

INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.
2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.
3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again -- beginning below the first row and continuing on until complete.
4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.
5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms

300 North Zeeb Road
Ann Arbor, Michigan 48106

75-25,335

HUTCHINS, Dorothy Lee, 1929-
EFFECT OF FEEDBACK ON TEACHERS' ASSESSMENTS
OF YOUNG CHILDREN'S DEVELOPMENTAL STATUS.

Iowa State University, Ph.D., 1975
Education, preschool

Xerox University Microfilms, Ann Arbor, Michigan 48106

Effect of feedback on teachers' assessments
of young children's developmental status

by

Dorothy Lee Hutchins

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major: Child Development

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1975

TABLE OF CONTENTS

	Page
INTRODUCTION	1
Developmental processes	4
Mental Health Services	9
Statement of the Problem	12
Hypotheses	15
REVIEW OF LITERATURE	16
Stability of Developmental Dysfunction	16
Teachers' Ratings	24
METHOD	39
The Research Setting	39
Research Design	41
Subjects	41
Instruments for Assessing Children's Abilities	45
Examiners	54
Teachers' Ratings	55
Feedback	56
Test Administration	61
Procedure	63
Statistical Treatment	65
RESULTS	66
Additional Findings	71

DISCUSSION	78
Research Results	78
Theoretical Implications	83
Methodological Implications	87
Suggestions for Further Research	93
SUMMARY	96
REFERENCES	99
REFERENCE NOTES	106
ACKNOWLEDGEMENTS	107
APPENDIX A. PERMISSION FOR TESTING CHILDREN	109
APPENDIX B. PRESCHOOL SCALE OF MOTOR DEVELOPMENT	112
APPENDIX C. TEACHER RATING SCALE	128

INTRODUCTION

Child development is an integrative field of knowledge drawing from other primary disciplines such as anthropology, biology, sociology, and psychology. The focus of concern is the nature of changes in structure, function, and behavior of children as they develop toward maturity. The field of child development organizes the findings of these basic disciplines and focuses them specifically on questions pertaining to the growth and behavior of children. As a basic science concerned with discovering trends in growth that are characteristic of children, child development provides a foundation for applied services for children, such as child rearing, child guidance, and educational programs.

The past decade has witnessed an expansion of services for young children, the impetus for these services coming from several sources. The federal government, through the Office of Economic Opportunity's Project Head Start, focused the attention of the nation on young children. Research from child development and related fields documents the importance of the early years for subsequent development (Caldwell, 1970; Denenberg, 1970; Hunt, 1961). Educators are considering the possibilities of extending public school programs to include children younger than those traditionally admitted to kindergarten (Lewis, 1973). Pediatricians, along with neurologists and other health care personnel involved in helping the developmentally disabled, encourage the early identification of developmental handicaps (Frankenburg, Dodds, & Fandal, 1970). The community orientation of mental health centers has prompted

the field of psychiatry to attend to the needs of all the community, including young children (Jones, 1972). Social welfare services providing aid to families have extended services by making day care facilities available for young children (Lazar & Rosenberg, 1971).

Ideally it is important not only that services provided for children should be of good quality, but also that the child is matched to the most appropriate program for his needs. The actual program provided for an individual child should follow on the assessment of that child's needs. Therefore, the assessment of the child's developmental status becomes crucial for the purpose of determining the type of program that is most appropriate for the child in order to meet his present needs and to enable him to attain subsequent steps in his development.

The majority of young children can be accommodated in standard types of early childhood programs. However, it is estimated that 10% of all children under six years of age need some type of special service. It is difficult to determine the prevalence of developmental problems in young children since estimates vary depending on the definition used, the population sampled, and the method of identification. Based on a review of major studies on the frequency of maladjustment, White and Harris (1961) suggested that 4% to 7% of public school children exhibit serious maladjustment, as derived from teachers' judgments. Bower (1969) indicated that approximately 10% of public school children in California are sufficiently disturbed to be labeled emotionally handicapped and that 2% to 3% of this group need immediate psychiatric

attention. Stennett (1966) found 22% of fourth, fifth, and sixth grade children moderately or seriously handicapped emotionally. From a 1968 review of research data, derived primarily from teachers' assessment of maladjustment in elementary school children, Glidewell and Swallow (1969) estimated that about 30% of elementary school children presented some problems of maladjustment. Most of these problems were not severe enough to require clinical attention, but about 10% gave evidence of problems that were severe. As many as 20% of children in a lower class Negro neighborhood may exhibit maladjustment (Kellam & Schiff, 1967). Even without an accurate count of how many children have some type of developmental difficulty, it is apparent that this is a sizeable problem.

It is important to devise appropriate techniques for identifying at an early age those children who need and can benefit from various forms of special services. Assessment is not equally difficult for all children. There are many children about whom developmental evaluation is comparatively easy and about whom decisions can securely be made (Flapan & Neubauer, 1970). However, there are other children about whom it is difficult to make decisions. Such difficulties raise a number of issues concerning assessment procedures.

In assessing the development of an individual child it is essential to take into consideration the general characteristics of development as a process. It is here that the basic science of child development can make an important contribution. Research in child development has led to the better understanding of the processes involved in development and the recognition of some fundamental principles that govern development

(Ausubel & Sullivan, 1970; Stone & Church, 1968). While recognizing that much of the research in the field of child development is oriented toward general growth patterns and not directly applicable to developmental deviation in individual children, the general principles formulated on the basis of the results of child development research do have implications for the solution of practical problems.

Developmental Processes

A comprehensive analysis of the processes of development from the point of view of the field of child development has been presented by Ausubel (1958) and Ausubel and Sullivan (1970). Their books integrated findings from other disciplines into an organized body of knowledge focused on the processes involved in the development of children. A consideration of a number of concepts regarding the processes of development is important as a basis for clarifying some of the issues involved in assessing the development of children. The concepts dealing with normative uniformities in development, the issue of continuities and discontinuities of developmental sequences, problems related to parallel and divergent development in different components of development, and the phenomenon of aperiodic regression are essential to the understanding of children's development.

Norms

It is because there are many uniformities which characterize development as a process that it is possible to formulate generalizations about the development of children. One concept of normality is based on a statistical approach in terms of the average. Developmental

norms may be used as a standard against which to evaluate the child's developmental status to provide clues for interpreting the child's behavior and degree of deviation. Comparison to a norm provides a quick measure of the child's development and is frequently used for judging deviation. However, there are a number of considerations relevant to the use of norms as a standard for assessing the development of an individual child. There is actually a wide range for normal development, and departure from the average for a particular child does not necessarily imply pathology. At times that which is not normal or not average may nevertheless be healthy (Flapan & Neubauer, 1970). Research on individual differences in temperament of infants (Thomas, Chess, & Birch, 1968) and the range of reaction (Hunt, Note 1; Hunt, Paraskevopoulos, Schickedanz, & Uzgiris, Note 2) suggest the need to allow for a wide range of variability in development which may be considered to be normal, a point which also has been emphasized by Gesell and Amatruda (1947), Flapan & Neubauer (1970), and Freud (1965).

Continuity

Development also is characterized by both continuity and discontinuity of the normative developmental sequences. Continuity refers to development that is a "process of gradual, quantitative, and continuous change;" discontinuity refers to development that is characterized by "abrupt, uneven, and discontinuous change" (Ausubel & Sullivan, 1970, p. 98). The degree of continuity is influenced by the rate at which development is progressing. A slow rate of change allows for the gradual assimilation of new patterns. During periods characterized by

a rapid rate of change, there is apt to be a greater impact of new variables on previously established patterns which may be abandoned or modified, resulting in a reorganization into a new pattern. However, established patterns are not easily disrupted and, in general, only minor adjustments are made. Continuity also is maintained by other aspects of the developmental process. Some developmental sequences are characterized by the giving up of certain patterns of behavior in favor of emerging new patterns. In such cases there is generally a long period of coexistence, with the patterns of earlier phases occurring concurrently with those of a succeeding phase until the newer pattern predominates, thus alleviating abrupt discontinuity.

While it might facilitate assessment to assume a sequence of phases of development, in some children it is not easy to discover clear, separate phases. For some children there may be an overlapping of phases which makes it difficult to determine whether an apparent disorder has lasted beyond the phase in which it first occurred (Flapan & Neubauer, 1970). It is important that this concept of continuous and discontinuous development be considered when assessing the developmental status of an individual child.

Parallelism

The concept of parallelism and divergence of various components of development has implication for the assessment of children. The inter-relatedness of various components of development is apparent; the child is a complex organism which, however, is not a collection of separate components but develops as a whole (Breckenridge & Murphy, 1963).

Development in one component affects development in all other components in a manner which may either limit or stimulate progression. A limiting effect is noted when progress in one component slows or ceases until growth in other components advances in order to facilitate further progress. On the other hand, acceleration in one component may stimulate the rate of development in another area. Within the various components of development there also may be a common factor which might account for apparent parallelism between them (Ausubel & Sullivan, 1970).

Although the components of development exert a mutual influence on each other, a relationship between them is not necessarily specified for a particular child. Assessment of the total development of a child involves integrating assessments of numerous components of development. It is important to take into account this interaction between components of development or the presence of a common factor within components.

Regression

Regression refers to development that is opposite in direction to the prevailing trend of growth (Ausubel, 1958). In this situation, the child exhibits behavior characteristic of an earlier level of development, after a more advanced level has been attained. If the developmental process is viewed as being one of ongoing progression, a diagnosis of regression would be judged as an interference with development and would not seem to be conducive to the child's further development. Therefore it is necessary to distinguish between true regression and other aspects of development which may be only superficially related to regression. One example of the latter is the presence of a new pattern

of behavior that is not completely stabilized. This occurs when an old and a new pattern are coexisting, with either one being dominant at a particular point in time. While this characteristic of development has already been noted as lessening the effects of discontinuity, it also makes it difficult to determine whether a child is exhibiting true regression. Apparent regression in one component of development may reflect also the individual's preoccupation with a new phase of growth in another component of development. As the child focuses on a newly acquired behavior, he may exhibit a relapse in a previously mastered behavior. However, developmental progress seldom proceeds evenly. Since development is simultaneously progressing in many aspects, interaction is constantly changing. Alternation between progression and regression is to be expected until equilibrium is finally established (Flapan & Neubauer, 1970).

Regression is often seen in periods of behavioral disequilibrium which are associated with normative developmental conflicts; regression may be a transitory reaction to situational stress (Freud, 1965). The problem lies in determining for an individual child whether an observed apparent regression at a specific point in time is true regression, which would be indicative of a developmental problem requiring some help, or is a temporary relapse which can be considered part of the process of normal development.

The explication of these issues forms one basis for the development of applied methods in the field of child development. The child development specialist can find opportunity for implementation of these methods

in a variety of settings, such as day care centers, public schools, and mental health centers. For the most part these organizations have evolved independently without any integrated plan for a generalized system of services for young children. It is at the interface of these organizations that the present study is focused.

Mental Health Services

In recent years, due in part to the Mental Health Act of 1963, the field of mental health has been moving toward a community approach and orientation, reflecting a hope that the effectiveness of mental health resources can be maximized if extended into the community. One component of this community focus is consultation with other caregiving agencies in the community. Mental health consultation provides opportunity for a relatively small number of mental health specialists to indirectly exert a widespread effect through other key individuals in the community.

Consultation services of mental health centers have been directed toward the personnel of various service agencies that are involved with the lives of children, most frequently the schools, but also welfare departments, juvenile and domestic relations courts, recreation departments, and clergymen (Glasscote, Fishman, & Sonis, 1972). The mental health professional serves an important role in helping the teacher gain a better understanding of the process of child development, including the emotional needs of children at particular ages. Increasing the teacher's awareness of any indications that a particular child is experiencing some undue difficulty in growing up is basic to the

teacher's determination of how the child might be helped best. A special approach to the child on the part of the teacher may be effective, or the teacher may recommend to parents that a mental health facility is better able to ascertain whether the child might benefit from specialized help.

In providing services for young children, community mental health centers increasingly focus on prevention. Cowen and Zax (1968) expand on the concept of prevention, delineating three levels. Primary prevention seeks to forestall dysfunction by reducing rates of occurrence in the population and implies also creating social conditions and influences which will promote psychological health and emotional well-being. Secondary prevention seeks to curtail the duration and impact of disorder through early detection and treatment, and may be defined either as prevention early in the life history of the child or early in the course of a given episode of disorder. Tertiary prevention is aimed at entrenched, essentially irreversible pathology, with the purpose of keeping impairment minimal. Consultation to community agencies appears to be an appropriate direction to proceed for all levels of prevention.

At present there is limited knowledge about how to effect primary prevention and encourage and create influences and conditions in society that will promote emotional well-being. It is intriguing to envision the extent to which mental health might be improved by enlisting the aid of mental health professionals in the planning of social programs. However, more immediate efforts of mental health specialists can be directed toward the adults with whom the child has contact. How a child

is treated by his parents and teachers does have an effect on the child's emotional health. There are diverse opinions about specific child rearing and child guidance techniques, making it difficult to enumerate general principles. However, there is substantial agreement about certain physical and emotional needs of children (Birch, 1971; Peters, 1971; Sigel, Starr, Secrist, Jackson, & Hill, 1971) and a belief that meeting these needs appropriately enhances the child's prospects for developing into an adequate adult.

The challenge to mental health professionals is to disseminate knowledge about the development of children to those persons who have direct responsibility for the care of children. Duffy (1970) acknowledged the significant contributions to child psychiatry that have been made by the field of child development, which serves as a basic science for efforts to deal with emotionally disturbed children. The role of a mental health professional and consultant appears to be appropriate for a child development specialist, and in fact early childhood programs can be a prime resource for promoting the mental health of young children (Allinsmith & Goethals, 1962; Caldwell, 1972; Westman, Rice, & Bermann, 1967).

Because of the increasing proportion of preschool children now in attendance in nursery schools, day care centers, and Head Start centers, these are the most likely places for the early identification of those children who may need special help. For many children, attendance at a nursery school or day care center may be the first opportunity for them to be assessed by someone who is knowledgeable about developmental

processes and may be the first time a developmental difficulty is noted. In addition, the opportunity for social interaction with other children often brings into focus or accentuates problems which would not otherwise be observed in the home situation. The recent growth of these preschool programs suggests that it is now possible to identify children who manifest developmental problems before they enter school.

A number of questions have evolved from the experiences of the child development staff of one mental health center program for preschool children. Various individual and group interactions between staff members and teachers in preschool centers led to concern about how these contacts could be more advantageous, both for the evaluation of individual children and for the continuing education of teachers. At the time a referral is made, information elicited from the teacher should facilitate the staff's assessment of the child's development. It is therefore necessary to determine what kind of information is most helpful and can be provided by the teacher. Are teacher's observations and reports valid and reliable sources of information? If the information is appropriate, it should be incorporated into the evaluation, diagnosis, and plan for treatment. If the information that is received from teachers is not appropriate, can teachers learn to make more valid reports of children's behavior? These questions, then, pose the central issue of this study.

Statement of the Problem

The present study evolved from a concern for how a community mental health center can facilitate preschool teachers' early identification of children who exhibit an interference in development, the focus being on

early identification as the first step in prevention. The focus of the present study relates to the abilities of teachers to assess developmental levels of young children. The variables under consideration are the child's cognitive, motor, and social development, as assessed both by the teacher and by standardized assessment instruments. Teachers are provided feedback regarding their assessment of the child's development compared to assessment by standardized instruments. The specific question concerns the effect of this feedback on the teachers' subsequent assessment of children's development.

For the purposes of the present investigation, the following definitions have been employed:

Teachers - those persons in day care centers who have primary responsibility for the care of the child during his period of attendance in the center. Since few states have licensing requirements for teachers in preschool programs, most teachers in day care centers do not have formal training in the provision of care for young children. The term teacher is therefore used here to encompass any adult who is employed by the day care center for the purpose of caring for the children and does not connote degree of training or experience. The day care center teacher is the primary subject of this study.

Cognitive development - the process of acquiring and using knowledge. Included are processes of perceiving, thinking, concept formation, problem-solving, and use of language.

Motor development - the process of acquiring use and control of parts of the body. This study focuses on spatial orientation, balance, and gross motor activities such as walking, jumping, and throwing as the activities which demonstrate motor development.

Social development - the process of acquiring the abilities and characteristics necessary for adequate functioning in interaction with other persons. Skill in initiating, responding to, and maintaining social contacts is an important component of the social development of preschool children.

Teacher ratings - the judgments made by the teacher as to her assessment of the child's status in each of the three aspects of development.

Feedback - information acquired through an individual conference held with each teacher by the principal investigator. The teacher is presented the results of her ratings of each child's development and the results of the child's performance on the standardized instruments. This conference includes an analysis of agreements and differences between the two types of assessments of the child's development.

The study is designed so that each teacher rates the cognitive, motor, and social development of half of the children in her class. Then the cognitive, motor, and social development of all children in her class will be assessed by standardized test instruments. The teacher then will be given feedback concerning the relationship between her

ratings of half the class and those children's test results. After the feedback intervention, the teacher will rate the other half of the children in her class.

Hypotheses

The following null hypotheses are formulated:

Primary hypotheses

1. Agreement of preschool teachers' ratings of children's cognitive development with objective test data is not changed by feedback on rating behavior and test results.

2. Agreement of preschool teachers' ratings of children's motor development with objective test data is not changed by feedback on rating behavior and test results.

3. Agreement of preschool teachers' ratings of children's social competencies with data from a standardized rating instrument is not changed by feedback on rating behavior.

Ancillary hypotheses

4. Agreement of preschool teachers' ratings of children's development with objective test data is independent of teachers' age, education, or experience.

5. Intercorrelations of teachers' ratings of specific developmental characteristics, i.e., cognitive, motor, and social, is not changed following feedback on rating behavior.

REVIEW OF LITERATURE

The purpose of the review of literature is to provide background for the present research project. Stability of early developmental dysfunction and teachers' ratings of children's behavior are considered to be relevant for the present study and will be surveyed in the review.

Stability of Developmental Dysfunction

Educational, psychological, and psychiatric literature in recent years has emphasized the need for identification at an early age of those children who exhibit difficulty in their developmental processes (Beach, Cowen, Zax, Laird, Grost, & Izzo, 1968; Flapan & Neubauer, 1970; Zax & Cowen, 1967). This need is based on the assumption that such children will continue to have problems and are likely to grow up to be adults with problems. The assumption raises the issue of whether difficulties noted in the early years of childhood are stable or transitory. Transitory difficulties may be remedied by the passage of time. If difficulties are stable, they will be ultimately damaging and help should be provided as soon as possible (Cowen, 1973). However, research findings on the stability of developmental dysfunction are not entirely consistent.

MacFarlane, Allen, and Honzik (1954) concluded from a study of the behavioral problems of normal children that most problems do not persist but rather tend to be transitory. While pointing out the methodological difficulties and limitations of retrospective studies, in which mentally ill adults are selected as subjects and their childhood histories reconstructed, Clarizio (1969) concluded from a review of a number of these

studies that they offer only moderate evidence that a maladjusted child grows up to be a maladjusted adult. In his review of follow-up studies of children who were earlier seen in child guidance clinics, he concluded that the extent of adult disturbance depended on the nature of the childhood problem and the child's environment. Aggressive, anti-social behavior and acting-out behavior were more likely to predict later adult disturbance, whereas shyness and withdrawn behavior tended to dissipate with increasing age.

On the other side of the issue, a number of longitudinal studies attest to the stability of early difficulties. Results of several such studies indicate that school maladaptation is not transitory. Only about one-third of a group of school children in kindergarten through fourth grade, identified by teachers as being maladjusted, improved spontaneously over a three-year period (Brownbridge & VanVleet, 1968).

Feldhusen, Thurston, and Benning (1969, 1970) had teachers of the third, sixth, and ninth grades identify four children (two boys and two girls) in their classes who consistently displayed aggressive and disruptive behavior in the classroom (a total of 568) and four who displayed socially approved behavior (a total of 982). Of these 1550 children selected by the teachers, 384 were randomly chosen for intensive study, which included interviews with the children and their parents, psychological testing, and analysis of school records. At the end of five years, academic development, as reflected in school achievement, was assessed for those children who were now finishing grade eight or eleven. Data were obtained for 240 who had been

selected by the teachers five years earlier and for 160 who had been studied intensively. Analysis of covariance indicated that those children who had been identified earlier as being disruptive and aggressive were achieving less well than the children who had exhibited socially approved behavior, based both on teachers' grades and on scores on Sequential Tests of Educational Progress. These achievement differences held for both boys and girls, at all three grade levels, and for both urban and rural children. The authors concluded that a host of frustrations generate and perpetuate aggressive-disruptive behavior which in turn has a deleterious effect on the acquisition of academic skills.

Zax and Cowen (1967) conducted a series of studies to look at the effects of early school maladjustment. They first identified children in the first grade in two elementary schools who, in terms of psychological test data obtained during social work interviews with mothers and actual observed behavior, seemed to have already manifested moderate to severe maladjustment, or who were considered to be in a group in which there was a high probability that such pathology was incipient. These children, 37% of the first grade classes, were designated as "Red-Tag" and compared with the other children in their class (Non-Red-Tag) at the end of the third grade. The indices used for evaluation were school absences, referrals to the school nurse, grades, SRA tests, Achievement-Aptitude Discrepancy Scores, teachers' ratings of behavior, peer evaluation, and psychological tests and overall adjustment ratings made by the research staff. Of a total of 20 criterion measures examined, differences favoring the Non-Red-Tag group emerged for fourteen of the

comparisons. By the end of the third grade, the child with an early diagnosed emotional disorder had suffered serious impairment in the academic, achievement, adjustment, and behavioral spheres.

These same children were studied in a follow-up investigation four years later when they were in the seventh grade (Zax, Cowen, Rappaport, Beach, & Laird, 1968). Of 47 comparisons made, 10 differences were significant beyond the $p=.05$ level, with the Red-Tags obtaining negative scores. For 30 of the 37 non-significantly different comparisons, the direction was for more negative scores by the Red-Tags. The children who manifested signs of maladjustment very early in their school career earned poorer grades, did less well on achievement tests, were referred to the school nurse more frequently, and were judged by their teachers and peers to be more poorly adjusted. These studies also demonstrated that the identification of potential for maladjustment can take place as early as the first grade.

To investigate the stability of more severe disorders, Havelkova (1968) conducted a longitudinal study of 71 children who had been diagnosed in their preschool years as psychotic, manifesting symptoms of early infantile autism or childhood schizophrenia of the autistic or pseudoneurotic form. Of the 71 children in her study, 29 were diagnosed as moderately affected and selected for treatment provided by an individual therapist in a day care center for disturbed children, 17 children diagnosed as severely affected were considered untreatable by the methods available in the center, and 25 mildly affected children did not receive treatment due to lack of space and the likelihood that

they could make progress in other settings, the majority being referred to ordinary nursery schools; some of these children later received treatment elsewhere.

Follow-up observation of these children and analysis of their placement (home, institution, or treatment center) and type of school class (normal, opportunity, or retarded) at the end of a four-year period indicated that the mildly ill children tended to improve spontaneously and generally were not admitted to a treatment program. Borderline cases seemed to improve with maturation, nursery school attendance, and an occasional follow-up interview. At the end of the four-year period, 41% of the children remained autistic.

A general pattern for the course of development for those children who improved was a change from the autistic to the pseudoneurotic form of the illness. Those children who made this change at an earlier age (before four-and-one-half years) had a tendency toward better intellectual functioning ultimately than those who changed at a later age. Of those children who began treatment early, one-half were able to function in normal classes. Although there was some doubt as to how the treatment changed the pattern of the illness, treatment that was begun early did improve the child's ability to attend school. Havelkova suggested that the treatment may have enabled better utilization of the child's original intellectual potential. The better outcome of the early-treated children as compared with the late-treated mildly ill children suggested that there is a critical period for treatment. Treatment appeared to be more successful for children who received it early. These results confirm

the stability of extreme forms of dysfunction, which appear to be partially alleviated by early and intensive treatment.

Data supporting the stability of early dysfunction is provided by retrospective reports. In a study by Bower, et al. (1960), adult schizophrenics had significantly poorer high school records than controls. A group of 44 hospitalized male mental patients, all of whom had been diagnosed as schizophrenic, was compared with a control group composed of high school peers. High school records and interviews with teachers and counselors who had known the boys provided information for characterizing the boys' behavior and performance in high school.

Differences in the school staff's perceptions of the preschizophrenic boys differed from their perceptions of the control group. The preschizophrenic boys were described as being more passive toward others and toward the school environment, with significant differences at the $p=.01$ to $.05$ level being obtained showing the preschizophrenic boys to be more apathetic, careless, dependent, irresponsible, depressed, submissive, and anxious, as well as less-well-liked by peers. Although in most cases the school staff was aware of beginning character disorders in these boys while they were in high school, only a small number of the boys were perceived by school personnel as being emotionally ill or having major problems. However, teachers and counselors did rate the preschizophrenic boys poorer than the control boys on their school adjustment and overall mental health at the $p<.001$ level. Only three of the 44 boys in the preschizophrenic group were rated by school faculty as above average in mental health as compared to 23 of 44 controls.

From middle class children who had attended a private nursery school, 130 were selected who could be followed through high school in the local school system (Westman, et al., 1967). Data were obtained from the records of the nursery school teachers and the cumulative school records of the children through the high school years. Three clinicians independently rated these records and evaluated each child's school adjustment. Information regarding the use of special mental health services also was obtained from the school pupil personnel division, the files of community psychiatric clinics serving children, and psychiatrists in private practice in the community. A correlation of .88 was found between a low school adjustment rating and later use of mental health services. Westman, et al. suggested that this evidence indicated that children did not outgrow behavior problems seen in early life but that children with adjustment problems in nursery school tended to have adjustment problems in later school life, and that the problems tended to be of the same order.

Robins (1966) obtained clinical records for over 500 children who were referred for evaluation to a child guidance clinic in the 1920's. A detailed study was reported for a follow-up of this group 30 years later when 90% of the children were located and 468 were interviewed individually. In addition, 100 adults who had not been referred as children were selected as subjects for a control group and interviewed. All interviews for the patient and control groups were analyzed by a psychiatrist to assess adjustment as adults.

As adults, 52% of the control group appeared to be free of psychiatric disease compared with only 20% of the patient group. At follow-up, 8% of the control group had seriously disabling symptoms compared to 34% of the patient group. The maladjustment of the patients was evidenced by higher incidence of arrests, mental hospitalization and psychiatric symptoms, alcoholism, divorce, alienation, use of welfare services, transmission of behavior problems to their children, and lower occupational achievement. Those children who had been referred for anti-social behavior differed from those children who had been referred for temper tantrums, learning problems, speech difficulties, sleep and eating disturbances, and all problems other than antisocial behavior, by exhibiting more serious disturbance as adults. The more severe the childhood antisocial behavior, the more disturbed was adult adjustment. Robins concluded that it was the nature and severity of the childhood behavior leading to referral which created the later problems in adulthood.

Although the clinic psychiatrist had recommended either out-patient treatment or psychiatric hospitalization for most of the children who had been evaluated, clinic records showed that only 28% of the children received any treatment. Treatment appeared to be most effective for children with moderate antisocial behavior. Prognosis was good for children who exhibited little antisocial behavior, whether they were treated or not. For children who showed much antisocial behavior, prognosis was poor whether they obtained treatment or not.

The findings presented by these studies are not consistent and attest to the complexity of the issue of the stability of early difficulties. The evidence suggests that it is not likely that the course of all types of early dysfunction is the same. Aggressiveness, for example, seems to be more stable than shyness (Clarizio, 1969; Robins, 1966). In general, the more deviant the dysfunction, the more apt it is to endure into adulthood (Robins, 1966). In addition, it does not appear likely that most early difficulties will dissipate spontaneously (Cowen, 1973). The emphasis on the need for early identification and intervention is justified by the evidence of the durability of serious early developmental dysfunction, a justification supported by Brownbridge and VanVleet (1968), Cowen (1973), Feldhusen, et al. (1969), Havelkova (1968), Robins (1966), and Westman, et al. (1967).

Teachers' Ratings

The use of rating scales has a long history in the fields of mental health and education but there has not been a full exploration of all the issues regarding the development and use of rating scales (Severson, Note 3). One question of interest is how teachers' ratings agree with other methods or sources of data about the behavior and characteristics of children. Conflicting results are reported for the reliability of judgments made by different raters.

The reliability of ratings made by two psychiatrists was reported in a detailed study of the mental health of a city population (Srole, Langner, Michael, Opler, & Rennie, 1962). The psychiatrists evaluated the same 1660 individuals (adults) by rating the descriptions made by

social workers based on a personal interview. The raters agreed on the behavior of individuals who exhibited either extreme disability or freedom from significant symptoms, but did not agree on pathology at the intermediate levels. There was 100% agreement on 47% of the cases. When a difference of only one step on the rating scale was allowed, there was agreement on 86% of the cases. It appeared that as the severity of symptoms increased, so did the inter-judge reliability.

Teachers' ratings were compared with ratings of other professionals in a study by Miller (1964). Using a Q sort technique, he obtained reports of specific behaviors of 36 children, ages six to ten years. Of these children, 18 had been referred by their parents to a child guidance clinic. For each of these, a same-sex sibling who had not been referred was selected for a control group. A Q sort of 80 statements, half referring to healthy traits and half referring to pathological traits, was completed for each child by five types of observers: fathers, mothers, teachers, psychologists, and psychiatrists. The only consistency of agreement was between parents (.60); the clinicians and teachers showed the least consensus (.24). There was a higher degree of agreement among the judges on the behavior of the healthy child (.49) as contrasted with the disturbed child (.28). Interrater agreement decreased as the mental health of the child was judged more pathological. Miller concluded that an observation could not be communicated with any confidence beyond the person making the observation.

Both educators and mental health specialists have been interested in the ability of teachers to identify children with poor mental health. In an early classic study, Wickman (1928) compared ratings made by teachers and mental hygienists as to the seriousness of a number of behavior problems of children. The investigation was first directed toward securing from teachers their opinions on what constitutes children's undesirable behavior. Teachers in two elementary schools in different communities were requested to list all the kinds of behavior problems which they had encountered at any time during their teaching careers.

The 27 teachers in the first school reported 428 specific problems of behavior. When duplications were eliminated, there were 185 separate items of undesirable behavior. The responses were classified according to similar types of problems into seven major groups, each containing sub-headings. The seven groups were arbitrarily determined and were established for convenience in analyzing the results and for obtaining a comprehensive classification of behavior problems in which synonymous and closely related descriptions were combined. The seven groups were:

- I. Violations of general standards of morality and integrity
- II. Transgressions against authority
- III. Violations of general school regulations
- IV. Violations of classroom rules
- V. Violations of school work requirements
- VI. Difficulties with other children
- VII. Undesirable personality traits

Comparison of the reports from this group of 27 teachers with the reports of 29 teachers in a second community showed the two lists to be in essential agreement. The items of problem behaviors reported by the second group of teachers covered all of the seven major groups and their sub-headings with one exception. No teacher in the second group described the "dissatisfied" or "unhappy" child. However, both of these descriptions were indicated only once in the first sample, and then by the same teacher. A few additional descriptions of behavior problems listed by the second group of teachers could be included as part of the seven major groups.

To assess the adequacy of the reports of behavior problems recognized by teachers, these lists of behaviors were compared with the records of 1123 cases in two Child Guidance Clinics to which parents, teachers, social agencies, and juvenile courts referred children with behavior problems. All of the items of undesirable behavior listed by the teachers were reported also in the case records of the children referred to the clinics. The clinical referrals included some additional behaviors which were not listed by the teachers. Those behaviors for the most part appeared in cases referred to the clinics by parents, and included such behaviors as suspicious attitudes, not confidential, being easily influenced by companions, fearlessness, lack of appreciation of danger, cowardliness in physical activities, muscular inabilities and incoordination. Parents also showed more concern than did the teachers for neurotic habits of children, such as enuresis, nailbiting, thumbsucking, fears, tics, nervousness, and problems of eating and sleeping.

On the basis of the teachers' lists of behavior problems, behavior rating scales were devised to assess the frequency with which various kinds of troublesome behaviors were recognized and reported by teachers. In one of the elementary schools, teachers rated each of their pupils on a scale which contained the 16 items of undesirable behavior which the teachers had enumerated most frequently in their lists of problems. In the second school, the scale was composed of 51 items, reflecting the entire range of problems which the teachers had previously described. On both scales, the teacher's response to each item was made with reference to four statements which defined the frequency of the occurrence of such behavior in the individual child: has never occurred, has occurred once or twice but no more, recurs with occasional frequency, has become almost habitual. On the more comprehensive rating scale, the teacher also indicated for each pupil the amount of difficulty the behavior produced. The teacher's view of the seriousness of a behavior problem thus was evaluated by having the teacher rate whether the behavior: is of no consequence, causing only slight difficulty, causing considerable difficulty, or causing very serious difficulty. A total of 801 pupils were rated in the first school, 874 in the second.

These reports of the teachers on the occurrence of undesirable behavior in their pupils indicated that the teachers were most aware of those problems which affected the child's school tasks and interfered with the purpose of teaching. Whispering was reported to occur in 75% of pupils and considered to be frequent or habitual for 40% of the pupils. Other overt types of behaviors which were reported most

frequently were: inattention, 60% of all pupils; careless work, 44%; disorderliness, 38%; failure to study, 36%; over-activity, 30%. These behaviors were reported to be frequent or habitual for 16% to 30% of all pupils. Problems of children which do not interfere directly with classroom procedures were reported less frequently. On the 51 item behavior rating scale, 14 items were reported to occur in more than 25% of the children. Of these 14 items, only three--shy and withdrawn behavior (35%), daydreaming (33%), and oversensitiveness (25%)--were not overt behavior or aggressive personality traits. These traits were rated by the teachers as causing little difficulty.

To secure the opinions of mental hygienists, 30 clinicians who were all actively engaged in the study and treatment of behavior disorders of children were selected. This group was composed of eight psychiatrists, four psychologists, 13 psychiatric social workers, and five experienced teachers with training in social case work and mental hygiene. The mental hygienists rated the same 51 behavior problems which had been incorporated in the teachers' rating scales. However, the mental hygienists were directed to rate the significance of each problem behavior in terms of its effect on the future life of the child.

Analysis of the results of the ratings showed no significant differences in the ratings submitted by the mental hygienists. The coefficient of correlation between the rank-order arrangements of the seriousness of the problems as rated by the mental hygienists and teachers was $-.22$ for the first group of teachers and $-.11$ for the second group. There was a complete reversal in estimating the seriousness

of problems describing the withdrawing, recessive personality and behavior traits. Whereas teachers considered shyness, sensitiveness, unsocialness, fearfulness, dreaminess among the least serious of all problems, the mental hygienists ranked them together with unhappiness, depression, resentfulness, physical coward, suggestible, and overcritical at the very top of the list as the most serious problems. These items in the mental hygienists' ratings completely replace the problems relating to sex, dishonesty, and disobedience which the teachers' ranked as most serious.

However, the problems rated most serious by the teachers are not entirely reversed in the mental hygienists ratings. Some of the problems considered serious by the teachers were also regarded as fairly serious by the mental hygienists (cruelty and temper tantrums). Problems relating to difficulties in adjusting to school work--lack of interest, laziness, carelessness, inattention--were all reduced in degree of seriousness by the clinicians. Problems designating failure of pupils to follow classroom discipline, like disorderliness and whispering, were also greatly discounted in importance by the clinicians and rated as the least serious of all problems.

Wickman concluded that the reactions of the teachers were personal and emotional responses to children who exhibited problems in behavior and suggested that the teachers appeared to be reacting primarily to the undesirable behavior of the child rather than to the child as an individual. Wickman pointed out that in interpreting the teachers' responses, consideration should be given to the nature of their teaching

responsibilities and their sensitivity to the overt types of behavior and aggressive personality traits that interfere with the purposes of teaching.

A number of later researchers assumed that Wickman's results indicated that teachers did not perceive behavior from a mental health point of view. However, as Wickman carefully pointed out, different instructions had been given to the teachers and to the mental health specialists. A list of the same behaviors was presented to each group, but teachers were asked to rate the seriousness of each behavior on the basis of its present effects, while the mental health specialists were asked to rate the seriousness of each behavior as to future effects. These differences in instructions, which appeared to cause some of the disagreement between teachers and mental health specialists, were often overlooked in analyzing the results of this and later studies.

Subsequent studies focused on this supposed discrepancy between teachers' and clinicians' ratings of the seriousness of certain behavioral difficulties. In general these subsequent studies reported greater teacher concern about withdrawal than Wickman indicated.

Peck (1935) had 175 teachers write a case study of an unadjusted pupil, telling why the pupil was considered maladjusted, explaining the factors tending to cause the maladjustment, and stating what should be done for the child. This method called for the teachers to report their reactions to the specific problems of actual children. Peck suggested that such a method might be expected to represent more truly than a behavior check list the real attitudes of the teachers toward the behavior

of their pupils. From the 175 children described, 698 problems of 90 different types were reported. In addition to classifying these problems into the seven groups listed by Wickman (1928), an additional subheading was included under the classification of "Undesirable personality traits" to allow for the grouping of a number of problems not suggested by the teachers in Wickman's study. This subheading was titled "A-typical mental traits" and included such items as: seeming dullness or mental deficiency, speech difficulties, epilepsy, birth injury to brain, mirror writing, specialized inability in one school subject, and ability seemingly too high for adjustment to ordinary school program.

Undesirable personality traits were reported as the chief reason for considering pupils maladjusted and included 53% of the problems. Further classification of this group of problems showed regressive traits as 31% of the total; aggressive traits, 9%; and a-typical mental traits, 13%. Sixteen percent of the 698 problems were classified as violations of school work requirements and 16% as violations of conventional standards of morality and integrity. The other categories had less than 7% each. Peck concluded that teachers judged aggressive behavior as less serious than withdrawal, emphasized the seriousness of undesirable personality traits, placed less emphasis on violations of moral standards and school work standards, and considered the various class of disciplinary offenses as least serious. These findings do not concur with Wickman's study.

In a study of the adjustment of elementary school students, Gordon (Note 4) compared ratings made by classroom teachers with ratings made

by a committee composed of the school principal, a counselor, a psychologist, and two special teachers. Of 80 children judged by the classroom teachers as maladjusted, 76 were rated by the committee as clearly disturbed. Six other children were judged to be maladjusted by the committee, but of these six missed by the teachers, five were children who presented no overt behavior problems at school.

It has been suggested that the results of rating scales are influenced by the directions given to the raters. Ellis and Miller (1936) compared the results secured by Wickman (1928) with results secured under changed conditions from 382 junior and senior high school teachers. In their study, the directions to the teachers were changed to secure ratings as to the seriousness of each problem for the future development of the child, essentially the same instructions that had been given by Wickman to the mental hygienists. Wickman had reported a rank order correlation of $-.80$ for the two sets of ratings by teachers and mental hygienists. Ellis and Miller reported a correlation of $.49$ between the ratings by the teachers in their study and the ratings by the mental hygienists in Wickman's study and a correlation of $.65$ between the teachers' ratings and the ratings made by Wickman's teachers. The most important difference found between the ratings of teachers in Ellis and Miller's study and those of Wickman's study is the increased realization of the seriousness of the withdrawing, recessive personality traits.

Other studies (Cutler, 1961; Schrupp & Gjerde, 1953) also suggest a trend during the thirty-year period following Wickman's 1928 report in the direction of more agreement between teachers and clinicians on

the relative importance of various kinds of problem behaviors. This change in view appeared to have occurred in both professions (Mitchell, 1942).

Kellam and Schiff (1967) had each of 3000 first grade children interviewed and observed in small groups in a play situation by two psychologists, who made independent ratings of the clinical status of each child on a six-point scale which ranged from "not clearly sick" to "extremely sick." Teachers rated the same children on their adaptation to the classroom, using six scales of adaptation: social contact, authority acceptance, maturation, cognitive achievement, concentration, and global adaptation. Children rated as sick by the clinician were viewed by the teachers as less adapted on the social contact, maturation, and global adaptation scales.

Other studies have compared reports of teachers with reports made by parents concerning behavior of children. In a survey of behavior difficulties of 900 public school children, ages 6 to 14 years, Griffiths (1952) collected symptom reports from parents, teachers, and children, and found a correlation of .90 between ratings of children's adjustment made by parents and teachers.

Schrager and Lindy (1970) screened 494 kindergarten children to select those who were hyperkinetic. They first surveyed pediatricians, teachers, psychologists, psychiatrists, and social workers to determine which specific behaviors were most commonly considered characteristic of the hyperkinetic child. Six behavioral attributes were selected: being fidgety and restless, inattentive, hard to manage, unable to sit still,

pay attention, or tolerate frustration. A checklist of 44 items was devised in which these six key behaviors were embedded. This was distributed to parents at the time of admission to school and to the eleven classroom teachers of the children four weeks later. Ratings on a composite of the six key behaviors were tallied for each child, resulting in an index of hyperkinetic behavior. All children were placed on a theoretical continuum from "nonhyperkinetic" to "high hyperkinetic." Chi square analysis indicated that hyperkinetic children had a significantly higher rate of absence from school without evidence of poorer health than their nonhyperkinetic classmates. They also exhibited poorer performance on standardized measures predicting readiness for school (Metropolitan Readiness Test and Lee-Clark Reading Readiness Test). It is interesting to note that, based on chi square analysis, there was strong agreement between teachers and parents for five of the six key behaviors, the exception being "unable to tolerate frustration."

Reports of parents and teachers about a sample of 36 children were compared by Del Solar (1949). In general parents expressed more satisfaction with the children than the teachers did. Parents and teachers were in close agreement about the children's adjustment to school, intellectual abilities, personality traits, social skills, and artistic ability. There was less agreement about problems of eating, sleeping, other routines, and difficulties with siblings--all behaviors which do not lend themselves to teacher observation.

Bower (1969) pointed out that the different roles of the teacher and the clinician could influence their interpretation of children's behavior. The teacher of necessity is involved with the total group of children in her class and the effect that the disturbing behavior of one child has on the group is of immediate concern.

The validity of teachers' reports as a mental health screening method has been examined (Bower et al., 1960; Glidewell, Mensh, & Gildea, 1957; Ullmann, 1952). These studies compared teachers' reports of children's behavior with ratings made by mental health specialists and reported substantial agreement between them (70%-80%). However, Goldfarb (1963) found less agreement (60%) between teachers and psychiatrists in their selection of school children who needed referral to the Division of Special Services and concluded that teachers could not adequately substitute for psychiatrists as case finders.

On the other hand, in screening 1200 kindergarten children for potential learning disabilities, Haring and Ridgway (1967) found that 48 kindergarten teachers were able to select those children with developmental retardation as confirmed by evaluation of the psychological staff. The teachers were provided with a structured guide to observation that specified areas of performance, such as eye-hand coordination, auditory and visual memory, and language. These investigators suggested that the results of the study confirmed the value of the teacher's role in the early identification of children with learning disabilities, and further suggested that an analysis of individual

performance behaviors was a valid basis for the selection of children who were developmentally retarded.

In an investigation of the use of the School Behavior Check List for elementary school children, Miller (1972) included teachers' judgments of the behavior of the children in their class. The teachers exhibited a "halo effect" in which all behavior was judged in terms of the child's academic competence. As competence decreased, more deviance in behavior was observed. The teachers appeared to be rating along an achievement-competence dimension.

Supporting evidence that teachers, utilizing symptom check lists, could make reliable identification of maladjusted children in the early grades was presented by Quay, Morse, and Cutler (1966) and Rubin, Simson, and Betwee (1966).

Eisenberg, Landowne, Wilner, and Imber (1962) studied the use of teacher ratings as a method for measuring the effectiveness of a therapeutic nursery program. With the use of a symptom check list and a health inventory, nursery school teachers were successful in detecting psychopathology in preschool children. On the basis of their study of nursery school adjustment, Westman, et al. (1967) suggested that observations by sensitive nursery school teachers were sufficiently accurate to indicate that nursery school teachers can play a key position in mental health screening.

It seems appropriate to conclude from the reported research that the teacher's knowledge of children's functioning can be utilized in assessing the development of children. This conclusion is supported

by Bower (1969), Eisenberg, et al. (1962), Glidewell and Swallow (1968), Haring and Ridgway (1967), Quay, et al. (1966), Rubin, et al. (1966), and Westman, et al. (1967).

METHOD

The purpose of the present study was to investigate the relationship between teachers' ratings of preschool children's development and the children's performance on standardized testing instruments. The specific question dealt with the changes teachers make in ratings after receiving feedback relative to the agreement of their ratings with objective assessment of the children's performance on standardized instruments. The variables studied were teachers' ratings of children's cognitive, motor, and social development, and children's performance on standardized instruments for assessing cognitive, motor, and social development. The experimental intervention consisted of a feedback discussion session with each preschool teacher in which the information focused on the discrepancies obtained between the teacher's rating in the three areas of development and the measures obtained by standardized instruments.

The Research Setting

The Child Development Unit of the Children and Youth Division of Mid-Missouri Mental Health Center is directed toward the prevention of preschool-age children's emotional problems and delayed personality development through a specialized treatment program designed for early intervention. The program is adapted to the preschool-age child who, because of organic or environmental influence, chronically displays any of the following: (a) inability to learn at a rate commensurate with his intellectual, sensory-motor and physical development; (b) inability to establish and maintain adequate social relationships; (c) inability to respond appropriately in day to day life situations; (d) a variety of

excessive behaviors ranging from hyperactive, impulsive behavior to depression and withdrawal.

In the first year and a half of operation of the Child Development Unit, 92 children were referred for evaluation of their developmental progress. Eighty of these children were between the ages of three and six years, the typical nursery school or kindergarten age range. Of these 80, 53 children were referred by a kindergarten, nursery school, or day care center. Thus it can be seen that over one-half of the referred children had contact with a preschool group. In addition, requests were frequently received by the Child Development Unit from teachers in kindergartens, nursery schools, or day care centers to consult with them concerning children about whom they had some concern.

For most referrals in which the child attended some type of group child care program, a visit by a member of the Child Development Staff to observe the child in the group situation and confer with the teacher was considered important. In some cases the teacher completed a school referral form which provided initial information concerning the child's behavior in the group.

In addition to consulting with the teachers of these various child care programs on specific cases, there was a number of contacts by the Child Development Staff with groups of teachers, as well as with university students in the fields of child development and education. The purpose of these meetings was to assist teachers and teachers-in-training in understanding the developmental needs of children and how teachers could provide a program designed to promote good mental health of

children. Efforts also were made to aid the teacher in the early detection of those children who exhibit a developmental delay.

The Mid-Missouri Mental Health Center offers a setting for research incorporating the important components of the processes necessary for early identification of children with developmental difficulties: a commitment to community service, contact with preschool agencies, and child development consultants. The principal investigator of the present research was one of the child development consultants.

Research Design

The investigation utilized a One-Group Pretest-Posttest Design (Campbell & Stanley, 1966). The subjects were preschool teachers. The pretest and posttest were comprised of the teachers' ratings of children's development. The independent variable was an experimental intervention between the pretest and posttest in the form of feedback to the teacher concerning her ratings and the children's performance on standardized instruments. The dependent variable was change in the agreement between the teacher's ratings and the objective test data as a function of the experimental intervention. In this design each subject serves as her own control.

Subjects

Subjects in this study were 14 teachers in five day care centers in the same community.

Selection of centers

In selecting centers to participate in this study, consideration was given to the number of children and teachers in the center and the likelihood of securing cooperation. To facilitate data collection,

it was preferable to select centers which employed several teachers. Thus there was the added advantage of further reducing variance attributable to environmental conditions. Also an attempt was made to select as a subject a teacher who had in her class eight to ten children between the ages of three and five-and-one-half years in order to have a minimum of four children for rating in both the pretest and posttest groups.

The staff of the Child Development Unit had previously consulted with the director of one of the potential centers and her support for the study was readily obtained. She not only enlisted the cooperation of the three teachers in that center but also suggested other center directors who likely would be interested in participating in the project. With the above considerations, five centers were selected. Table 1 presents the composition of these centers as to number of adults and children.

Table 1. Composition of Day Care Centers

Center	Director	Full-time Teachers	Part-time Teachers & Aides	Children	Type of Support
1	1	3	yes	60	Community
2	1	3	yes	40	Community
3	1	2	no	30	Private
4	1	3	yes	70	Community
5		2	no	25	Private

Organization of centers The three larger centers are organized so that each teacher is assigned specific children, based on the age of the child. For example, in Center 1, Teacher 1 has children ages two-and-one-half to three-and-one-half years, Teacher 2 has children ages three-and-one-half to four-and-one-half years, and Teacher 3 has those children over four-and-one-half years of age. The teachers are responsible for providing activities for children in their group and the children usually stay together as a group for eating and napping. More than one group may be on the playground at the same time and there is interaction between the groups in the less formal activities. Therefore, each teacher knows most of the children in the center but is more familiar with the characteristics and behavior of the children assigned to her class.

The two centers with smaller numbers of children (Centers 3 and 5) operate on an open system in which both teachers help with the supervision of all of the children in the center. In Center 5, the owner-director is also a full-time teacher. In these two centers, teachers are familiar with all the children.

Selection of teachers

Teachers selected for this study were those who had day-to-day contact with children between the ages of three years and five-and-one-half years. On this basis, directors were excluded, except the director-teacher in Center 5, who had direct daily contact with the children. In one center a teacher was excluded who had only begun working there on the day of testing. Another teacher was excluded because she was

resigning before the data collection period could be completed. Also excluded were aides who were in the center a few hours each week and did not have direct responsibility for children. Some centers used teachers on a part-time schedule and these were included.

Selection of children

After selection of teachers had been made, the roster of children enrolled in each center was examined to select those children who were to be rated and tested. Children who were less than three years of age at the time of testing or who were attending kindergarten were excluded in order to maintain homogeneity in the sample. For each teacher, a list was compiled of the children in the class (or in the total center for the two schools with the open plan) who were available as subjects for rating and testing.

It was then necessary to obtain consent of the parent or guardian of each child. The teacher gave to the parent or guardian of each child on the list a letter explaining the project and enclosing a consent form for the parent to sign and return to the teacher (Appendix A). No children were tested for whom parental consent was not obtained.

On the day scheduled for the initial testing, the list of children in the center was revised on the basis of the children for whom consent forms had been returned and who were present that day. From the names on the revised list, children were randomly assigned to one of two groups (Pretest Group or Posttest Group).

Instruments for Assessing Children's Abilities

The research design of this study required the assessment of pre-school children's abilities in the domains of cognitive, motor, and social development for the purpose of presenting feedback information to the teacher and to assess the adequacy of the teacher's perceptions relative to objective data.

Selection of instruments

In selecting the assessment instruments to use, consideration was given to the age-appropriateness of the test for the children included in this study. The following criteria were judged to be important in selection of the tests:

1. The test should be administered individually.
2. Test administration time should be of a reasonable length that would not unduly fatigue the child to the extent that it would reduce the reliability of the test.
3. The tasks required by the test should be of interest to young children.
4. The test should sample a variety of the specific skills of the ability being measured.
5. The test should have been standardized on a sample of children that included the ages of the children in this study (three to five-and-one-half years).
6. There should be a standardized scoring procedure.

Description of test instruments

On the basis of the above criteria, the following three tests were selected: the McCarthy Scales of Children's Abilities (McCarthy, 1972) for assessing cognitive development; the Preschool Scale of Motor Development (Bayley, Note 5) for assessing motor development; and the California Preschool Social Competency Scale (Levine, Elzey, & Lewis, 1969) for assessing social development.

McCarthy Scales of Children's Abilities The McCarthy Scales of Children's Abilities (MSCA) (McCarthy, 1972) was selected for use in assessing cognitive development. The test manual describes this test as providing a comprehensive evaluation of intellectual and behavioral development for children ages two-and-one-half to eight-and-one-half years. Designed for use with normal children as well as for the early evaluation of exceptional children, the content of the tasks was selected so as to be suitable for both sexes, as well as for children from various ethnic, regional, and socio-economic groups. Fifteen separate tests to assess the child's abilities in a variety of areas are grouped into six scales: Verbal, Perceptual-Performance, Quantitative, General Cognitive, Memory, and Motor.

The tests constituting the Verbal Scale assess the child's ability to express himself verbally and the maturity of his verbal concepts. Mental processes such as short- and long-term memory, divergent thinking, and deductive reasoning are tapped by asking the child to respond with one-word answers, phrases, and sentences to a variety of items.

The Perceptual-Performance Scale assesses the child's reasoning ability through his manipulation of materials in game-like tasks which do not require the child to speak. Skills such as imitation, logical classification, and visual organization are assessed by a variety of spatial, visual-perceptual, and conceptual tasks.

The Quantitative Scale measures the child's facility with numbers and his understanding of quantitative words. Each item requires only a single step, rather than a sequential process, for solution. The purpose of this scale is to assess the child's number aptitude rather than to explore the upper limit of his computational skills.

The General Cognitive Scale is composed of all the tests in the Verbal, Perceptual-Performance, and Quantitative Scales. Each task in these scales is cognitive in nature and thus the General Cognitive Scale as a whole provides a measure of the child's overall cognitive functioning. Only three of the eighteen tests in the MSCA, Leg Coordination, Arm coordination, and Imitative Action, which involve gross motor rather than cognitive ability, are not included on the General Cognitive Scale.

The Memory Scale assesses the child's short-term memory. Auditory stimuli only is presented by the Verbal and Numerical Memory tasks; auditory and visual stimuli are presented simultaneously by the Pictorial Memory and Tapping Sequence tasks. Since scores on memory tasks are partially a function of the child's ability to deal with the specific content to be memorized, each test of memory has also been placed on the Verbal, Perceptual-Performance, or Quantitative Scale.

Motor Scale tests assess the child's coordination as he performs a variety of gross and fine motor tasks. Gross motor ability is assessed by the Leg Coordination, Arm Coordination, and Imitative Action tests. Fine motor coordination is assessed by Draw-A-Design and Draw-A-Child, in which the child demonstrates his level of hand coordination and finger dexterity. Since these tests have a strong cognitive component they also are included on the Perceptual-Performance and General Cognitive Scales.

For each of the six Scales, the child's raw score is converted to a scaled score, called an Index, according to his chronological age. The General Cognitive Index has a mean which has been set at 100 and a standard deviation of 16. The scores for each of the remaining five scales (e.g., Verbal Index, Perceptual-Performance Index) have an arbitrarily chosen mean of 50 and standard deviation of 10.

The child's General Cognitive Index shows his cognitive level in relation to other children of his chronological age. McCarthy stresses that although this index has a mean of 100 and a standard deviation of 16, essentially the same parameters used to define intelligence quotients obtained from many mental tests, the term IQ was deliberately avoided in the McCarthy Scales because of the many misinterpretations of that concept. The General Cognitive Index is an index of the child's functioning at a given point in time and represents the child's ability to integrate his accumulated learnings and adapt them to the tasks of the MSCA.

The standardization of the MSCA was based on a nationwide sample of 1032 children. The sample was stratified on the variables of age, sex, color, geographic region, and father's occupation. The goal of this stratified sampling was to produce norms that would be representative of the United States population of children ages two-and-one-half through eight-and-one-half years.

Scoring the MSCA The total of the scores on the individual items of each of the scales of the MSCA provides the child's raw score for each scale. The raw score is then converted to a scale score, called an index, by reference to the age-appropriate table in the test manual. The General Cognitive Index, which is the scale of primary concern in the present study, has a mean of 100 and a standard deviation of 16.

Reliability of the MSCA Reliability coefficients for the component tests based on split-half correlations corrected by the Spearman-Brown formula are reported by McCarthy (1972). Average coefficients for ten age groups (two-and-one-half to eight-and-one-half years) were obtained by using Fisher's z transformation. The General Cognitive Scale has an average reliability coefficient of .93 for the total age span for which the test is designed. The Motor Scale has an average reliability coefficient of .79. The average reliability coefficients for the other scales range from .79 to .88. For the ages tested in the present study (three to five-and-one-half years) average reliability coefficients are .93 for the General Cognitive Scale and .81 for the Motor Scale.

Stability over time of the MSCA was assessed by retesting after a three to five week interval 125 children in the normative sample. For three age levels (three to three-and-one-half years, five to five-and-one-half years, seven-and-one-half to eight-and-one-half years) the average coefficients obtained were .90 for the General Cognitive Scale and a range of .75 to .89 for the other four cognitive scales. For the ages tested in the present study (three to five-and-one-half years) stability coefficients were .90 for the General Cognitive Scale and .76 for the Motor Scale.

McCarthy (1972) concludes that the reliability coefficients and standard errors of measurement give evidence that the six MSCA scales are both internally consistent and stable, and that a child's obtained indexes, especially the General Cognitive Index, are quite accurate indicators of his ability on the tasks of the MSCA.

Preschool Scale of Motor Development The test selected for assessing motor development was the Preschool Scale of Motor Development (PSMD) (Bayley, Note 5). This scale is an upward extension of the 1935 California Infant and Preschool Scale of Motor Development. It is based on longitudinal tests of approximately 50 children tested at three or six-month intervals between the ages of two and six years. Because the PSMD has not been published, its use in research is on a restricted basis. Current normative data on these scales are to be incorporated in a continuation of the revised form of the Scale of Motor Development in the First Two Years (1961 Research Form).

The PSMD consists of 20 tests of motor activities, such as standing, walking, jumping, hopping, and ball play, for which credit is given for the level of accomplishment (Appendix B). Special equipment is necessary for the administration of this test. This includes stairs, a walking board, hurdle, and bean bag toss target. Specifications for the construction of the equipment are provided by Bayley (Appendix B).

Scoring the PSMD The total of the number of tasks performed represents the child's motor ability raw score, which is converted to a standard score according to a conversion table provided with the test. Another table is included from which to derive the child's motor age, which is the chronological age of those children in the norming sample who performed the same number of tasks as the child being scored.

Reliability of the PSMD No information about reliability of the Preschool Scale of Motor Development was available at the time of this research. However, reliability had been established for an earlier scale, parts of which were included in the extended scale. For the motor scores of the 1935 Bayley Scales of Infant Mental and Motor Development, Werner and Bayley (1966) reported test-retest reliability at a one-month interval for infants one to 15 months of age of .55 to .89.

California Preschool Social Competency Scale Instruments for assessing a child's social development require special consideration. Social competency assumes the acquisition of abilities and

characteristics that allow the individual to function adequately in society. This usually involves interaction with other persons. It is more difficult to observe these behaviors and abilities in a short span of time. Considerable observation of a child in a variety of situations is needed in order to make any valid statement about the child's skill in initiating, responding to, and maintaining social contacts. Since it was not deemed feasible for the examiners to spend the amount of time that would be necessary to make a valid assessment of the child's social abilities based on observation, it was decided to use the teacher's knowledge of the child's behavior for this purpose. Therefore, a structured rating scale to be completed by the teacher was selected as the standardized instrument for assessing the children's social competency. This permitted use of the information about the child's behavior by the person who had opportunity to observe the child and also allowed that rating to be compared with the standardization sample.

The teacher's perceptions of the child's social development were obtained by the California Preschool Social Competency Scale (CPSCS) (Levine, et al., 1969). This is a 30-item scale designed for use in evaluating the social competency of children ages two-and-one-half to five-and-one-half years. The items represent samples of the critical behaviors in the preschool child's social functioning that are observable within the context of a preschool program. The scale was explicitly developed to be used by teachers in preschool programs to measure the adequacy of the preschool child's interpersonal

behavior and the degree to which he assumes social responsibility. The nature of the items requires the rater to have had considerable opportunity to observe the child in a variety of situations and ratings are based on cumulative observations of the child in his normal pre-school environment. For each of the 30 items, there are four descriptive statements which represent varying degrees of competence in the behavior assessed by that item. The rater selects one of the four statements that best describes the child's typical performance.

The norms for the CPSCS are based on teacher ratings of children within the age range of two-and-one-half to five-and-one-half years attending preschool or nursery school programs.

Scoring the CPSCS The sum of the ratings for each of the 30 items represents the child's social competency raw score, which is then converted to a percentile score by reference to the age-appropriate table of norms in the test manual. To facilitate analysis for purposes of this study, the percentile score was then converted to a standard score.

Reliability of the CPSCS Reliability coefficients (Pearson r 's) are available for three studies in which ratings were obtained from independent observers. Ratings for 24 children made by classroom teachers, program director, and program consultant yielded reliability coefficients of .76 for teacher and consultant, .76 for teacher and director, and .86 for consultant and director. For 15 children in another study, a reliability coefficient of .78 was obtained for ratings made by teacher and director. A reliability

coefficient of .79 was reported for ratings obtained on 71 children from teachers and assistant teachers working in six summer Head Start programs. In reporting these reliability coefficients, Levine, et al. (1969) pointed out that these coefficients are conservative estimates as interjudge differences in the use of the scale and knowledge of the children being rated were not taken into account.

Odd-even reliability coefficients computed by age, sex, and occupational level of parents ranged from .90 to .98 after correction by the Spearman-Brown formula (Levine, et al., 1969). The CPSCS was used during the second year (1970-71) of the 1970 Head Start Planned Variation sample with 3857 children. Odd-even reliability coefficient for that sample was .96 (Walker, Bane, & Bryh, 1973).

Measurement validity was rated "poor" by Hoepfner, et al. (1971). Correlations of the scale with the other tests in the 1970 Head Start Planned Variation Study were low; the largest correlation was .39 with the 64-item Preschool Inventory (Walker, et al., 1973).

Examiners

The principal investigator and four additional examiners participated in administering the cognitive and motor tests. All examiners had training in psychological testing and experience in testing children. Four examiners administered the cognitive test. These four examiners met together to review the testing procedures and to observe each other administering the tests to young children. Administration techniques and scoring procedures were then discussed to insure that comparable routines were used for all children.

Three examiners were trained in a similar manner for administering the motor test. All social competency rating scales were administered to the teachers by the principal investigator.

Teachers' Ratings

The research method utilized information obtained from the pre-school teachers concerning the development of children in their classes in the day care center. A rating scale was devised for soliciting this information (Appendix C).

In constructing the rating scale for use by the teachers, the following criteria were considered:

1. The scale should cover the domain of behavior under study in the research design, i.e., cognitive, motor, and social behavior.

2. The scale should provide a separate rating for each child.

3. To assure that all teachers are rating the same type of behavior, the instructions should include a description of the behavior to be rated.

4. The scale should have enough points to allow for discrimination at several levels among the children.

5. The scale should be easily scored.

6. The teachers were being asked to complete the rating scale during their time at work in a busy and demanding setting. It seemed unreasonable to expect that they would be able to leave their duties for a long period of time. Therefore, the rating scale should be as short and simple as possible to obtain the desired information.

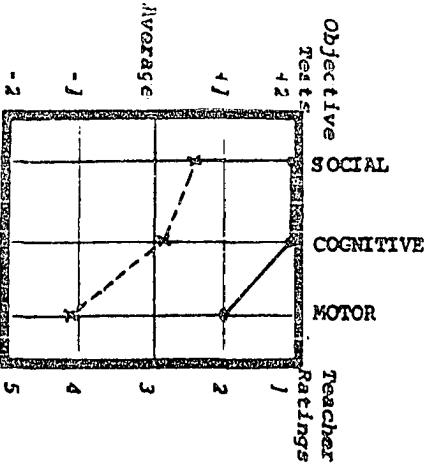
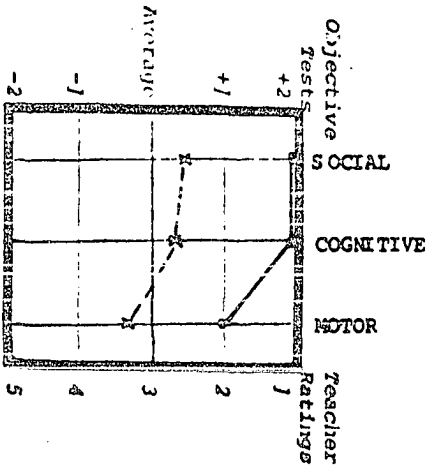
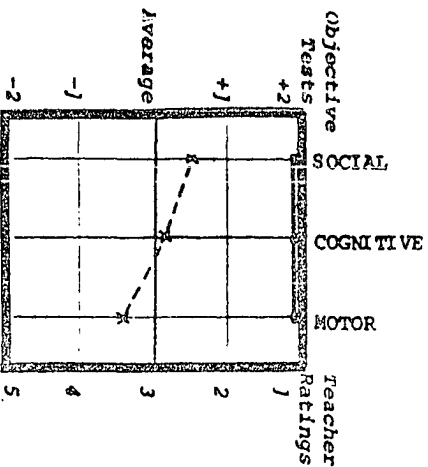
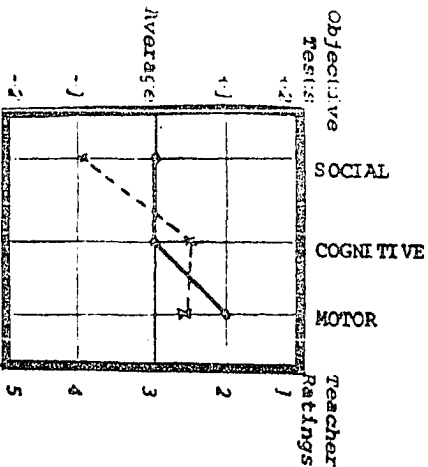
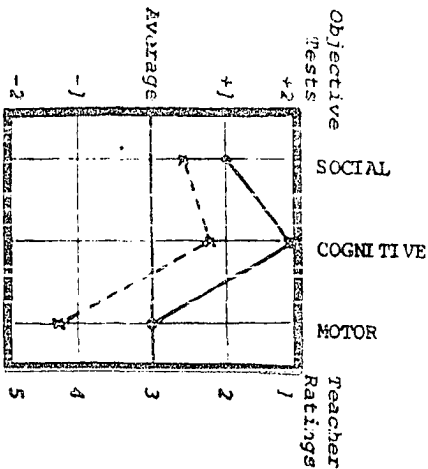
On the basis of the above criteria, the rating scale devised was a five-point Likert-type scale with which the teacher rated each child's level of cognitive development, motor development, and social development. "One" indicated the child had attained a high level of development, "five" indicated the child had attained a low level of development.

Feedback

Feedback was given by the principal investigator in an individual session with each teacher. The feedback information was based on the ratings made by the teacher for one-half of the children in her class (the Pretest Group) and on the results of the examiners' testing of these children. Information was oriented toward providing both an idiographic and a normative profile for each child.

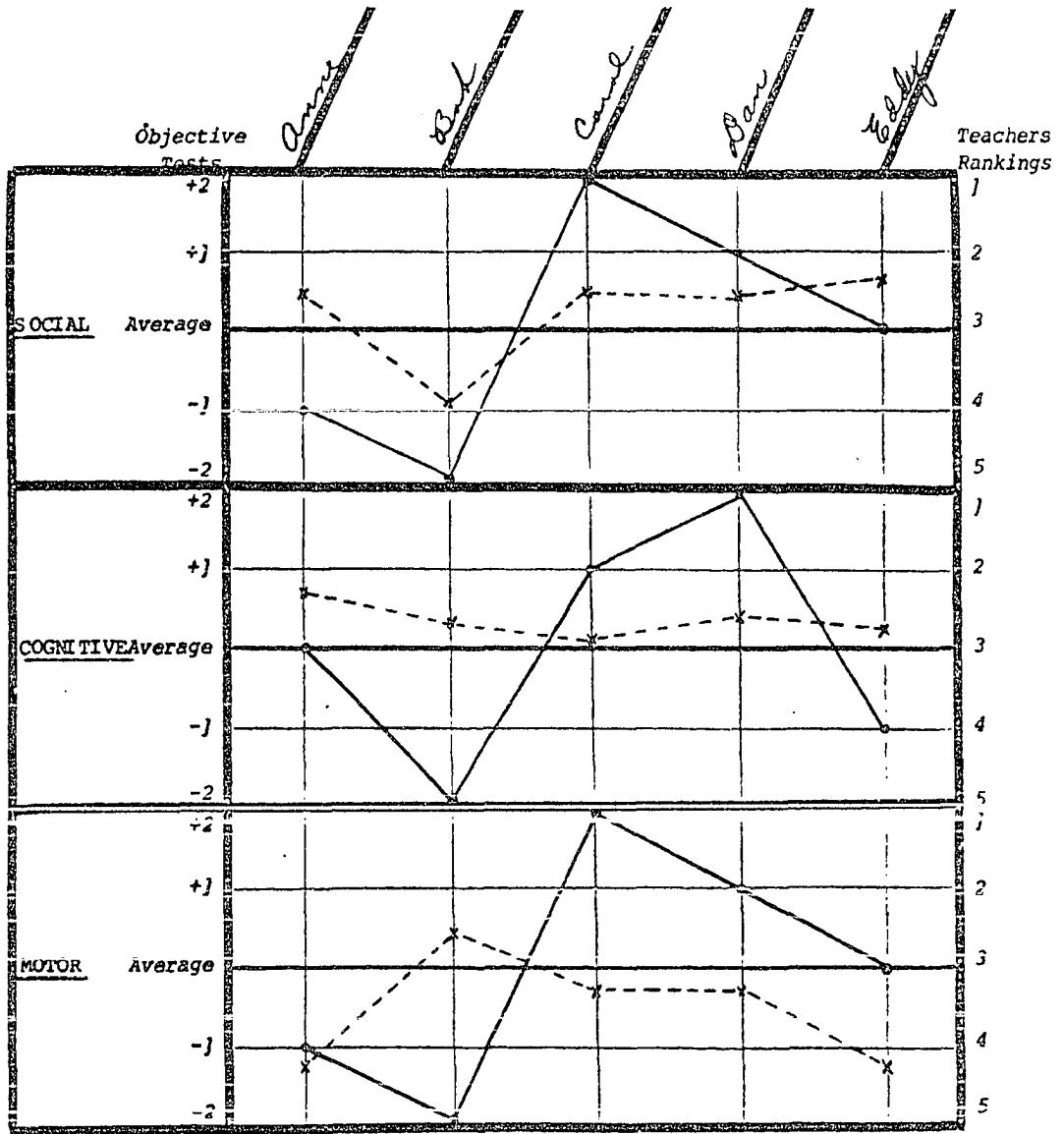
In order to provide an equivalent feedback session for each teacher, a standard method for informing teachers of feedback data was utilized. Comparison of ratings and test results was facilitated by construction for each child of a graph showing both the rating the teacher made for each of the three areas of development and the child's score on the test instrument (Figure 1). A second presentation of the same data was arranged in a profile format which allowed comparisons between children in a normative, as opposed to an idiographic, manner. In this normative graphic summary, the data for each set of children were presented separately for each of the three instruments (Figure 2).

Figure 1 Example of idiographic feedback chart



Key: — Teacher's ratings
 ---- Test scores

Figure 2. Example of normative feedback chart



Key: — Teacher's ratings
 ---- Test scores

Teacher Jones

The standard feedback procedure consisted of the following:

1. A description and explanation of the areas of development being assessed in this research, i.e., cognitive, motor, and social. Examples of each of these developmental aspects were drawn from the typical behavior of three year to five-and-one-half year old children.
2. A review of the teacher's ratings for each child in each of the three areas of development, including a graphic presentation.
3. Interpretation of the results of the three standardized instruments for each child. This information was presented graphically. Test results did not include a numerical score but showed the child's performance as compared to other children in the norming sample.
4. Comparison of the teacher's ratings for each child and the child's performance on the standardized instruments. The comparison was facilitated by use of the graph.
5. Discussion of the analysis of each child's developmental status as indicated by the teacher and by the test results. Also included was an analysis of sub-tests.
6. Discussion of the relationships between cognitive, motor, and social abilities. Examples of the behavior of three year to five-and-one-half year old children were used, with emphasis on the integration of all areas of development.

Test Administration

Since assessment of the children by the examiners took place in five different centers, attempts were made to standardize the test

administration as much as possible. The cognitive test required a quiet room, a table, and chairs for the child and examiner. The motor test required a large area for setting up the equipment and permitting some seclusion from the general nursery school activities. The day care centers were housed in facilities which did not provide separate rooms for testing. In any event generally noisy conditions in the centers precluded an optimal testing situation within the centers. However, factors were taken into consideration which denied the possibility of taking the children away from the center for testing: the reluctance of the center supervisors to permit the children to leave the premises and the belief that testing young children is more effective if done in familiar surroundings.

To meet these requirements, a mobile house trailer was used for the testing room for the cognitive test. A testing room was set up at each end of the trailer with a divider between so that two children could be individually administered the cognitive test simultaneously by two examiners. The trailer was parked in the driveway or parking lot at each center.

In four centers the motor testing took place out of doors in the yard or parking lot. One center was in a new building which did not yet have a lawn and outdoor space was not adequate for the equipment without interfering with the school routine. There was a large room used for nap time which was adequate for the motor test apparatus. Although some variation was necessary in the arrangement of the motor equipment, the materials used and the format of each task were identical for each center.

No special equipment other than the regulation testing kit was required for the cognitive test. The motor test did require special equipment which was constructed according to the specifications furnished by Bayley (Appendix B).

In order to interfere as little as possible in the routine of the center, attempts were made to minimize the actual time period when children would be out of their classes. This was facilitated by having three examiners administering tests at the same time, two using the cognitive test and one using the motor test. Time required for testing each child was 30 to 45 minutes for the cognitive test and 20 to 30 minutes for the motor test.

Procedure

Data were collected during a six week period in May and June, 1973. The procedure was to move the mobile unit and set up the apparatus for the motor test early in the morning so that testing could begin as soon as children arrived at the center. In order to minimize interruption of the scheduled program for the center, testing was not conducted during the lunch hour. In some centers all children were required to rest after lunch, so testing was resumed after nap time. In other centers, the director permitted children who were not sleeping (usually the older children) to be excused for testing. If all testing was not completed in one day, the same procedure was followed the next day. When testing was completed at one center, the mobile unit was moved to another center and the same procedure followed.

While the examiners were administering the cognitive and motor tests to the children, the principal investigator interviewed the teachers to obtain biographical information and the teacher's ratings of the children in the Pretest Group.

The feedback session with each teacher was held the day after all the Pretest Group children in the center had been tested and rated. One week later, the Posttest Group children were rated and tested. After the children in the first two centers had been tested, the procedure was modified to preclude attrition due to the absence of children, as it became apparent that a number of children assigned to the Posttest Group were not available for testing one week later. Under the new procedure, after random assignment to groups, children from both the Pretest Group and the Posttest Group were tested on the same day, or consecutive days, although ratings for the Posttest Group children were not made by the teacher until a week after the feedback session. Teachers had no prior knowledge as to which group a child was assigned or of the results of testing the Posttest Group children prior to their ratings of the children.

Although it was not essential for the experimental design of the research study, a second session was scheduled for each teacher to provide feedback information on the teachers ratings and the test scores for the Posttest Group.

Statistical Treatment

Scoring

Although each cognitive test was scored by the examiner who administered it, all tests were reviewed by the principal investigator and all scores recomputed to insure accuracy. All motor test scores and social competency scale scores were computed by the principal investigator.

In order to standardize the format for the feedback sessions and to facilitate analysis, all scores from the three developmental tests were converted to standard scores. In this procedure, scores were equated for chronological age. Age, therefore, was not a factor in subsequent statistical treatment of the data.

Coding

All test scores, teacher ratings, and demographic data were coded on data processing cards.

Statistical analysis

Under the direction of Dr. Leroy Wolins of the Iowa State University Department of Statistics, analyses of variance and correlation matrices were compiled for the dependent variables.

RESULTS

The purpose of this study was to investigate the effect of feedback on the adequacy of teachers' ratings of preschool children's development as compared with the children's performance on standardized test instruments. Each teacher was asked to make a rating of each of the children in her class for cognitive, motor, and social development. For each child, the McCarthy Scales of Children's Abilities was used to obtain an index of cognitive ability, the Preschool Scale of Motor Development was used to assess motor development, and the California Preschool Social Competency Scale was used to evaluate social competency. The experimental intervention consisted of information fed back by the investigator to the teachers about their ratings relative to the children's performance on the three objective scales. The feedback session occurred between the testing of children randomly assigned to pretest and posttest groups.

The following null hypotheses were formulated for this study:

1. Agreement of preschool teachers' ratings of children's cognitive development with objective test data is not changed by feedback on rating behavior and test results.

2. Agreement of preschool teachers' ratings of children's motor development with objective test data is not changed by feedback on rating behavior and test results.

3. Agreement of preschool teachers' ratings of children's social competencies with standardized rating data is not changed by feedback on rating behavior.

4. Agreement of preschool teachers' ratings of children's development with objective test data is independent of teachers' age, education, or experience.

5. Interrelationships among the individual teacher's ratings of specific developmental characteristics, i.e., cognitive, motor, and social, do not decrease following feedback on rating behavior.

For Hypotheses 1 through 4, the dependent variable was the degree of agreement between the teachers' ratings and objective test data. The degree of agreement was operationally defined as the correlation between these two sets of measurements. Hypothesis testing was accomplished by determining the significance of the difference between the correlations obtained for the Pretest Group and the Posttest Group.

To test the primary hypotheses concerning the effect of feedback, the correlations between ratings made by the teachers and the test scores obtained by the children were examined for the Pretest Group and the Posttest Group. Table 2 presents the pretest and posttest pooled within group correlations¹ of teachers' ratings with test scores for each of the three areas of development. The comparisons for testing the hypotheses are pretest versus posttest differences between the pooled within group correlations. A z transformation was used to test the significance of the difference between the pretest and posttest correlations (Edwards,

¹ The pooled within group correlations effectively partial out differences between groups of children clustered by teachers. In essence, a correlation between rating behavior and test data was obtained separately for each teacher. These correlations were then pooled to yield a single estimate of the degree of relationship for all teachers free of biases deriving from between group differences.

1960). The correlation of teachers' ratings with objective test scores for cognitive development increased from pretest to posttest; correlations of teachers' ratings with test scores for social competency decreased from pretest to posttest. Neither of these changes was significant. Therefore, Hypotheses 1 and 3 fail to be rejected on the basis of the findings presented in Table 2.

Table 2. Tests for significance of differences between Pretest and Posttest Groups based on pooled within group correlations of teachers' ratings and test scores.

Pooled within group correlations				
Scale	Pretest (n=63)	Posttest (n=66)	z	p one-tailed test
Cognitive	.35	.45	+ .576	.28
Motor	.35	.61	+1.650	.05
Social	.63	.58	- .379	.65

$z = 1.64$, $p = .05$, one-tailed test

A significant z was obtained for differences between pretest and posttest correlations of teachers' ratings of motor development with motor test scores in the direction of a higher correlation for the posttest. Therefore, Hypothesis 2 is rejected in its null form, since teachers did change their ratings of children's motor development after the experimental intervention, indicating that feedback had an effect in this area of assessment.

To test Hypothesis 4, relationships were examined between the dependent variable (rating and test score correlation) and three teacher characteristics (age, education, and experience). Table 3 presents

Table 3. Rank order correlations of teachers' age, education, and experience and correlations between teachers' ratings and test scores for cognitive, motor, and social development, Pretest and Posttest Groups.

	Pretest (N=14)			Posttest (N=13)		
	Cognitive	Motor	Social	Cognitive	Motor	Social
Age	.03	.34	.01	.04	.66	-.28
Education	.26	-.38	.43	.26	-.36	.42
Experience	-.09	.15	-.07	.01	.29	-.71
rho = .53, p = .05, 12 df. rho = .66, p = .01, 12 df				rho = .55, p = .05, 11 df. rho = .68, p = .01, 11 df.		

the rank order correlations. These coefficients were computed by ranking, for example, the teacher's age as one correlate and the coefficient serving as the dependent variable as the other correlate. The degree of relationship was then assessed by computing rho. Using this method served to avoid assumptions concerning the nature of the distribution on the various measures for this small sample of teachers. Two of the 18 coefficients were significant, both in the posttest matrix. In an exploratory matrix of this nature, yielding 18 coefficients, one such correlation could be expected to be significant at the .05 level by chance alone. Thus Hypothesis 4 fails to be rejected.

Table 4 presents intercorrelations of the teachers' ratings of the

Table 4. Pooled within group intercorrelations of teachers' ratings for three areas of development, Pretest and Posttest Groups.

Pooled within group intercorrelations				
	Pretest (n=63)		Posttest (n=66)	
	Cognitive	Motor	Cognitive	Motor
Social	.48	.52	.58	.47
Cognitive		.80		.67

children's abilities in the three areas of development. There is no significant difference in the intercorrelations between the pretest and posttest ratings. Therefore, Hypothesis 5 cannot be rejected. As a comparison, the intercorrelations for the objective test data are presented in Table 5. While one would not hypothesize any significant

Table 5. Pooled within group intercorrelations of objective test scores for three areas of development, Pretest and Posttest Groups.

Pooled within group intercorrelations				
	Pretest (n=63)		Posttest (n=66)	
	Cognitive	Motor	Cognitive	Motor
Social	.40	.31	.31	.27
Cognitive		.43		.49

change in the pattern of these intercorrelations due to the experimental intervention, it is of interest to note the apparent overlap in the tests, connoting a general developmental factor common to all three areas. Such a general factor would preclude sharp and independent discriminations among the three areas by teachers.

Additional Findings

Differences among the classes of children were explored for each of 13 variables. Twenty-six one-way analyses of variance were run, 13 each for the Pretest Group and Posttest Group. Significant differences were found for age of children among the various classes for both the Pretest Group and the Posttest Group. The analyses of variance presented in Table 6 indicate that random assignment to the classes on the basis of

Table 6. One-way analyses of variance between classes for age of children in both the Pretest and Posttest Groups.

Pretest					Posttest				
Source	SS	df	MS	F	Source	SS	df	MS	F
Between	2749	13	211.50	4.988	Between	3931	12	327.60	6.561
Within	2077	49	42.39		Within	2646	53	49.93	
Total	4826	62			Total	6577	65		
F = 2.52, p = .01, 13/49 df					F = 2.54, p = .01, 12/53 df				

age cannot be assumed. This, however, is not a factor in previously presented analyses, since all of the test data utilized to test the hypotheses of the study do accommodate the age factor in the normative scoring of the test. All test scores used in the analyses were corrected for age of the child.

The age correction is illustrated in a comparison of the analyses of variance of the cognitive raw score (Table 7) with analyses of variance

Table 7. Analyses of variance between classes for raw scores of the McCarthy Scales of Children's Abilities of children in both the Pretest and Posttest Groups

Pretest					Posttest				
Source	SS	df	MS	F	Source	SS	df	MS	F
Between	27,990	13	2153.0	4.088	Between	48,230	12	4019.0	5.573
Within	25,810	49	526.8		Within	38,220	53	721.1	
Total	53,800	62			Total	86,450	65		
F = 2.52, p = .01, 13/49 df					F = 2.54, p = .01, 12/53 df				

of the cognitive scale score (Table 8) obtained from the McCarthy Scales of Children's Abilities. The raw score yielded a significant F whereas the scale score, corrected as it was for age, did not yield a significant difference between classes. This held true for both the Pretest Group and the Posttest Group.

Table 8. Analyses of variance between classes for scale scores of the McCarthy Scales of Children's Abilities of children in both the Pretest and Posttest Groups.

Pretest					Posttest				
Source	SS	df	MS	F	Source	SS	df	MS	F
Between	4662	13	358.6	1.865	Between	2052	12	171.0	1.166
Within	9423	49	192.3		Within	7774	53	146.7	
Total	14,085	62			Total	9826	65		
F = 1.93, p = .05, 13/49 df					F = 1.94, p = .05, 12/53 df				

Similar data are presented in Tables 9 and 10 for the raw scores

Table 9. Analyses of variance between classes for raw scores of the Preschool Scale of Motor Development of children in both the Pretest and Posttest Groups.

Pretest					Posttest				
Source	SS	df	MS	F	Source	SS	df	MS	F
Between	4703	13	361.8	3.269	Between	2456	12	204.7	2.162
Within	5423	49	110.7		Within	5017	53	94.7	
Total	10,126	62			Total	7473	65		
F = 2.52, p = .01, 13/49 df					F = 1.94, p = .05, 12/53 df				
					F = 2.54, p = .01, 12/53 df				

Table 10. Analyses of variance between classes for standard scores of the Preschool Scale of Motor Development of children in both the Pretest and Posttest Groups

Pretest					Posttest				
Source	SS	df	MS	F	Source	SS	df	MS	F
Between	283,000	13	21,770	1.075	Between	496,100	12	41,340	3.539
Within	992,000	49	20,240		Within	619,100	53	11,680	
Total	1,275,000	62			Total	1,115,200			
F = 1.93, p = .05, 13/49 df					F = 2.54, p = .01, 12/53 df				

and standard scores on the Preschool Scale of Motor Development. The same finding prevails in a comparison of these two tables as was obtained on the cognitive test for the Pretest Group only. The Posttest Group yielded a significant between groups variance for both raw scores and standard scores.

The findings suggest the importance of testing for age differences across groups in subsequent research. This will be especially true in studies utilizing test instruments which are not standardized according to age norms.

The analyses of variance for the raw and standard scores of the California Social Competency Scale yielded no significant differences between the classes for either the Pretest Group or Posttest Group (Tables 11 and 12).

Table 11. Analyses of variance between classes for raw scores of the California Preschool Social Competency Scale of children in both the Pretest and Posttest Groups

Pretest					Posttest				
Source	SS	df	MS	F	Source	SS	df	MS	F
Between	3,703	13	284.9	1.030	Between	3,160	12	263.3	1.160
Within	13,550	49	276.5		Within	12,040	53	227.1	
Total	17,253	62			Total	15,200	65		
F = 1.93, p = .05, 13/49 df					F = 1.94, p = .05, 12/53 df				

Table 12. Analyses of variance between classes for standard scores of the California Preschool Social Competency Scale of children in both the Pretest and Posttest Groups

Pretest					Posttest				
Source	SS	df	MS	F	Source	SS	df	MS	F
Between	136,100	13	10,470	1.125	Between	76,690	12	6391	0.946
Within	455,800	49	9,302		Within	358,800	53	6755	
Total	591,800	62			Total	434,700	65		
F = 1.93, p = .05, 13/49 df					F = 1.94, p = .05, 12/53 df				

Means and standard deviations of the test results for the children in the Pretest Group and Posttest Group for each of the three test instruments are presented in Table 13. There is a significant difference

Table 13. Means and standard deviations of test scores for Pretest and Posttest Groups

Scale	<u>Pretest (n=63)</u>		<u>Posttest (n=66)</u>		t
	Mean	SD	Mean	SD	
Cognitive	92.65	14.95	99.76	12.20	2.94
Motor	-.27	1.42	-.34	1.30	.29
Social	-.24	.97	-.04	.81	1.26

t = 2.62, p = .01, 127 df, two-tailed test

between the Pretest Group and Posttest Group for the cognitive test but not for the motor or social tests.

The descriptive statistics on the biographical data collected in the personal interview with each teacher are presented in Table 14. A significant F test was obtained for differences in the amount of contact the teacher had with the children in her class. This finding was significant in both the Pretest Group and the Posttest Group (Table 15).

Table 14. Means and standard deviations of biographical characteristics of teachers (N=14)

Characteristic	Mean	SD
Age	27 years	12.93
Education	14 years	1.95
Hours per week in center	35 hours	8.06
Length of employment in center	5 years	7.40

Table 15. Analyses of variance between classes for amount of contact between the teacher and children in her class, Pretest and Posttest Groups

Pretest				
Source	SS	df	MS	F
Between	44,530,000	13	3,425,000	18.83
Within	8,912,000	49	181,900	
Total	53,442,000	62		

F = 2.52, p = .01, 13/49 df

Posttest				
Source	SS	df	MS	F
Between	16,280,000	12	1,357,000	12.32
Within	5,837,000	53	110,100	
Total	22,117,000	65		

F = 2.54, p = .01, 12/53 df

DISCUSSION

The present investigation studied the effect that information given teachers on their assessment of children has on their ability to subsequently improve the accuracy of these assessments. The study was designed to compare teachers' ratings of children's development to the children's scores on standardized test instruments before and after the experimental intervention of feedback. Specifically, feedback consisted of providing information to the teachers concerning the comparison of their performance in rating the children in their classes with the objective test scores. The findings and implications of this study will be discussed in this chapter. The discussion will focus on the research results, theoretical implications, methodological implications, and suggestions for further research.

Research Results

Effect of feedback

The results of the study indicate that the experimental intervention was effective in increasing the agreement of teachers' ratings with test scores in the assessment of motor development but not in the assessment of cognitive and social development. One interpretation of these results would be that the effect of feedback depends on the aspect of development being assessed. No significant change between pretest and posttest in the agreement between teachers' ratings and test scores was found for the assessment of children's cognitive development or social development. However, there was significantly closer agreement between teachers' ratings of the children's motor development and the motor test scores for the posttest than for the pretest.

The findings indicate that the feedback session provided the teachers with information that they used in their subsequent assessment of the children's motor development even though no instructions were given to the teachers telling them that they should use the feedback information for their subsequent ratings, nor was it suggested to them that they observe specific behaviors of children.

Consideration should be given to factors which might differentially affect the teachers' utilization of the feedback information in the different areas of development. One possibility for the improved teachers' ratings for motor development is that motor behavior may be more specific, objective, and observable, whereas cognitive and social development involve inferences about less obvious behaviors. Since the motor testing equipment was set up on the play ground, there is the possibility that teachers may have observed some of the testing sequences. No comparable opportunity was available for the cognitive test. Any such observation of the motor tasks may have supplemented the feedback information and provided more meaningful information for the teachers to use in their ratings for the Posttest Group of children. Also motor activities are easily described and may be more objectively observed by teachers. Between the feedback session and their ratings of the Posttest Group, the teachers had only a few days in which to observe the children and apply the information they received to their assessment of the children's development. Motor skills may have been more apparent to the teachers during these few days. Because children are physically active, there are numerous times during the nursery school day when children can be observed walking, jumping, and throwing.

However, cognitive skills are more subtle and there may be fewer opportunities for the teachers to naturally observe and make inferences about cognitive processes unless they are making a special effort to do so. The ability of the teachers to assess development by focusing on specific behaviors has support from an investigation by Haring and Ridgway (1967). Their research reported that kindergarten teachers provided with a structured guide to observing specific areas of performance were able to effectively screen children for potential learning disabilities. An analysis of individual performance behavior appeared to be a valid means for early identification of children who were developmentally retarded.

As Glidewell and Swallow (1968) pointed out, it seems reasonable to expect that the training and experience of a teacher would tend to influence the validity of her ratings. Hunter (1957) reported that formal teacher training tends to increase teachers' agreement with clinicians. Other studies have found that specific mental hygiene courses have irregular effects on teachers' performance (Cutler, 1961; Ellis & Miller, 1936). Increase in experience appears to increase agreement of teachers' ratings with mental health professionals in progressively smaller increments up to a certain point. After about ten years of teaching experience, agreement between teachers and mental health specialists begins to decrease (Ellis & Miller, 1936; Hunter, 1957).

Thus, the present study also examined the relationships between the dependent variable (rating and test score correlation) and the

teachers' age, education, and experience. The computation of rank order correlations indicated only two of 18 coefficients to be significant. For the Posttest Group only, age was positively correlated with the agreement of teachers' ratings with motor test scores; experience was negatively correlated with the agreement of teachers' ratings with social test scores. Since these give contradictory evidence for the general hypotheses and since by chance alone one of the 18 coefficients would be expected to be significant, no definite conclusions can be drawn concerning the effect of teachers' age, education, or experience on the agreement of their ratings of children's development with test scores.

Examination of the interrelationships among the individual teacher's ratings of the children's abilities in the three specific components of development (cognitive, motor, and social) showed no significant differences in the intercorrelations of these measures for the Pretest and Posttest Groups. Highest correlations (Pretest $r = .80$, Posttest $r = .67$) were noted for the teachers' ratings of cognitive and motor development. The intercorrelations of teachers' ratings (Table 4) are higher than the intercorrelations of the objective test scores (Table 5). The intercorrelations of objective test scores suggest that there are apparent overlaps in the tests. The higher intercorrelations of teachers' ratings suggest the teachers are influenced by a general developmental factor in their assessment of children to a greater extent than would be warranted by the apparent overlap in the tests.

Miller (1972) found that elementary school teachers exhibited a "halo effect" when judging children's behavior. All behavior appeared

to be judged in terms of the child's academic competence; as competence decreased, more deviance was observed. The teachers in Miller's study appeared to be rating along an achievement-competence dimension.

Additional findings

The additional findings compare the classes of children. The finding of a significant difference for age of children among the various classes in both the Pretest and Posttest Groups is not unexpected. The age difference occurs because a child is assigned to a class on the basis of his age, e.g., one class may be organized for three year old children and another may be organized for four year old children. The standardized tests used in the study were scored on the basis of the children's age so that the observed differences in age of children among classes is not a factor of concern. However, this finding does suggest the importance of testing for age differences across groups in future research, especially if test instruments are not standardized according to age norms.

Means and standard deviations of the scores of the three tests were examined for any differences between the Pretest and Posttest Groups. There was no significant difference for either the motor or social tests. However, the cognitive test did show a higher mean (99.76) for the Posttest than for the Pretest (92.65). This difference, significant beyond the $p = .01$ level of confidence, should not affect the results testing the main hypotheses since the dependent variable was the correlation between teachers' ratings and objective test scores. The correlation coefficient, used in this case as a dependent score,

is a dimensionless statistic which in essence deals with a distribution with a mean of zero and a standard deviation of one regardless of the actual raw score distribution.

A questionnaire provided data on the amount of contact that the teacher had with each child in her class. The amount of contact was determined by the number of hours each week the child was in the teacher's class times the number of weeks the child had been in the class. An analysis of variance indicated significant differences among the classes for amount of contact between teacher and children. However, since this finding was true for both the Pretest and Posttest Groups, it does not influence the results of the experimental intervention.

Theoretical Implications

The present investigation has implications for a number of concepts that should be considered in the assessment of the development of children. Developmental norms are often used as a standard against which to evaluate a child's developmental status. They provide clues for interpreting the child's behavior as well as the degree of any deviation. Initially, each teacher undoubtedly used her own concept of norms as a basis for making her ratings of the children's development. Of general concern is the accuracy of the teachers' concepts of norms and the range of variability (Flapan & Neubauer, 1970; Thomas, et al., 1968) a teacher might allow before viewing the child's behavior as deviant. In the present study, feedback included a normative profile to allow the teacher to compare each child in her class to children in the norming sample for each test. The results indicate that this

feedback was useful to the teachers in their application of normative criteria in assessment of motor development.

The concept of parallelism refers to the child as a complex organism developing as a whole person rather than as a collection of separate component parts (Breckenridge & Murphy, 1963). However, growth in one component may have a limiting or stimulating effect on growth in another component, or as Ausubel and Sullivan (1970) point out, there may be a common factor which might account for parallelism among components. The intercorrelations of the three standardized tests used in the study, as shown in Table 5, suggest a general factor in the three tests. The presence of a general developmental factor that is common to all three components of development increases the difficulty of making independent discriminations among these three areas.

The assessment of the total development of a child involves integrating the assessments of these various components of development. Teachers may be more aware of the child developing as a whole and thereby focus less specifically on any one of the three areas of development studied in the present research, whereas the tests used isolate the three areas of development. For assessment purposes it is important to isolate the components of development in order to specify the child's strengths or weaknesses. A child who has a weakness may have found a means to compensate for it and a global appraisal of the child may overlook this. However, the specific attention given to this area of development by a standardized test can focus on an area that may have the

potential for a future problem. Feedback included an idiographic profile showing the child's performance in each of three components of development. A description and explanation of each of the components was included. The results of the study suggest that this feedback approach helped teachers to focus on specific components as evidenced by their improved rating ability for motor development. A number of research studies support this approach to assessment. Quay, et al. (1966) and Rubin, et al. (1966) give evidence that teachers made reliable identification of maladjusted children with the use of symptom check lists which focused on specific behaviors. Haring and Ridgway (1967) reported good results when they provided teachers with a structured guide to observation that specified areas of performance.

The issues of developmental regression and the continuity and discontinuity of developmental processes are important considerations in the developmental assessment of children. Both of these concepts require assessment over a period of time. While the time span of the present research was not designed to study these issues, the feedback helped teachers become more perceptive of motor development. Increased awareness of specific components of development should allow teachers to better judge whether a change in behavior over time is true regression or a temporary relapse to an earlier pattern. The experiment could be meaningful for investigating the teachers' assessment of regressive behavior by a repeated measures design, extending it over a sufficient period of time so that opportunities for observation, assessment, and feedback on discontinuities and regression could occur. Flapan and

Neubauer (1970) emphasize that a longitudinal view is essential for the assessment of developmental progression or interference.

The results of the present study indicate that the services of a mental health center can be effectively used in working with preschool centers. A child development consultant from a mental health center's program for preschool children, by providing feedback to teachers in day care centers, was able to effect change in the teachers' assessment of children in the area of motor development. Further efforts by mental health centers would seem to be appropriate for encouraging the cooperation of the personnel from both mental health centers and agencies providing services to young children. Previous research (Eisenberg, et al., 1962; Westman, et al., 1967) supports the efficacy of cooperation between nursery schools and mental health services.

A worthwhile direction for mental health services could be in providing more extensive educational services to teachers in preschool centers. The feedback intervention in the present study was of limited content and short duration; more extensive educational programs over a longer period of time, perhaps in the form of in-service workshops for preschool personnel, might be more effective. Teachers improved after feedback in their assessments of motor development. This may be an appropriate area of children's development with which to begin educational programs, since it may be easier for teachers to understand and observe the specifics of children's motor behavior. Success in learning to more objectively observe motor development might then be transferred to cognitive and social development.

Provision of consultative services by mental health centers is also appropriate. Cowen (1973) strongly urges consultation to community agencies as an important function of mental health centers. In the present study the investigator noted an interest on the part of the teachers in discussing cases, in learning more about specific parts of the tests, and in relating the tasks (especially on the cognitive test) to the preschool curriculum. It appears that teachers would make use of the services of a child development consultant from a mental health center. It is also possible that mental health centers which have a program for working with developmentally disabled children could provide a setting for bringing teachers in to the center in order to sensitize the teachers to developmental problems.

Methodological Implications

Campbell and Stanley (1966) distinguish between the internal and external validity of experimental designs. Internal validity asks *whether the experimental treatment did in fact make a difference.* External validity refers to generalizability, that is to what populations, settings, treatment variables, and measurement variables the effect can be generalized. The present investigation utilized a One-Group Pretest-Post design. Campbell and Stanley caution that in this design internal validity may be jeopardized by several confounding extraneous variables offering alternate hypotheses other than the hypothesis relating to the effect of the experimental treatment. Each of the rival hypotheses suggested by Campbell and Stanley will be discussed relative to implications for the present research design.

History refers to other events that may have occurred between the first and second observations in addition to the experimental treatment. Another event becomes a more plausible rival explanation of change the longer the time lapse between the observations and if the event has occurred to most of the subjects. In the present study, one week elapsed between the teacher's first and second ratings. It is not known what other events may have occurred during this week which would affect the teachers' second ratings. However, since the teachers were drawn from five different centers, it is unlikely that they would all have been exposed to the same change-producing events in such a short period of time. In addition, the intervening time period was staggered due to the testing schedule and therefore did not fall on the same days for all teachers.

A second source of rival hypotheses deals with maturation, referring to processes within the subject which vary with the passage of time but are independent of the specific events. Thus the difference between observations may be a reflection of the process of maturation rather than the experimental treatment. Here, too, the time lapse was so short as to preclude maturation as a serious rival hypothesis.

The effects of taking a pretest on the scores of a second test may also pose a threat to internal validity. This testing threat may be controlled for by having a second group of subjects who take the pretest and posttest but do not receive the experimental treatment. The present study lends itself to criticism, however, because a control group was not used. Lack of adequate control groups is one of the weaknesses of

research in a field setting. It would have been extremely difficult to obtain cooperation from double the number of subjects and centers in order to have a control group also. In another sense this rival hypothesis is not so serious in the present study because the very nature of the experimental intervention involved the development of "test sophistication" on the part of the teachers. In a strict experimental sense the testing effect could not be partialled out from the experimental intervention.

Instrumentation is another rival hypothesis and refers to changes in the measuring instruments which might account for the difference between two sets of observations. The rating scales used by the teachers and the standardized tests administered to the children did not change between the pretest and posttest. However, consideration should be given to the human factors within the examiners who tested the children and the teachers who made the ratings. Since several different examiners were involved in administering the tests, precautions were taken, as previously described, to insure that comparable administration techniques were used by all examiners. The examiners participated in both pretest and posttest assessments, but did not know to which group, pretest or posttest, a child was assigned. The use of ratings made by teachers might be affected between pretest and posttest by processes within the teachers. There is no way of knowing with certainty whether those processes within the teachers were a function of the experimental treatment, but this does not represent a strong rival hypothesis.

Statistical regression may be another confounded variable, especially in experiments which select subjects on the basis of their extremity. If the teachers in the study were very low in agreement on their ratings, they would be expected to do better on the posttest even without experimental intervention. However, teachers were not selected on the basis of extremes and their average initial correlations were not extreme. Regression, therefore, is not a threat in this study.

Biases resulting in differential selection of respondents for the comparison groups were controlled in the study because the same group of teachers comprised the pretest and posttest subjects, thereby acting as their own controls.

The threat of mortality, differential loss of respondents from the comparison groups, is also controlled by a one-group design in which the subjects act as their own controls.

It would be hoped that results of the study could be generalized to other preschool teachers. A number of threats to generalizability (external validity) are of concern in the use of the design for the present study. Factors jeopardizing external validity are the reactive or interaction effect of testing, the interaction of selection biases with the experimental variable, and reactive effects of experimental arrangements. None of these can be considered of major concern in the present study.

The main limitation relative to generalizability of the study would stem from the nature and limits of the sampling and the uniqueness of the setting. In this regard the sample and choice of setting were

largely determined by the opportunities available. The teachers used as subjects were not selected because of prior knowledge about any of their attributes. They were selected because of their willingness to cooperate in the study. The director of each center agreed to have the center participate in the project and the teachers' cooperation was enlisted by her. The fourteen teachers then did work with the investigator in the rating and feedback sessions. The threat of the interaction of this selection bias with the experimental intervention is more likely if there are some characteristics of the centers that caused the feedback to be more effective than it would be in other centers. Interaction would be more likely if the centers differed markedly from other centers to which generalizations could be made. The five centers used in this study were not all the same. They varied in size, type of support, facilities, and training and experience of the directors and teachers. This variability reduces the threat of the interaction of selection biases with the experimental intervention to the external validity of the study.

In summary, it appears that the methodological design was adequate for the present study while at the same time suffering from some of the limitations inherent in field research. Research outside an experimental laboratory is subject to the willingness of subjects to participate and the lack of control of all variables. In the present study, there was the advantage of working with both children and teachers in a setting familiar to them. The results of research in a field setting may be more generalizable than if the research had been conducted in an experimental laboratory.

The method used to determine the accuracy of the teachers' ratings was in terms of "convergent validity," which Glidewell and Swallow (1968) define as the agreement between teachers' reports and other methods or sources of data about the behavior of the same children. In the present study, the validity of the teachers' ratings was operationally defined as the correlation between the teacher's ratings of the children in her class and their performance on objective, standardized tests. Methodological problems inherent in this approach include the issue of the reliability and validity of the external data sources, i.e., the standardized tests. While the reliabilities and validities of the tests used appear adequate for certain diagnostic purposes, it is quite possible that the error variance in the instruments themselves combined with the difficulties of control in a field setting in a manner that resulted in nonsignificant results for two of the three areas explored.

Other considerations derive from the assumption that there is a body of knowledge about the assessment of children's development and that this knowledge can be helpful to teachers in their evaluation of the development of children in their classes. What constitutes effective procedures to disseminate relevant knowledge about the development of children to those persons who have direct responsibility for the care of children? There are a variety of methods that could be used, varying in time span, degree of teacher involvement, method, content. The experimental intervention (feedback) in the present study only required from 15 to 30 minutes. It focused on the teacher's ratings of the children's development and an interpretation of the children's

performance on standardized tests. The three areas of development (cognitive, motor, and social) were described and explained with examples of typical behavior of preschool children. There were no specific instructions to the teachers as to how they could use this information. Nevertheless, the information was both relevant and of specific interest to the teachers, constituting as it did the application of information the teacher had been involved in generating on children in her direct care. Under these circumstances it is encouraging that the hypothesized change occurred, at least in the assessment of motor performance. While significant changes were not observed in the assessment of cognitive and social development, a more intensive training effort extended over a longer period of time might generate significant effects in these latter two areas.

Suggestions for Further Research

A difficult problem in research in a field setting is the generation of an experimental intervention that is of sufficient power to overcome other sources of random variance. Further research on the topic under investigation in the present study might effectively use a more extensive experimental intervention by increasing the length of the feedback session and expanding the scope of the content. This method would provide the teachers with more information about their assessments and the children's test performance. The intervention would perhaps be found to be more effective if explicit directions were given to the teachers as to how they could use the feedback information in their subsequent observation and assessment of children. Other

opportunities for strengthening the experimental intervention could be in-service workshops, continuing training sessions on a regular basis, and teacher observation in clinic settings.

The experiment could be extended over a longer period of time with the opportunity for repeated measures of teachers' ratings, thus providing an opportunity for teachers to assess the children over a period of time. This method would also provide for the evaluation of teachers' rating behavior over time.

Selecting preschool centers which have a higher frequency of children with developmental disabilities would provide a means for more accurately evaluating the ability of the teachers to discriminate among children and identify the children who have developmental problems. This would expand the variance in the group of children for which the teachers are required to diagnostically discriminate, thus, in turn, expanding the opportunities for variance among teachers on the dependent variable.

Further research could be conducted in an experimental setting in which preschool teachers could be exposed to children with a variety of developmental disabilities. This method would provide a clinical approach which would give teachers the opportunity to assess the syndromes of developmental disabilities. Direct observation and interaction with children with identified developmental problems should increase the teachers' sensitivity to the behavior of children under their care.

Another research opportunity might be the development and field testing of a curriculum component for teaching the early identification

of developmental problems. Using a competency-based approach to curriculum construction would make the approach readily useable in both the traditional setting of a child development degree program and in a field setting similar to that employed in the present research.

SUMMARY

The present research derived from a concern for the way in which a community mental health center's program for preschool children could aid the personnel of preschool programs in the identification at an early age of those children who exhibited an interference in their normal developmental processes and thus were in need of some type of special help. The specific focus of the present research was on the adequacy of information that teachers in day care centers could provide about the children in their classes. If information elicited from the teacher regarding her observations of the child's development is valid and reliable, it should be incorporated into the evaluation, diagnosis, and treatment plan for the child. If the information provided by teachers is not appropriate, then we are concerned with helping the teachers learn to make more valid reports of children's behavior. With these considerations in mind, a research study was designed to investigate the effectiveness of feedback to teachers in their evaluation of children's cognitive, motor, and social development.

The following null hypotheses were proposed:

1. Agreement of preschool teachers' ratings of children's cognitive development with objective test data is not changed by feedback on rating behavior and test results.
2. Agreement of preschool teachers' ratings of children's motor development with objective test data is not changed by feedback on rating behavior and test results.

3. Agreement of preschool teachers' ratings of children's social competencies with standardized rating data is not changed by feedback on rating behavior.

4. Agreement of preschool teachers' ratings of children's development with objective test data is independent of teachers' age, education, or experience.

5. Interrelationships among the individual teacher's ratings of specific developmental characteristics, i.e., cognitive, motor, and social, does not decrease following feedback on rating behavior.

Five day care centers were selected for the study of these hypotheses by means of an experimental design in a field setting. Each of 14 preschool teachers rated the cognitive, motor, and social development of half of the children in her class. Each child's cognitive, motor, and social development was assessed by means of a standardized test instrument. Feedback information to the teacher included a comparative analysis of the results of the child's performance on these tests and the teacher's ratings of the child. One week later, each teacher rated the other half of the children in her class and these children were tested.

The correlation between the teacher's ratings and the standardized test results served as the dependent variable in the study and was obtained for both Pretest and Posttest Groups to determine the effect of the experimental intervention of feedback. A z test yielded a value of 1.65 ($p=.05$, one-tailed test) indicating a significant increase from pretest to posttest in the correlation between teachers' ratings and

test scores for assessment of motor development. Pretest to posttest differences for cognitive and social development were not significant. Therefore, Hypothesis 2 was rejected but Hypotheses 1 and 3 failed to be rejected. Hypothesis 4 and Hypothesis 5 also failed to be rejected. In summary, the findings suggest that the experimental intervention of feedback was effective in increasing the agreement of teachers' ratings with test scores in the assessment of motor development.

In recent years there has been an increase in the proportion of preschool children in attendance in nursery schools and day care centers. These preschool programs appear to be likely places for the early identification of those children who may need special help. Attendance at a nursery school or day care center may be the first opportunity for a child to be evaluated by someone who is knowledgeable about developmental processes. The opportunity for social interaction with other children and adults may bring into focus or accentuate problems which would not otherwise be observed in the home and this may be the first time that a developmental difficulty is noted. A review of earlier research studies concerning the effectiveness of nursery school teachers in assessing the development of children supports the proposal that preschool programs are appropriate places for early identification of developmental problems. The present study suggests that preschool teachers can be effective in early identification. Taken in the context of the recent growth of preschool programs, the research results suggest that it is now possible, through specific training and increased involvement of teachers in day care centers, to identify at an early age those children who manifest developmental problems.

REFERENCES

- Allinsmith, W., & Goethals, G. W. The role of schools in mental health. New York: Basic Books, 1962.
- Ausubel, D. P. Theory and problems of child development. New York: Grune & Stratton, 1958.
- Ausubel, D. P., & Sullivan, E. V. Theory and problems of child development (2nd ed.). New York: Grune & Stratton, 1970.
- Beach, D. R., Cowen, E. L., Zax, M., Laird, J. D., Grost, M. A., & Izzo, L. D. Objectification of screening for early detection of emotional disorder. Child Development, 1968, 39, 4, 1177-1188.
- Birch, H. G. Malnutrition and early development. In E. H. Grotberg (Ed.), Day care: Resources for decisions. Washington, D. C.: Office of Economic Opportunity, 1971.
- Bower, E. M. Early identification of emotionally handicapped children in school (2nd ed.). Springfield: Thomas, 1969.
- Bower, E. M., Shellhamer, T. A., & Dailey, J. M. School characteristics of male adolescents who later became schizophrenic. American Journal of Orthopsychiatry, 1960, 30, 712-729.
- Breckenridge, M. E., & Murphy, M. N. Growth and development of the young child (7th ed.). Philadelphia: Saunders, 1963.
- Brownbridge, R., & VanVleet, P. (Eds.). Toward community action: Investments in prevention. The prevention of learning and behavior problems in young children. San Francisco: Pace I. D. Center, 1968. (ERIC Microfiche No. ED 024 997)
- Caldwell, B. M. The rationale for early intervention. Exceptional Children, 1970, 36, 717-726.

- Caldwell, B. M. The rationale for early intervention. Exceptional Children, 1970, 36, 717-726.
- Caldwell, B. M. Mental health aspects of early intervention. Journal of Operational Psychiatry, 1972, 3, 7-13.
- Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1966.
- Clarizio, H. F. Stability of deviant behavior through time. In H. F. Clarizio (Ed.), Mental health and the educative process. Chicago: Rand McNally, 1969.
- Cowen, E. L. Social and community interventions. In P. H. Mussen & M. R. Rosenzweig (Eds.), Annual Review of Psychology (Vol. 24). Palo Alto: Annual Reviews, Inc., 1973.
- Cowen, E. L., & Zax, M. Early detection and prevention of emotional disorder: Conceptualizations and programming. In J. W. Carter (Ed.), Research contributions from psychology to community mental health. New York: Behavioral Publications, 1968.
- Cutler, R. L. A research evaluation of an action approach to school mental health. American Journal of Orthopsychiatry, 1961, 31, 339-346.
- Del Solar, C. Parents and teachers view the child: A comparative study of parents' and teachers' appraisals of children. New York: Teachers College, Columbia University, 1949.
- Denenberg, V. H. (Ed.). Education of the infant and young child. New York: Academic Press, 1970.
- Duffy, J. C. Child psychiatry--whither goest thou? Child Psychiatry and Human Development, 1970, 1(1), 3.

Edwards, A. L. Experimental design in psychological research.

New York: Rinehart, 1960.

Eisenberg, L., Landowne, E. J., Wilner, D. M., & Imber, S. D. The use of teacher ratings in a mental health study: A method for measuring the effectiveness of a therapeutic nursery program. American Journal of Public Health, 1962, 52, 18-28.

Ellis, D. B., & Miller, L. W. Teachers' attitudes and child behavior problems. Journal of Educational Psychology, 1936, 27, 501-511.

Feldhusen, J. F., Thurston, J. R., & Benning, J. J. Aggressive classroom behavior and school achievement. Journal of Special Education, 1969, 4, 431-439.

Feldhusen, J. F., Thurston, J. R., & Benning, J. J. Longitudinal analysis of classroom behavior and school achievement. Journal of Experimental Education, 1970, 38, 4-10.

Flapan, E., & Neubauer, P. B. Issues in assessing development. Journal of the American Academy of Child Psychiatry, 1970, 9(4), 669-687.

Frankenburg, W. K., Dodds, J. B., & Fandal, A. W. Denver Developmental Screening Test Manual (Rev. ed.), Denver: University of Colorado Medical Center, 1970.

Freud, Anna. Normality and pathology in childhood. New York: International Universities Press, 1965.

Gesell, A., & Amatruda, C. S. Developmental diagnosis (2nd ed.). New York: Harper & Row, 1947.

Glasscote, R. M., Fishman, M. E., & Sonis, M. Children and mental health centers. Washington, D. C.: American Psychiatric Association, 1972.

- Glidewell, J. C., Mensh, I. N., & Gildea, M. C.-L. Behavior symptoms in children and degree of sickness. American Journal of Psychiatry, 1957, 114, 47-53.
- Glidewell, J. C., & Swallow, C. S. The prevalence of maladjustment in elementary schools. Chicago: University of Chicago, 1968.
- Goldfarb, A. Teacher ratings in psychiatric case findings. American Journal of Public Health, 1963, 53, 1919-1927.
- Griffiths, W. Behavior difficulties of children as perceived and judged by parents, teachers, and children themselves. Minneapolis: University of Minnesota Press, 1952.
- Haring, N. G., & Ridgway, R. W. Early identification of children with learning disabilities. Exceptional Children, 1967, 33, 387-394.
- Havelkova, M. Follow-up study of 71 children diagnosed as psychotic in preschool age. American Journal of Orthopsychiatry, 1968, 38, 846-857.
- Hoepfner, R., Stern, C., & Nummedal, S. G. CSE-ECRC Preschool/kindergarten test evaluations. Los Angeles: UCLA Graduate School of Education, 1971.
- Hunt, J. McV. Intelligence and experience. New York: Ronald Press, 1961.
- Hunter, E. C. Changes in teachers' attitudes toward children's behavior over the last thirty years. Mental Hygiene, 1957, 41, 3-11.
- Jones, R. S. A child development health center: Problems and potentials. Missouri Medicine, April, 1972, 273-274.

- Kellam, S. G., & Schiff, S. K. Adaptation and mental illness in the first grade classrooms of an urban community. American Psychiatric Association Psychiatric Research Reports, 1967, 21, 79-91.
- Lazar, I., & Rosenberg, M. E. Day care in America. In E. H. Grotberg (Ed.), Day care: Resources for decisions. Washington, D. C.: Office of Economic Opportunity, 1971.
- Levine, S., Elzey, F., & Lewis, M. California Preschool Social Competency Scale Manual. Palo Alto: Consulting Psychologists Press, Inc., 1969.
- Lewis, E. The real California Report: A new approach to education. Phi Delta Kappan, April, 1973, 54, 558-559.
- MacFarlane, J. W., Allen, L., & Honzik, M. P. A developmental study of the behavior problems of normal children between twenty-one months and fourteen years. Los Angeles: University of California Press, 1954.
- McCarthy, D. McCarthy Scales of Children's Abilities Manual. New York: The Psychological Corporation, 1972.
- Miller, L. C. Q-Sort agreement among observers of children. American Journal of Orthopsychiatry, 1964, 34, 71-75.
- Miller, L. C. School Behavior Check List: An inventory of deviant behavior for elementary school children. Journal of Consulting and Clinical Psychology, 1972, 38, 134-144.
- Mitchell, J. C. A study of teachers' and of mental hygienists' ratings of certain behavior problems of children. Journal of Educational Research, 1942, 36, 292-307.

- Peck, L. Teachers' reports of the problems of unadjusted school children. Journal of Educational Psychology, 1935, 26, 123-128.
- Peters, A. D. Health support in day care. In E. H. Grotberg (Ed.), Day care: Resources for decisions. Washington, D. C.: Office of Economic Opportunity, 1971.
- Quay, H. C., Morse, W. C., & Cutler, R. L. Personality patterns of pupils in special classes for the emotionally disturbed. Exceptional Children, 1966, 32, 297-301.
- Robins, L. N. Deviant children grown up. Baltimore: Williams & Wilkins, 1966.
- Rubin, E. Z., Simson, C. B., & Betwee, M. C. Emotionally handicapped children and the elementary school. Detroit: Wayne University, 1966.
- Schrager, J., & Lindy, J. Hyperkinetic children: Early indicators of potential school failure. Community Mental Health Journal, 1970, 6, 447-454.
- Schrupp, M. H., & Gjerde, C. M. Teacher growth in attitudes toward behavior problems of children. Journal of Educational Psychology, 1953, 44, 203-214.
- Sigel, I., Starr, R., Secrist, A., Jackson, J. P., & Hill, E. Social and emotional development in young children. In E. H. Grotberg (Ed.), Day care: Resources for decisions. Washington, D. C.: Office of Economic Opportunity, 1971.
- Srole, L., Langner, T. S., Michael, S. T., Opler, M. K., & Rennie, T. Mental health in the metropolis. New York: McGraw-Hill, 1962.
- Stennett, R. G. Emotional handicap in the elementary years: Phase or disease? American Journal of Orthopsychiatry, 1966, 36, 444-449.

- Stone, L. J., & Church, J. Childhood and adolescence (2nd ed.). New York: Random House, 1968.
- Thomas, A., Chess, S., & Birch, H. Temperament and behavior disorders in children. New York: New York University Press, 1968.
- Ullmann, C. A. Mental health screening of school children. Public Health Report, 1952, 67, 1219-1223.
- Walker, D. K., Bane, M. M., & Bryh, A. S. The quality of the Head Start Planned Variation data. Cambridge, Mass.: Huron Institute, 1973.
- Werner, E., & Bayley, N. The reliability of Bayley's revised scale of mental and motor development during the first year of life. Child Development, 1966, 37, 39-50.
- Westman, J. C., Rice, D. L., & Bermann, E. Nursery school behavior and later school adjustment. American Journal of Orthopsychiatry, 1967, 37, 725-731.
- White, M. A., & Harris, M. W. The school psychologist. New York: Harper, 1961.
- Wickman, E. K. Children's behavior and teachers' attitudes. New York: Commonwealth Fund, 1928.
- Zax, M., & Cowen, E. L. Early identification and prevention of emotional disturbance in a public school. In E. L. Cowen, E. A. Gardner, & M. Zax (Eds.), Emergent approaches to mental health problems. New York: Appleton-Century-Crofts, 1967.
- Zax, M., Cowen, E. L., Rappaport, J., Beach, D. R., & Laird, J. D. Follow-up study of children identified early as emotionally disturbed. Journal of Consulting and Clinical Psychology, 1968, 32, 369-374.

REFERENCE NOTES

1. Hunt, J. McV. Heredity, environment, and class or ethnic differences.
Paper presented at the Invitational Conference on Testing Problems, Educational Testing Service, 1972.
2. Hunt, J. McV., Paraskevopoulos, J., Schickedanz, D., & Uzgis, I. C. On the range of reaction for age of achieving object permanence. Paper presented at the annual meeting of the American Psychological Association, 1972.
3. Severson, R. A. The rationale and development of teacher-completed behavior rating scales. Mimeo., Madison: University of Wisconsin, 1971.
4. Gordon, S. Emotionally disturbed children in the public elementary school--a Philadelphia Report. Mimeo., Samuel S. Fels Fund, 1963.
5. Bayley, N. A scale of motor development for preschool children (2-6 years). Restricted Research Form, given out by direct permission of the author.

ACKNOWLEDGMENTS

I wish to express my gratitude to Dr. Damaris Pease for her supervision of the research and for her guidance and encouragement during my graduate study. I am also grateful to Dr. Leroy Wolins for his aid in the design and statistical analyses of the study and for his encouragement throughout my graduate work. Appreciation for guidance and service on my committee is extended to Dr. Gerald Klonglan, Dr. Sam Clark, Dr. Dahlia Stockdale, Dr. C. Arthur Sandeen, and Dr. Margueritte Scruggs.

The support provided by the staff of the Mid-Missouri Mental Health Center and the cooperation of the directors and teachers in the day care centers are gratefully acknowledged. The collaboration of Gilbert Hutchins in the construction of the motor test equipment and in the utilization of the trailer as a mobile testing site at the various centers facilitated significantly the collection of data. Terry Hedrick, Mary Ann Barr, Sharon Schneeberger, and Janet Becherer ably assisted in the administration of the tests to the children. The young children who participated in the testing program made the data collection a delightful experience.

I would like to take this opportunity to give special thanks to my family for willingly cooperating in my educational endeavors. My husband wholeheartedly supported my decision to change my former status as a college drop-out and his continuing moral support has been an unfailing source of strength for me. My sons learned early on that household chores are not always women's work as they encouraged my

efforts to pursue an academic interest in the development of all children. My mother's thoughtful concern for others and my father's dedicated service to children have provided inspiring examples for their children and grandchildren. I wish to express my most sincere appreciation and affection to my family.

APPENDIX A. PERMISSION FOR TESTING CHILDREN

May 1, 1973

Mrs. Dorothy Hutchins, Child Development Specialist and consultant to the Preschool Center, is undertaking a project concerned with assessment of the growth and development of young children. The Preschool Center is participating in this project.

Your child has been one of the children selected. This will require only about one hour during your child's regular time at the center. During this time your child will work with materials appropriate to young children which are game-like in nature (blocks, puzzles, ball playing, etc.) Only those persons who work with your child in this session and the teacher will have knowledge about the results of this session. The results will be held in the strictest confidence.

We hope that you will grant permission for your child to participate.

Mrs. Janice Jones
Director, Preschool Center

Mrs. Dorothy Hutchins
Child Development Specialist

_____ has my permission to participate
in this child development project.

Parent or Guardian

APPENDIX B. PRESCHOOL SCALE OF MOTOR DEVELOPMENT

A SCALE OF MOTOR DEVELOPMENT

for
Preschool Children (2 - 6 years)
Restricted Research Form, given out only by
direct permission of the author

Nancy Bayley

This scale is an upward extension of the California Infant and Preschool Scale of Motor Development (1935). It is based on tests of approximately 50 children born in 1928 and 1929, tested repeatedly at three or six months intervals over this age range (The Berkeley Growth Study). As presented here, the items overlap somewhat those in the scale for infants, in order to make the two scales continuous. The item numbers and system of cumulative point scoring match those of the original 1935 version of that scale. The general directions for that scale apply to this one, with suitable adjustment for age-appropriateness.

It is hoped to secure current normative material on these tests, and then to publish them, as a continuation of the revised form of the Scale of Motor Development in the First Two Years (1961 Research Form).

Materials Required for Administering Items 47 through 130.

The tests should be given in a room which allows ample space for a three meter runway, the walking board, and the high-jump tests.

Equipment includes two flights of three stairs, to be placed back to back so the top stairs form a platform. The stairs have a lift of 6½ inches, are 10 inches deep, and 24 inches in width. The stairs should be placed so there is a wall at one side for the child to reach for support if needed.

A walking board, eight feet long, and $2\frac{1}{2}$ inches wide, with top surface four inches from the floor. This is made according to the specifications given by Baldwin and Stecher. See accompanying specifications.

A tennis ball (preferably one which has lost most of its elasticity) and "basket" to throw the ball into. This basket is a circular hole eight inches in diameter cut in a board which is 28 inches high and 18 inches wide. The hole is placed so its upper edge is center $2\frac{1}{2}$ inches from the upper edge of the board. A cloth bag is attached at the back of the hole, to catch the ball. The board is placed so that the lower edge of the opening is 28 inches from the floor, and with braces attached so that it tilts back at an angle of about 15° . The board is painted with a two-inch band of contrasting color circling the hole.

Two "high-jump" standards, one meter high made of 4 x 4 boards nailed to a base to keep them upright. Small ($\frac{1}{4}$ inch) holes are drilled at 2cm. (.8 inch) intervals (from center to center of the holes) up one side of each standard. Two pegs (one for each standard) which fit into these holes and may be moved to adjust the height of the cross-stick. The pegs should extend out about three cm. ($1\frac{1}{4}$ "") when in place. A lightweight (bamboo or other light material) cross-stick, about $\frac{1}{2}$ to $\frac{3}{4}$ inches thick, and 1.5 meters long.

A stop-watch, strong twine, chalk, measuring tape, individual record forms.

Directions for Administering and Scoring Test Items

The following directions are given in the order of difficulty of the lowest level at which each test is scored.

47, 48. Stands on one foot with help. While holding one of the child's hands, endeavor to persuade him to lift one foot, to show you his shoe, or to touch with it an object (toy) held just off of the floor. Repeat for the other foot. Credit for each foot on which he stands, with help.

49, 50, 53, 54, 66, 76. Walking up and down stairs. Ask the child to walk up a flight of three steps, and then down again. The younger children may be motivated by a toy, etc. Record whether he marks time or alternates, and whether he holds on to the wall for support. Credit at item 49 (20.3 mo.) if he walks up the steps, holding on to the wall for support. Credit at item 50 (20.5 mo.) if he walks downstairs holding to wall for support. Credit at 53 (24.3 mo.) if he walks upstairs without using his hands for support, standing on each step with both feet before stepping up to the next. Credit at 54 (24.5 mo.) if he walks downstairs by the same method. Credit at 67 (35.4 mo.) if he walks upstairs, stepping to each successive step with alternating feet without balancing with both feet on a step. Credit at 74 (45.0 mo.) if he walks down the stairs by the same method.

51, 55, 60, 70, 82, 89, 98, 116. Walking Board. Place the walking board in the center of the floor where he can reach nothing for support with his hands, and ask him to walk the length of it. Demonstrate if necessary, walking the length of the board, alternating the forward foot. If he succeeds in walking the full length of the board, record the time, telling him how long he took, as a hint for speed, and ask him to walk the board two more times. Credit at 51 (22.5 mo.) if he tries to stand on the board; at 55 (27.6 mo.) if he walks on it one foot off, one foot

on; at 60 (31.0 mo.) if he succeeds in standing on the board with both feet for a few seconds; at 70 (38.0 mo.) if he takes two or more alternating steps on the board before stepping off (alternating steps are left foot in front of right, right in front of left, and so on). Credit at 82 (56.0 mo.) if he walks the full length of the board, alternating the forward foot; at 89 (59.5 mo.) if he does this in 9" or less; at 98 (66.0 mo.) if his best score of three trials is 5" or less; at 116 (80.0 mo.) if his score is less than 3".

44, 52, 63. Aufstehn pattern. Have the child lie flat on his back on the floor, then ask him to stand up quickly. (This test is often resisted by the younger children, so should be introduced with caution). He may be playfully tossed in the air and then put down gently on his back. Credit one point (44, 14.0 mo.) if he rolls completely onto his stomach before getting to his knees and standing; 2 points (52, 22.5 mo.) if he turns half-way, resting on one arm; and 3 points (63, 32.7 mo.) if he pulls directly forward to a sitting position before standing.

56, 73, 77. Jumping over string. Fasten one end of a heavy cord about one meter long, to the wall at a height of 20 cm. from the floor. First allow the string to lie along the floor in a straight line, and demonstrate jumping over it with both feet together. Encourage the child to do the same. Credit at 56 (28.0 mo.) if he succeeds in jumping off the floor with both feet together. If he is successful in this, raise the string from the floor, holding it loosely so that the center part of the string is about 5 cm. from the floor, and encourage him to jump over the string, jumping with both feet together. Demonstrate if necessary. Credit success at 73 (41.5 mo.). If he succeeds in jumping over the string

at 5 cm. (3 trials) raise the string by about two centimeter intervals, until a height is reached which he cannot clear in three trials, until the string is horizontal at 20 cm. Credit at 77 (50.0 mo.) for success at 20 cm.

57, 58, 85, 94, 99, 106, 107, 111, 112, 117, 122, 123, 127, 129.

Stands on one foot. Demonstrate standing on one foot on a chalk line.

Then ask the child to do the same. Start the stop watch as soon as he is in position, holding where he can see it, and saying "Let's see how long you can stay on one foot. See if you can stay a whole minute."

(This last sentence should be omitted for the younger children.) Stop the watch as soon as the child puts his foot down, or at the end of one minute, and record the time. Then ask him to stand on the other foot, and repeat as above. Credit at 57 (29.2 mo.) if he stands alone momentarily on his left foot, and at 58 (29.3 mo.) if he does the same on his right.

Credit for time scores as follows:

<u>Right foot</u>			<u>Left foot</u>		
Item	Age	Time	Item	Age	Time
85	57.2	5"	86	25.5	5"
94	63.5	10"	99	66.1	10"
106	71.0	15"	107	71.5	15"
112	74.3	20"	111	74.2	20"
122	84.0	30"	117	80.0	30"
129	92.0	40"	123	84.5	40"
			127	90.2	60"

59, 67. Walks on tiptoe. Demonstrate walking on tiptoe on a chalk line 3 meters (10 feet) long, drawn on the floor. (Alternatively an inch-wide strip of colored adhesive (e.g. mystic tape) may be placed

permanently on the floor.) Encourage the child to do the same. Credit at 59 (30.1 mo.) if he succeeds in walking a few steps without his heels touching the floor; credit at 67 (36.2 mo.) if he walks tiptoe three meters (not necessarily on the line.)

61, 71, 84. Walks on a line. Using the three-meter (10 foot) line demonstrate to the child by walking on it the entire distance, always stepping on the line. Ask the child to do the same. Credit at 61 (32.1 mo.) if he succeeds in walking close to the line (usually stepping on it with one foot, or astride it) for the entire length. Credit at 71 (38.0 mo.) if he walks the entire length of the line without stepping off. Then demonstrate walking with eyes closed, hands held in front to touch the wall at the end of the line, and ask the child to do the same. Credit at 84 (57.0 mo.) if he walks the entire 3 meters without opening his eyes. Allow considerable leeway in direction, as a child often veers a foot or more to the right or left.

62, 68. Jumps from a height; and 69, 72, 74, 100, 124. Distance jump from height. Ask the child to jump off the first stair (6½ inches high). Demonstrate if necessary. Credit at 62 (32.1 mo.) if he jumps with both feet together. Then demonstrate jumping from the second step, jumping off the side (not over the lower step) with both feet together, and encourage the child to do the same. Credit at 68 (37.1 mo.) if he succeeds in jumping from this height with both feet together. Then ask him to jump from the same height, as far as he can, pointing to a spot some distance away and asking "Can you jump clear out to here?" After he has jumped, measure with a tape the distance from the base of the

platform to the point where the nearest heel struck the floor. Give three trials, recording all three distances, and credit the greatest distance jumped, according to the following:

Item	Age	Distance	
		cm.	inches
69	37.3	10	4
72	39.7	36	14
74	48.4	60	24
100	66.2	86	34
124	87.5	111	44

65. Walks backward three meters. Demonstrate walking backward on the three-meter line, and encourage the child to do the same. Credit (33.0 mo.) if he walks backward the entire distance. (It is not necessary to keep the feet on the line.)

75, 80, 87, 90, 93, 95. Hopping on one foot. Using the three-meter (10 foot) line, ask the child to hop on one foot the entire distance. Demonstrate if necessary. Then ask him to repeat with the other foot. Credit for each foot as follows:

<u>Right foot</u>			<u>Left foot</u>		
Item	Age	Distance	Item	Age	Distance
75	49.3	2 or 3 hops	80	54.5	2 or 3 hops
87	58.0	6 feet	93	62.5	6 feet
90	60.3	10 feet	95	63.6	10 feet

73. (50.0 mo.) Jumps to tiptoe from second step. Using the second step of the stairs, demonstrate jumping from the side with both feet together, and landing on the toes. Encourage the child to do the same, calling attention to the tiptoes. Allow 3 trials. Credit if he succeeds in landing on his toes without his heels touching.

79, 91, 103. Ball throw and catch. Place the "basket," so there is at least 8 feet in front of it, and with the chalk draw a cross on the floor 5 feet from the basket. Ask the child to get the ball out of the little boy's basket, and then direct him to stand on the chalk mark and throw the ball into the basket. Give three trials. Note which hand is used, and whether the throw is over- or under-hand. Then stand 5 feet from the child who is standing on the chalk mark, and start a game of toss and catch, by asking him to throw the ball to you so you can catch it. After warning him to be ready to catch the ball, toss it, underhand, so it will fall directly in front of him at the level of his hands, when they are held up to catch it. Score on the first three throws which fill this requirement. If the child is successful in catching the ball in two hands, demonstrate catching with one hand, and ask him to do the same, giving three additional trials. Credit at 79 (54.0 mo.) if he succeeds in catching the ball in his arms, 1 of 3 trials. Credit at 91 (61.5 mo.) if he throws the ball into the opening of the basket in 1 of 3 trials. Credit at 103 (68.5 mo.) if he catches the ball in two hands. Catching with one hand is scored at a level of difficulty above the standards in the present scale.

81, 96, 105, 108, 119. Stands toe to heel, eyes open. Draw a straight chalk line about two feet long on the floor (or use the tape line). Demonstrate standing on the line while saying: "See, I stand on the line, with my feet pointing straight, one foot in front of the other, and my toe touching to my heel." Then ask the child to stand on the line in the same way. As soon as he is correctly placed, start the stop watch, holding it so he can see it, and say: "Let's see how long you can stay

that way, without moving your feet. See if you can stay a whole minute." Stop the watch as soon as a foot moves more than an inch out of its original position, or at the end of 60 seconds. Record the time: Credit if he can keep his feet in position for at least 10 seconds. Credit as follows:

Item	Age	Time, Seconds
81	55.5	10
96	64.0	20
105	70.5	30
108	72.2	40
119	80.5	60

83, 92, 101, 109, 115, 118, 128. High jump. If the child passes item 77, place the wooden high-jump standards in position, giving him a runway of about 1 meter, and turned so that a failure will allow the cross-stick to fall easily to the floor. Place the cross-stick at a level low enough that the child can be sure of success in jumping over it, and ask him to jump over with both feet. Demonstrate if necessary. (A step-jump, one foot in front of the other, is permissible). With each successful jump raise the cross stick one notch until a height is reached which he cannot jump over in three trials. Credit the best of three trials, as follows:

Item	Age	cm	Height approx. inches
83	56.5	20	8
92	62.0	24	9.5
101	66.5	28	11
109	72.3	32	12.5
115	79.0	36	14
118	80.5	40	16
128	92.0	44	17.5

88, 97, 110, 125. Jump-and-reach. Ask the child to stand facing a smooth wall, and reach as high as he can while standing with his toes touching the board, heels on the floor. This point is marked with chalk and then he is asked to jump as high as he can. Mark with chalk the point reached by the tip of his fingers. Demonstrate if necessary. Measure the distance between his standing reach and each jumping reach. Credit the highest jump as follows:

Item	Age	Jump	
		cm	inches
88	59.0	6	2.5
97	64.0	10	4
110	74.0	14	5.5
125	89.5	18	7.5

102, 114, 120, 130. Stands toe to heel, eyes closed. Ask the child to stand as in 81, but as soon as he is in position say, "Now close your eyes and see how long you can stay with your eyes shut tight so you can't see." Start the watch as soon as he closes his eyes, stop it when his eyes open or when he moves his feet off the line, or at the end of one minute. Watch his eyes closely, make a remark from time to time. Tell him how much time has elapsed at the end of each quarter minute. Record the time. Credit as follows:

Item	Age	Time, seconds
102	67.0	5
114	78.0	10
120	81.2	20
130	95.5	30

104, 113, 121, 126. Stands on toes, feet together. Ask the child to place his feet together, toes and heels touching (demonstrate) and then go up on his toes, saying "Let's see how long you can stay on your toes." Start the stop watch as he goes up on his toes, and hold it where the child can watch it. Stop the watch if his heels touch the floor, if his feet move apart, or at the end of one minute. Record the time. Credit as follows:

Item	Age	Time, seconds
104	69.0	10 or more
113	77.0	20
121	82.0	30
126	90.0	40

Directions for Scoring the Test

Using as a basal score the number of the successful test items below which there are no failures, add to this score one point for each item passed above it. This point score may be converted into a Sigma Score, using the means and standard deviations as given in the table, by the formula:

$$\text{Sigma score} = \frac{\text{Child's score} - \text{mean}}{\text{standard deviation}}$$

The mean and standard deviation for the age nearest the child's chronological age, as given in the table, should be used. A sigma score of zero means that the child's performance is average for his age; a plus sigma score means that he is above average; a minus sigma score means that he is below average. His motor age is the age in the table (not on the record form) nearest the point score attained by him.

If it has been necessary to omit one or more of the test items, the point score may be computed according to the Thurstone method by counting the

number of failures, starting with the easiest item failed and counting up toward the more difficult ones; then counting the number of passes, starting with the most difficult item passed and counting down toward the easier items. Of course, do not count the items omitted. Find the point at which the number of items passed is equal to the number of items failed. The item number at this point is the point score to be assigned, and to be used in computing the sigma score.

APPENDIX

Walking Board Specifications

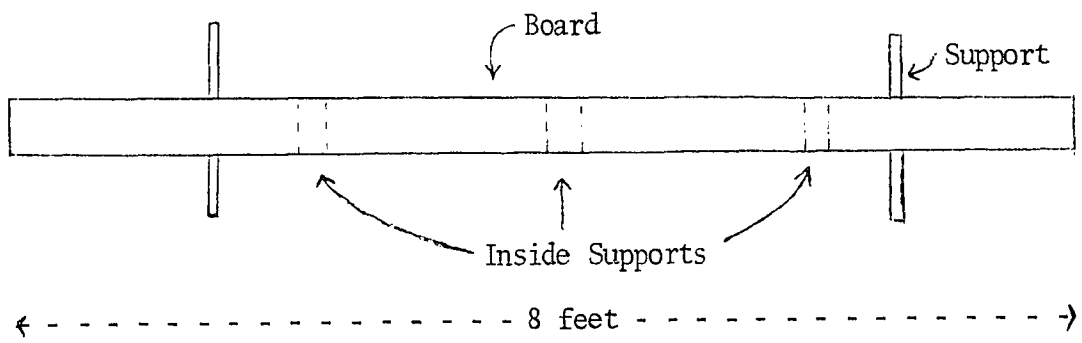
The Walking Board was used originally by Buford Johnson, and is described by Baldwin, B. T., and Stecher, L. I., Psychology of the Pre-school Child. New York, Appleton, 1924.

This board was originally made of an eight-foot length of two-by-four-inch plank. It is turned so the narrower side forms the top surface, and is held in position by two cross pieces, one placed near each end, at a distance of 1.5 feet from the end.

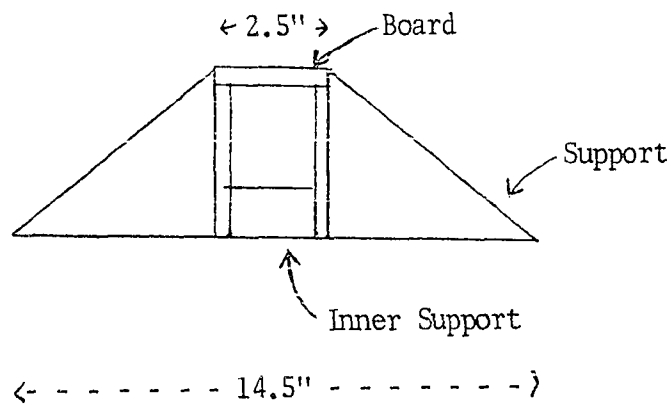
The board used in the California Infant Scale of Motor Development (drawing A) was made lighter, for ease in handling, by nailing together half-inch planks in such a way as to form a hollow, three-sided structure open at the bottom with outside dimensions $2\frac{1}{2}$ " by 4" by 3 feet. In cross section the board would look like drawing B. It should be made stable by placing 1.5 inch blocks inside at the bottom edge at 2-foot intervals. The positioning supports can be triangular pieces, cut from 4" by 6" rectangles of 1" by 4" wood. If they are nailed or glued securely to the side pieces before the top piece is attached the entire structure can be made firm.

BAYLEY MOTOR DEVELOPMENT SCALEWALKING BOARD

A. Top View



B. Cross Section



References

1. Baldwin, B. T., & Stecher, L. I. Psychology of the preschool child. New York: Appleton, 1924.
2. Bayley, Nancy. Development of motor abilities during the first three years. Monog. Soc. Res. in Child Developm. Vol 1, No. 1, 1935. pp 26.
3. Thurstone, L. L. The scoring of individual performance. J. Educ. Psychol., 1926, 17, 446-457.

APPENDIX C. TEACHER RATING SCALE

TEACHER APPRAISAL OF PRESCHOOL CHILDREN

Teacher's name _____

Center _____

Date _____

Dorothy Hutchins
May, 1973

For each of the following children, indicate the date he entered your class and how many hours per week you work with him.

<u>Name of Child</u>	<u>Entered Class</u>	<u>Hours per week</u>
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

Comments:

When we are working with children, we often compare the growth and development of a specific child to that of other children we have known. Based on your knowledge of and experience with children, think about what you consider normal development. Then rate each of the following children in your class on the basis of his overall development. Remember, this is a comparison of this child to all other children.

(1) indicates the child has attained a very high level of development

(3) indicates the child has attained an average level of development

(5) indicates the child has a low level of development

(2) and (4) may be used for in-between levels

Please circle one number for each child.

<u>Name of Child</u>	<u>Rating</u>
_____	1 2 3 4 5
_____	1 2 3 4 5
_____	1 2 3 4 5
_____	1 2 3 4 5
_____	1 2 3 4 5

In looking at the growth and development of a child, we usually think that how he gets along with others is important. Rate each of the following children in your class on the basis of his social development.

- (1) indicates the child has attained a high level of social development
- (3) indicates the child has attained an average level of social development
- (5) indicates the child has a low level of social development
- (2) and (4) may be used for in-between levels

Please circle one number for each child.

<u>Name of Child</u>	<u>Rating</u>
_____	1 2 3 4 5
_____	1 2 3 4 5
_____	1 2 3 4 5
_____	1 2 3 4 5
_____	1 2 3 4 5

The California Scale of Preschool Development assesses the child's motor development. This includes balance, walking, skipping, playing ball, etc. Rate each of the following children on the basis of how you think he will perform on this test.

(1) Indicates the child will do very well

(3) Indicates the child will do average

(5) Indicates the child will do poorly

(2) and (4) may be used for in-between levels

Please circle one number for each child.

<u>Name of Child</u>	<u>Rating</u>				
_____	1	2	3	4	5
_____	1	2	3	4	5
_____	1	2	3	4	5
_____	1	2	3	4	5
_____	1	2	3	4	5

The McCarthy Scales of Children's Abilities is used to evaluate a child's cognitive development. This includes his use of language, use of number concepts, memory and general intellectual ability. Rate each of the following children on the basis of how you think he will perform on this test.

(1) indicates the child will do very well

(3) indicates the child will do average

(5) indicates the child will do poorly

(2) and (4) may be used for in-between levels

Please circle one number for each child.

<u>Name of Child</u>	<u>Rating</u>				
_____	1	2	3	4	5
_____	1	2	3	4	5
_____	1	2	3	4	5
_____	1	2	3	4	5
_____	1	2	3	4	5