DOCUMENT RESUME

ED 068 161 PS 005 896

AUTHOR Sigel, Irving E.; And Others
TITLE Psycho-Educational Intervention Beginning at Age Two: Reflections and Outcomes.
INSTITUTION State Univ. of New York, Buffalo.
PUB DATE 72
NOTE 50p.; Paper presented at Second Annual Blumberg Symposium on Compensatory Education, Johns Hopkins Univ., 1972

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Behavioral Objectives; Child Development; *Cognitive Development; Cognitive Objectives; Comparative Testing; *Compensatory Education; Concept Formation; Curriculum Development; *Early Childhood Education; Evaluation Techniques; Experimental Programs; Inservice Teacher Education; *Intervention; Low Income Groups; Negro Youth; Rating Scales; Research; Sex Differences; *Social Development; Speeches; Test Results
IDENTIFIERS Bayley; Kagan; Piaget; Stanford Binet

ABSTRACT

The Early Childhood Education Project (ECEP) is an experiment in the effects of educational intervention of two-year-old first-born children from impoverished black families. Twenty ECEP children were compared to a control group on the basis of a battery of pre- and post-tests after six months of intervention training. Teachers attempted to develop social-emotional adjustment and cognitive behavior. Training occurred in both group and tutorial sessions. Narrative observations, rating scales and observation of test-related behaviors were measures of social-emotional development and a battery of tests measured cognitive skill. Preliminary findings indicate that the intervention program was effective in the development of cognitive skills, especially in language, pantomime, seriation and train building. Only sex differences are compared for social-emotional development. Boys were more cooperative, defended their possessions and were more aggressive to people and objects than girls. Girls were more compliant to demands and initiated more interactions with teachers. (DG)

PSYCHO-EDUCATIONAL INTERVENTION BEGINNING AT AGE TWO: REFLECTIONS AND OUTCOMES[1]

Irving E. Sigel

Ada Secrist George Forman

State University of New York at Buffalo

Buffalo, New York 14214

Since the advent of Head Start the renewed activity of programming for preschool children and the evaluation of the effects of such programming have reopened Pandora's fabled box. We have learned that the issues in programming and in evaluation are complex, taxing our theoretical and methodological ingenuity and forcing us to reexamine our research and teaching strategies. The state of affairs at this juncture is not only the product of ignorance or naivete on the part of the investigators, but is equally a result of failures in finding expected outcomes. Beller (in press) has pointed out that this is a good time for reassessment. Reexamination should allow us to profit from past mistakes and more importantly may yield clues concerning underlying mechanisms that could explain the variety of results and give shape and continuity to past and future work.

We present this paper with those hopes in mind. Our purpose is to present one model of preschool programming, the Early Childhood Education Project (ECEP) with a brief evaluation of the first year's work. Perhaps more important than our results at this point, is a description of the problems we encountered and the successful and unsuccessful solutions we found to them.

The Early Childhood Education Project is an experiment in educational intervention begun with two-year-old first-born children from impoverished black families in the inner city of Buffalo. This simple statement immediately raises some basic questions. What is intervention? Why intervene? Why work with two-year-old black children?

(Presented at the Second Annual Blumberg Symposium on Compensatory Education, Johns Hopkins University, 1972. Dr. Julian Stanley, Chairman.)

The frequently used concept of intervention rivals compensatory education in ambiguity and emotionality. It is difficult to find a broader concept. After all, intervention is what all education is about. When illiterate, egocentric, middle-class five-year-olds become literate and able to communicate and take the perspective of others, one can argue for the efficacy of the school's intervention. Intervention if so broadly conceptualized leads to heated arguments concerning the true meaning of intervention, e.g. good vs. poor school systems. Nevertheless, the right of the school system to exist and to intervene is taken for granted.

Intervention must be defined more carefully especially for research purposes. For us intervention is considered as a conscious and purposeful set of actions intended to change or influence the course of development. The ECEP program has operationalized what was meant by "conscious and purposeful set of actions", and specifically stated what change we wanted in the course of development. To deliberately set out to change someone immediately engages several crucial moral and value questions. Who has the right to change anyone? What or who gives him that right? These are but some examples of the multi-faceted questions which need clarification.

It seems to us that intervention programs operate with some of the assumptions of the medical model. Someone who is sick generally elects to go to a physician because the disease is not expected to subside without his intervention or treatment. Elimination of the disease is considered desireable and essential for healthy, future development. While there may be several factors contributing to the disease, e.g, a specific infection, particular work and housing conditions or the quality of a personal relationship within the family, the physician will generally concentrate on the factor he feels will do most to eradicate the illness.

His judgment is generally not questioned, or if it is another physician is sought. His right to treat the patient is certainly not questioned unless a legal issue arises.

Most preschool intervention programs have focused on the impoverished groups, e.g., blacks, rural and urban whites, Mexican-Americans, Indians, and Puerto Ricans. Observations of the past decades have shown that many individuals who have lived in poverty have often had difficulty in acquiring in schools the skills necessary for entering the economic mainstream. Poverty has been assumed to be a major determinant of ineffective school performance and hence, later on to an unproductive adaptation to a complex, technical, urban environment. The assumption is made that education is essential to break the poverty cycle. There are of course, other factors that could be changed; e.g., the economic system, the social structure, the climate of social and political opinions or the system of national priorities.

So far the similarity to the medical model seems uncomplicated. The individual who is bogged down in poverty and who cannot progress up the economic ladder is like the patient with a disease which prevents him from working and from continuing to develop. We choose that aspect of the disease to work on, namely education, which we believe will provide the most relief with the least amount of social change.

There are some important differences however, between the medical model and the intervention program. The patient as a rule chooses his physician. The question of choice in an intervention program is a moot one.

Secondly, the physician's confidence in his choice of treatment is frequently backed up the years of research, observation and proven change. The intervention

doctor has no such comfort. Are we preparing children only to conform to the so-called Establishment? Are we homogenizing rather than individualizing our children? Are we overvaluing intellectual achievement and ignoring other skills? In order to deal honestly with these problems, it seemed to us that constant, careful and continuing analysis of our aims and goals was essential. Furthermore, we also felt that involved participation by parents was necessary in order to provide a mandate for intervention.

After much analysis and soul searching, we chose two major objectives that we felt were consonant with the assumptions of intervention. The first objective concerned intellectual activity. There is no denying that "making it" in middle class America requires competence in conceptual and symbolic behavior. But "making it" from other perspectives we believe, requires similar competences. The second major objective concerned social and emotional variables. We decided simply to pay considerable attention to these factors in terms of the individual and the classroom climate.

Unless program descriptions reported in the literature (Beller in press) are discrepant from the actual social reality of the program, one is struck by the preoccupation with I.Q. to the virtual exclusion of all else. The interaction and interdependence of cognitive and social-emotional factors have been given much lip service and been little studied. It seemed almost self-evident to us that intervention in one of these areas is bound to influence the other. We decided, therefore, that our program would cultivate an environment that maximized opportunities for enhancing a sense of personal competence in both intellectual and social areas.

Given these considerations, the attendant problem was the question of evaluation. Why do we want assessment procedures? What kind of assessment is useful? Granted, such data are needed in order to broaden our base of understanding of young children, but we felt that our assessment program should evaluate developing competencies growing out of our program. Therefore, rather than heavy reliance on an I.Q. score, which in reality tells us very little about particular cognitive processing skills, we have put together a battery of tasks rationalized in terms of our program objectives. We can't describe gain in I.Q. points but we can discuss changes in sequential memory and various analytic and perceptual skills.

As a result of our conviction that the subject-participants of an intervention program should provide a mandate for the program, discussions were held with the parents prior to enrollment of the children. Some parents rejected our rationale on various grounds. A few felt that the children were too young for such a group program; others believed that the mixing of children's backgrounds was inimical to the development of their own children.

Our position is summarized as follows: The course of development for poverty children untouched by remedial programs is predicted to become socially pathological and dysfunctional. The remedial program we suggest emphasizes the acquisition of basic cognitive and social competencies whose interaction within the individual might be expected to extend the options a child has in his adaptation to environmental demands.

It does not seem reasonable to us to place the entire burden of modification on preschool education. This is but one link in the educational chain. Other efforts, such as the Follow Through program for example, would be an important adjunct (Bissell 1971).

Conceptual Framework

In a series of studies beginning in 1965, we discovered that a large proportion of black children from impoverished backgrounds were less competent when dealing with representational material than their more privileged counterparts (Sigel, Anderson, and Shapiro 1966; Sigel and McBane 1967; Sigel and Olmsted 1970; Sigel 1972). These studies revealed that children from impoverished backgrounds had greater difficulty in classification tasks when the task items were colored or black and white photographs in contrast to three dimensional objects. Such discrepancies were found in spite of the fact that the children could label and recognize the pictorial stimuli. The issue becomes, then, not one of recognition, but knowledge that an object can be represented in several modes and still be a member of the same conceptual class. It is easy to confuse recognitory behavior as equivalent to "knowing" or comprehension of equivalence. Not only did these children have difficulty in responding equivalently to pictorial stimuli and their three dimensional counterparts, they also had difficulty with other tasks involving representational skills, e.g., Motor Encoding from the Illinois Test of Psycholinquistics; Dramatic Play (Sigel and Perry 1968).

A review of the literature does not reveal many data specific to this issue, but some can be reinterpreted to provide further evidence concerning the problem. For example, Kamii and Radin (1969) examined the types of Stanford Binet items which children in the Perry School Project found difficult, found them to be conceptual items as compared to the more rote perceptually concrete types. It may be that the reason I.Q. scores, so often used as an evaluation measure for intervention programs, show a later decline is because the test items involve more conceptual-verbal items than rote-concrete ones. These data suggest then, that

the generic problem is development of representational competence, i.e., the ability to deal with representational material. This is a capacity basic to any symbolic activity and even considered by Piaget (1962) to be intimately related to thought.

Examination of life experiences of children from impoverished and deprived environments, suggests that these life experiences reduce the opportunity for representational competence to flourish. For example, Hunt writes that "children of poverty lack opportunities to develop cognitive skills..." (Hunt 1969 p. 204). In referring to parent-children interactions he says "what these parents talk about is also lacking in such conceptual constructions as prepositional relationships, casual explanations, and concepts of space, time, and justice... The parents of the slums not only talk less with their children than do parents of the middle class, but they seldom undertake to discuss with their children matters which prompt them to discern various kinds of relationships among things and people or to use language to describe these relationships." (Hunt 1969, p. 205-206). It is reasonable to assume that these kinds of experiences may be critical antecedents for the development of representational competence.

If the capacity to represent the environment or to symbolize, is a generic human ability and if competence can be affected by certain kinds of life experiences, than a careful examination of experiences that facilitate or preclude representational competence seems in order.

We hypothesize that the most relevant experiences are those that would involve orientation to the environment by anticipation or objectification of temporal, spatial, and causative relationships. It seems to us that certain classes of socializing experiences function to "demand" employment of representational activity. It may well be that the necessary and/or sufficient conditions that set the processes of representation in motion are those that serve to create psychological, spatial, or

temporal distance between person and object. These behaviors, referred to as distancing behaviors, are the class of events that create psychological distance between ostensive reality and its reconstruction (Sigel 1970). Thus, the basic hypothesis emerged, namely, that exposing children to distancing behaviors should enhance the development of representational skills.

Distancing behaviors, however, will only be effective stimulants if the recipient is motivated to engage and to interact with the significant person or event. Exposure without the prerequisite willingness to participate precludes any effect. To create such a climate necessitates an environment in which significant adults are warm, accepting and sensitive to the child's status. Adults should be able to listen, to appreciate the child's perspective and be "tuned" into his comprehension level. Thus, the adult must be able to assess the child's developmental level and respond appropriately.

When Intervention Should Occur

Representational thought begins to emerge somewhere between 18 and 24 months or the period between the sixth stage of the sensorimotor period and the beginnings of pre-operational thought (Piaget 1955; Inhelder and Piaget 1964).

Representational thought is characterized as the mental activity of evoking objects and events which are outside the immediate field of perception. Thus, representational activity extends the perceptual field of the child from the observable present to the past and the future. It involves anticipatory behavior as well as hindsight. The products of representation are symbols and signs which are differentiated from their concrete palpable referents. Somewhere within the 18 and 24 month period children become capable of re-enacting past events, such as re-enacting home experiences, e.g., preparing meals in the nursery's housekeeping corner, putting a doll to sleep, etc.

Space does not permit a detailed description of this developmental period and the reader is referred to Overton (1972, pp. 97-100), Ginsburg and Opper (1969), Piaget (1955, 1962) and Inhelder and Piaget (1964). Suffice it to say, that this period is of particular significance since it defines the origins of what becomes adult thought.

Development is a cumulative process. What happens at one stage or period influences the direction that development takes at subsequent stages. It is subsequent to this period that differences in cognitive development seem to occur (Golden and Birns 1968). It seems reasonable to assume that providing additional experiences which help foster and encourage representational thought should contribute to the development of the semiotic function. We made the additional assumptions that an appropriate environment is a necessary condition for fostering representational thought and that the appropriate environmental experiences can be hypothesized as distancing behaviors. The intervention program becomes in effect, the opportunity to extend the frequency and quality of behaviors deemed relevant to activating and maintaining representational thinking.

These were the reasons that age two was selected as the target period.

Mode of Intervention

There are many options for intervention but the choices are basically between individual and group settings. We opted for the latter in spite of the dearth of information about group educational programs for two-year-old children.

The appropriateness of a group setting was based on a number of premises both theoretical and practical. First, we believe cognitive growth is enhanced by a broad experiential base, with experiences in various contexts and with various

materials. A nursery school setting seems appropriate. Furtherm ore, a nursery school can provide a more intensive and cumulative contact with the social and non-social environment than the home. Finally, it provides the opportunity for setting up a sequential set of experiences that can be used to reinforce acquired gains.

Teacher Training and Curriculum Planning

The next task involved devising a program to meet our objectives. This involved the simultaneous tasks of training teachers and creating the curriculum.

None of the teachers had direct experience in working with two-year-olds in a nursery school that had explicit curriculum objectives. It was necessary to orient the teachers to the theoretical system and this required spelling out the concept of representational thinking and defining strategies which would exemplify "distancing". Such an orientation was important in order to provide coherence to the program, to enable the teachers to participate actively in curriculum development and thereby to employ appropriate teaching strategies.

In-service training is a continual process built directly into the program. Regular staff meetings are held weekly as well as daily brief reviews of the morning sessions. It is at these meetings that curriculum units are planned and explicated, management problems discussed, and theoretical points argued.

Fostering both group and individual participation became a major challenge because of the array of individual differences. Some children were very articulate and had relatively long attention spans whereas others demonstrated a fairly low level of socialization. Management procedures, therefore, had to be coordinated to cognitive objectives. The control techniques that were given special emphasis came from research conducted several years ago at the Merrill-Palmer Institute

(Hoffman 1960; Sigel 1960). This work basically systematized the influence techniques used by parents to modify the ongoing behavior of their children. The value for our program was the consonance of these techniques with the concept of imposing distance on the child.

The Curriculum

The curriculum contains two coordinated but separate programs. The first is the daily classroom program and the second consists of tutorial sessions.

Classroom Program: Social and Emotional Objectives. The basic objective is to raise the level of socialized behavior. This is done primarily by preventing aggression to others, and encouraging sharing and cooperation. Individual and group interactions are used to make the child aware of others and aware of his environment. This implies that the child is simultaneously being encouraged to anticipate future events, to develop higher levels of frustration tolerance, and to verbalize his feelings rather than to demonstrate them by hitting, biting or pinching. Anticipation and verbalization are essential to imagery and symbolic behavior so that this social objective clearly serves the cause of improving representational competence.

Classroom Program: Cognitive Objectives. All our objectives concern the acquisition of specific relationships and concepts that we believe to be central to representational competence. Furthermore, all objectives demand a particular kind of teacher-child interaction. The teacher must be sensitive to the cognitive possibilities of any situation and she must implement the concept of "distance" when directing the child in his activities.

1. The child must consider things that are not present. The teacher must challenge the child to think of the not present. This may take the form of a question, such as asking the child what would happen if an over-sized block were placed on an already unsteady tower, or asking the child to imitate the action of a sight seen the day before. The teacher may make requests which force the child to think in past and future terms. For example, she may ask the child to find a spoon to use in the pudding she is going to mix.

The content varies widely, as does the child's mode of response. As often as possible, the teachers attempt to encourage verbal expression rather than simple motor or gestural expression.

2. The child must search for alternative actions when unable to solve a problem successfully. Varying the response to a difficult problem situation maximizes the probability of success.

3. The child must recognize that one object has several different properties. This objective is intimately related to the second objective. Learning to explore the multiple attributes of a given object is a precursor to problem solving and possibly of classificatory skills. The timing of the teacher intervention is critical and must enable the child to persist and shift to alternatives.

4. The child must be made aware of temporal and physical relationships, whenever possible. Smaller-bigger, higher-lower, full-empty, dark-light, for example, are simple relationships which categorize objects of all kinds. These relationships can also be part of simple cause-effect conditions. For

example, if you flick a switch a light goes out and it becomes dark. If you pour material into a container it is no longer empty but full. Categorizations and cause-effect conditions seem to us to be essential components of representational competence.

Incorporating training of this kind into our program was most effectively done by planning mini-lessons suitable for the activity areas of the classroom, such as the kitchen area, sand and water play, transportation and blocks, painting, and manipulative toy area. The following miniature lesson is an example.

The lessons suggested for water play vary according to what is placed in the water bin. Cups and funnels of graduated sizes make more-than less-than comparisons easy to elicit. For example, if a full large cup of water is poured into a small cup, the small cup will overflow but the reverse action leaves the large cup unfilled. Teachers will demonstrate how the sound of a stream of water changes with a change in the height of the fall, or they may help a child discover how a lidded cup can be inverted without loss of water, how an empty cup can float, how a small cup can be inserted into a large cup. Throughout the lesson the teacher emphasizes relationships by pointing, using words for action or states, requesting observations and verbal responses from the children.

In order to maintain our basic classroom atmosphere of learning by self-discovery and self-initiated action, the teacher plans a list of possible forms of intervention which she enacts only after noting the child's interest. The planned methods give the teacher general objectives and particular suggestions which prevent unnecessary lapses of activity on her part, but she is free to invent when cued by the child's interest.

All major objectives are carried out within the schedule of routine activities such as free play, juice and rest time and group activities.

Tutorial Sessions. Each child was seen individually in a one-to-one instruction that the child is exposed to the curriculum units which best represent the features unique to the ECEP program. These units once again are designed to provide experiences conducive to the development of representational competence.

Several sets of materials are given to the child in sequence during the course of the year. The teachers' general approach to all materials does not vary greatly.

The child is invited, for example, to play with a set of geometric blocks. No particular response is requested. Generally the child will start some type of exploration, e.g., stacking, standing on edge, etc. The teacher watches approvingly and considers her own moves carefully before intervening. When she does so, she tries to elaborate slightly on something the child has already started. For example, the child may roll a large and a small ring, one in each hand. The teacher might show him that the smaller one fits inside the larger ring and both can be rolled together. At that point the child may ignore the teacher's elaboration of his response; he may initate the elaboration once, or he may initiate the elaboration and then generalize it to other materials.

The objectives of the teachers in these sessions is to focus on the following: conflict inducement, timely presentation of contrasting material with the intention of stimulating flexibility of thought, orientation to a product, development of reflectivity, sensitivity to the negative instance, analysis and synthesis and the developing awareness of cause-effect relationships.

ASSESSMENT

An attempt was made to evaluate both cognitive skill and social-emotional development. Cognitive skills were assessed by means of a battery of tests and social and emotional variables were studied by means of narrative observations, rating scales, and observation of test-related behaviors.

Research Design

Cognitive Assessment. ECEP children were compared with non-ECEP children in order to evaluate the effectiveness of the intervention program. A simple pre-post test design was used. A battery of tests was assembled in order to measure several aspects of cognitive skill. ECEP children were tested twice with essentially the same battery. The first testing (Battery I) was administered upon entry into the program and the second testing (Battery II) was given at the end of the first academic year. Control subjects were not tested twice, however, because of the high dropout rate in the programs from which the control subjects were chosen. Control subjects were compared to ECEP children on the basis of age at time of testing only.

Subjects. Twenty, two-year-old first born children, 10 boys and 10 girls, were enrolled in the ECEP program. They were recruited from the inner city of Buffalo, New York and bussed to the nursery school four mornings a week for two and one half hours each morning.

Both white and black boys and girls, matched in age to the ECEP children, and coming from approximately the same socio-economic level, were used as control subjects. They were all enrolled in other day care programs and nursery schools in order to make them as comparable as possible to the ECEP group.

Test Battery.[2] A battery of 25 tests was assembled to measure as many areas of intellectual functioning as was feasible. The areas tested included form perception, classification, language, memory for images, imitation and pantomime, number and seriation. Most of the tests came from the Bayley and Stanford-Binet scales. Several Piaget classification and sorting tasks were included and a group of tests especially devised by us completed the battery. The complete list of tests used will be found in Table 1.

Social and Emotional Assessment.[3] This involved the ECEP group only. Rating scales, teacher observations, narrative observations of classroom behavior, and observations of child behaviors occurring concomitantly with test performance were used to obtain information. Frequently data were used and comparisons between boys and girls were made as well as correlations between observed social and emotional variables and test performance. (Insert Table 1)

Some Preliminary Findings

The results of our first year of data gathering are not completely analyzed. Control subjects are still being examined and much of the fine-grained analysis remains to be done. The comparisons with control groups are still very tentative.

Cognitive Assessment. Test Results.

The ECEP boys and girls were compared on the initial testing (Battery I), the second testing (Battery II), and on the difference scores between Battery I and II. Each sex was compared to controls of the same and opposite sex on Battery I and II. Since the white control group for both sexes was still very small for Battery I, no statistical comparisons for these groups have been made.

Test results are given in Tables 2 through 5. The means for each group and the significance level of the difference between these means are presented for all the tests in Batteries I and II. Significance was tested with the "t" test. Difference score means for the ECEP group are not shown. (Insert Tables 2 - 5)

1. LANGUAGE TESTS

ECEP boys vs. ECEP girls. There were no significant differences in performance between the ECEP boys and ECEP girls in either Battery I or II but differential improvement was seen in the difference scores. Both boys and girls showed significant improvement in the Stanford Binet Picture Vocabulary. However, in the other tasks, the girls showed improvement in the Bayley Names Pictures and the boys improved in the Stanford-Binet Objects by Use. There is a qualitative stimulus difference in these tests; the Bayley test consists of a two-dimensional picture whereas the differences between the ECEP boys and girls was not significant overall, the boys showed greater general improvement in the language area.

ECEP vs. Control Ss. ECEP boys were comparable to the black control boys in both batteries but scored lower than the black control girls in Battery I in the Stanford Binet Picture Vocabulary and the Bayley Z Scale (Understanding Prepositions). This discrepancy was overcome by Battery II. However, they scored lower than both white boys and girls in vocabulary and the Bayley Z Scale on Battery II. ECEP girls initially scored lower than other black girls on vocabulary but caught up to them by the second testing. However, like the ECEP boys, they scored lower on Understanding Prepositions than both male and female white control groups.

2. IMITATION AND PANTOMIME TESTS

ECEP boys vs. ECEP girls. The girls were superior to the boys in the Bayley M Scale on both batteries.

The pantomime task consisted of asking the child to pantomime a simple action under four sets of instructions, each instruction giving the child a more concrete cue concerning the object to be used in the pantomime. For example the task, "Show me what you do with a pencil" was given as a verbal request, with a life size picture of a pencil present on the table, with the actual object placed on on the table but out of the child's reach, and finally the child was given the actual object to use. Neither sex responded well to the verbal instruction but the girls showed some tendency to better performance when shown pictures or when looking at the object. The boys, on the other hand responded more favorably only when given the actual object to use.

The only significant improvement in performance between Battery I and Battery II is shown by the girls in Pantomime with Picture as Cue. Overall, the girls showed slightly greater improvement in this area.

ECEP vs. Control Ss. There were no differences between ECEP and control groups of either sex in the imitation task (Bayley M Scale). ECEP boys performed better than other black females in Pantomime Using Appropriate Object. ECEP females who were lower than other black females in Battery I in that task, superceded them in that task by Battery II. However, white males were better than ECEP females in Pantomime Using Appropriate Object.

Again, boys and girls seem to be responding differentially to a stimulus quality. Girls are able to perform when clued by a two-dimensional picture. Boys show better performance when clued by an actual object.

3. PERCEPTION AND CLASSIFICATION TASKS.

ECEP boys vs. ECEP girls. There were no differences in the initial testing, but in the second testing the boys were significantly better at the Piaget Large-Small Sorting task. Looking at the difference scores between Battery I

and II, boys showed significant improvement in the Bayley R Scale on the second testing, whereas the girls improved significantly in the Stanford-Binet Three Hole Form Board. The girls also showed greater improvement in the Kagan Embedded Figures task than did the boys.

ECEP vs. Control Ss. The only significant difference in Battery I is between ECEP females and black control females. The control Ss scored higher on the Piaget Large-Small Sorting.

In Battery II, ECEP males and females are significantly better than black male control Ss in the Stanford Binet Rotated Form Board. ECEP males are better than other black males on the Piaget Large-Small Sorting but they did not perform as well as white males. The white females scored significantly higher than ECEP females on this same task, and both white boys and girls scored significantly higher in Identity Matching than the ECEP group.

Better performance on the Piaget Large-Small Sorting and Identity Matching may be confounded with language skill. A preliminary factor analysis of the test batteries indicates that both these tests have a sizeable language component.

The Piaget Classification Task was given to the ECEP group only. A system of notation has only recently been devised which will allow comparable scoring for the control groups.

The Piaget Classification Task consisted of presenting the child with an array of blocks varying in size, shape and color. The child was asked to group these and his behavior, following instructions was videotaped and scored from the tape. Four categories containing 24 variables were systematically scored. These categories were: 1) Non-Grouping Responses which included such behaviors as holding, banging, pushing, random clustering of blocks; 2) Grouping Responses

which involved any indication that the child was putting objects together on the basis of similarity or forming a group by stacking (piling one block on another), aligning (lining blocks up horizontally) taking two blocks and putting them together in consisten fashion; 3) Decision Responses which included exchanging one block for another in order to create a design, scanning and holding blocks to exchange for others, and hesitations; 4) Completion Responses which included any indication that the child was finished with his arrangement. Since this category contained less than 5% of all responses it was not included in the statistical analysis. Table 6 shows the percentage of each category of response for each sex in both batteries. (Insert Table 6)

Clearly, Non-Grouping responses decrease and Grouping responses increase in Battery II for both sexes but particularly for boys. Decision Making responses remain fairly constant. Each of the three major response categories showed some significant changes from Battery I to Battery II for both sexes.

Non-Grouping Responses: Holding, Touching and Hold-Release, decreased significantly for boys. The girls decreased significantly only in Hold-Release.

Grouping Responses: Both girls and boys increase significantly in stacking.

Decision Making: While both sexes generally decreased responses in this category, boys decreased significantly in Visual Examination but girls increased significantly in Adjustments.

The results, while complex, allow us to offer some tentative conclusions. We can infer stages in this classification behavior. Initially children examine objects and function as if they are learning and becoming acquainted with them as evidenced by the frequency of hesitation responses. At the second testing, six months later, familiarity with similar forms or perhaps the memory of these forms, enables the children to be less hesitant and more certain in their productions. The most perplexing phenomenon was the prevalence of stacking as the major Grouping Response, as opposed for example to aligning, or pairing.

4. MEMORY FOR IMAGES

The tasks in this category involve several capacities, language and classification for example. However, image retention seemed to be an underlying similarity that was required for correct performance.

ECEP boys vs. ECEP girls. There were no differences between ECEP boys and girls in the first battery, but in the second testing the boys were significantly better at the Delayed Response Task and the girls at Seriation.

ECEP vs. Controls. ECEP males scored significantly higher in the Delayed Response Task and in Seriation in Battery I than the black male control group, but this difference disappeared in Battery II. The boys scored significantly better than white males in Seriation and Train Building (Battery II) but both white boys and girls did better than ECEP boys in the Sigel Memory Matching Task. Black control girls were significantly higher in Seriation than ECEP boys as were the ECEP girls.

There were no differences in Battery I between ECEP girls and black control girls but in Battery II, ECEP girls were significantly better in Seriation than this control group and were better in tower building than white boys.

5. NUMBER CONCEPT

There were no within group differences. White boys and girls scored higher than ECEP children in both the concept of one and two.

OVERVIEW OF TEST RESULTS

Since the control groups are small and much of the data is still exploratory, it seems reasonable to examine trends and look at differences and changes without taking customary significance levels into account. If one compares the total number of tests in which ECEP scored higher than any given group, relative to the total number of tests in which they scored lower than that given control group, a general picture of improvement becomes clear.

ECEP females showed the greatest improvement when compared to black control females. In Battery I, ECEP scored lower than black control females in 21 tasks and higher in 5 ($X^2 = 10.9$ $p < .001$). In Battery II girls in 14 tests and lower in 8, a proportion not significantly different from chance.

Similarly, ECEP boys improved in Battery II relative to control black females. In Battery I, ECEP males scored higher in 6 tasks and lower in 20 than the control females. ($X^2 = 6.50$ $p < .02$). In Battery II, ECEP males scored higher in 14 and lower in 15 tasks, a proportion not significantly different from chance.

Comparisons with white control groups were not done for Battery I, but there was no overall difference between ECEP females and white females or males in Battery II. ECEP males, on the other hand, were not different from black male control groups over the entire Battery I or II, but performed significantly lower than white males. In Battery II, ECEP males scored higher in 7 tasks and lower in 22 than the white male control group ($X^2 = 6.75 < .01$).

ECEP children showed significant strides in Battery II as evidenced by their catching up to or overtaking the black control groups in Language, Pantomime, Seriation, Train Building. It should be noted that pantomime, seriation and imitation are given special emphasis in the classroom curriculum and these are the major areas, other than language, where improvement was noted.

In general, ECEP children did not perform as well as the white control group in Language, Identity Matching, the Sigel Memory Matching task and in Number Concept. An argument might be made that language skill may be the basic differentiator since a preliminary factor analysis of the test batteries indicates that both the Identity Matching, and the Sigel task are highly correlated with verbal ability.

There may be an interesting sex difference concerning differential response to three vs. two dimensional stimuli. Both sexes improved in vocabulary but girls raised their scores in Naming Pictures (two dimensional stimuli) and boys improved in Objects by Use (three dimensional stimuli). Furthermore, ECEP males, who usually scored below black female controls were superior to the black females in both batteries in Pantomime Using Appropriate Object (3 dimensional stimulus). Again, ECEP females, did better than black female controls in Pantomime Using Appropriate Object, a result consistent with the classroom program. However, they scored lower than white males in that same task. This may be related to a difference in developmental rate. Response to the concrete three dimensional stimulus may appear before appropriate responses to the more abstract and less immediate two dimensional stimulus. While the evidence is conjectural at this time, it seems to be a difference worth further investigation.

At this time, the test results support the conclusion that the classroom intervention program was effective. The improvement over the black control groups shown in Battery II, argue to that point.

SOCIAL AND EMOTIONAL ASSESSMENT, PRELIMINARY RESULTS.

I. Test Related Behaviors.

As the child was being given the tests of Battery I, a variety of attentional, motivational, verbal and manipulatory behaviors which occurred

as a function of specific tester behaviors were recorded. Thirteen tester behaviors involving giving instructions, presenting test materials, giving the child reinforcement and personal attention, were systematically observed. Thirty-six child behaviors such as looking at the tester or material, handling items, wandering, staring, etc. were observed in relation to the tester behaviors which were presumed to be probable elicitors. One tester behavior and two child behaviors were recorded every five seconds throughout the test sessions. Two to four observers recorded these behaviors with an overall reliability of 85%.

Of these tester-child behaviors, only six tester and twenty child behaviors occurred with enough frequency to warrant statistical treatment. Three types of responses were analyzed: 1) Attentional and Interpersonal Responses, 2) Orientation and Manual Responses to Materials, 3) Verbal Responses. Comparisons along these dimensions were then made between the children who scored in the first and fourth quartile of the test battery.

A. Attentional and Interpersonal Responses.

(Differences at .05 and .01 probability levels): It was found that high and low scorers seem to differ in their orientation to the material and to the tester. Low scorers stare away while being watched, sit still during instructions and throughout the test sessions; high scorers smile when material is presented and smile at the tester throughout the session, more than low scorers do.

(Differences at .20 and .10 probability levels): Boys appear more distracted than girls as evidenced by staring away from either the tester or the task, but there are some indications of greater dependence as seen in going to the teacher's lap while girls went close to the tester and boys

seem to smile more during instructions. This latter observation is particularly interesting since it is the high scorers rather than the low scorers who smile more during instructions and it will be recalled that the high scoring group contains more females than males.

B. Orientation and Manual Responses to Test Materials.

(Differences at .05 and .01 probability levels): The differences in this category, reflect the nature of high versus low scoring groups and indicate one possible source of this difference. Obviously correct and incorrect gestural responses such as pointing to the correct picture or placing a form into the correct slot differentiates high from low scoring groups. In addition, low scorers show greater absence of responding. Low scorers engage in more tactile manipulation of material during instructions and throughout all tester-child combinations. This may indicate some impulsivity since instructions were given with the command to not touch the material until told to do so. It may also be an indication of a more primitive mode of response to the material since tactile manipulation was defined as random touching and fingering rather than any exploratory constructive handling of the materials.

(Differences at .20 and .10 probability levels): The differences in this area that are at lower levels of statistical significance, but also very suggestive are as follows: high scorers looked at the materials more than low scorers, although boys appear to look more than girls as a group, even though boys require more orientation to the materials. This may be a confounding factor in the looking response. While boys respond to the initial instructions with the correct gesture more often than girls, the girls improved very significantly when their instructions were repeated. Girls were also better at piling blocks and aligning them laterally.

C. Verbal Responses

(Differences at .05 and .01 probability levels): Girls and high scorers clearly verbalize more in general, and respond correctly to test questions, whereas boys remain silent rather than make errors.

In summary then, the high scorers who are predominantly female, are a more active group enjoying the task as evidenced by the smiling differences or perhaps reflecting greater interpersonal ease. They sit still more vis-a-vis the low scoring group only when instructions are repeated indicating once again perhaps a response to a personal demand from an adult or possibly need for achievement. The low scorers who were predominantly male, seem to show less motivated interaction with the materials, stare away and engage in tactile manipulation of materials as opposed to an exploratory response.

When we just look at the boys' versus girls' behaviors, irrespective of their scoring status, we found that the boys are quiet, more unresponsive, less verbal, less drawn to test materials and require more orientation. The girls are more restless, more verbal, better at piling and aligning blocks.

Thus, we find that there is some obvious relationship between performance and attendant behaviors. As one examines these results, one should be cautioned to realize that some of the attendant behaviors and performance are confounded. However, there are indications that the attendant test behaviors seem to be in the serive of test performance rather than just as a concomitant of this.

II. Rating Scales.

Teachers were asked to rate the child's behavior in the classroom according to a bipolar scale of social-emotional dimensions, originally devised by Emmerich (1971). The bipolar scales include twenty-one pairs of adjectives

such as withdrawn, involved, expressive, restrained, sensitive to others, self-centered, active-passive, etc. The teacher rated the child on a point scale with a rating of four indicating no difference between the two poles of the dimension.

The teachers were randomly assigned a different group of five children to rate each week during the school year. The ratings done from November 1969 to January 1970 and from February to May 1970, were pooled and designated as Scale 1 and 2. The reason for the division is that we wanted to see if any differences existed between the beginning and the end of the year and this arbitrary division divided the data into two approximately equal groups of scales.

The median rating for boys and girls in Scale 1 and Scale 2 was not significant when individual dimensions were examined, but very interesting changes in teacher ratings between Scale 1 and 2 should be noted.

In Scale 1, that is the early part of the year, the girls were very decidedly seen as more productive, powerful and active than boys. Girls outranked boys in nineteen of the twenty-one dimensions which is a significant difference. Girls were ranked as more involved, vulnerable to frustration, rebellious, expressive, sensitive to others, dominant, active, constructive, purposeful, aggressive, academically motivated, socially secure, energetic, stable, socially assertive, independent, rigid and happy. Boys outranked girls only very slightly as more relaxed and as more masculine. Of the nineteen categories in which girls were seen as possessing more of that particular quality, only four are usually construed as undesirable attributes, i.e. vulnerable to frustration, rebellious, rigid, aggressive to others. In terms of stereotyped sex-role typing, males are usually seen as more dominant, active, rebellious, vulnerable to frustration, energetic and assertive. All of these characteristics were given to the girls in Scale 1.

In Scale 2, however, the position of the girls vis-a-vis the boys changed quite considerably. Now the girls were seen as ascendant in only eleven categories, a difference which is no longer significant. Eleven of the categories in which girls were ranked higher than the boys are the same as Scale 1: dominant, constructive, purposeful, academically motivated, stable, social, assertive, independent and rigid. They have become more tolerant of frustration and more relaxed than the boys judging by Scale 2.

The boys, at least in the view of teachers, have made considerable strides in adjusting to the school situation. By Scale 2, they outranked the girls in being more involved, rebellious, expressive, sensitive to others, active, aggressive, socially secure, energetic and happy. Only two of these categories, rebellious and aggressive, are construed as negative qualities and both of them, at least in stereotyped sex-role typing are attributed to males more often than females.

Some preliminary analyses of the relationship between performance on each of the test batteries and rating on social-emotional dimensions were made. The X^2 contingency test between high and low scorers on Test Battery I and II and scoring below or above the median on social-emotional dimensions indicate several significant relationships.

Both males and females who scored above the median on the constructive and socially secure dimensions, scored significantly higher on cognitive tasks in the first testing. Furthermore, in the second testing, males who were identified as above the median on social dimensions, received higher test scores. At .10 and .20 significance levels, males and females who were seen as happier, more academically motivated and purposeful in the classroom, performed better in the initial testing. The more relaxed and

flexible males also obtain higher scores in Test Battery I. The girls who were viewed by the teachers as being energetic, assertive, independent, aggressive, dominant, social and involved, also received higher scores on this initial battery.

III. Teacher Observation Schedules.

The teachers were also required to make certain specific observations of classroom behavior and to record these observations once a week. Most of the behaviors that were observed centered around imitation, performance in puzzles and games, signs of independence or dependence, and social and emotional variables similar to those on the rating scale but asked in terms of specific classroom behaviors. For example, "Did child ask for help to solve problems or to complete a task?" "Did child imitate your gestures in an imitation game?"

Teachers were randomly assigned five different children each week, so that all the children were observed at least once a week by one of the teachers. Observations were scored "frequently", (seen daily or several times a day), "sometimes", (see 2 or 3 times a week), "rarely", (seen not at all or once a week).

The observations for November to March were averaged to obtain one score for Winter. Scores for March to May were averaged to obtain one score for Spring.

Results:

Three categories of behaviors were analyzed for male-female differences: 1) Social and Emotional Behaviors, 2) Imitation, and 3) Performance.

1) Social and Emotional Behaviors. The pattern of results for boys girls was similar, although the frequency of the behaviors was higher for girls than for boys in both Winter and Spring Observations.

In the Winter Observation, all categories of response except two, occurred with greater frequency for girls. Those two exceptions were Throwing Objects and Watches Others.

In the Spring Observation, the frequency of behaviors in this category decline. Both sexes decreased significantly in Watches Others and Follows Children. Boys decreased significantly in Throwing Objects and took a sizeable drop in Takes Turns and Follows Teacher.

There were only two instances in which boys and girls went in reversed directions in the Spring Observation. Boys increased the frequency of Seeks Help, and girls decreased. Boys increased in the frequency of Helps Others and girls decreased. Attentiveness and Aggressiveness were the most stable behaviors for both sexes, but again girls were considered to be slightly more aggressive and more attentive than boys.

2) Imitation Behaviors. The pattern of results repeats that of the Social Emotional Behaviors as far as the Winter Observation is concerned. Girls imitate more often than boys. However, there is an almost complete reversal at the end of the year. In the Spring Observation, boys increase imitative behavior in all categories - Imitates Teacher, Imitates Language, Imitates Actions. Girls decreased in all categories, significantly so for Imitates Actions. Boys showed the greatest increase in Imitates Language.

3) Performance. Once more, a consistent pattern emerges. Girls are more interested in puzzles and put them together correctly more frequently than boys. Both sexes increase in interest and accuracy in the Spring Observation but the girls still have the higher frequency of behaviors in this area.

Summary:

Girls were rated by their teachers as more active in almost all areas of classroom activities in both the Winter and Spring Observations. The major exception was Throwing Objects and this declined for boys in the Spring Observation.

Both sexes seem to gain in independence as evidenced by reduced following of other children and the teacher, decreased watching of others, and for girls, a decrease in seeking help.

Boys increase in frequency of imitation of the teacher and other children while girls decreased their behaviors in the imitative category.

IV. Narrative Data.

Narrative observations of the child's classroom behaviors were tape recorded every other month from late October to the end of June. All of the regular school activities, such as free play, juice, rest time, guided games, were considred to be a set of observations. Each child was observed in each of the school activities for a ten minute period and a complete set of observations was made on each child every other month.

A coding system for analyzing these reports was devised. The codes categorized the child's behavior into Interaction with People and Interaction with Objects. Six major kinds of behavior were coded concerning Interaction with People: General Social Interaction (parallel versus solitary play; cooperation, etc.), Defense-Offense Behaviors, Help Seeking, Help Giving, Imitation and Pretend. Four major kinds of behaviors were coded concerning Interaction with Objects: Gaining Objects, Handling Objects, Construction, Puzzle Completion and Seriation Tasks, and Location of Objects in Space. Occurrences of anticipation, persistence and language were looked for in all categories.

Only one analysis has been completed so far. Ninety-one items subsumed

under the Interaction with People and Interaction with Objects categories were compared for male-female differences. Comparisons were made for four time periods: November-December, January-February, March-April, and May-June. An overall comparison across all time periods was also done.

The data consisted of frequency of a particular behavior per minute of observation. The t test was used to compare the groups.

Results:

Reciprocal Behaviors. Boys clearly are more cooperative. They begin the year significantly more cooperative and remain so.

Defense-Offense Behaviors. Boys defend possessions and direct their aggression to objects and things more than girls. They defended their possessions more than girls in three of the four time periods, going from a passive defense (remonstrace but no physical contact with the other) to active physical defense of possessions in time period four. Averaged over the year, boys were significantly more aggressive to others. The only item in this category in which girls showed a higher frequency than boys was in Defends Self. Although only differences were significant at the .05 or .01 levels, boys had the higher mean in 8 out 9 differences starting at the .20 probability level.

Help Seeking and Help Giving. In the beginning of the year boys gave help spontaneously and sought help for interpersonal problems and in order to complete a task more often than girls. These behaviors showed no appreciable sex difference after March, although boys maintained a slight superiority in seeking help for interpersonal reasons across the four time periods. There was some indication that girls were more task oriented than boys earlier in the year.

Teacher-Child Interactions. In interactions that were initiated by the teacher, a difference in complying to teacher demands stands out clearly. Girls complied with demands, instructions, requests more than boys; boys refused to comply more often. This difference was maintained throughout the year. Although significance levels of the differences ranged from .20 to .10, girls had the higher mean for compliance in 9 out of 9 comparisons and boys had the higher mean for refusals in 9 out of 9 comparisons.

There was some indication of teacher by child's sex interaction. In March-April, the teacher gave significantly more positive reinforcement to girls, yet refused offers of help from girls more often. This is confounded by the fact that girls offered to help more often, however,

Child-Teacher Interactions. These interactions were child initiated and girls initiated all the interactions where any male-female differences occurred. They made more bids for attention, offered more help and had their demands acceded to more often than boys.

Interaction with Objects. Boys interacted with objects at fairly low levels of sophistication. They collided with objects and engaged in more random tactile manipulation, but in the last observation period boys completed puzzles or seriated without error more often than girls.

Anticipation and Language Behaviors. Any indication that the child planned ahead or was aware of an outcome before it occurred, was coded as an instance of Anticipation. All differences in this area showed boys to have the higher mean. They had the higher mean in 6 out of 6 comparisons and the anticipation was observed in all modalities: visual, motoric, and verbal.

In the beginning of the year, boys used inarticulate screams, gestures, cries, more often than girls and in the fourth observation period, girls were observed as having more articulate speech and using words and sentences.

Summary.

There were clear sex differences in classroom behavior. Boys more cooperative, defended their possessions and were more aggressive to people and objects than girls. Boys sought help for interpersonal reasons and to complete a task more than girls, but both sexes showed a decline in help seeking behaviors. Boys exhibited more negative responses to teacher instigated interactions of all kinds. Boys appeared to have more immature interaction with objects earlier in the year than the girls, but completed puzzles without error more often than girls at the end of the year. This may reflect interest rather than skill. A difference in language was also noted; the boys were more inarticulate than the girls in the early part of the year and the girls used more articulate speech by the end of the year.

Girls on the other hand, were more compliant to demands, requests, or offers of assistance made by the teacher. These responses were consistent throughout the year. Girls also initiated more interactions with teachers and had more of their demands complied with than boys.

SUBJECTIVE IMPRESSIONS

The quantitative results provide only some of the outcomes of our program. Various aspects of the program could not be systematically evaluated because of limitations of time and money as well as availability of adequate methodology. Eventually some of these gaps will be filled by case history analysis.

Few, if any, of our children had difficulty separating from their mothers. They entered this new environment joyfully and enthusiastically. Although the mothers were in the classroom the first few days of school, the children gravitated toward the materials in the classroom. The significance of this case of separation is difficult to interpret. It may be due to varied experiences with adult caretakers, the high valence of the room and the other children. The fact is that, with this group the separation was not an overt issue.

Engagement with the materials varied from intense concentration to flighty sporadic involvement. In spite of limited use of expressive language, the children seemed to understand the language of the teachers. Further, the children made their wants known.

The relationship with the teachers tended to be comfortable but not intense. The children seemed relatively self-sufficient and would seek comfort and help when hurt or frustrated. We did not notice strong ties to adults. In fact, toward the middle of the year, we experimented with teacher withdrawal from some of the usual stations and observed the children's unsupervised behavior. We were impressed with the length of time constructive play behavior was sustained under these conditions.

Relations of children to each other also were highly varied with consistent and stable friendship patterns emerging. There were enough instances of cooperative play to suggest that stereotyping of two-year-olds' social interaction as individualized and non-cooperative is an over-generalization.

Many of the children were highly variable in the degree of which they would respect productions of peers. There were those occasions when a child would destroy the production of another, grab materials and disregard the "rights" of other children. This type of imperialism subsided but what is most interesting is the observation that these behaviors were not consistent patterns.

Some children tended to hoard materials, as if the possession of an object functioned as a security blanket. Reluctance to surrender an object even if not used was observed. This possessiveness raised problems of management. If the object functions as a "security blanket", how does the teacher handle this without provoking undue anxiety in one child and still facilitate the concept of possession and surrender of what is, in effect, public property.

The relationship of the individual to his possession in a nursery school environment is of particular significance when one of the objectives of a program is to facilitate the understanding of the value of another's product and possession.

Over the year, the children not only increased in fluency and comprehensibility of language expression, but also in the complexity of their messages. Our program encouraged language production in the service of reporting events and actions and in requesting materials from children and adults. However, no corrections or explicit imitations of adult language was done. Hence, gains in accuracy, fluency and complexity seem to have occurred in relation to relevant encouragements and inducements to use language in lieu of grabbing, hitting or

other physical methods. Of course, increase in age also accounts for these changes. The program, however, did contribute when we take note of the fact that the opportunity and encouragement for verbal expression was consistently expected in our program.

An interesting sidelight on this issue is reflected in the teacher's reporting that the children talked more on the way home than they did in the classroom. We pondered this and concluded, after more observation, that the reason may well be that in the classroom, the children were very task oriented, involved with actions on the materials whereas in the bus materials were not available and the children did not have such distractions.

Another area in which considerable variability was observed was in attention span and involvement with materials. The children could and did get involved in activities and persisted from some length of time. These activities included small manipulative objects, building with blocks or even symbolic play, especially in the doll corner.

In effect, the experience with this two year old group demonstrated that our knowledge base for working in a psycho-educational context with these children is sparse. Solace or guidance could not be found in the literature. Experimental work within a group psycho-educational context for children of this age is virtually non-existent. Consequently, we depended heavily on our observations for teaching us what to do.

The previous discussion highlighted some areas in which our expectations and our observations differed.

Were the goals we set for ourselves realistic? At this point and with the data we have the answer, is an unqualified yes. It is feasible to establish group

nursery half-day programs for children as young as two. This is in contrast to the traditional concept which raised not only questions about age at entrance, but number of days per week. We have everything to indicate that our children successfully handled a four day week and have every reason to believe they could assimilate a five day week as well.

To conclude on a note of smug self-satisfaction would be misleading. There are a number of problem areas that need identification and further study, and so let us turn to these.

Problem Areas

One of the major problems facing us involved evaluation. Techniques of assessment and assessment tasks do not exist in great abundance. Most of them are embedded in intelligence tests, many are perceptual or motor types and do not allow for explicit testing of behavioral outcomes predicted from our program. This problem faces many of us in research in this area. We were not interested in measures limited to I.Q. scores. Hence, we had to construct some tests specifically for this project. Admittedly, these measures are new and do not have the necessary validation to argue for their psychometric purity. Rather, they are first efforts. The need for measurement procedures continues to be great.

In this regard, however, our experience has raised considerable doubt as to the validity of traditional psychometric procedures. We have found that at least in the numerical and language areas, performance in the classroom does not relate to performance in the testing situations. Children frequently failed items in the formal situation, e.g., concept of one, but understood this in the classroom context. This suggests an interesting problem concerning the contextual basis of knowledge and the degree to which measurement for very young children should follow the same format as for older children.

When do we know when a child knows is the basic issue that emerges. If he can respond only in one context and not another, what does this tell us? We are not convinced that this issue is resolved with psychometric reliability, but rather that it indicates the contextual basis of knowledge. In effect, the child is not able to transcend the situational supports that are integral to his knowledge base. From a theoretical and practical perspective, such information is very telling of the child's cognitive status.

Another set of evaluation issues relates to the tasks and the formal test situations. Many of the problems are self-evident, boredom, fatigue, strangeness. What of the particular variables themselves? For us, the decision was to test for variables that were directly relevant to the primary mission of the study. The time for developing procedures was not there, nor were there measures already specifically developed for our program. Thus, the procedures employed are in fact crude efforts. The only solution was to create all new measures prior to the onset of the program. The problems are horrendous not only in terms of operationalizing our conceptualizations but in demands of time and money. Thus, there is no question in our mind that our assessments are approximations containing error. To employ only the true and the tried, e.g., Stanford-Binet, shifts the nature of our assessment to global procedures and this is a choice we did not want to make.

Form testing however, also leaves much to be desired. Such procedures do not provide any explanatory bases for obtained change. What is needed is a careful monitoring of transactions that occur from T_1 to T_2. To carry out such a process involves a heavy input of observational data gathering. And what are we observing? A complex, systematic network in which the teacher, the child,

other children and materials interact in varying degrees and on various levels. The complexity of the setting demands a very careful conceptual analysis defining the relevant variables. We are in the process of analyzing data gathered over a one year period. To date, we can say nothing except that the effort was strenuous and we hope it will prove productive. It is, we believe, a necessary task if we are to provide a rational base for making statements about sources of behavioral change.

Since children are rapidly aging organisms, disentangling developmental effects from program effects becomes a major question. Employment of control groups is one method of dealing with the problem. Control groups in other settings, should be monitored with equal degrees of intensity, however. Few programs are able to mount this effort and few have done it with some success. There is much more work to be done here because all programs have method problems for which they cannot be faulted. This seems to be the challenge for this field of research. Our methods are not geared to such multi-variabled problems.

A particularly confounding element is one that is characteristic of the populations usually involved in such evaluations. Many middle class situations have some stability of family and living arrangements. In many of the social groups serviced by these programs, however, shifts and changes in family and living patterns are frequent and contribute sources of error beyond the control of research techniques.

That such problems are characteristic of all longitudinal studies is obvious, but when the project under study is involved in defining sources of change, then the research task is even more complicated. It is not always feasible to have random assignments, control over psychological and sociological factors, and cohorts of children and of teachers for tightly controlled replications.

Another potent source of error is the relationship of research personnel and research objectives to the educational personnel and program. Mutual trust and confidence, ability to surrender autonomy and maintenance of flexibility are among the characteristics that are essential ingredients for carrying out the mission.

These are but some of the issues we pondered and struggled over with varying degrees of success. The solutions often are partial and ephemeral, especially those related to interpersonal factors. Ongoing monitoring seems to be the most effective solution.

The more precise research issues, however, are still before all of us. We are not pessimistic but neither are we naive to believe that simplistic solutions are possible. Rather, the solution resides in a lot of hard work and thinking embedded in the firm conviction that we are engaged in a mammoth task that will tax our patience, endurance, and creativity. Since the experimental research model of the laboratory is not appropriate, we must be on the search for a model that is realistic and tuned into the social realities, and this is often far from the ideals of our research design textbooks. For all researchers in this field, the first problem is to work at conceptual model building. Models that integrate discrepant data and that generate provocative as well as practical hypotheses are essential to progress.

Research in preschool education poses a major challenge to social scientists. It is difficult, taxing and rewarding. We believe it has a unique set of problems that cry for solution.

References

Beller, E. K. In press. Research on teaching organized programs in early education. Chicago: Rand McNally.

Bissell, Joan S. 1976. Implementation of planned variation in Head Start. Washington, D. C.: U. S. Office of Child Development, Department of Health, Education and Welfare, April.

Emmerich, W. 1971. Disadvantaged children and their first school experiences: Structure and development of personal and social behavior in pre-school settings. Princeton, New Jersey: Educational Testing Service, November.

Ginsburg, H., and Opper, S. 1969. Piaget's theory of intellectual development: An introduction. Englewood Cliffs, New Jersey: Prentice Hall.

Golden, M., and Birns, B. 1968. Social class and cognitive development in infancy. Merrill Palmer Quarterly 14: 139-149.

Hoffman, M. L. 1960. Power assertion by the parent and its impact on the child. Child Development 31: 129-143.

Hunt, J. McV. 1969. The challenge of incompetence and poverty. Urbana, Illinois: University of Illinois Press.

Inhelder, Barbel, and Piaget 1964. Early growth of logic in the child: Classification and seriation. New York: Harper and Row.

Kamii, C., and Radin, N. L. 1969. The retardation of disadvantaged Negro pre-schoolers: Some characteristics found from an item analysis of the Stanford-Binet Test. Psychology in the Schools 6: 283-288.

Overton, W. F. 1972. Piaget's theory of intellectual development and progressive education. In J. R. Squire, (Ed.), A new look at progressive education. Association for Supervision and Curriculum Development, Washington, D. C., pp. 88-116.

Piaget, J. 1955. The child's construction of reality. London: Routledge and Kegan Paul.

Piaget, J. 1962. Play, dreams and imitation in childhood. Trans. by C. Gattegno and F. M. Hodgson. New York: Norton.

Sigel, I. E. 1960. Influence Techniques: a concept used to study parental behavior. Child Development 31: 799-806.

Sigel, I. E. 1972. Distancing hypothesis: Revisited. An elaboration of a Neo-Piagetian view of the development of representational thought. Paper presented a third annual Western Symposium on Learning. Bellingham: Western Washington State College.

Sigel, I. E., Anderson, L. M. and Shapiro, H. 1966. Categorization behavior of lower and middle class Negro preschool children: Differences in dealing with representations of familiar objects. Journal of Negro Education 35: 218-229.

Sigel, I. E. and McBane, Bonnie 1967. Cognitive competence and level of symbolization among five year old children. In J. Hellmuth (Ed.), The disadvantaged child, Vol. 1. Seattle, Washington: Special Child Publications.

Sigel, I. E. and Olmsted, P. 1970. Modification of cognitive skills among lower class black children. In J. Hellmuth (Ed.), Disadvantaged Child, Vol. 3. New York: Bruner, Mazel.

Sigel, I. E. and Perry, C. 1968. Psycholinguistic diversity among culturally deprived children. American Journal of Orthopsychiatry 38: 122-126.

Zigler, E. and Butterfield, E. C. 1968. Motivational aspects of changes in I.Q. test performance of culturally deprived nursery school children. Child Development 39: 1-14.

Footnotes

[1]The research reported herein is supported by the Office of Economic Opportunity, Grant #CG8547. Dr. I. Sigel is Project Director, Dr. A. Secrist, Research Coordinator, and Dr. G. Forman, Curriculum Coordinator.

[2]Descriptions or copies of all tests and instruments used for assessment purposes may be had upon request.

[3]Descriptions or copies of all tests and instruments used for assessment purposes may be had upon request.

TABLE 1

TESTS USED IN ASSESSMENT OF ECEP AND CONTROL CHILDREN

LANGUAGE TASKS

1. Bayley V Scale (Naming Parts of a Doll).
2. Bayley Names and Points to Pictures.
3. Bayley Z Scale (Understanding Prepositions).
4. Stanford-Binet Picture Vocabulary, Year II.
5. Stanford-Binet Understanding Objects by Use. (Vocabulary measure)
6. Stanford-Binet Identifying Parts of the Body. (Vocabulary measure)

IMITATION AND PANTOMIME TASKS

7. Bayley M Scale (Paper Folding and Copying a Line Task).
8. Bayley, Building a Train.
9. Pantomime with Four Cue Conditions: (Verbal, Pictorial, Object Present, Object to Use).

PERCEPTION AND CLASSIFICATION TASKS

10. Bayley Pegboard.
11. Bayley R Scale (Blue Form Board. Circles and squares are alternately inserted into form board.)
12. Stanford-Binet Three Hole Form Board (Circle, square, triangle).
13. Stanford-Binet Three Hole Form Board, Rotated.
14. Kagan Embedded Figures (A car, cat and flower are embedded in backgrounds of varying difficulty, 18 trials.)
15. Piaget Large-Small Sorting.
16. Piaget Identity Matching. (Blocks are matched according to color and size.)
17. Piaget Free Arrangements. (Rings, cubes, circles, triangles, squares, rectangles in red, green and yellow are
    He arranges them ad lib.)

MEMORY FOR VISUAL AND AUDITORY IMAGES

18. Delayed Response, Simple. (A toy kitty is hidden in one of three boxes.)
19. Delayed Response, Invisible Transportation. (After kitty has been placed under a box, the position of the box is changed.)
20. Delayed Response, Complex. (The box containing the kitty is moved twice before child is asked to find the kitty.)
21. Sigel Memory Matching. (The child is shown the picture of an object which he is to remember and pick out from an array of three objects.)
22. Auditory Sequential Memory. (I.T.P.A. task)
23. Seriation. (Five pegs of varying heights, 1" gradation are to be placed in ascending order in pegboard.)

NUMBER TASKS

24. Bayley Concept of One, Two.
25. Bayley Building a Tower. (Cubes are stacked as high as possible.)

Sigel TABLE 2

Means of Each Test In Battery I. ECEP Males Compared to ECEP Females and to Male and Female Control Groups.

Tests	ECEP Males N=9 $\bar{x}$	ECEP Females N=10 $\bar{x}$	Non-ECEP Black Females N=8 $\bar{x}$	Non-ECEP Black Males N=6 $\bar{x}$
LANGUAGE.				
Bayley V Scale	5.00	5.40	6.25*	5.00
Bayley Name Pictures	3.67	3.50	5.75**	2.83
Bayley Z Scale	1.00	1.60	2.25**	1.17
Stanford-Binet Picture Vocabulary	4.44	4.60	7.50**	3.17
Stanford-Binet Objects by Use	1.44	2.40*	2.75*	2.50
Stanford-Binet Identifying Parts of Body	5.11	5.70	5.88	3.83
IMITATION, PANTOMIME.				
Bayley M Scale	2.00	3.20**	2.25	1.83
Bayley-Building a Train	1.11	1.10	1.00	1.00
Pantomime with Four Cues:				
Verbal Cue	1.00	0.00	3.00	.00
Picture Cue	.56	0.00	2.38	3.33
Object Cue	0.00	0.00	0.00	0.00
Using Object	2.00	2.20	.25**	1.33
PERCEPTION, CLASSIFICATION.				
Bayley Pegboard	3.00	3.00	3.00	3.00
Bayley R Scale	8.11	7.80	8.38	7.67
Stanford-Binet Three Hole Form Board	1.67	1.50	1.88	1.83
Stanford-Binet Three Hole Form Board-Rot.	1.89	1.70	1.88	1.33
Kagan Embedded Figures	6.89	7.50	9.25	6.33
Piaget Large-Small Sorting	.22	.40	.75*	.17
Piaget Identity Matching	1.44	1.60	1.75	1.50
MEMORY FOR VISUAL AND AUDITORY IMAGES.				
Delayed Response (Invisible Transportation)	2.00	1.60	1.75*	1.17***
Sigel Memory Matching	2.00	2.50	3.25*	1.33
Delayed Response, Simple	1.78	1.90	1.63	2.50*
Delayed Response, Complex	1.22	1.50	1.63*	1.17
Auditory Sequential Memory	1.33	1.70	2.25	.67
Seriation	.44	1.40**	1.63**	1.91**
NUMBER.				
Concept of One	0	0.10	0.25	0
Concept of Two	0	0	0	0
Tower Building	4.88	5.80	6.38*	4.00

* Difference significant at $<.20$.
** Difference significant at $<.05$.
*** Difference significant at $<.001$.

"t" test was used to compare groups.

TABLE 3.

Mean of Each Test in Battery I. ECEP Females Compared to Male and Female Controls.

TESTS	ECEP Females N=10 $\bar{x}$	Non-ECEP Females (Blk.) N=8 $\bar{x}$	Non-ECEP Males (Blk.) N=6 $\bar{x}$
LANGUAGE.			
Bayley V Scale	5.40	6.25	5.00
Bayley Name Pictures	3.50	5.75	2.83
Bayley Z Scale	1.60	2.25	1.17
Stanford-Binet Picture Vocabulary	4.60	7.50**	3.17
Stanford-Binet Objects by Use	2.40	2.75	2.50
Stanford-Binet Identifying Parts of Body	5.70	5.88	3.83*
IMITATION, PANTOMIME.			
Bayley M Scale	3.20	2.25**	1.83**
Bayley-Building a Train	1.10	1.00	1.00
Pantomime with Four Cues:			
Verbal Cue	0.00	3.00*	0.00
Picture Cue	0.00	2.38*	3.33
Object Cue	0.00	0.00	0.00
Using Object	2.20	.25**	1.33
PERCEPTION, CLASSIFICATION.			
Bayley Pegboard	3.00	3.00	3.00
Bayley R Scale	7.80	8.38	7.67
Stanford-Binet Three Hole Form Board	1.50	1.88*	1.83
Stanford-Binet Three Hole Form Board-Rot.	1.70	1.88	1.33
Kagan Embedded Figures	7.50	9.25**	6.33
Piaget Large-Small Sorting	.40	.75	.17
Piaget Identity Matching	1.60	1.75	1.50
MEMORY FOR VISUAL AND AUDITORY IMAGES			
Delayed Response (Invisible Transportation)	1.60	1.75	1.17*
Sigel Memory Matching	2.50	3.25	1.33*
Delayed Response, Simple	1.90	1.62	2.50*
Delayed Response, Complex	1.50	1.63	1.17
Auditory Sequential Memory	1.70	2.25	.67*
Seriation	1.40	1.63	1.92
NUMBER			
Concept of One	0.10	0.25	0
Concept of Two	0	0	0
Tower Building	5.80	6.37	4.00*

* Difference significant at $<.20$
** Difference significant at $<.05$
*** Difference significant at $<.001$

"t" test was used to compare groups.

Sigel

TABLE 4.

Means of Each Test in Battery II. ECEP Males Compared to ECEP Females and to Male and Female Controls.

TESTS	ECEP Males N=9 $\bar{x}$	ECEP Females N=10 $\bar{x}$	Non-ECEP White Males N=17 $\bar{x}$	Non-ECEP Black Males N=4 $\bar{x}$	Non-ECEP White Females N=8 $\bar{x}$	Non-ECEP Black Females N=5 $\bar{x}$
LANGUAGE.						
Bayley V Scale	5.88	6.30	6.35	6.75	5.50	7.00
Bayley Name Pictures	5.44	6.40	7.30*	6.50	6.62	5.00
Bayley Z Scale	2.11	2.70	3.35*	2.25	4.25**	2.00
Stanford Binet Vocabulary	10.78	9.20	10.12	10.25	9.63	9.00
Stanford Binet Objects by Use	3.44	3.40	3.41	3.00	3.88	3.00
Stan. Binet Ident. Parts of Body	6.11	6.70	6.41	6.25	5.25	6.80
IMITATION, PANTOMIME.						
Bayley M Scale	2.00	3.50**	3.06*	2.00	2.63	3.20*
Bayley Block Train	1.56	1.50	1.06***	1.00*	1.25	1.40
Pantomime with Four Cues:						
Verbal Cue	8.00	7.20	6.88	0.00	10.13	2.40
Picture Cue	1.11	5.50*	3.17	0.00	1.50	1.40
Object Cue	.33	.90	.82	0.00	0.00	0.00
Using Object	2.11	.80*	2.59	.75*	1.13	1.60
PERCEPTION, CLASSIFICATION.						
Bayley Pegboard	2.89	3.00	3.00*	3.00	3.00	3.00
Bayley R Scale	9.00	8.80	8.82	7.50*	8.50	8.20
Stan.Binet Three Hole Fm. Board	1.89	2.00	1.94	1.75	2.00	1.60
Stan.Binet Three Hole Fm.Bd.-Rot.	2.89	2.60	2.47*	1.00**	2.63	2.60
Kagan Embedded Figures	9.00	11.20	11.65*	7.75	10.63	6.40
Piaget Large-Small Sorting	.11	.50	.94**	0.00	2.00***	.40
Piaget Identity Matching	2.00	.60**	2.29	1.50	2.25	2.20
MEMORY FOR VISUAL AND AUDITORY IMAGES.						
Delayed Response (Inv. Transport.)	1.77	1.70	1.88	1.75	1.50	2.00
Sigel Memory Matching	1.44	2.80*	2.71*	2.50	3.00*	3.60**
Delayed Response, Simple	2.22	2.00	2.59	2.20	2.38	2.00
Delayed Response, Complex	1.44	1.30	1.47	1.25	1.25	1.00
Auditory Sequential Memory	2.44	3.4	3.70	3.75	2.36	3.40
Seriation	2.17	1.65	.92***	1.50	2.38	3.50**
NUMBER.						
Concept of One	0	0	0.36**	0	0.62**	0.40
Concept of Two	0	0	0.12**	0	0.38**	0.20
Tower Building	4.33	5.60	5.24	6.00	6.88**	5.60

* Difference significant at $<.20$.
** Difference significant at $<.05$.
*** Difference significant at $<.001$.

"t" test was used to compare groups.

Sigel

TABLE 5.

Means of Each Test in Battery II. ECEP Females Compared to Male and Female Controls.

TESTS	ECEP Females N=10 $\bar{x}$	Non-ECEP Wht. Males N=17 $\bar{x}$	Non ECEP Wht. Females N=8 $\bar{x}$	Non ECEP Blk. Females N=5 $\bar{x}$	Non ECEP Blk. Males N=4 $\bar{x}$
LANGUAGE.					
Bayley V Scale	6.30	6.35	5.50	7.00	6.75
Bayley Name Pictures	6.40	7.29	6.62	5.00*	6.50
Bayley Z Scale	2.70	3.35	4.25*	2.00	2.25
Stanford Binet Picture Vocabulary	9.20	10.18	9.63	9.00	10.25
Stanford Binet Objects by Use	3.40	3.41	3.88	3.00	3.00
Stan. Binet Ident. Parts of Body	6.70	6.41	5.25	6.80	6.25
IMITATION, PANTOMIME.					
Bayley M Scale	3.50	3.06	2.63	3.20	2.00**
Bayley Block Train	1.50	1.06**	1.25	1.40	1.00*
Pantomime with Four Cues:					
Verbal Cue	7.20	6.89	10.13	2.40	0.00
Picture Cue	5.50	3.18	1.50*	1.40	0.00*
Object Cue	.90	.82	0.00*	0.00	0.00
Using Object	.80	2.59***	1.13	1.60	.75
PERCEPTION, CLASSIFICATION.					
Bayley Pegboard	3.00	3.00	3.00	3.00	3.00
Bayley R Scale	8.80	8.82	8.50	8.20	7.50*
Stan.Binet Three Hole Fm. Board	2.00	1.94	2.00	1.60	1.75*
Stan.Binet Three Hole Fm.Bd.-Rot.	2.60	2.47	2.63	2.60	1.00**
Kagan Embedded Figures	11.20	11.65	10.63	6.40*	7.75*
Piaget Large-Small Sorting	.50	.94	2.00**	.40	0.00
Piaget Identity Matching	.60	2.29***	2.25**	2.20**	1.50
MEMORY FOR VISUAL AND AUDITORY IMAGES.					
Delayed Response (Invisible Transp.)	1.70	1.88	1.50	2.00	1.75
Sigel Memory Matching	2.80	2.71	3.00	3.60	2.50
Delayed Response, Simple	2.00	2.59*	2.38	2.00	2.00
Delayed Response, Complex	1.30	1.47	1.25	1.00	1.25
Auditory Sequential Memory	3.40	3.71	2.38	3.40	3.75
Seriation	1.65	.92*	2.38	3.50***	1.50
NUMBER.					
Concept of One	0.10	0.36**	.62**	0.40**	0
Concept of Two	0	0.12**	.38**	0.20**	0
Tower Building	5.60	5.24	6.88	5.60	6.00

* Difference significant at $<.20$.
** Difference significant at $<.05$.
*** Difference significant at $<.001$.

"t" test was used to compare groups.

TABLE 6.

Percentage of Grouping, Non-Grouping and Decision Making Responses in Piaget Free Arrangement Task.

	Response Category					
	Non-Grouping		Grouping		Decision Making	
Battery	I	II	I	II	I	II
Males	38%	18%	33%	57%	24%	20%
Females	35	21	43	53	22	21