DOCUMENT RESUME

ED 035 453 PS 002 766

AUTHOR Stern, Carolyn
TITLE Assessing Process and Product with Young Children in School Settings.
PUB DATE Mar 69
NOTE 7p.; Paper presented at the biennial meeting of the Society for Research in Child Development, Santa Monica, California, March 1969

EDRS PRICE EDRS Price MF-$0.25 HC-$0.45
DESCRIPTORS *Behavioral Objectives, Behavior Change, Class Management, Classroom Observation Techniques, *Preschool Programs, *Program Evaluation, *Student Behavior, Time Factors (Learning)
IDENTIFIERS Head Start, Observation of Substantive Curricular Input

ABSTRACT

What needs to be done with preschool programs is to define specifically their desired product (the behaviors we want in the children) and to determine the optimal process by which this product can be brought about. Using an instrument called the Observation of Substantive Curricular Input in 1967-68, a coding system based on a series of 3-minute scans of classroom activity, data were compiled for 151 Head Start classes. Analysis of the data indicated that while Head Start classes differed among themselves, certain generalizations could be drawn. Comparatively little time was spent on caretaking activities, aimless wandering, and teaching of pre-academic skills. A considerable amount of time was spent on structured waiting for the teacher, language development, dramatic play, and social interaction. Group vs individual activity differentiated classes and seemed dependent on teacher control. The next step will be to relate the obtained classroom characteristics to specific changes in children. With the causal relationships between process and product more clearly understood, more effective programs can be devised based on data rather than intuition. (MH)

ASSESSING PROCESS AND PRODUCT WITH YOUNG CHILDREN IN SCHOOL SETTINGS[1]

Carolyn Stern

For a long time, the field of early childhood education, as represented by nursery school environments, operated under the sacrosanct cloak of a mystique. Principles and procedures were accepted on faith and without question. It was assumed that the young child was a tender plant, needing only the sunny warmth of love and the milk of human kindness to be nourished, and that he would then unfold and flourish as a flower. These precious early years were to be devoted to creative and dramatic play, games, and art, and socialization, with the harsh demands of "school work" relegated to that traumatic moment when the child entered the "real world" of the school. Fortunately for the child, his middle-class parents usually supplied him with the "hidden curriculum" which made it possible for him to cope with school tasks and expectations. The high incidence of school success with this nursery school graduate nurtured the mystique since there was no reason to challenge the efficacy of this ideology.

The overwhelming evidence for the critical importance of early stimulation, paired with the clear demonstration of cumulative decrement in the performance of children from poor families as compared to that of the advantaged child, created the press for a national intervention program. It seemed only natural to turn to the nursery school as the ideal resource by which the cycle of ignorance and poverty could be effectively interrupted. For the first time, the traditional nursery was tested - and found wanting. Even before the advent of Head Start, compensatory and intervention programs were being designed to speak specifically to the needs of the disadvantaged child. Experimental comparisons were carried out and demonstrated, many with p less than .001, that cognitively oriented programs were distinctly superior to the traditional approach. Unfortunately, it is now evident that this solution is not the complete panacea that had been anticipated. After six or eight months in the regular school these children often become indistinguishable from their classmates who have not had the benefits of special training.

Depending upon one's political orientation and personal bias, there are two possible interpretations. One can go the genetic route proposed by Professor Jensen in the Harvard Educational Review and say, "Let's stop kidding ourselves. You can't make a silk purse out of a sow's ear. These kids don't have what it takes. They're just not as smart as us educated, middle class types." Or one can do a little constructive soul-searching and ask two simple questions: First, what is it we are trying to do? and second, what is the best way of doing it?

If the opening remarks seemed overly critical of the traditional nursery, it is not because of any disagreement with the basic premises. As a matter of fact, there is even less support for those who, after years of fighting white segregationists, now find themselves supporting black separ-

[1]This paper was part of a symposium at the biennial meeting of the Society for Research in Child Development, March 1969, Santa Monica, Calif.

atism, and who pejoratively categorize basic human virtues as "middle class values." The major objection to the nursery mystique is that it makes assumptions about the relationship between activities and behavior based on intuition and unsupported by evidence.

In his book on Evaluative Research, Suchman (1967) has offered a model for investigations for testing hypotheses in "real life" settings. Evaluative research starts with the premise that there is a product which is desirable and worth while. These are the goals which the research intervention is attempting to achieve. The second element in this paradigm is the assumption that a program can be designed and implemented to bring about the desired objectives. The process is thus intimately related to the product to be developed. The final component is the criterion test to determine whether or not the expressed objectives have been attained. If they have not, further action must be taken to examine the intervention program to determine why it failed to have the desired effect.

Applying this paradigm to research in preschool settings, the first requirement is to spelling out in objective and measurable terms the behaviors which constitute the outcomes or products of the intervention. Until quite recently, these seemed to be quite straight forward: Preschool activities for disadvantaged children should foster the development of intellectual competencies and personality characteristics, and remediate deficits and handicaps, so as to optimize success in school. Appropriate subgoals would include the acquisition of enriched language facility as well as other preacademic cognitive skills, especially those related to problem solving and logical operations. In the affective domain, the child would be expected to develop positive self-concept, sense of worth, high self-esteem, motivation to achieve, satisfaction in attaining competence, independence, and self reliance. In the interpersonal area, there would be the ability to form warm human relationships, to trust and have confidence in adults, to share and cooperate.

An instrument to measure teachers' expectations for achievement of children in Head Start has been developed at UCLA. This instrument attempts to list in precise operational terms a variety of behaviors by which the attainment of the desired outcomes can be demonstrated. This scale has been given to a wide variety of professionals and paraprofessions in the field of early childhood, who rated each item in terms of relative importance. There was a surprising consistency with which the more middle-class professionals selected the affective outcomes, while the paraprofessionals and parents tended to emphasize the academic skills. However, there are many who object to these outcomes as being "middle-class" goals, and who maintain that they are inappropriate to the reality and life-style of the child in the ghetto. What then are the products and behaviors we should specify as the objectives of our intervention program? Are we to teach the child to be distrustful and suspicious? To look out for himself first and devil take the hindmost? To become a better thief, a more successful arsonist? Until someone can provide a valid list of non-middle class goals to serve as the foci of a program to effect change, the middle-class values will have to do.

The next consideration is designing a program which will produce the behaviors desired. A number of people have constructed systems describing factors or conditions which theoretically provide the setting for the desired growth. For simplicity, these can be subsumed under three basic areas or dimensions. The first relates to the characteristics of the adults. The teacher or aide should be reinforcing, capable of recognizing the needs of individual children, and act as a model for all the desired behaviors and attitudes. Care should be taken to provide adults of varied ethnic and socioeconomic backgrounds so that there will be some one from the child's own culture with whom he can identify.

The second dimension is concerned with the classroom ambience, the organization and structure of the learning environment, the management of adult and peer relationships, and the general emotional climate which encourages trust and sustained responding over varied conditions. Finally, there is the program content, the subject matter input and organization to provide a stimulating learning environment, with sufficient pattern to develop confidence in the continuity of events and fulfillment of promises, yet flexible enough to be open to change, encourage exploration, and use of materials.

No planned intervention with young children can neglect to consider important components outside the classroom. The Head Start program is a particularly comprehensive one, concerned with the total child, and includes parent involvement, medical and dental care, nutrition, and a variety of social services. If one considers the total Head Start program as the process for achieving the specified outcomes, a multifaceted, multivariate assessment model must be designed. This is the scope of the National Head Start Evaluation program. However, this paper will be limited to assessment of the input variables in the classroom alone. While it is taken for granted that there are many different types of preschool programs, somehow most people think of Head Start as a unitary concept, a sort of unique entity about which generalizations can be made. Thus there is a tendency to measure changes in children and compare those "with Head Start" and those "without Head Start," with no attempt to specify the content of the program to which the child has been exposed. The assessment of early intervention programs were subject to this same criticism. Another and more basic criticism of impact studies is the tendency to assume that certain activities or inputs are intrinsically valuable and will by some kind of magic transform the Head Start child into a competent, achieving kindergartner. The relationship between activity and outcome behavior is based on some type of face validity. For instance, it has been assumed that experience in matching geometric shapes and forms will improve the child's readiness to discriminate letter forms. This commonsense assumption has never received any conclusive experimental support. Although the example cited involves a fairly simple skill and is susceptible to direct test, the majority of valued activities are far more subtly related to the outcomes they are presumed to implement.

Testing the relationship between process and product requires the evaluative research design advocated by Suchman. In a sense it is the one adopted by the National Evaluation Program, with the exception that the goals are more implicit than explicit. One of the major problems in carrying out the research is the inadequacy of measurement instruments. In most Head Start classes, teachers tend to place most value on affective, social-emotional growth. Yet it is in this area that reliable and validated instrumentation is practically nonexistent. Even if one knew which procedures were most effective, there is no way of obtaining confirmation of growth in adaptive social-emotional behaviors with a reliable instrument. Because of the inadequacy of measuring tools, and because it is so difficult to express affective goals in terms of specific behavior, most programs are ultimately evaluated on the basis of cognitive measures. With most teachers clearly rating cognitive growth as of low priority, it should be no surprise that the amount of change produced on such tests is not too impressive.

A much more rigorous test of the process-product relationship can be made if there is some definitive evidence as to the actual program input. Instruments to analyze and categorize classroom events are available for older students. A number of models, such as the interactional analysis of Amidon and Flanders, look at the ratio of teacher-student verbalizations, and other components which are presumed to produce diverse effects in students. For the preschool environment the major reliance has been an anecdotal records of classroom events. The problem with such data is of course that of reducing it to some sort of analyzible form. Abby Sher, working with Professor Edmund Gordon, has been engaged in an extensive series of anecdotal observations with the objective of obtaining meaningful descriptors . John Pierce-Jones at the University of Texas has developed a teacher evaluation instrument called the Observer Rating Form, which was used in the 1966-7 Head Start Evaluation. This looked at characteristics of teacher input, but was difficult to interpret. At Syracuse, Bettye Caldwell has developed a sensitive observational instrument called APPROACH, but this system requires complex coding and is not easily exportable. Don Medley has been developing a system called PROSE which will be used in the E.T.S. Longitudinal study, and Virginia Stern at Bank Street is also developing a classroom observation instrument. Electronic technology vis-a-vis videotaping is also being adapted for observational techniques. There is no lack of activity in the instrument development area and presumably the future holds promise of unlimited data for computer consumption.

However, as of 1967-68, there were no observational systems geared to the preschool level. The responsibility for the development of an objective yet viable system for collecting classroom input data was assigned to the UCLA Center in June 1967, with the requirement that it be ready for national use by the beginning of September 1967. At that time, it was expected that the UCLA instrument would be concerned with the curricular input alone. It was therefore called the Observation of Substantive Curricular Input, or OSCI.

The OSCI is a coding system based on a series of three-minute scans of on-going activity. It requires that two trained observers be present to provide adequate coverage in classrooms where simultaneous activity occurs in different areas, or where some children may be playing outdoors while others are inside. During each three-minute scan, the largest group is located, and four major codes recorded for this group: group size, context of the activity, content of the activity, and locus of control. The context of the activity is the overall setting or situation, such as eating or building; the content code describes the quality of the input taking place within the context. Thus, eating could be a routine, mechanical affair, with children required to sit quietly and eat, or it could be an active learning experience with verbal communication and both sensory experience and content. For example, feeling textures, naming colors, counting pieces of vegetables, talking about food value of vegetables, how they grow, etc. Thus the same context could conceivably have considerably different input value.

All context codes are indicated by single capital letters and content codes by two lower case letters. While the system requires a training period and reliability checks over observers, the coding is closely related to the code meanings, e.g. "B" stands for Building, "la" for language.

The materials used in the activity, whether the child is active or passive, and where the activity is located, either indoors or outdoors, are also coded. Within that same three-minute period, the observer then locates the next group, makes the same records, and proceeds in the same manner until the last individual child unit possible in the time permitted has been recorded. A three-minute record could potentially consist of from one unit (indicating all children were occupied in the same activity) to as many units as there are children, presuming that each child is doing his own thing.

These three-minute scans are repeated on a schedule of seven scans each half-hour, followed by a nine-minute rest period for the observers, then another series of seven scans, until five half-hour periods have been completed. Each daily observation covers approximately the total day for most Head Start classes. To assure sampling across days of the week as well as some seasonal variation, five observation days, each on a different day of the week and approximately four weeks apart, are scheduled. This provides a total of 175 three-minute classroom scans or records for each of the sample classes.

In the 1967-68 evaluation, complete OSCI data were available for 151 classes, providing geographic, ethnic, and population variations typical of the national range. These data have been analyzed for reliability, frequency of occurrence, and factor clusters. The internal consistency reliabilities of most of the basic codes are above .90. The problem of obtaining observer reliability is a difficult one since programs change over time, making test-retest reliability meaningless, and no two observers watch the same events although they may be present simultaneously, thus making cross-observer reliability difficult to estimate, except in special testing situations. This type of reliability check was obtained during the training period, and tended to average between .70. and .90

The data provide frequency distributions of classroom activities which do speak to the question of the diversity of Head Start programs. These data corroborate the fact that Head Start is not a uniform treatment or experience. The classes differ substantially and substantivally in many important ways. The Head Start experiences of children in one class are likely to differ markedly from those of children in another. What impact these differences have is as yet untested, but that the classes do differ is indisputable.

Certain salient features of Head Start can be inferred from the frequency distributions. First of all, comparatively little time was sent on caretaking activities such as clean-up, dressing, and toileting. Over 75% of the classes spent less than 5% of the observed time on these activities. Because categories with low frequency tend to have more than their proportionate effect, these related, low frequency items will be summed into a single factor for the next factor analysis of the data.

Surprisingly enough, virtually no time was observed in undifferentiated activities, that is where the child is wandering aimlessly about and not actively involved. Of the sample, 93% spent less than 5% in such unfocussed activity. However, there was a considerable amount of time spent by certain classes in structured waiting for the teacher to initiate the next activity.

Comparatively little time was spent in teaching preacademic skills such as auditory or perceptual discrimination, other than visual, nor was there much recorded cognitive input in the area of mathematics, social studies, or science. Over 70% of the classes spent less than 5% of their total time on any of these activities. However, considerable time was spent on language development either in formal, structured lessons or in informal verbal communication.

Social interaction and dramatic play are considered to be wide-scope developmental activities, facilitating verbal development, peer relationships, the ability to take the role of the other and thus fostering the transition out of the egocentric, developmental stage. The data indicate that the modal class spent 10% to 25% of the time in dramatic play and 50% in social interaction situations.

The amount of time spent in group vs individual activity also differentiated classes, and seemed to be related to the extent of teacher control. Classes which were primarily structured as whole groups would tend to have a higher degree of teacher control.

In the preliminary factor analyses, there is some indication that sample classes from the same E & R Center tend to have some common component. Eight of the eleven UCLA classes fell into one cluster, as did ten of the thirteen Syracuse classes, eight of the eleven Columbia classes, nine of the nine Kansas classes and six of the nine Chicago classes. Again, this may be due to the fact that a particular activity with low frequency pulled classes together on a relatively non-meaningful variable.

After several more factor analysis runs to produce the most meaningful discriminators, the next step will be to relate the derived classroom characteristics to specific changes in children. For instance, does a class in which there is a high emphasis on language development produce different gain patterns on the verbal sections of the Binet? Do what are traditionally considered to be rich, stimulating environments really make differences in children, compared to those in impoverished settings in which the largest percent of time is spent waiting for the teacher? Do children in classes which are under teacher control, with a great deal of whole or large group activity, demonstrate less growth in social-emotional development? Is role-playing and dramatic play producing any measurable effect, or is this just part of the old nursery mystique? Would the child's time be more profitably spent in cognitive rather than such role-playing activities? Some clue as to this would be gained from looking at changes in children who had been exposed to the cognitive structured programs.

There are many exciting questions to be asked, and hopefully some of the data collected by the OSCI will permit a meaningful answer. Unfortunately, no direct measure of the teacher was obtained, other than that of locus of control, and it is quite possible that the one most important variable may be the teacher. A great deal of demographic information on the teacher is available. Data as to training, age, sex, race, experience, etc., has been collected, but there is no record of the types of controls individual teachers use with the children in the class. In the revision of the OSCI being used for the 1968-69 data collection, this deficiency will be remedied. Alternate three-minute scans look at the total class, then the teacher, and the teacher aide. Information on the types of reinforcing events used, frequency, degree of involvement, and other seemingly important characteristics of teacher-class interactions are being obtained.

If the causal relationships between process and product can be more precisely delineated, it may be possible to plan intervention programs with greatly increased effectiveness, providing guidelines for future curricular planning based on verifiable data rather than on intuitive mystique.