The purpose of the study was to evaluate the usefulness of teacher rating and observational techniques in the evaluation of a preschool program for emotionally disturbed children. Eleven children (nine boys and two girls) participated in the project occurring within the preschool program during one academic year. Teacher ratings were felt to reveal that children enrolled in the program improved in overall functioning during the course of the year. Both teacher ratings and observational techniques were judged to reflect marked gains in social development of the children, and were generally felt to be useful tools for evaluation of such a program. (Author/CD)
INTERIM REPORT

Project No. 422001
Grant No. OEG-2-7-070701-4249

EVALUATION OF A SPECIALIZED NURSERY SCHOOL PROGRAM FOR EMOTIONALLY DISTURBED CHILDREN

Linda Bilsky
Research and Demonstration Center for the Education of Handicapped Children
Teachers College, Columbia University

August 1970

Department of Health, Education and Welfare

U.S. Office of Education
Bureau of Education for the Handicapped
INTERIM REPORT
Project No. 422001
Grant No. OEG-2-7-070701-4249

EVALUATION OF A SPECIALIZED NURSERY SCHOOL PROGRAM FOR
EMOTIONALLY DISTURBED CHILDREN

Linda Bilsky
Teachers College, Columbia University
New York, New York
August 1970

The research reported herein was performed pursuant to a grant with the Bureau of Education for the Handicapped, U.S. Office of Education, Department of Health, Education, and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official positions of the Bureau of Education for the Handicapped.

Department of Health, Education, and Welfare
U.S. Office of Education
Bureau of Education for the Handicapped

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
1970 - 3
The present project consisted of an attempt to acquire information on the usefulness of teacher rating and observational techniques in the evaluation of a preschool program for emotionally disturbed children. Teacher ratings revealed that children enrolled in this program improved in overall functioning during the course of one academic year. Both teacher ratings and observational techniques reflected marked gains in social development.
EVALUATION OF A SPECIALIZED NURSERY SCHOOL PROGRAM
FOR EMOTIONALLY DISTURBED CHILDREN*

Linda Bilsky
Teachers College, Columbia University

Recently, the problem of evaluating preschool programs has become an issue with far-reaching political, as well as educational, implications. Not only have the results of evaluations been expected to provide guidelines for improving educational practices within existing programs, but also these results have been expected to provide justification for the existence of the programs themselves. The situation has been complicated by the absence of appropriate evaluation techniques of established effectiveness. In the case of the present project, which consisted of the evaluation of a fairly unique type of preschool program, existing techniques were strikingly inadequate. Consequently, it was necessary to develop evaluation techniques which were specifically appropriate for this type of preschool program.

The program constituted the central thrust of a multidisciplinary effort to improve the functioning of preschool children who were experiencing serious emotional and social difficulties. A variety of techniques were employed with small groups of children in an attempt to provide a therapeutic nursery school program. The need to evaluate the effectiveness of this program led to the present project which endeavored to develop techniques which (a) would detect behavior changes occurring within this program during one academic year and (b) would be appropriate for the evaluation of similar programs.

Previous evaluations of preschool programs have generally described progress in terms of measures of cognitive development, such as IQ, language

*The work presented or reported herein was performed pursuant to a grant from the U.S. Office of Education, Department of Health, Education and Welfare.
development, or reading readiness (Di Lorenzo & Salter, 1968; Gray & Klaus, 1965). However, the nature of the present program demanded the measurement of a much broader range of performance, which would include social and emotional as well as cognitive behavior. Teacher rating techniques have often been employed as measures of social and emotional adjustment in the classroom. However, although teacher observations are an extremely important source of information about the child's day to day behavior, they are inevitably influenced by the teacher's relationship with each child and the teacher's involvement with the goals of the program. Thus, the present project sought to supplement teacher ratings with an additional, more objective, measure of behavior. Wright (1967) outlined a method which employed trained observers to record fairly complete descriptions of an individual child's behavior during a specified period of time. Although Wright's method was intended to provide ecological data on child behavior, similar techniques have been adapted to provide relatively objective information for use in the evaluation of classroom programs. Werry and Quay (1969) successfully employed an observational technique in the evaluation of an elementary school program for emotionally disturbed children. Werry and Quay's technique consisted of recording the frequency of behaviors in certain pre-selected categories during a specified period of time. Raph, Thomas, Chess, and Korn (1968) trained observers to record the behavior of children in a nursery school program for specified periods of time. They subsequently analyzed these narrative accounts in terms of types of social interactions. The present project sought to employ a combination of rating scale and observational techniques which would provide an objective, as well as comprehensive, record of behavior occurring within a preschool program.
METHOD

Subjects. Eleven children (9 boys and 2 girls) participated in the present evaluation project. All children were experiencing severe emotional and social difficulties associated with manifestations of brain damage, mental retardation and/or other complicating factors. The group ranged in chronological age from 47 to 73 months with a mean of 57.18 months. Four of the children were participating in their first year of the nursery school program and the remaining seven were participating in their second year. Psychological evaluations were administered to all children during the first month of the academic year. Four children were judged to be untestable. The remaining children obtained IQs ranging from dull normal to superior on the Merrill Palmer Scale.

The Program. The present evaluation project was chiefly concerned with performance within a nursery school program which was designed as the central part of a multidisciplinary treatment effort. The nursery school program consisted of two classes, each conducted by one teacher and one assistant teacher, which met for five hours four days a week and for two and a half hours on the fifth day. Three children were enrolled in one class and six children were enrolled in a second class throughout the year; one child was transferred from the first class to the second class after five months; another child was transferred from the first class to a normal nursery school class after six months.

Both classes followed a relatively structured nursery school routine which included activities, described by Fouracre, Connor, and Goldberg (1962); some of these activities were arts and crafts, music, listening to stories, free play, gym, outdoor play, swimming, rest periods, and lunch. Gym and outdoor play periods were conducted jointly with several normal nursery school classes. This
curriculum was utilized to provide opportunities for therapeutic, emotional and social experiences. Activities were structured so that the children could (a) learn to relate to others through shared experiences and (b) increase their personal competence by acquiring a variety of new skills. Moreover, the nursery school activities provided a setting for the implementation of the chief program goal: to improve social and emotional functioning.

The two classes were part of an overall treatment effort designed and directed by a psychiatrist. Additional services were provided by a multidisciplinary staff which met as a group at least twice weekly. This staff included a psychologist who provided evaluational services, a language therapist who provided regular language training for most of the children, and a psychiatric resident who provided individual therapy for one of the children. Several social workers provided counseling for all parents at least once a week and, in addition, conducted a mothers' group which met weekly.

Evaluation Techniques. The evaluational strategy for the present project consisted of an attempt to describe behavioral changes occurring within the preschool program during one academic year. Descriptions of each child's behavior were obtained from two sources, (a) rating scales completed by the child's teacher and (b) observation reports provided by a team of outside observers.

The rating scale techniques involved the use of two rating scales. These scales, designed to measure a child's typical classroom behavior as observed by his teacher, were (a) the Behavior Rating Scale and (b) the Curriculum Guide Rating Scale. The Behavior Rating Scale, which emphasized social and emotional behavior, was developed for the present project in order to supplement the Curriculum Guide Rating Scale, which provided a more general description of behavior. The Behavior Rating Scale contained fourteen items reflecting selected behavioral objectives of the present nursery school program. Each item contained
five statements descriptive of typical behavior related to these objectives. These statements were rank ordered from one to five in terms of increasing social desirability. The rater indicated which of the five statements best described the child's typical behavior in each instance. The Curriculum Guide Rating Scale was modified for the present project from that described by Four-acre, Connor, and Goldberg (1962). After modifications, it consisted of six sections classified as Intellectual Development (26 items), Imaginative and Creative Expression (12 items), Social Development (16 items), Manipulative Development (17 items), Motor Development (12 items), and Self Help (27 items). Each item was rated on a scale from zero to five, with five representing optimal performance on an item. Both the Behavior Rating Scale and the Curriculum Guide Rating Scale were completed for each child in the Fall and again in the Spring. The Spring ratings were completed without reference to the Fall ratings.

Information based on teacher ratings was supplemented by information based on a series of observations which took place in the Fall and Spring of the same academic year. All observers were graduate students enrolled in Psychology or Education programs. The 22 Fall observers participated as a course requirement, whereas the 14 Spring observers participated on a volunteer basis. Ten minute observation sessions were scheduled, according to the availability of observers, so that they occurred within approximately seven weeks in the Fall and in the Spring. An attempt was made to schedule observations during as wide a variety of daily activities as possible. Observers were assigned to children randomly without replacement, with the restriction that an observer should observe the same child no more than twice. Observation reports with missing information were discarded, leaving a pool of 148 Fall reports and 83 Spring reports from which six Fall reports and six Spring reports for each child were randomly
selected. These constituted the 132 reports used in the analyses.

All observers were instructed to unselectively record as much as possible of an individual child's behavior for ten minutes. Interpretations were to be avoided. Observers were required to submit an initial practice observation report which was screened for procedural deviations before the remaining observations took place. Observations of classroom behavior were conducted from observation booths with one-way vision screens and sound amplifying equipment. When observations took place outside the classroom, observers were instructed to remain as unobtrusive as possible.

All reports selected for analysis were summarized (by non-participant raters) according to a classification system developed for the present project. First, each report was divided into units of behavior, defined as behavior segments involving a change of subject, object, or type of behavior. Units were then classified as representing one of the following types of behavior:

1. E-A Emotional behavior with apparent cause
2. E-NA Emotional behavior with no apparent cause
3. I-V Social interaction involving verbal contact only
4. I-C Social interaction involving physical contact only
5. I-VC Social interaction involving both verbal and physical contact
6. A-P Activity which appears purposeful or goal-directed
7. A-HIP Activity which does not appear purposeful or goal-directed

Behaviors in each of the above categories were then rated as either self-initiated or other-initiated; either positive, negative, or neutral in emotional tone; and occurring either with self only, a peer only, an adult only, or a group (i.e., more than one peer and/or adult).

The classification system was applied by two raters; the first rated 22 randomly selected reports which consisted of one Fall report and one Spring
report per child; and the second rated the 110 remaining reports which consisted of five Fall reports and five Spring reports per child. Both raters initially rated several practice reports independently and then negotiated any differences. In addition, each rater independently rated ten randomly selected reports which had been rated by the other rater so that reliability could be established. Dates and names of children had been removed from reports so that this information would not influence ratings.

RESULTS

Teacher Rating Scales. Sign Tests (two-tailed) were applied to each of the fourteen Behavior Rating Scale items in order to determine the significance of Fall-Spring changes as reported by the teachers. Positive changes, significant at the .05 level, were obtained on two items: interaction with others, and participation in group activities.

Sign Tests were also applied to each section of the Curriculum Guide Rating Scale to test for Fall-Spring differences as seen by the teachers. Significant differences were revealed for all six sections of the Curriculum Guide Rating Scale. Improvement, significant at the .01 level, was observed in the areas of imaginative and creative expression, social development, and motor development. Improvement, significant at the .05 level, was observed in the areas of intellectual development, manipulative development, and self help.

Observations. Inter-rater reliabilities were based on 20 observation reports which had been rated independently by both raters. Spearman Rank Order correlation coefficients between raters were obtained for each major behavior category. The inter-rater correlation for the category of negative behavior was significant at the .05 level (\(\rho = .53\)). Correlations for all other categories were significant at the .01 level: emotional behavior (\(\rho = .85\)), social interaction (\(\rho = .89\)), activity (\(\rho = .91\)), self-initiated behavior (\(\rho = .96\)),
other-initiated behavior ($\rho = .88$), positive behavior ($\rho = .85$), neutral behavior ($\rho = .90$), behavior with self ($\rho = .96$), behavior with peer ($\rho = .87$), behavior with adult ($\rho = .93$), and behavior with group ($\rho = .92$).

It was initially determined that the number of behavior units rated in the Spring did not differ from the number of behavior units rated in the Fall ($t = .55$, df = 10, $p > .05$). Subsequently, Sign Tests were performed to test for Fall to Spring changes in the frequency of behaviors classified in each major category. An increase, from Fall to Spring, in behaviors classified as social interactions was significant at the .05 level. None of the other categories revealed significant changes in frequency from Fall to Spring.

**DISCUSSION**

The results of the present project indicated that a number of positive behavior changes had occurred within the specialized nursery school program for emotionally disturbed children. The most striking changes, as seen by both the teachers and the outside observers, occurred within the area of social development. These findings were especially encouraging, because an effort to improve social functioning had constituted a central emphasis of the nursery school program. Additional results, based on the Curriculum Guide Rating Scale, completed by the teachers, indicated general improvement in several other areas including intellectual development, imaginative and creative expression, manipulative development, motor development, and self help skills.

Overall, these findings suggested that the children within the nursery school program were making significant progress in a number of important areas. However, because of the unavailability of a suitable control group, the possibility cannot be ruled out that factors other than the teaching approaches employed in the nursery school program may have contributed to the behavior
changes occurring within the program. On the other hand, the information obtained should be useful in (a) describing behavior which typically occurs within this type of program, (b) indicating behaviors which tended to increase in frequency and indicating behaviors which tended to remain fairly stable during an academic year, (c) providing guidelines for the establishment of teaching priorities within the present preschool program as well as similar programs, and (d) providing information on potentially useful evaluation methods for future use in the present program and similar programs.
REFERENCES


FOOTNOTES

Linda Bilsky is Research Associate, Research and Demonstration Center for the Education of Handicapped Children, Teachers College, Columbia University, New York. The work reported herein was performed pursuant to a grant from the U.S. Office of Education, Department of Health, Education and Welfare. The writer is grateful to Dr. Jerry M. Wiener, Program Director, and to Miss Joan Miller and Mr. Joel Kannengeiser, nursery school teachers, for their extensive cooperation throughout the project. Gratitude is also extended to Miss Carol Gottlieb and Miss Pamela Cantor for their help in developing a system for classifying observational data. The writer also wishes to thank Mr. Thomas Stribling and Miss Lucy Gilbert for their help in the preparation and analysis of the data. Special thanks are extended to Dr. Leonard Blackman and Dr. Ross Evans for consultation throughout the project. The writer also wishes to thank the many observers who served on the project.