term and also to empirically describe a classroom which is more open than traditional, the way the teacher might behave, and the sorts of child behavior one might see. Some of the most complete and helpful discussions which attempt to define characteristics of open or informal classrooms are those by: Barth (2), Bussis and Chittenden (6); Featherstone (12,13), Flurry (14); Gross and Gross (17); Katz (24); Rogers (28); and Weber (40). Although these authors vary somewhat in the approaches they take, they tend to agree on the central characteristics of open classrooms. They recognize that open classrooms in the United States have been inspired by the informal methods of British teachers (8).

Certain basic assumptions about the way children learn are made by open educators. They accept the developmental view that children

learn through direct experience and are curious and motivated to learn. They accept the idea that children differ in the ways they learn and the times at which various aspects of development occur. Open educators believe that children have interests which are expressed in their activity and play and that these interests should generally be respected and nurtured in the school. The developmental perspective adopted by open educators is based on the psychological theories and research of Susan Isaacs (21), Nathan Isaacs (20) and Jean Piaget.

Beatrice and Ronald Gross (17) present a list of principles which seem to reflect the views in the literature I have seen on open education. They say:

> There are four operating principles of the open classroom. First, the room itself is decentralized: an open, flexible space divided into functional areas, rather than one fixed homogeneous unit. Second, the children are free for much of the time to explore this room, individually or in groups, and to choose their own activities. Third, the environment is rich in learning resources, including plenty of concrete materials, as well as books and other media. Fourth, the teacher and her aids work most of the time with individual children or two or three, hardly ever presenting the same material to the class as a whole (p. 10).

The notion of the "integrated day" is an important one in the British literature on informal education. The core idea is that subject matters and contents of learning are not dealt with singly but that the child's behavior and activity can lead to the use of various subject matters in a given activity. Learning through the development of an interest in something, rather than by going from page 10 to page 11, is favored by many practitioners of informal methods.

Thus, the school day is seldom divided into blocks of time for certain subject matters, although there may be divisions of certain types.

It might be worth emphasizing that interactions between the teacher and the children and among the children are a salient feature of open classrooms. Certain proponents of this general approach emphasize the development of cooperative learning and group learning more than others (for example, the Bank Street model of open education stresses child-child interactions). Virtually all open educators, however, consider interpersonal interactions to be a significant feature of the informal school experience.

Most authors describing open classrooms and the philosophy of open education comment on the very demanding role of the teacher in this approach to education. In order to have an educationally satisfactory open classroom, the teacher must be able to prepare materials which can be used by a variety of children, he/she must be able to extend and help develop children's interests, and must be able to tolerate more busyness and noise than one encounters in the traditional classroom. The teacher must be an astute observer of the children and must in some fashion keep track of their work so that she can decide when and if to intervene in the day of a child. It is usually agreed that it takes at least two years for a teacher to feel "on top of" the demands of the teacher role and that an open classroom teacher must still expect to work harder than a traditional teacher (see Andreae, 1).

It would be easy to get the impression that with children following their own interests, the best open classroom is one in which the teacher intervenes least. This position has been taken by some but seems to reflect a basic misunderstanding of informal principles. According to Berlak and his colleagues (3), open educators do not believe that a child will always pursue the most educationally beneficial path on his own. They believe the teacher has an important responsibility to set standards for work and conduct, to direct, extend and develop student interests, and at times to impose certain requirements on children. Thus, although there is a good deal of confusion on this point, open educators are concerned with the development of academic skills and standards. They do not see, however, the necessity for the development of these skills on a fixed timetable and in a fixed mode. Furthermore, they do not believe that academic skills are the sine qua non of the elementary school.

I trust this discussion adequately conveys the general characteristics of the schooling referred to in the open education literature. One of the problems, to which we will return, is that the term open education has been used to refer to many varieties of innovation other than the ones just described. In reviewing the existing research on open classrooms, an important first step is to clarify the actual nature of the program which is being studied. Frequently the program being studied lacks some of the critical features of open education as

I have described it here. For example, there has been some construction of open plan schools, schools which have open spaces that can be flexibly used. These open plan facilities are suitable for open education programs, but can also house a variety of other programs. There are studies which have looked at open plan schools that do not house open education programs. Analagously, many individualized instruction programs are being instituted in schools around the country, but, individualization alone does not signify an open education program.

For purposes of this paper, I will attempt to base my discussion on the conditions presented in the literature as rather essential to an open education program. Open educators are not doctrinaire and do not have a single prototype to implement, but they do seem to have an essential approach toward which they feel a teacher should be oriented.

If we understand generally what is meant by open education, we can next move to the questions of evaluating such programs. I will attempt to distinguish two major points of view about the way to evaluate open programs and will review the limited amount of literature which is available. I will also try to detail some of the major problems to be overcome in this area and will suggest some possible strategies for arriving at useful solutions.

The Classic View of Program Evaluation

The existence of open education programs, as well as the problem of evaluating preschool programs of diverse complexion (35) has contributed to the development of new ideas about evaluation. The classic view of evaluation focuses on the collection of evidence about a program's effectiveness in achieving certain objectives or goals. Classically designed evaluations are studies of outcomes. With the emergence of large-scale projects which require evaluation data, many studies have been conducted which attempt to compare one program with another or one program with a"control group".

The classic evaluation paradigm for evaluating open education programs poses many problems; yet the approach has a valid place and deserves further effort.

In the application of the classic evaluation paradigm to open education programs, an important step (often omitted) is a careful consideration of the objectives of open education. Four approaches to relate outcome measures to objectives are possible. In my opinion, an evaluation of any educational program should include evidence about the achievement of the intended outcomes of the program. At times, an evaluation study will include data which measures atrributes thought to be of value whether the program sponsor has specified them or not. These atrributes may be considered valuable on an a priori basis and are not necessarily aims of program itself. If we had consensus about a set of desirable outcomes which should result from educational programs

for children, we could utilize measures of that set of outcomes to assess all programs, but I do not think we are in that position. In conducting comparative evaluation studies, one might include measures of attributes which are a collection of the intended outcomes of diverse programs and thus make some measures probably more appropriate for one program than for the other. Last, the evaluator may attempt to measure some attributes which might be unintended outcomes of a program.

Douglas Pidgeon (26) very ably expresses the classic point of view with respect to the evaluation of informal education. He says:

> With all the extensive changes now taking place, it is clear that only an evaluation specifically designed to discover the extent to which a variety of objectives is being met is likely to produce the kind of evidence that will satisfy protagonists and antagonists alike. Such a study would have to be carried out on a carefully selected sample of schools which represented the best and the worst of both the traditional and newer informal approaches to teaching and learning. The aims and objectives of the sample schools should over not only those aspects of education which are given more emphasis with the informal approach, but also those which the critics of change hold in high esteem (pp. 19-20).

There are some examples in the literature of studies which have selected outcomes to be measured in the four ways I have mentioned. Until very recently, the most common practice was the use of measures which seemed valuable in their own right and which did not necessarily reflect the objectives of highest priority for a given program or even objectives being attempted. Most of the Head Start evaluations were of this sort, with the ubiquitous intelligence test serving as a major component. I have reviewed elsewhere the general nature of the findings

of preschool evaluation programs (21, 34). Two points merit reiteration here because they are equally germane to the work in open education. First, the pattern of results in the existing literature generally show more successful outcomes for structured than for unstructured programs. Second, this appears to be the case because the measures used in the studies tend to reflect the objectives of the structured programs better than the unstructured ones. It is also the case because the structured programs are more homogeneous with respect to the experiences of the children. Thus, the structured programs meet the important methodological standard of being a "treatment" to be evaluated, whereas more unstructured programs tend to be much less uniform. In unstructured or open situations, children engage in diverse behaviors. While diversity is a core commitment for open educators, it greatly complicates any attempt to study the effects of schooling.

Why has there been so little assessment of the intendend outcomes of open education? I believe one central difficulty has been the lack of procedures for assessing the kinds of behaviors of interest to open educators.

It should be clear from the earlier discussion of open education that the objectives of highest priority lie outside the easily measured domain of skill learning. In fact, many of the central objectives of open education are not in the cognitive domain at all. Even for those educators who have recognized the goals of open classrooms as valid, a serious problem still remains in locating or developing

or developing methods for studying such characteristics in children.

Jennifer Andreae (1) captures many key objectives in the following discussion of evaluation.

> The test of true learning is in the ability to use appropriate thinking skills to meet such situation. Standardized achievement tests do not show this; nor do they show initiative; responsibility: ability to relate to others; or level of intelligent thinking; or attitudes toward learning; or carryover at home; or where and what the gaps are in a child's understanding, ability to solve problems, use of imagination, creativeness, and ingenuity; or awareness or sensitivity to a child's interest and learning style. At the present time, there are no standardized tests to measure these vitally improtant factors, and yet without this information, the picture presented of a child is a shallow and nearly meaningless one (p. 57).

There have been some attempts to get at certain aspects of self concept, attitudes toward schooling, achievement motivation and locus of control (10, 18, 26, 28, 29, 30, 37). Methodologically, the quality of these studies is varied and the range of instruments narrow.

Lack of Instruments in the Affective Domain

While virtually all workers in the affective domain have lamented the lack of adequate instrumentation, there have been very few projects which have focused on the development of psychometrically reliable and valid instruments of socioemotional functioning, curiosity, initiative, interest and the like. A recent book by Deborah K. Walker (38) does provide a very helpful listing and critique of existing socioemotional measures for preschool and kindergarten children. Presumably, the instrument construction problem is somewhat less difficult when one deals with children of elementary school age. Where

it is easier to interview young children and get meaningful responses, it becomes possible to consider the use of group measures and paper and pencil devices.

Certainly one reason for the lack of instruments in this broad area has been a feeble commitment to the importance of such objectives, coupled with the very serious problems of the fairness of evaluating children with respect to attitudes, values, and other personal attributes. Our American values of individuality, privacy, lack of indoctrination and so on, have led many school people to shy away from the collection of evidence about affective behaviors. While concerns about values should not be glossed over lightly, the lack of evidence about affective and attitudinal development seriously limits our ability to adequately judge the consequences of educational experiences. Furthermore, one can collect such information without necessarily using it in any potentially punitive context. There is a difference between information to be used for student evaluation and information to be used for program evaluation. While overlap in these two types of data would be efficient, it may not always be possible. At least it would be helpful to have data on socioemotional and affective variables for program evaluation.

Few sustained efforts at affective instrument construction for school children have ben launched, and those instruments which have been developed often have low validity and cover but a fraction of the richness of behavior in which we are interested. While continued

and renewed efforts hould be supported, I have some recurrent doubts about the potential of such efforts. The use of criterion-referenced rather than norm-referenced instruments may be helpful (15, 37). Yet my work with children in different kinds of classrooms makes me question whether the psychometric approach will prove relevant to many of the constructs which interest us. I find it somewhat consoling that other colleagues involved in the evaluation of educational programs have voiced similar doubts. Herb Zimiles (42), in a discussion of personality measurement, suggests that Allport's idiographic view may have merited more support by psychologists than it received.

> . . . different traits are differentially salient for different children (or adults). Across-the-board measurement of a particular trait generates a hodge-podge of data. The data gathered from those for whom the trait is salient may be quite telling, but a large portion of the data obtained from the rest of the sample may have no functional significance (p. 5).

Recently, I have been attempting to understand the choices children made in classrooms which were moving toward an open philosophy. I used a series of psychometric instruments which I thought might help differentiate children in conjunction with observations of children's classroom behavior. I was interested in finding out what sorts of children select what sorts of activities and exhibit different styles in an open setting. There appear to be relationships between some of the characteristics I measured (intelligence, cognitive style, locus of control, and associational fluency) and certain aspects of the child's functioning in the classroom--for example, the extent to which he pursues activities alone. However, various aspects of child

functioning seem to relate to somewhat independent clusterings of child variables (and, of course, some relate to none of the variables). Furthermore, our clinical impressions of the children lead us to very different hypotheses about the correlates and etiology of similar behavior patterns in different children.

What I am trying to illustrate here is the perennial tension between the recognition of individuality and the attempt of measurement to provide comparability of data under standard conditions. While both endeavors are necessary, we may be least successful in applying the psychometric approach in the socioemotional area. Nor am I certain that a quantitative view is appropriate for many aspects of personality and affect. Many of the constructs which interest open educators are more qualitative in nature and depend upon certain conditions for their expression. For example, I am not sure that it is meaningful to ask whether a child has a more positive or negative self concept than another child rather than knowing the nature of his views of himself. Similarly a construct like initiative seems to come in many behavioral forms which are difficult to assess in a uniform manner across children.

An extremely insightful article by Edna Shapiro (31) contains a discussion of the relative utility of tests and classroom observations as sources of data. She says:

> The rationale of the test, . . . is that each child is removed from the classroom and treated equivalently, and differences in response are presumed to indicate differences in what has been taken in, made one's own, that survives the shift to a different situation.

> . . . This requires reconsideration of the role of classroom data, individual test situation data, and the relation between them. If we minimize the importance of the child's behavior in the classroom because it is influenced by situational variables, do we not have to apply the same logic to the child's responses in the test situation, which is also influenced by situational variables (pp. 533-34)?

Shapiro does not consider classroom data and test data to be equally useful for all evaluation purposes, but she seriously questions whether the sort of test data we are able to obtain will be any more useful for program evaluation than classroom observations would be. She eventually argues that evaluation should play its chief role as a contributor to program development and as such would justify more flexible uses of observations and experimental instruments. While I agree with this general point, I am also sympathetic to policy makers who want evaluation data in order to make policy decisions about the effectiveness of programs. Perhaps a consideration of some suggestions and studies made by proponents of open education and evaluators who do not hold the classic view will assist us to reach some useful conclusions about the best evaluation strategies.

Alternative Views of Evaluation

A number of persons involved in the field of evaluation have questioned the appropriateness of measuring outcomes as the basis for judging program effectiveness. Many open educators advocate a process of documentation of child growth and development based on children's work and actual classroom behavior (see Carini, 7, Chittenden and Bussis, 9, and Dean, 10). Some helpful examples are presented in a recent article by Hawes (18) which reports the work of Ruth Aldrich at Marcy School, Lillian Weber, Mary Stitt and others. A number of

evaluators, including myself (31, 33, 35), have also recommended that the classroom behavior of children be considered valid and important evidence about a program.

There are several positive aspects to this general strategy. First, it is important to better understand the actual classroom experiences and processes of children in diverse educational settings. Data which details the salient features of various classroom environments and of different children's experiences in different classrooms can be used for a number of highly significant purposes.

Most germane here is the use of such descriptive data to make judgments about the quality of an educational program. Zimiles (42) persuasively argues for:

> . . . systematic and comprehensive evaluation of the child's school environment, to be followed by a theoretical analysis of the potential impact of his school experience.
> . . . It is a paradox that we have the responsibility and the capacity comprehensively to describe and record the essential character of an educational program, yet do not do so; and at the same time, we do not know how to assess the impact of a comple set of experiences on the psychological functioning of a developing child, yet we persist in trying to do so (pp. 7-8).

Zimiles believes that we can look at educational environments and make value judgments about them. Such judgments would be based on our theories about the educational conditions which lead to certain desirable ends. Zimiles stresses the necessity for a conceptual framework which would allow us to make connections between what we see in classrooms and our ideas about the "potential impact" of these

conditions on the child. I believe that heads of English schools and inspectors who routinely evaluate and advise British teachers implicitly use this strategy.

A related point of view is presented by Chittenden, Bussis, and Amarel at ETS (see Hawes, 18). They suggest that the quality of the classroom experience should form the basis for evaluation of open classrooms. They would evaluate diverse aspects of both student and teacher behavior and products as well as aspects of the physical environment of a classroom. In contrast to conventional evaluation which judges student responses to some uniform condition, they are trying to develop standards for judging diverse student productions and activities. The ETS approach to evaluation focuses on standards of quality regarding the "process, content and context of the child's learning" (Hawes, 18).

Both the ETS group and Zimiles endorse the need for evaluation evidence to judge programs. They are searching for new ways of producing evidence which will in some sense be objective and comparable across educational environments. I share their concern about data which will allow one to make judgments about different educational experiences, but I believe that program evaluation per se may be less important for the educational enterprise than evaluation which will feed back into student learning and studies which will begin to explicate the nature of the empirical relationships between various educational experiences, child characteristics, and child outcomes.

It may not be necessary to pit one type of work against the other. Yet I believe large scale classic evaluations have absorbed too much of our limited resources without adequately answering the key questions.

Open education presents a real challenge to evaluators and educational researchers because it broadens the scope of educational environments and experiences to be understood. We have begun to develop methods for systematically describing some of the relevant features of classroom structure and process. Promising observational procedures have been developed in response to the need for systematic descriptions of educational environments (5, 16, 23, 24, 32, 34).

There appear to be some promising starts on the development of an empirical base for understanding how educational environments differ (4, 15, 32, 34). Some possible consequences of the educational differences are also being studied. For example, in a recent paper using Follow Through data, Stallings (33) related differences in classroom conditions to differences in children's classroom behaviors and to tested outcomes. With respect to open classrooms, she says:

> In the more open, interdisciplinary classrooms, where a wide variety of activities are occuring, a wide variety of materials are available, children can select their own groupings part of the time, and can engage in activities without adults, children have higher scores on the Raven's perceptual problem-solving test. They are also absent less often, and they take more responsibility for their success as measured on the Intellectual Responsibility Scale. They are more independent, cooperate more often, and ask more questions (p. 13).

Replication and extension of thi ort of work should be a high priority.

My students and I have been attempting to contribute in this area by intensive small-scale studies which followed individual children over time in free-choice classrooms. We have been able to use samplings of children's behavior to define variations in the children's experiences which occur in a given classroom. These variations in activity have, in turn, been related very strongly to differences in certain outcome measures (see Karlson, 22). Since students clearly do pursue unique paths within an open classroom, it seems educationally important to know more about the types of choices they make and the types of behaviors they exhibit. We hope to unearth possible correlates of choices and styles both in terms of individual differences in the students and in relation to the situational determinants of choice. Welch (41) has been able to show at the nursery level that cognitive style differences are strongly predictive of differences in the free play behavior of children, particularly with respect to the pacing of activity and patterns of attention. My most recent work with seven- and eight-year-olds also seems to be revealing that some systematic variation in child behavior in open environments is related to measurable individual differences.

In addition to process-oriented studies, we should shift from evaluations of programs themselves to evaluations of student learning and development for feedback purposes. Teachers continuously observe

students and analyze their activity and products in an effort to ascertain the best strategies for facilitating educational development. Systematic and long-term efforts of this type would contribute both to the immediate process of education and to our understanding of the consequences of educational programs. Obviously, evaluation directed toward a feedback or formative function is designed primarily, if not exclusively, in terms of the intended goals and objectives of a given program.

Recommendations

In order to improve the classic approach to evaluation of open education, I would recommend the following actions:

1) Empirical description of the nature of the educational environment under study should be a necessary step in all evaluations. It is essential to verify that a program is actually being implemented and to record the salient features of the program.

2) Where technically feasible, students should at least be measured with respect to the high priority objectives that the program itself endorses. Where tests are not available for this purpose, consideration should be given to using classroom observations and other situationally complex methods for assessment.

3) Long-term studies of children and environments are essential. The objectives of open educators are not expected to be easily obtainable. Only longitudinal evidence can effectively address the question of outcomes for open education (I believe a similar argument could be made with respect to many educational programs.)

4) Instrument construction, particularly in the socioemotional and affective domains, should be attempted with validity as the key criterion in judging the utility of new instruments.

5) Rating and judgment methods should be further explored as possible alternatives to test situations.

6) The use of children's own reports through interviews, logs, and other devices should be considered valid evidence about program process as well as student development. Similarly, the views of parents and teachers can be crucial ingredients in an evaluation plan.

Outside the domain of classic evaluation, I would recommend the following:

1) Systematic efforts should be made to collect work samples and behavior observations of children in open classrooms. Where possible, the standard design considerations should be employed to avoid bias, lack of randomization, and so on.

2) Efforts should be made to develop and explicate a theory which would permit inferences about the potential impact of an educational environment on child growth and development.

3) As the above theory is developed, empirical validation, wherever feasible, should be attempted.

4) The collection of case studies on children over long periods of time would provide helpful information for making judgments both about the quality of an educational experience and its possible consequences. Such materials might also add to our understanding of human variation in development. When possible, such case studies should contain information about the actual experiences of the child in the educational setting so that we may better understand his educational "treatment."

5) Comparative studies of process in classrooms, preferably at a small-scale, intensive, rather than large-scale, extensive level, should be supported. To the extent that open education presents important variants in educational environments, understanding the components of open programs with respect to other programs should clarify the different contributions to be made by these varieties.

During the writing of this paper, I have felt discomfort with the task. I strongly feel the lack of highly specific, concrete suggestions in what I have written. Hopefully the more general considerations will prove stimulating to those more intimately involved with the day-to-day process of education in the classroom. I have tried to provide through references a key to those helpful methods and ideas which are currently available in print. Sadly, the field has not moved very far yet. Nevertheless, the challenge of open education represents an important and sustained impetus for change--which should both please and perplex us.

REFERENCES



1. Andreae, J. Open education: ESEA Title I. New York State Education Department, Albany, December, 1970. ED 059 334

2. Barth, R.S. Open education and the American school. New York: Agathon Press, 1972.

3. Berlak, A.C., Berlak, H., Bagenstos, N.T., & Mikel, E.R. A participant-observation study of teaching and learning in English primary schools. School Review, February, 1975, in press.

4. Bissell, J. The cognitive effects of preschool programs for disadvantaged children. Unpublished Ph.D. dissertation, Harvard University, 1970.

5. Bowman, G.W. Analysis of communication in education: Bank Street follow through compared with Bank Street school and children and fourteen non-follow through programs. Mimeographed. New York: 1973. Bank Street College of Education.

6. Bussis, A.M., and Chittenden, E.A. Toward clarifying the teachers role. In E.G. Nyquest & G.R. Hawes (Eds.), Open education: A sourcebook for parents and teachers. New York: Bantam Books, 1972. Pp. 117-136.

7. Carini, P.F. Documentation. In evaluation reconsidered. New York: The Workshop Center for Open Education. May, 1973, 15-24.

8. Central Advisory Council for Education. Children and their primary schools (the Plowden Report). London: Her Majesty's Stationary Office, 1967.

9. Chittenden, E.A., & Bussis, A.M. Open education: research and assessment strategies. In E.G. Nyquist & G. R. Hawes (Eds.), Open education: a sourcebook for parents and teachers. New York: Bantam Books, 1972. Pp. 360-374.

10. Cline, M.G., et al. Education as experimentation: evaluation of the follow through planned variation model. Volume IA: early effects of follow through. Cambridge, Mass.: Abt Associates Inc., March 1, 1974.

11. Dean, J. Informal schools in Britain today: recording children's progress. New York: Citation Press, 1972.

12. Featherstone, J. The British and us. Open schools I. New Republic. 165, No.11, September 11, 1971, 20-25.

13. Featherstone, J. Tempering a fad. Open schools II. New Republic. 165, No.11, September 25, 1971, 17-21.

14. Flurry, R.C. Open education: what is it? In E.G. Nyquist & G.R. Hawes (Eds.), Open education: a sourcebook for parents and teachers. New York: Bantam Books, 1972. Pp. 102-110.

15. Glaser, R., & Nitko, A.J. Measurement in learning and instruction. In R.L. Thorndike (Ed.), Educational measurement. (2nd ed.) Washington, D.C.: ACE, 1971. Pp. 625-670.

16. Grannis, J.C. Final progress report: Columbia Classroom Environments Project, Contract No. OEC-0-71-0593. Teachers College, Columbia University, February, 1973.

17. Gross, B., & Gross, R. A little bit of chaos. Saturday Review, May 16, 1970.

18. Hawes, G.R. Managing open education. Nation's Schools, 1974, 93, No.6, 33-47.

19. Heimgartner, N.L. A comparative study of self-concept: open space vs. self contained classroom. University of Northern Colorado, 1972. ED 069 389

20. Isaacs, N. Piaget: some answers to teachers' questions. London: National Froebel Foundation, 1965.

21. Isaacs, S. The children we teach. London: University of London Press, 1963.

22. Karlson, A.L. A naturalistic method for identifying behavioral aspects of cognitive acquisition in young children participating in preschool programs. Unpublished Ph.D. dissertation, University of Chicago, 1972.

23. Karlson, A.L., & Stodolsky, S.S. Predicting school outcomes from observations of child behavior in classrooms. Paper presented at Annual AERA meeting, New Orleans, February, 1973.

24. Katz, L.G. Research on open education: Problems and issues. In D.D. Hearn, J. Durdin, & L. Katz (Eds.), Current research and perspectives in open education. Washington, D.C.: American Association of Elementary-Kindergarten-Nursery Educators, 1973.

25. Mayer, R.S. A comparative analysis of preschool curriculum models. In R.H. Anderson & H.G. Shane (Eds.), As the twig is bent: readings in early childhood education. Boston: Houghton Mifflin Co., 1971. Pp. 286-314.

26. Pidgeon, D.A. Informal schools in Britain today: evaluation of achievement. New York: Citation Press, 1972.

27. Reynolds, R.N. A comparative evaluation of the effects of an open classroom instructional program and a traditional instructional program. Pennsylvania Department of Education, February, 1974.

28. Rogers, V. (Ed.) Teaching in the British primary school. New York: Macmillan Company, 1970.

29. Ruedi, J. & West, C.K. Pupil self concept in an "open" and in a "traditional" school. 1972. ED 066 217.

30. Scheiner, L. An evaluation of a pilot program to assess the introduction of the modern english infant school approach to learning with second and third year disadvantaged children. Philadelphia School District, Pennsylvania Office of Research and Evaluation, October, 1969. ED 034 595.

31. Shapiro, E. Educational evaluation: rethinking the criteria of competence. School Review, 1973, 81, 523-549.

32. Soar, R.S., & Soar, R.M. An empirical analysis of selected follow-through programs: An example of a process approach to evaluation. In I.J. Gordon (Ed.), Early childhood education: the seventy-first yearbook of the National Society for the Study of Education, Part II. Chicago: University of Chicago Press, 1972, 229-260.

33. Stallings, J. Follow through program classroom observation evaluation 1971-72. Menlo Park, Calif.: Stanford Research Institute, 1973.

34. Stallings, J. What teachers do does make a difference—A study of seven follow through educational models. Paper presented to Early Childhood Conference on Evaluation, Anaheim, California, August, 1974.

35. Stodolsky, S.S. Defining treatment and outcome in early childhood education. In H. Walberg & A. Kopan (Eds.), Rethinking urban education. San Francisco: Jossey-Bass, Inc., 1972. Pp. 77-94

36. Stodolsky, S.S. Final report to national institute of education. Observational studies of variation in child behavior in classrooms, in preparation, Fall, 1974.

37. Stodolsky, S.S. What tests do and don't do. Pamphlet prepared for Association of Childhood Education International, forthcoming.

38. Vogel, F.X., & Bowers, N.D. The relationship of forms of school organization to pupil behavior. February, 1969. ED 028 505.

39. Walker, D.K. Socioemotional measures for preschool and kindergarten children. San Francisco: Jossey-Bass, Inc., 1973.

40. Weber, L. The English infant school and informal education. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

41. Welch, L.R. Individual differences in tempo: a study of the free play behavior of reflective and impulsive nursery school children. Paper presented at Biennial SRCD meetings, Philadelphia, 1973.

42. Zimiles, H. A radical and regressive solution to the problem of evaluation. Paper presented at Minnesota Round Table in Early Childhood Education, Wayzata, Minnesota, June, 1973.