DOCUMENT RESUME

ED 085 407 TM 003 348

TITLE	A Report on Two National Samples of Head Start Classes: Some Aspects of Child Development of Participants in Full Year 1967-68 and 1968-69 Programs. Final Report.
INSTITUTION	Research Triangle Inst., Durham, N.C.
SPONS AGENCY	Child Development Services Bureau (DHEW/OCD), Washington, D.C. Project Headstart.
PUB DATE	Dec 72
NOTE	220p.
EDRS PRICE	MF-$0.65 HC-$9.87
DESCRIPTORS	Achievement Gains; Child Development; *Classroom Environment; *Cognitive Development; Compensatory Education Programs; Culturally Disadvantaged; Emotional Development; *Family Characteristics; Family Environment; Parent Child Relationship; Parent Influence; Personality Development; *Preschool Children; Preschool Evaluation; Program Evaluation; *Social Development; Teacher Characteristics
IDENTIFIERS	*Project Head Start

ABSTRACT

Some characteristics of children, their families, the programs they attended, and the relationship of these experiences to their development are covered in this national report on the immediate changes associated with participation in full year Head Start programs. Major emphasis was placed on the child's cognitive and social development in the context of the classroom experience. Data was collected through the use of 24 forms, questionnaires and tests. Major findings were: (1) life circumstances associated with family economic status were associated with the child's performance on measures of cognitive development and achievement upon entering a Head Start Program. (2) There were significant overall gains on cognitive development, preacademic readiness, the ability to learn a new task, and achievement motivation. The gains were greater than expected at usual maturational rates. (3) As age increased, achievement motivation increased and adjustment problems decreased. (4) Teacher demographic variables were not related to the progress of the children. (5) Programs did make a difference and the highly structured focused, and well-implemented compensatory programs brought about greater immediate cognitive gains than low structured, diffuse and less well implemented programs. (6) There was substantial variation in cognitive gains among classes and a significant amount of this variation could be predicted from class and teacher-based observation. (EH)

RESEARCH TRIANGLE INSTITUTE

Project No. 22U-581

December 1972

Final Report

A REPORT ON TWO NATIONAL SAMPLES OF HEAD START CLASSES: SOME ASPECTS OF CHILD DEVELOPMENT OF PARTICIPANTS IN FULL YEAR 1967-68 AND 1968-69 PROGRAMS

Prepared for

The Bureau of Head Start and Early Child Development
Office of Child Development
Department of Health, Education, and Welfare

Under

Contract Number HEW-OS-70-207

by

Research Triangle Institute
Post Office Box 12194
Research Triangle Park, North Carolina 27709

ACKNOWLEDGMENTS

The analyses summarized in this report involved the efforts and cooperation of a number of people both within and outside the Research Triangle Institute (RTI).

The RTI project staff consisted of Dr. John Bailey, Jr., Mr. Philip Cooley, Dr. George Dunteman, and Dr. A. V. Rao. Dr. Dunteman, project director, and Dr. Richard Endsley, a consultant from the University of Georgia, were the principal authors of this report. Dr. A. V. Rao and Dr. John Bailey, Jr. also contributed to sections of this report. Appreciation is extended to Mr. Philip Piserchia and Dr. Babu Shah for statistical advice and to Mr. James Wright and Mr. James Batts for computer programming assistance. Mrs. Frances Heald had the major responsibility for typing this report.

Both consultants to RTI and Office of Child Development (OCD), Review Panel members provided guidance to the RTI project staff. The consultants to RTI were Dr. Anthony Conger, Dr. Gertrude Cox, Dr. Richard Endsley, Dr. James Gallagher, Dr. Lyle Jones, Dr. John McDavid, and Dr. Jo Anna Williams. The OCD Review Panel members were Dr. John Bartko, Dr. Lois-ellin Datta, Dr. Robert Egbert, Dr. Charles Gershenson, Dr. Earl Pollack, Dr. Earl Schaefer, and Dr. Leon Yarrow.

Dr. Lois-ellin Datta, Chief, Evaluation Branch, OCD provided valuable guidance and made detailed contributions to the contents of this report.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES - TEXT

LIST OF TABLES - TEXT (continued)

LIST OF TABLES - TEXT (continued)

LIST OF TABLES - TEXT (continued)

LIST OF TABLES - TEXT (continued)

LIST OF TABLES - APPENDIXES

LIST OF TABLES - APPENDIXES (continued)

LIST OF TABLES - APPENDIXES (continued)

LIST OF TABLES - APPENDIXES (continued)

LIST OF TABLES - APPENDIXES (continued)

A REPORT ON TWO NATIONAL SAMPLES OF HEAD START CLASSES: 1967-68 and 1968-69

CHAPTER I

SOME ISSUES IN EVALUATING THE EFFECTS OF HEAD START ON CHILD DEVELOPMENT

This is a report on one aspect of Project Head Start: some characteristics of children, their families and the programs they attended, and the relationship of these experiences to their development. It is the first national report on the immediate changes associated with participation in full-year programs. Earlier national studies have examined summer programs (Temp and Anderson, 1967) and the achievement of children several years after leaving Head Start (Cicirelli et al., 1969).

It is, however, not a comprehensive report; the focus is on the child's cognitive and social development in the context of the classroom experience. Such a focus is limited because early human development is a complex function of the child's own characteristics (for example, sex, physical appearance, health, energy/activity level, temperament--all of which differentiate babies at birth), the characteristics of his family (for example, family structure and status, and the personal attributes of the immediate and extended family) and the social and institutional milieu into which he is born. To isolate classroom experience as the primary associate of change is an oversimplification if we claim to be studying human development with a full appreciation of its richness. This complexity was recognized quite fully when Head Start was developed. Head Start objectives have included from the beginning:

a) Improving the child's physical health and physical abilities.

b) Helping the emotional and social development of the child by encouraging self-confidence, spontaneity, curiosity, and self-discipline.

c) Improving the child's mental processes and skills with particular attention to conceptual and verbal skills.

d) Establishing patterns and expectations of success for the child which will create a climate of confidence for his future learning efforts.

e) Increasing the child's capacity to relate positively to family members and others while at the same time strengthening the family's ability to relate positively to the child and his problems.

f) Developing in the child and his family a responsible attitude toward society, and fostering constructive opportunities for society to work together with the poor in solving their problems.

g) Increasing the sense of dignity and self-worth within the child and his family.

Only some aspects of these objectives are assessed. The assumptions on which Head Start was based did not suppose these objectives could be achieved in a linear or an independent manner, in pursuing by more-and-more of one activity, one goal at a time. The Head Start model is interactive, requiring support of the whole child in his whole life space. It is this model and this policy that uniquely distinguish Head Start from other early education programs. Any evaluation of Head Start that is less than that is incomplete and potentially misleading. It may provide a test of a more simplified view of human development, but it will not provide a test of Head Start as it was originally conceived and as it has in policy been operated. The major implication of this limitation is that a great deal of initial variance (individual differences), final variance and change variance will not be accounted for by classroom factors. Some of this unaccounted-for variance will be due to the less-than-perfect reliability of measures. More is likely to be due to those institutional, social, contextual and constitutional factors which are not taken adequately into account.

To point out this limitation early in the report is not to disparage in any value the importance of the findings. Classroom experiences are of great importance, and particularly when the child is making the first transition from the home to the classroom. The nature of this experience is surely of profound significance; the usually vivid recall of one's first day at school and of the kindergarten year, while later grades are more blurred, becomes a personal "verification" in addition to the significance predicted from psychology theory. Identification of how the nature of this experience is related to changes in the child is a task which is in itself complex, and worthwhile. The findings thus say something of much importance to Head Start and to early child development; they do not tell the complete story.

The study design is directed primarily to the question of interaction: do different kinds of programs have different effects on different children? If so, what are the optimum interactive circumstances? If not, what single pattern of experiences "works best?" There are no control groups of children who have not attended Head Start: the comparison is of development in different kinds of Head Start programs.

To those who have thought of Head Start as a preschool compensatory education program with a uniform curriculum, these statements may be a first introduction to what Head Start set out to accomplish, how the program developed, and what it is like. The data reported here are for 1967-68 (FY68) and 1968-69 (FY69), the first fully operational years for the nine-to-twelve month programs. Other data, both earlier and more recent, will be used to supplement the findings from the present analyses.

A. The History and Objectives of Project Head Start

The Office of Economic Opportunity (OEO) was created in 1964 by President Lyndon B. Johnson as part of the War on Poverty; the agency was to develop and administer new solutions to the problems of the poor in an effort to eliminate or reduce institutionalized poverty and the suffering of millions of Americans caught in the poverty cycle. Central to many OEO efforts was the belief that effective solutions required the meaningful participation of those affected in program development and administration. The creation of a new role for the poor--the role of valued participants in decisionmaking and management--was a significant step for many persons, both economically privileged and deprived. The tide had long been running toward emancipation and self-determination, and participation by minorities in events affecting them: OEO became one of the swiftest parts of this current.

Early OEO programs were directed to the needs of adults: job training, legal assistance, health, and other locally determined community action projects. In December 1964, R. Sargent Shriver, then Director of OEO, requested Dr. Robert E. Cooke, Pediatrician-in-Chief of Johns Hopkins Hospital, to develop a plan for children of the poor. The program was to be of immediate benefit for them and, as part of the overall OEO program, contribute to breaking the poverty cycle at as many points as possible.

A summary of the report of the Panel brought together by Dr. Cooke is reproduced as Appendix A. The Panel recommended creation of a special program for young children. In recognition of the fact that the influences of poverty on the child are manifested in many ways, the program was to be comprehensive, attacking health, nutritional, family and developmental problems as these might influence the child's development. While a successful comprehensive intervention was expected to give the child a better chance in school, the conferees recognized that institutional changes would be necessary to sustain positive effects and that bettering the child's life would be no simple task. The goals of the program, cited previously from the Cooke memorandum, reflect this view.

Not infrequently, conversations and articles reflect a common misconception of Head Start as "an eight week summer program to prepare children for school." In rereading the Cooke memorandum, it is evident that while the panelists expected that environmental intervention could have significant effects on development, they appreciated the complexity and magnitude of the task. The eight-week summer program was seen as a start, on a small scale, to assess the feasibility of the approach and gauge its effects. There was, however, what can be regarded as over-optimism on how easily the approaches and ideas developed in studies of 20 children or so could be melded with the spirit of community action on a national scale, and under-anticipation of the immediate popularity of the program, which resulted in a shift in scale from a relatively small effort to a large national program.

After approval of the Cooke memorandum, Head Start was scheduled to begin in summer 1965 serving about 100,000 children. The responses of Community Action Agencies (CAA's) and Local Education Agencies (LEA's) such as the public schools were overwhelming. The hastily gathered staff, working late at night and weekends, managed to process grants authorizing service to 561,000 children. All of the support and training services were put together under similar pressure by pediatricians, social workers, educators, psychologists, and administrators under the guidance of Dr. Julius B. Richmond, first Director of Head Start, and Mr. Jules Sugarman, the Deputy Director.

Local evaluation studies, national reports of services provided, and the reports from consultants and communities generally indicated that despite the believable mix-ups, confusion and delays, the program was resoundingly successful in reaching poor children and their families, people long thought to be unreachable by "the Establishment" or its representatives. The myth that poor families were indifferent to their children's welfare was exploded by the initial response to Head Start and the continued support and virtually unaminous support found in every survey of those served, and evident in the action initiated by parents at the local level to support the Head Start program.

A few summer programs continued to operate in 1965-1966, while the agency caught its breath and began preparing for what clearly was going to be the next year. During this period, program objectives became more clearly defined, the role and function of different components were elaborated, training programs were planned for staff at all levels, and work was begun on the Head Start manual and the series of pamphlets (the "Rainbow Series") that provide the national framework of expectations and approaches. A national research and evaluation program was also developing under the leadership of Dr. Edmund Gordon.

Summer 1966 thus was the "shake-down" year for the newly organized project; more of the summer 1966 programs continued to operate in 1966-67, and new full-year grantees were funded throughout FY67. In part, the decision to fund the more expensive full-year programs was based on the belief derived from theory and early evaluation data that longer, more intensive experiences were needed for children and their families to benefit as fully as intended in the objectives. Summer 1967 was the first fully operational program, with most staff trained and most support systems in place. The 1967-68 full-year programs similarly were the first where implementation of the Head Start idea could be expected to approach the quality intended.

In 1967-68, two new programs were developed, again based on theory and evaluation evidence. The Parent and Child Centers served extremely deprived families with children from 0 to 3, in the belief that for many children, Head Start was "too little and too late." Project Follow-Through extended Head Start upward in K, 1, 2, and 3rd grades, in the belief that if the gains associated with Head Start were to continue, the public schools as institutions would have to change, becoming more like

Head Start in parent participation and in offering comprehensive services and enriched classroom experiences. By summer 1968, data from studies of Head Start and other preschool programs were reasonably consistent in showing immediate benefits, usually greater for the longer programs, and a leveling off in performance after the children entered public schools.

The Westinghouse Report on the public school achievement of children who had attended summer 1965, 1966, 1967, and 1968 programs and full year 1965, 1966, and 1967 projects confirmed these earlier reports of a "catch-up" by non-Head Start children after the second year of public school (Cicirelli et al., 1969). The Westinghouse Report led (a) to conversion of many, presumably less effective summer programs to full-year programs, (b) to an emphasis on finding maximally effective curriculums for Head Start and Follow-Through, and (c) to a widespread feeling that "Head Start was a failure." In subsequent years, losses of funds have meant cutbacks in the program that have been slowly restored as some new funds became available. (Head Start Fact Sheet, Appendix B.) Only a few new programs have been funded since 1968.

In 1969, operation of Head Start was delegated by OEO to the Department of Health, Education and Welfare to be administered by the newly created Office of Child Development (OCD). Under the guidance of Dr. Edward F. Zigler, first Director of OCD, and Mr. Richard E. Orton, Associate Director for Project Head Start and Child Service Programs, Head Start has turned increasingly to experimental programs examining innovative ways of serving children and their families. The present report is thus to a certain extent a national statement of accountability on some aspects of the immediate impact of the 1967-68 and 1968-69 Head Start programs and is part of an on-going effort, begun in 1968-69 and now scheduled for completion in 1974, to find out what kind of classroom experiences can best contribute to Head Start's objectives for the psychological development of low-income children.

This brief history is summarized in Table I.1 which shows by year funds appropriated, programs funded and children served.[1/]

B. Some Research and Evaluation Issues Associated with Head Start

Initially, two issues received considerable attention in Head Start evaluations, both those funded by Head Start and the many initiated by other agencies and individuals: (1) is the program delivering the services it is obligated to deliver to eligible children and families? and (2) is the program of benefit as expected to children and their families?

1/ References contributing to this section include: Stearns (1971); Datta (1969); Grotbert (1969); Cicirelli et al. (1969).

Table I.1

HEAD START FUNDS, PROGRAMS AND CHILDREN, 1965 THROUGH 1973
(Dollars in Millions)

Year	FY	Funds	Grants	Children
Summer 1965	66	$85.0	2,397	561,000
Full Year 1965-66		--	--	--
Summer 1966	67	98.0	1,645	573,000
Full Year 1966-67		81.9	470	160,000
Summer 1967	68	116.6	1,249	466,300
Full Year 1967-68		210.4	750	215,100
Summer 1968	69	91.0	1,185	476,000
Full Year 1968-69		192.0	709	217,700
Summer 1969	70	90.2	1,100	446,900
Full Year 1969-70		212.3	700	216,700
Summer 1970	71	26.1	504	117,461
Full Year 1970-71		298.7	1,152	264,714
Summer 1971	72	22.0	450	89,600
Full Year 1971-72		317.5	1,225	278,880
Summer 1972	73	20.0	425	77,600
Full Year 1972-73		335.1	1,240	281,280

To these questions have been added three others:[2/]

1) Does the success, or failure, of Head Start viewed solely as a compensatory education program mean compensatory education has succeeded, or failed?

2) How durable are the benefits of very early, early, primary school and later interventions?

3) What program (curriculum model or approach) works best?

2/ Identification of these only implies that these are popular questions; it does not imply that they are necessarily the most appropriate or the most relevant from the point of view of an ideal evaluation of a program such as Head Start.

The first two questions may seem heavy burdens to bear for a new agency, creating and training its staff in new careers, often caught in the yearly turmoil of budget justifications and re-organizations. They should not, however, be indefinitely deferred. Program documentation, analysis and effects can contribute much to program planning however complex the skein of idea, implementation, and assessment. However fragile available measures or limited possible designs, evaluation data are collected, analyzed, reported and used for program decisionmaking.

Where these are child and family data, interpretation of such evaluation is often difficult. Collection of data on child and family development should be, in our view, among the very last steps in a social change process that begins with identification of a social problem, continues with the conceptualization of possible solutions to the problem, moves to small-scale "tryouts" of the idea, then develops a delivery or managerial system for implementing the idea, which finally--if all of these prior steps have gone well--will have major effects on children and families. In this paradigm, a "positive" finding is likely to indicate "success" at every prior stage. A "negative" finding is uninterpretable if the study involves only "outcome" assessment: the failure could be at any one or at several of the stages between identification of social need and outcome.

Like most assessments, almost all Head Start evaluations are of outcome; only recently have evaluations begun to study the whole process to identify where success and the breakdowns, if any, have occurred. By referring to other studies and to some of the history outlined earlier, it will be possible to place the 1967-68 and 1968-69 findings in a broader perspective than the study design itself provides.

C. A Brief Overview of Head Start Evaluations

Head Start evaluations can be grouped in four categories: descriptive studies of program compliance with the guidelines, one-site research and evaluation reports, summative (overall) national impact studies, and national intervention and interactive model studies.

1. Descriptive Studies. Between summer 1965 and summer 1970, the Bureau of the Census conducted, at the request of Head Start, surveys of random samples of Head Start programs, centers, classes, staff and children. The questionnaires were designed by Head Start evaluation staff and program specialists to assess compliance with national guidelines and to describe Head Start programs. A series of reports prepared by Barbara Bates of the Office of Child Development provides an extensive description of each summer and full-year program since 1965. In general, the findings indicate compliance with Head Start guidelines, but also detail the substantial variation within the guidelines of child, family, staff and other

program characteristics. These reports indicate that Head Start is not a homogeneous program: the diversity matches and perhaps exceeds that of the population served, since variations in program elements themselves overlay population characteristics. In addition, the Census surveys have identified areas where programs are not providing the required level of service. These data suggest that while most programs are in compliance with most guidelines, program quality is (a) uneven within components and (b) likely to vary from extremely good to very poor, from class to class and from project to project. Variations seem, predictably, to be greater in the early programs than in more recent years; trend analyses have shown steady progress in many areas of significance to Head Start with provision of extensive in-service training and election of parents to Policy Councils.

2. One-site Research and Evaluation Reports. Most of these reports are available through the ERIC Early Childhood Education Clearinghouse (University of Illinois: Urbana, Illinois). Many were funded by Head Start to assess the immediate effects of program participation on child development and to follow the children after they entered public school. Some have control groups of non-participating children; some involve pre and post measurement; some are after-only studies. Several have rigorous experimental designs comparing the immediate and longer-range effects of program participation in "traditional," and "experimental" Head Starts. Something of this variety exists, it should be noted, in the research studies of experimental preschool programs, not all of which have pre and post measures with eligible children assigned at random to experimental and control conditions. The findings from these studies have been reviewed by Grotberg (1969), Datta (1969), Stearns (1971), Butler (1970), Miller (1968), Weikart (1967) and others.

In general it has been found that most intervention programs have a statistically reliable effect on cognitive and linguistic performance, and that the gains are greater when (a) the program is directed toward specific educational objectives, and (b) the children participate for an eight- to nine-month period as contrasted to summer only.

Programs selecting children with marked developmental lags, programs in which the educational objectives are clearly identified and daily activities painstakingly matched to these objectives, and programs which are taught by the researcher or are supervised by him are likely to show very large gains and average or above average final levels of attainment. In these programs, few interactions with child characteristics are reported. Programs with children who are initially more competent, (as measured by the tests) programs with broadly stated objectives, and programs with little supervision directed to the classroom experience are likely to show modest benefits and may report more interactions.

It is of interest that most reports have either failed to find direct psychometric evidence of personal-social development or have not attempted to measure it, although "increasing the self-concept" is a frequently stated goal. This reflects in part the fact that there are few entirely satisfactory measures of cognitive, perceptual-motor, linguistic and academic skills for preschool children and, at present, no published measures of personal-social growth rated as satisfactory by reviewers. One important exception is Emmerich's report (1971) that Head Start children show substantial gains in socialization, cooperation and task-orientation and decreases in aggressiveness and timidity within the first six months of the program. This finding (based on actual observation of children, not tests) is consistent with similar changes reported in Head Start children by Dittman et al. (1971) in the clinical case histories also derived from observations and interviews, rather than tests. These studies lend support to the Zigler and Butterfield (1968) conclusion that changes in personal-social attitudes such as trust in adults and motivation may be readily achieved by intervention and may actually account for many so-called changes in "cognitive ability" as measured in the usual testing situation.

With due regard for methodological and other limits to present knowledge, most researchers conclude that immediate changes in the development of low-income preschool children are typically consequences of intervention programs, consequences that are well within the current state-of-the-art for small-scale programs and at least theoretically possible with proper supervision and effort in larger-scale projects.

Why then, have many other researchers concluded that "compensatory education has been tried and apparently has failed" (Jensen 1969)? For one thing, the large effects with above-average final levels are not that frequently reported. To date, only classes taught by the researchers themselves (Sprigle, Englemann) have yielded mean performances at the superior or gifted levels. Only classes under the direct supervision of the researcher (and not always then) yield large gains and average final levels (Weikart, Karnes). Most reports find statistically reliable gains and/or slightly below-average final levels.

Secondly, most--but not all--followup studies have found a "catchup effect." The experimental (E) children do not gain as rapidly after they enter public school although earlier gains are not lost. Control (C) children show a spurt which, while slight, is often enough to reduce the E-C difference to statistical non-significance. Almost all one-site Head Start followup studies and almost all followup studies of experimental preschools have reported this phenomenon for most measures.

While there have long been important exceptions (Beller, 1971), until very recently (Sprigle, 1971) there has been no instance of substantial effects persisting for three or four years after the children enter regular public schools. The recent report of Sprigle's Learning to Learn programs shows large, high final level and durable differences between E and C two years after the children have left the program. Understanding why this program has been so effective may, in conjunction with analyses of other data, constitute a "break-through" in the development of child development programs.

3. National Impact Studies. Only one national impact study has been reported for Head Start where the objective has been to assess overall effects rather than relate program characteristics to outcome. The Westinghouse Report (Cicirelli et al., 1969) was funded by the OEO Division of Research, Program Planning and Evaluation in summer 1968. The design involved comparison of the performance of Head Start and non-Head Start children in 104 sites who were in the first, second and third grades in October-November 1968. About two-thirds of the children had attended summer programs. Measures included the Metropolitan Readiness and Achievement tests, the individually administered Illinois Test of Psycholinguistic Abilities, and three experimental self-concept, scholastic motivation and attitudes toward school inventories. Teacher ratings and parent interviews were also collected. Although aspects of the study are methodologically controversial (Smith and Bissell, 1970), the findings are similar to those of the earlier one-site reports. Children tested soon after leaving Head Start achieved higher scores than controls; the second and third graders showed a "catchup" pattern. There was no evidence of motivational benefits or of benefits from attendance in the summer Head Starts, as measured by the criterion tests. Both the initial analyses of Cicirelli et al. and the reanalyses of Smith and Bissell showed that in some groups (Southern; black; urban) the Head Start children were substantially better off than the nonparticipant controls. These results are also similar to the findings of the 1965 summer programs reported by Coleman et al. (1966).

While the findings of the Westinghouse Report were similar to those of other studies, the conclusions drawn were not. Most researchers interpreting similar findings had concluded that programs should be extended through primary school. Cicirelli et al. concluded that most Head Start programs did not have the effects expected and recommended curtailing the ineffective summer projects and improving the educational program by installing known successful models. An experimental approach to Head Start was strongly urged. While many writers, bolstered by early reports from Follow-Through and other studies where programs are extended

upward in whole or in part, continue to advocate comprehensive long-term interventions, other writers (Bereiter, Rohwer) have questioned the "mystique" of early intervention and recommended interventions directed to career training at the late primary or junior high school ages. Wargo et al. (1971) found little evidence, however, that later interventions were more successful or durable than preschool programs; if anything, preschools appeared more "successful." From many points of view, these alternative recommendations and interpretations help maintain a perspective on development as a continuous process, whose "important periods" occur throughout the life of an individual: recognition of the importance of early childhood should not mean an under-emphasis on the opportunities and needs of later childhood and youth.

4. National Intervention and Interaction Studies. Most national evaluations funded by Head Start are based on an interactive assumption: the belief that different programs have different effects on different children. Prompted in part by Hunt's description of the importance of the proper "match" between the child's competencies and the program's challenges and in part by the diversity of Head Start children which seemed too great to be equally well served by one approach, this belief has been shared by such researchers as Di Lorenzo and Salter (1969) and Karnes (1969), who have investigated the different effects of several well-known curricula.

Head Start national evaluations have included a review of the summer 1965 reports (Planning Research Corporation), an analysis of the associates of the first full year (1966) program experience (Planning Research Corporation), an evaluation of the summer 1966 Head Start programs (Educational Testing Service), a study of the effects of Head Start on community institutions (Kirschner Associates), the Head Start/Follow-Through Planned Variation Longitudinal Study (Bissell, 1971), and the ETS Longitudinal Study (Shipman, 1970). From 1966 to 1969, the major evaluation program was undertaken by the 14 Head Start Research and Evaluation Centers. (See Table I.2.)

In all of these studies, assessments of program and teacher characteristics and of classroom process variables have received as much attention as assessment of initial and final levels of performance. The analyses in these projects have been directed to identifying the conditions associated with greatest gain for different children.

For summer 1966 and full year 1966 teacher and program characteristics were not in general reliably associated with changes or differences in language, cognitive or personal-social development. A number of correlations has been identified, however, as worthy of further investigation; for example, the negative relation between SB gains and structured classes and positive

correlations with informational-playful aspects. There has been to this date relatively little evidence from the Head Start data that child, family or program characteristics are differentially associated with magnitude of gain during the Head Start experience. While in every study, some differential effects and interactions appear, the small size of the effects and the infrequency of the interactions are more striking. Methodological problems inherent in these earlier data mean that interpretations must be made cautiously: one explanation was that the range of classroom characteristics was not sufficiently large in the 1965 and 1966 programs to permit a sensitive test of the interaction hypothesis, and that the measures were too academically oriented. The need to test the interactive hypothesis as sensitively as possible has prompted the evolution of Head Start national evaluations from simple random sample designs (1966 summer) to planned variations or interventions intended to influence what the classes look like (1969, 70 and 71).

The evaluation data reported here are intermediate stages in this evolution. In 1967-68, naturally occurring variations were sampled; in 1968-69, about one-third of the classes were "regular" Head Start programs and two-thirds were experimental programs.

Table I.2

HEAD START: NATIONAL RESEARCH AND EVALUATION STUDIES, 1965 THROUGH 1972

Year	Project
Summer 1965	Local research studies; PRC report; Census Study
Summer 1966 Full Year 1966-67	ETS national evaluation, Census survey E & R center study; Census survey
Summer 1967 Full Year 1967-68	Local research studies; Census survey E & R center study; Census survey
Summer 1968 Full Year 1968-69	Local research; Census survey E & R center study; Census survey; ETS longitudinal study, Year 1; (Westinghouse report, spring 1969)
Summer 1969 Full Year 1969-70	Census survey Planned Variation, Year 1; ETS longitudinal study, Year 2; Census survey; community impact study
Summer 1970 Full Year 1970-71	Census survey Planned Variation, Year 2; ETS longitudinal study, Year 3; Census survey
Summer 1971 Full Year 1971-72	Health Start Year 1 Planned Variation, Year 3; Home Start Year 1; ETS longitudinal study Year 4; parent participation study
Summer 1972 Full Year 1972-72	Health Start Year 2 Home Start, Year 2; ETS longitudinal study, Year 5; staff mobility study

CHAPTER II

THE DESIGN OF THE STUDY AND SELECTION OF MEASURES

A. Methodological Issues

Evaluation methodology has changed rapidly since 1965 when the first Head Start programs were funded. So rapid are these changes that in the two or three years between study design, data collection, analysis, and preparation of the final report, the design is likely to be obsolete. Studies likely to be designed in 1970-73 as compared with evaluation studies designed in 1965-68 are liable to be:

1) Less over-optimistic about how long it takes to implement an idea. 1972 studies of new programs such as Health Start and Home Start involve assessment of the planning process, of staff selection and training, of program management and administration, and will allow about three years before impacts on the target population are assessed.

2) Closely linked to program objectives. 1972 evaluators work with program management to identify objectively how the manager will know if the program is successful and when short-term, intermediate, and longer-term effects are expected. Evaluators in 1972 are more conscious of the need for prior development of sensitive, reliable measures, or when the program is launched, alert those concerned to the fact that some objectives cannot be measured satisfactorily. Measures are more likely to be criterion-referenced indicators of specific behaviors to be changed and less likely to be exclusively indicators of "transfer" or "generalized" effects.

3) More efficiently designed. The limitations of quasi-experimental designs, of post hoc matches, of gain scores, and of covariance adjustments to compensate for playing fast and loose with the design required for rigorous statistical inference have become more apparent. The problem is not only one of assessing change, but also of being able to ascribe the change to a program intervention, which is what most evaluation is required to accomplish, or to impute relative change to different kinds of treatments or programs. The minimum design required, unless there are nationally standardized measures of effect and the program is administered to a simple random national sample, involves a control group which does not receive the experimental treatment, one selected at random from applicants for the experimental treatments. Many evaluation theorists are now advocating adherence to minimal design criteria even if this means smaller samples for nationally important studies.

The studies reported here were designed with some but not perfect foresight. Among the limitations (from a 1972 perspective) are:

1) Lack of controls who have not attended Head Start.

2) Non-random assignment of children to different program types within Head Start.

3) Use of "transfer" tests rather than criterion-referenced measures and "transfer" tests.

4) Non-uniformity of time and conditions of data collection.

Among the advantages of these 1967 and 1968 studies are:

1) Extensive descriptions of what actually happened in the classes.

2) Data on child, family and program characteristics for large national samples over a two-year period so replication is possible.

3) Concern with personal-social as well as academic development.

4) Initial and final measurement.

In 1966-67, 470 full-year programs serving 160,000 children were funded. Funding began in the summer of 1966 and continued throughout FY66. By January 1967, only about 50 percent of the programs that were to be funded during the year were actually operating. Evaluation of these programs was primarily vested in 14 regional University-based Evaluation and Research (E & R) Centers. Not all of these Evaluation and Research Centers could be funded in September; the last Center did not begin operation until early in 1967. The overlap of the beginning of full-year Head Start classes with operation of E & R Centers was not complete so that the 1966-67 data do not provide an ideal "pre" and "post" study. The average interval between first and second testing was only four months, and some children were enrolled for as long as seven months before they were first tested. The study was designed and measures selected by the 14 Evaluation and Research Center Directors and Head Start national office staff. Data were collected by the centers and analyzed under the direction of the national office.

The Evaluation and Research Centers were in operation from 1966 to 1969 when the planned variation project was begun, and the Centers phased out. The designs and measures developed in the three-year period are summarized in Table II.1. (See also Datta, 1969).

Table II.1

HEAD START NATIONAL EVALUATION DESIGNS AND MEASURES
1966, 1967, AND 1968

	1966	1967	1968
Design			
. N classes	226	177	148
. N children	1806	1889	1989
. Selection of classes	variability by category	variability by class process	experimental interventions
. Selection of children	random within class	all in each class	no prior H/S
Common Measures			
. Classroom activity	ORF	OSCI I	OSCI II, POT, teacher questionnaire
. Teacher characteristics	questionnaire	questionnaire	questionnaire
. Class resources and facilities	observer form	observer form	observer form
. Family background	interview	interview	interview
. Administration	---	---	checklist
. Cognitive performance	Binet	Binet	Binet
. Achievement	PSI	---	PSI, GUMP
. Motivation	teacher rating	tester rating	tester rating
. Social adjustment		tester rating; SIO	tester rating, sociometric
Special Measures	---	four clusters	Special for each experiment
Data Collection			
. Mean week of Pretest	12	---	7
. Mean week of Posttest	28	---	32
. Mean interval of weeks	16	---	25
. % with prior Head Start	30	---	29

1. Design

In 1966, classes were selected to provide variability by the delegated agency operating the program (Community Action Agency or Local Educational Agency), location (rural/urban), and child ethnicity (black, white, Spanish-surnamed, Polynesian). As many classes as possible were studied. Samples of 6-8 children within each class were selected.

In 1967, information on classroom style and process for individual teachers as provided by regional office staff was used to select classes. A distribution on the 1966 characteristics was also sought. Fewer classes were studied, and all children within each class were to be included. This design might possibly, but not inevitably, provide a less representative sample of Head Start programs than in 1966. However, it offered a greater opportunity to assess differential effects within a classroom.

In 1968, many E & R directors felt that the "natural variation" approach meant that too few sites were exemplars of new ideas in early childhood education and some believed program effectiveness was not uniformly high, if performance on achievement and cognitive tests was the criterion. Each E & R director developed a research design in an area of particular interest with comparison (regular) Head Start classes and intervention classes. Insofar as possible the interventions were diverse including projects emphasizing language, basic cognitive processes, parent education, community participation, setting of objectives, and facilities. The designs are summarized in Table II.2. There were no sampling requirements other than the selection of children with no more than four months of prior Head Start experience and the national geographic dispersion due to location of the E & R Centers. Most Center directors chose sites relatively near to their universities, with attention to the rural/urban distributions. This design would permit greater inference about curriculum and interactive effects but would not necessarily be programmatically representative of Head Start classes. Since nationally representative Census samples were drawn in 1968-69, divergence from many national child/staff characteristics could be checked, and, it was expected, either appropriate statistical adjustments made, or conclusions properly limited.

Table II.2

EVALUATION AND RESEARCH CENTER DESIGNS, 1968-69

Center[2]	No. of Classes C[3]	No. of Classes E[4]	Objective of Intervention or Study
Boston	4	8	Effects of community participation
Syracuse	14	-	Teacher praise/blame behavior
Bank Street	8	8	Comparison of Bank Street and "structured" HS
Temple	6	6	Effects of greatly enriched materials
South Carolina	-	12	Language development program[1]
Tulane	-	10	Language development program[1]
Southern	1	11	Language/parent education[1]
Texas	2	11	Language development program[1]
Kansas	4	7	Behavior modification training
Michigan	-	9	Piagetian training-cognition
UCLA	8	10	Teacher feedback on progress of evaluation
Hawaii	-	8	Language/parent education[1]

1/ Tulane, South Carolina and Texas all had one design and program; Hawaii/Southern also had another design and program.

2/ The University of Chicago (Dr. V. Shipman) and Teachers' College (Dr. R. Thorndike) participated in the 1966-67 and 1967-68 studies only.

3/ C = Comparison, regular Head Start classes.

4/ E = Experimental Head Start classes.

2. Measures

In every evaluation, parent and class data have been collected and efforts made to develop and collect child data appropriate to the several Head Start objectives. It has, however, proved difficult to create reliable, readily interpreted and easily administered measures which are also psychologically meaningful for different age and cultural groups. In many studies, the evaluator is caught between investing in (a) new measures that look meaningful but may be very costly to use or prove to be insensitive or unreliable, and (b) available measures that may be inappropriate developmentally or culturally, and thus uninterpretable--or worse, misleading because the usual interpretation is well known but invalid for the Head Start population.

The 1966 measures included the Stanford-Binet, interpreted as a measure of general cognitive performance; the Preschool Inventory, interpreted as a measure of achievement and school readiness; and the Head Start Behavior Rating Scale, completed by the teacher for each child. The Observer Rating Form (ORF) describing teacher style in the classroom completed the main battery.

In 1967, a large investment was made in collecting very detailed information on child/child and child/adult classroom interactions. The Social Interaction Observation (SIO) measures were collected pre, mid, and final, and involved 45 minutes of free play observation in 10 second intervals, about 2900 records per child each of which could contain between three and ten discrete "bits" of information. A tester rating form to describe the child's adaptation to the testing situation was added, and a detailed record of classroom activity (OSCI) was obtained for five days spread throughout the year. In addition, E & R centers formed four clusters which collected finer-grained measures of the child's (a) personal-social development, (b) language, perceptual motor and cognitive development, (c) teacher and classroom, and (d) family characteristics.

In 1968, the OSCI was revised to record both classroom activity and teacher behavior, a checklist of administrative variables and a rating form for teacher style/sanctions were developed, the parent interview was augmented, and a sociometric measure and a new test of achievement motivation for preschool children (Gumpgookies) were adopted. National training for OSCI observers and parent interviewers, and a national coordinating center to help ensure uniformity of data collection and coding were provided.

The 1967-68 and 1968-69 measures and the rationale for their selection are described in the following section. The 1966 measures are not included since this report is directed to the 1967 and 1968 program years. In addition, the cluster measures (1967) and special measures appropriate for the E & R Center intervention (1968) are not discussed. The cluster measure data have not been analyzed, in part because some of the experimental tests proved unreliable and in part because the common core data received priority. Information on the special measures is available in the individual reports on their experiments submitted by E & R Center directors to the national Head Start office.

B. Measure Selection and Psychometric Characteristics

For 1968-69, child-family and teacher-program information was collected through the use of 24 forms, questionnaires or test instruments. Information available on the child included pre- and posttests on the following instruments: Stanford-Binet, an Inventory of Factors Affecting the Stanford-Binet, Gumpgookies, Hertzig-Birch Response Style, Sociometric Play Situation, Animal House Scale of the Wechsler Preschool and Primary Scale of Intelligence (AH-WPPSI), and the Preschool Inventory. In addition, the Child Master Data card used in 1967-68 data collection was used again.

1. Child and Family Variables (1968-69)

The Stanford-Binet was developed as an intelligence test, that is, a measure of the individual's ability to learn from experience. This interpretation requires that the individual tested have been exposed to experiences similar to those of individuals on whom the measure was standardized. In the 1960 revision of the Binet, the standardization sample is considered to overrepresent white, and more advantaged children. Even if the sample had been representative of the ethnic and economic circumstances of children in the general population, the performance of any group whose life experiences diverged from the standardization sample as a whole would have to be interpreted as both an indication of experience per se and ability to learn from experience. The high IQ's of more advantaged children therefore indicate a greater ability to perform certain tasks but the ability is due both to possible differences in learning capacity and certain differences in life circumstances that favor performance or test items. The lower performance typically achieved by a child from economically constricted circumstances indicates less ability on the skills measured by the Binet but the level of performance is due both to possible differences in learning capacity and certain differences in life circumstances that do not favor performance on test items. In addition, willingness to perform and adaptation to the test situation influence performance on any measure in this study, and further attenuate interpretations that might reflect "native capacity."

As a measure of performance, the Binet absolute level has repeatedly been found sensitive to intervention of many kinds; relative position within groups exposed to the same interventions changes much less. Where the intervention is very effective, the individual differences within groups may in fact be increased, as children more able to learn respond to the opportunities they have not previously had. The Binet frequently has been reported to predict school achievement as well as rate of learning a variety of new tasks.

The Inventory of Factors Affecting the Stanford-Binet is a modification of the face sheet of the Stanford-Binet (SB). The tester rates the child on his observed behaviors during the SB testing situation. The adaptation used was prepared by Dr. Herbert Zimiles of the Bank Street College of Education and Dr. Carolyn Stern of UCLA. The ratings reflect traits such as impulsivity, concentration, and activity level.

The Gumpgookies (GUMP) is an experimental test of achievement motivation (Ballif and Adkins, 1968) designed for preschool children. The pre form consisted of 100 items while the post form consisted of 55 items. Since the original form was considered too long to administer to young children, 45 of the 100 items were discarded on the basis of E & R Center critiques of the 100 items and psychometric data. Each item consists of a card containing two "Casper the Friendly Ghost"-like creatures, one of which is intended to exhibit more achievement-oriented behavior than the other. For example, one figure might be successfully completing a task such as building a block tower while the other figure might have fallen blocks all around him. The child is then asked which "Gumpgookie" is his "Gumpgookie." Internal consistency reliability for the Gumpgookies is .76.

The Hertzig-Birch Response Style Instrument, like the Inventory of Factors Affecting the Stanford-Binet, is given in conjunction with the SB (Hertzig et al., 1968) and is designed to measure various cognitive styles that the child employs when responding to SB items. For example, a distinction is made between a wrong response where a child "works" at a problem solution and a wrong response where a child does not "work" at a problem solution. Hertzig, Birch et al. (1968) have demonstrated differences in these cognitive styles among various ethnic and SES groups.

In the Sociometric Play situation developed by Dr. Robert Boger, each child selects (a) three play situations depicted on six choice cards and then (b) the photograph of a preferred classmate to put into each of the three play situations. Measures of each child's popularity and isolation can be derived from this instrument. The structure of

the observations resulted in artifactual pre-post correlations; that is, if some children's popularity increased from midsession to post, the remaining children's popularity has to decrease due to constraints imposed upon the children's ratings. Analyses with other measures given at the same time, or a categorical use of the data with particular attention to the isolate, seem metrically promising under such circumstances.

The Animal House Scale of Wechsler Preschool and Primary Scale of Intelligence (AH-WPPSI) is similar to a Digit-Symbol or Coding Test. A key at the top of a board has pictures of different animals, each with a differently colored cylinder ("house") under it. The task involves inserting the correctly colored cylinder in the hole beneath each animal on the board. The score is a function of the time, errors, and omissions. The WPPSI is also a standardized test of ability to learn. This subtest was selected as a relatively culture fair performance measure of ability to learn a new task.

The Preschool Inventory (PSI) was originally developed in 1965 to assess the level of cognitive maturation of Head Start children. It was designed to measure skills that are believed necessary for success in entering elementary school. There are four subscales and a total scale for this instrument. The four subscales are the Social Responsiveness Scale, the Associative Vocabulary Scale, the Concept Activation-Numerical Scale, and the Concept Activation-Sensory Scale. These scales are highly intercorrelated. The total score is derived from the responses to 76 items. The PSI seems to be measuring cognitive functioning at a more concrete level than the SB, and is interpreted as a measure of achievement or school readiness. The form used in 1968-69 was an experimental version developed for this study which has since become the standard measure commercially available (ETS). Performance data from the 1967-68 E & R sample contributed to item selection and revision and the 1968-69 E & R pre data are now the national standardization sample data for low-income children.

Information on the family was available from the following questionnaires and subitems: Parent Interview-Pre, Parent Interview-Post, First Day of School question, and the Adult-Child Influence Techniques items. The latter two were part of the parent interviews.

The Parent Interview-Pre contained questions pertaining to socio-economic level such as family income, father's occupation, mother's education, and father's education. There were also measures of the mother's expectations, aspirations, attitude toward education, and attitude toward life in general. These latter items are interpreted as locus-of-control measures.

Locus-of-control (internal/external) reflects the individual's sense of power (internal control) versus helplessness (external). The individual with a sense of internal control believes he has the

power to control his own life and its circumstances and to influence the lives of others and the behavior of authorities. The individual with a sense of external control believes he has little command over what happens to him and his life circumstances and little ability to change, influence or modify the lives of others. I/E has been found to predict which individuals among those sharing constricted life circumstances are able to overcome these circumstances. In the Coleman (1966) survey, for example, the black children with high test performance also had high I scores. In other studies, I/E changes have been associated with change in achievement. Direction of causality is open to question: Are the high I individuals more accurate in rating their influence and power because they are more competent, or are they better able to take advantage of whatever opportunities they have because they believe they can?

Information was also available on family structure, amount of reading to the child, living space, parent participation in Head Start, and various parental child control techniques. The Parent Interview-Post administered at the end of the child's participation in Head Start contained many of the same questions as the Pre-Interview so that it was possible to assess changes in the families' life styles and attitudes from the beginning to the end of the child's participation in the Head Start program. In most instances, the mother provided the parent information.

The First Day of School items (Hess et al., 1969) involved reporting a mother's response to the open-ended question "What would you tell your child to do on the first day of school?" The item measures how the mother views the teacher, classroom, social situations, achievement, home-based preparation, and personal safety. A trained rater places thought expressed by mother's "messages" into a number of a priori categories. The ratings were done under the supervision of Dr. Robert Hess. The measures have been found to be sensitive to social class differences within black samples, to predict preschool achievement, and to be reliable predictors of later school achievement.

There were four Adult-Child Influence Techniques items. One question asks the mother what is one of the worst things that her Head Start child does and how does she (the mother) respond to it. A second question asks the mother what are little annoying things that her Head Start child does and how does she respond to them. The third and fourth questions relate to what the child does that the mother likes and likes a great deal, and how the mother responds when the child does these favored things. The responses are then categorized by the parental means of modifying or reinforcing the child's ongoing behavior. The broad categories along the main power dimension are non-intervention, intervention,

qualified power assertion, and unqualified power assertion. This measure was developed by Dr. Irving Sigel and Dr. Bela Fehrer. Coding was completed under Dr. Fehrer's supervision. The first two items also have been found to predict preschool achievement and later school achievement. The last two were experimental.

2. Teacher-Program Variables (1968-69)

Head Start teacher and program variables were measured by a Post Program Teacher Interview, the Checklist of Administrative Variables, the Observation of Substantive Curricular Input, the Post Observation Teacher Rating Scale, the Characteristics of Teaching Staff Form, and a Pre and Post Class Facilities and Resources Inventory.

The Teacher Interview questions explored program focus, goals considered important by the teacher, description of teaching approach, and packaged programs used by the teacher. It was developed as an independent reliability measure for the intervention studies of 1968-69 by Dr. Lois-ellin Datta from inputs by all of the Head Start E & R Center Directors.

The Checklist of Administrative Variables was designed to determine th person(s) responsible for each of the 31 key functions or activities the Head Start full-year program. Some sample questions are: Who articipated in planning the parent orientation meeting program? Who was involved in making out the agenda for teachers' meetings? Prepared by Dr. Shuell Jones (Tulane E & R director), it assessed, by teacher report, whether or not events that should happen in a well-run Head Start program did happen and if the people who should be involved in making decisions about these events (teachers and parents as well as the H/S director) were involved. It is thus a measure of the quality of program administration as seen by the Head Start teachers.

The Observation of Substantive Curricular Input (OSCI) was developed by Dr. Carolyn Stern. The OSCI involves a complex schedule that is used by a trained observer to code the behavior of the class and teacher. The rater's task consists of using a system of codes for describing the wide variety of activities and program inputs found in Head Start classrooms.

Briefly, the OSCI coding system is based on a series of three-minute scans of ongoing activity in the classroom. In some of the three-minute scans the observer looks at what the children are doing and in other three-minute scans, the observer looks at what the teacher and the aide are doing. For each three-minute classroom scan, the largest group of children is located and a number of major codes is recorded for this group. The major codes reflect the context of the activity (e.g., building), content of the activity (e.g., mechanical), the forms of control (e.g., teacher controlled or child initiated), and teacher involvement (e.g., present and watching children). Within the three-

minute period, the observer then locates the next largest group, and makes the same recordings, proceeding likewise until the last group or individual child is observed in the three-minute time permitted.

A three-minute record could potentially consist of one unit if all children were occupied in the same activity or as many units as there are children if each child was working individually. For each three-minute teacher scan, the observer records, among other things, the teacher context, teacher content, and teacher involvement for each of six 30-second intervals comprising the three-minute scan period. The behavior of the teacher aide is also recorded in a similar manner in other three-minute scans.

These teacher and classroom observations were collected on five separate days spread through the midpoint and end of the program year. Each daily observation covered most of the day for the majority of Head Start sample classes. Three-day training sessions for the observers were held and the reliability of each observer was checked. Observer quality was also checked by reviewing each OSCI record at UCLA before the next observation was made. Additional training was given if necessary. The frequency of occurrences for selected content-context categories and control categories were obtained for both the class-centered and teacher-centered observations for most Head Start sample classes and teachers. The simple class frequencies for the various categories were then intercorrelated, factor analyzed, and the factors rotated. Likewise, the simple teacher frequencies for the various categories were intercorrelated, factor analyzed, and the factors rotated. For both the class and teacher analyses, six rotated factors were retained for interpretation and factor scores were calculated for analysis purposes.

The six teacher factors were labeled as follows:

TI Social-Emotional Interaction

TII Structured Lessons--Large Group

TIII Art Activities

TIV Creative Instruction--Small Group

TV Routines

TVI Receptive Learning

Teacher Factor TI (Social Emotional Interaction) was partially defined by emotional and social behavior of the teacher. Teacher Factor TII (Structured Lessons--Large Group) was partially defined by the teacher's use of

Table II.3

DESCRIPTION OF EACH OF 6 FACTORS BASED ON 66 TEACHER VARIABLES WITH A LISTING OF THE VARIABLES LOADING .30 AND ABOVE*

Variable Number	Factor Loading	Variable Description
Factor TI:	Social-Emotional Interaction	
43	.66	emotional/All Contexts
32	.65	social/Physical Contact
46	.64	nothing/Not Applicable
16	.59	discrimination/Routines
66	.58	content directed to other child
34	.52	social interaction/All Contexts
25	.51	verbal communication/Routines
64	-.50	no content directed
47	-.39	nothing/Nothing
4	-.37	language/Structured Lesson
57	.39	medium group (5-10)
42	.39	rules/Routines
56	.38	small group (1-4)
33	.36	social phys./Activities, Lessons, Routines
65	.36	content directed to child
55	-.36	programmed materials
58	-.34	large group (11-20)
61	-.32	teacher setting up
17	.30	visual discrimination/Activities
Factor TII:	Structured Lessons-Large Group	
4	.85	language/Structured Lesson
18	.80	visual discrimination/Structured Lesson
55	.76	programmed materials
15	.71	auditory discrimination/Activities
20	.60	visual motor/Auditory Discrimination
47	-.59	nothing/Nothing
41	.55	rules/Activities-Language
58	.49	large group (11-20)
65	.46	content directed to child
62	-.42	teacher uninvolved
60	.38	teacher supervision
63	-.38	teacher not present
42	.36	rules/Routines
59	.35	teacher/Active

(continued)

* A few variables with loadings under .30 have been retained as factor descriptors. They had maintained their position on those factors on other solutions and with higher loadings. These factor loadings were derived by Dr. Carolyn Stern and her colleagues at the UCLA Head Start Evaluation and Research Center.

Table II.3 (continued)

Variable Number	Factor Loading	Variable Description
Factor TIII:	Art Activities	
13	.82	art/Painting
23	.68	verbal communication/Arts
52	.65	dramatic play equipment
45	-.54	nothing/Interval
31	-.50	social/Interval
28	.33	social verbal/Performing Arts
40	.27	skills/Activities-Language
Factor TIV:	Creative Instruction--Small Group	
56	.63	small group (1-4)
50	.61	large muscle equipment
19	.50	visual motor/Activities
51	.49	small muscle equipment
27	.46	social verbal/Activities
48	.45	science materials
22	.44	verbal communication/Activities
58	-.42	large group (11-20)
59	.38	teacher active
21	.35	perceptual-other/Activities-Language
28	.35	social verbal/Performing Arts
29	.35	social verbal/Structured Lessons
9	.35	dramatic/Performing
38	-.30	mechanical/Eating
10	.28	dramatic/Activities
14	-.26	dance/Performing
39	-.25	mechanical/Waiting
Factor TV:	Routines	
35	.56	mechanical/Activities
49	.55	large muscle equipment
27	.50	social verbal/Activities
37	.48	mechanical/Routines
30	.45	social verbal/Routines
38	.45	mechanical/Eating
47	-.42	nothing/Nothing
59	.40	teacher active
36	.40	mechanical/Clean Up
11	.39	music/Musical Activity
12	.38	music/Other Activities
5	.34	cognitive/Routines

(continued)

Table II.3 (continued)

Variable Number	Factor Loading	Variable Description
Factor TV:	Routines (continued)	
65	.33	content directed to child
53	.32	music & drama materials
57	.31	medium group (5-10)
25	.30	verbal communication/Routines
Factor TVI:	Receptive Learning	
2	.67	language/Watching-Listening
17	.60	visual discrimination/Activities
54	.55	language materials
44	.48	not applicable/Watching
24	.48	verbal communication/Watching-Discussion
65	.44	content directed to child
33	.39	social phys./Activities, Lessons, Routines
26	.39	verbal communication/While Waiting
59	.34	teacher active
29	.31	social verbal/Structured Lessons
56	.31	small group (1-4)
22	.30	verbal communication/Activities
6	.29	quantitative/In all possible contexts
1	.29	language/Activities

structured language lessons, structured visual discrimination lessons, and the use of programmed materials. Teacher Factor TIII (Art Activities) was partially defined by art content in a painting context and verbal communication in an art context. Teacher Factor TIV (Creative Instruction--Small Group) was partially defined by visual motor activities, small groups and the presence of large and small muscle equipment. Teacher Factor TV (Routines) was partially defined by the teacher using mechanical devices in an activity and routine context. Teacher Factor TVI (Receptive Learning) was partially defined by language content in a watching-listening context, visual discrimination in an activities context, and the presence of language materials.

The six class factors were labeled as follows:

CI Structured Lessons

CII Group Activities and Routines

CIII Social-Emotional Interaction

CIV Verbal Communication

CV Instruction in Creative Arts

CVI Language and Discrimination Learning

Class Factor CI (Structured Lessons) had high loadings from the presence of programmed materials, language through structured lessons, head teacher in control, and large groups. Class Factor CII (Group Activities and Routines) was defined primarily by music, drama, and art activities in a social context. Class Factor CIII (Social-Emotional Interaction) was partially defined by emotional content in all contexts, social content in a physical contact context, and child being in control of the classroom. Class Factor CIV (Verbal Communication) was primarily defined by verbal communication content in a number of diverse contexts such as activities, arts, and routines. Class Factor CV (Instruction in Creative Arts) was primarily defined by visual motor content in an activities context, child in control, small group, presence of dramatic play equipment, and dramatic content in a performing context. Class Factor CVI (Language and Discrimination Learning) was primarily defined by the presence of language materials, language content in a watching-listening context, and visual discrimination in an activities context. Table II.3 and II.4 give the teacher and class OSCI factor loadings. These OSCI Factors will assume more meaning in later data analysis and interpretation sections of this report.

Table II.4

DESCRIPTION OF EACH OF 6 FACTORS BASED ON 60 CLASS VARIABLES WITH A LISTING OF THE VARIABLES LOADING .30 AND ABOVE*

Variable Number	Factor Loading	Variable Description
Factor CI:	Structured Lessons	
53	.91	programmed materials
4	.90	language/Structured Lesson
58	.82	Head Teacher in control
56	.81	large group (11-20)
18	.80	visual discrimination/Structured Lesson
20	.68	visual motor/Auditory Discrimination
15	.62	auditory discrimination/Activities
54	-.51	small group
49	-.46	small muscle equipment
31	.45	social phys./Activities, Lessons, Routines
9	-.39	dramatic/Performing
46	-.34	science materials
60	-.34	other audit in control
38	.31	mechanical/Waiting
Factor CII:	Group Activities and Routines	
30	.64	social phys./Activities, Lessons, Routines
28	.60	social/Interval
11	.58	music/Musical Activity
55	.57	medium group (5-10)
32	.55	social phys./Activities, Lessons, Routines
59	.54	assistant teacher in control
34	.52	mechanical
44	.51	nothing/Interval
36	.48	mechanical/Routines
47	.47	large muscle equipment
37	.46	mechanical/Eating
10	.45	dramatic/Activities
51	.41	music & drama materials
3	.38	language/Discussion
38	.36	mechanical/Waiting
56	.34	large group (11-20)
39	.31	skills/Activities/Language
60	.30	other audit in control

(continued)

* A few variables with loadings under .30 have been retained as factor descriptors. They had maintained their position on those factors on other solutions and with higher loadings. These factor loadings were derived by Dr. Carolyn Stern and her colleagues at the UCLA Head Start Evaluation and Research Center.

Table II.4 (continued)

Variable Number	Factor Loading	Variable Description
Factor CIII:	Social Emotional Interaction	
42	.63	emotional/All Contexts
45	.57	nothing/Nothing
29	.56	social/Physical Contact
57	.55	child in control
54	.46	small group (1-4)
33	.39	social interaction/All Contexts
38	.37	mechanical/Waiting
28	.35	social/Interval
44	.30	nothing/Interval
Factor CIV:	Verbal Communication	
22	.72	verbal communication/Activities
26	.71	verbal communication/While Waiting
24	.60	verbal communication/Watching/Discussion
23	.57	verbal communication/Arts
25	.50	verbal communication/Routines
59	.31	assistant teacher in control
1	.30	language/Activities
Factor CV:	Instruction in Creative Arts	
19	.55	visual motor/Activities
57	.51	child in control
54	.50	small group (1-4)
50	.47	dramatic play equipment
48	.45	large muscle equipment
30	.42	social phys./Activities, Lessons, Routines
21	.40	perceptual-other/Activities/Language
9	.40	dramatic/Performing
32	-.38	social phys./Activities, Lessons, Routines
49	.37	small muscle equipment
13	.36	art/Painting
41	-.35	rules/Routines
46	.33	science materials
26	.32	verbal communication/While Waiting
28	.31	social/Interval
Factor CVI:	Language and Discrimination Learning	
52	.81	language materials
2	.80	language/Watching/Listening
17	.68	visual discrimination/Activities
16	.39	discrimination/Routines
8	.35	science/In all possible contexts
40	.32	rules/Activities/Language

The Post Observation Teacher Rating Scale involved 33 different ratings on a five-point scale of teacher behavior obtained on the same day that the OSCI was given. Consequently, the teacher in the classroom was rated on five different occasions. The teacher was rated on such behavior as her reliance on ongoing activities, her attention to groups, and her awareness of pupil frustration. This measure was developed by Dr. John Dopyera of the Syracuse E & R Center.

The Characteristics of the Teaching Staff form (by Dr. Carolyn Stern) was administered at the end of the Head Start program and provided information regarding the head teacher, teacher, and teacher aide's age, sex, ethnicity, education, training, and experience with preschool children and disadvantaged preschool children.

A Class Facilities and Resources Inventory-Pre gathered class information including the operating length of the program, length of class day, number of children in class, and the various types of educational and physical resources available in the class. The Class Facilities and Resources Inventory-Post contained information similar to the pre instrument and was intended to check on changes in facilities and resources of the class during the year. Much of this form is identical to the 1968-69 Head Start national census survey forms and can be used to assess similarity of the E & R sample programs to the national sample.

Information on all of the above instruments, forms, questionnaires and tests were available on most 1968-69 sample children. Since the 1967-68 data was used mostly for replication, the discussion of the 1967-68 instruments, forms, questionnaires and tests that were analyzed in this study is briefer than that for 1968-69.

3. Child-Family Variables (1967-68)

SB IQ Scores, both pre and post, were available on most of the 1,898 children in the 1967-68 sample. Pre and post scores were also available on the Inventory of Factors Affecting the SB. A Child Master Data Card also contained information regarding the child's class such as ethnic mix and stability of class, length of daily program, and number of class meetings per week.

A Parent Interview administered at both the beginning and end of the Head Start program contained information similar to the pre and post Parent Interviews for the 1968-69 sample families, e.g., attitudes, expectations, aspirations, socioeconomic, and "quality of life" information.

4. Teacher-Program Variables (1967-68)

A Characteristics of Teaching Staff Instrument, similar to the 1968-69 instrument, contained information regarding head teacher, teacher, and teacher aide training and experience. A center and classroom composition instrument similar to the 1968-69 Pre and Post Class Facilities and Resources Inventory contained information on the physical and educational resources of the Head Start centers and classrooms.

The 1967-68 Observations of Substantive Curricular Input (OSCI) was similar to the 1968-69 OSCI but gathered class information only. Analysis of the intercorrelations of the frequencies of the various class content-context categories across classes yielded only four rotated Class Factor Scores, while the more comprehensive 1968-69 OSCI yielded six Class and six Teacher Factor scores.

The four Class Factor Scores were as follows:

CI High Cognitive Activity in a Low Structure Situation

CII Routines and Rules

CIII High Cognitive Activity in a High Structure Situation

CIV Child-Centered, Unstructured

Class Factor CI (High Cognitive Activity in a Low Structure Situation) was mostly defined by high loadings of verbal instruction during learning activities and verbal instruction during routines. Class Factor CII (Routines and Rules) was characterized by high loadings on rules emphasized during routines, whole group activities, and social interaction during learning activities. Class Factor CIII (High Cognitive Activity in a High Structure Situation) was partially defined by positive loadings of visual discrimination and cognitive input during watching or listening, and a negative loading for individual activity. Class Factor CIV (Child-Centered, Unstructured) was partially defined by high positive loadings for social interaction during learning activities and high negative loadings for mechanical performance of routines.

Also available were data from the Child Influence Techniques items and the First Day of School items. Both of these are similar to the corresponding instruments for 1968-69 and were coded under the supervision of Dr. Hess, Dr. Sigel and Dr. Fehrer.

Child Interactional information unique to the 1967-68 data was obtained from the Kansas Social Interaction Observation procedures (SIO). This instrument was child oriented and attempted to measure the amount

and types of interaction among the sample children and adults. Verbal and nonverbal initiations from the child to an adult, the number of adult initiations responded to by the child, and the number of child initiations responded to by the adult are available. Developed at the University of Kansas E & R Center, it reflects a behavior modification orientation to social development and adaptation.

CHAPTER III

SELECTION OF VARIABLES FOR ANALYSIS; ANALYTIC AND METHODOLOGICAL CONSIDERATIONS

The information collected in 1967-68 and 1968-69 was a minimal set of basic data constrained by the 30 minutes or so attention span in a testing session before the child fatigued, the problems of scheduling repeated sessions, and the desire to avoid overtesting. The risks of using only standard scales (few such measures are available and these are mostly in the areas of cognitive, linguistic and academic achievement) were balanced against the risks of using only experimental measures (loss of expensively collected data due to unreliability or uninterpretability). To a certain extent the measures deliberately oversampled, collecting a rich data base for use by many scholars, but one which could not be fully utilized by a single analytic project.

The first task of the analytic effort was therefore to examine the data base, conceptually and empirically. The objective was to identify a set of variables that:

1) Were reliable,

2) Showed good variation and distribution,

3) Were developmentally meaningful for an analysis of change,

4) Included dimensions of child, family and program characteristics of significance in the current child development literature, in emerging theory, or to Head Start policy,

5) Were reasonably interpretable, or for which other studies, reports and manuals provided a context for understanding presently observed data, and

6) Were feasible with regard to manipulation of the data bank.

The variables eventually selected may be disappointing in that the personal-social data, particularly the SIO (social interaction) and Hertzig-Birch response style were not extensively analyzed. This in no way reflects a policy decision reducing the importance of this area. Personal-social development was--and is--among the greatest concerns of Head Start. It is also among the most difficult to measure, analyze and interpret. More effort was placed on analyzing classroom process and family variables among the other measures, because of the quantity of available references on these measures and the prior analytic work, relative to scarcity of available information on the SIO and Hertzig-Birch.

The variables and analyses reported here therefore should be regarded as early findings from the Head Start data. Subsequent analyses will doubtless expand our knowledge of other aspects and may modify initial conclusions. While the rest of this report will not dwell on what was not analyzed, the fact that these are selected data mined from a remarkably complete and rich lode should be kept in mind.

The overall question of the analyses was to identify relative benefits of different program experiences for different children. Two broad classes of variables were accordingly identified: outcome variables (which in some analyses also served as predictor variables) and explanatory variables. For these analyses, the individual child was the focus. Changes in families, teachers and classes per se are touched upon briefly. The direct concern is the developmental status of the child: Does the program sustain a presumably desirable characteristic (e.g., adjustment to testing with no or very few behavior problems) and does the program support the development of the child who may have initially shown growth retardation or adjustment problems? Which programs are most effective in this regard, and for which children?

A. Outcome and Explanatory Variables - 1968-69

1. Selection of Variables*

The 1968-69 variables will be described within the framework presented in Table III.1. Originally, 90 explanatory and outcome variables were considered for analysis; this list can be found in Appendix C of the present report. The variables presented in Table III.2 are the final list considered for analysis. Variables from the preliminary list were dropped for a variety of reasons. The most common reasons were: (a) computational and cost factors in deriving an index from the raw data provided on tape, (b) undesirable marginal distributions (e.g., no variation in an explanatory variable), (c) low internal consistency reliability of derived or scaled variables as measured by Kuder-Richardson Formula 20, and (d) apparent uninterpretability or extreme redundancy as shown in analyses of the intercorrelations and inter-dependencies among many sets of preliminary outcome and explanatory variables.

---

* Consultants involved in selecting variables were Dr. Lyle Jones (University of North Carolina), Dr. James Gallagher (University of North Carolina), Dr. John McDavid (Georgia State University), Dr. Richard Endsley (University of Georgia) and Dr. Anthony Conger (University of North Carolina). Dr. Endsley and Dr. Conger worked closely with Dr. George Dunteman and Dr. A. V. Rao of RTI in preparing final selections.

Table III.1

CLASSIFICATION OF 1968-69 EXPLANATORY AND OUTCOME VARIABLES

I. Outcome Variables (Raw Gain and Gain Scores Adjusted for Pre Scores)

- A. Cognitive Behavior (e.g., Stanford-Binet performance)
- B. Behavior Ratings (e.g., Behavior Problem Scale)
- C. Achievement Motivation (e.g., Gumpgookies)

II. Explanatory Variables

- A. Initial Child Status
  - 1. Demographic Variables (e.g., age, race, and sex)
- B. Family Influences
  - 1. Socioeconomic Influences
    - a. Current family condition (e.g., index of socioeconomic status)
    - b. Family background (e.g., family structure)
  - 2. Psychological Influences
    - a. Mother-child interaction (e.g., extent of physical control)
    - b. Mother's expectations and aspirations (e.g., how far would mother like to see child go in school)
    - c. Mother's behavior and attitude toward educational experience (e.g., attitude toward school)
    - d. Mother's attitude toward life in general
- C. Head Start Program Variables
  - 1. Social-psychological
    - a. Classroom (e.g., OSCI class factor scores)
    - b. Teacher (e.g., OSCI teacher factor scores)
  - 2. Teacher demographic (e.g., age, experience, training)
  - 3. Physical
    - a. Educational materials resources (e.g., learning games)
    - b. School plant resources (e.g., indoor floor space)
    - c. Class characteristics (e.g., size of class)

Table III.2

FINAL VARIABLES IN 1968-69 ANALYSIS

Name	Abbreviation
I. Outcome Variables (Raw Gain and Gain Scores Adjusted for Pre Scores)	
A. Stanford-Binet IQ Score	SB
B. Preschool Inventory Total Raw Score	PSI
C. Wechsler Preschool and Primary Scale of Intelligence--Total Raw Score for Animal House Subtest	AH-WPPSI
D. Behavior Problem Scale	BP
E. Motivation Problem Scale	MP
F. Feelings of Inadequacy Scale	FI
G. Gumpgookies	GUMP
II. Explanatory Variables	
A. Initial Child Status	
1. Ethnicity	ETH
2. Sex	SEX
3. Age	AGE
B. Family Influences	
4. Socioeconomic Status of Family	SES
5. Family Structure	
6. Mother's Employment Status	
7. Primary Source of Income	
8. Ratio of Rooms to People	
9. Accessibility of Adults	
10. Number of Moves in Past Three Years	

(continued)

Table III.2 (continued)

Name		Abbreviation
11.	Extent of Parent's Reading to Child	
12.	Extent of Physical Control	
13.	Mother's Aspirations for her Children	
14.	Mother's Expectations for her Children	
15.	Frequency of Parent Visits to Head Start Center	
16.	Extent of Parent Volunteering in Classroom	
17.	Parental Membership on Policy Council	
18.	Parent Opinion Scale	
C.	Head Start Program Variables	
19.	Observation of Substantive Curricular Input--Classroom [Structured Lessons]	OSCI C-I
20.	Observation of Substantive Curricular Input--Classroom [Group Activities and Routines]	OSCI C-II
21.	Observation of Substantive Curricular Input--Classroom [Social Emotional Interaction]	OSCI C-III
22.	Observation of Substantive Curricular Input--Classroom [Verbal Communication]	OSCI C-IV
23.	Observation of Substantive Curricular Input--Classroom [Instruction in Creative Arts]	OSCI C-V
24.	Observation of Substantive Curricular Input--Classroom [Language and Discrimination Learning]	OSCI C-IV
25.	Observation of Substantive Curricular Input--Teacher [Social-Emotional Interaction]	OSCI T-I

(continued)

Table III.2 (continued)

Name		Abbreviation
26.	Observation of Substantive Curricular Input--Teacher [Structured Lessons-Large Group]	OSCI T-II
27.	Observation of Substantive Curricular Input--Teacher [Art Activities]	OSCI T-III
28.	Observation of Substantive Curricular Input--Teacher [Creative Instruction-Small Group]	OSCI T-IV
29.	Observation of Substantive Curricular Input--Teacher [Routines]	OSCI T-V
30.	Observation of Substantive Curricular Input--Teacher [Receptive Learning]	OSCI T-VI
31.	Region	
32.	Previous Head Start Experience	
33.	Teacher's Age	
34.	Teacher's Ethnicity	
35.	Teacher's Years of Education	
36.	Teacher's Preparation in Early Childhood Education	
37.	Teacher's Residence in Head Start Neighborhood	
38.	Index of Educational and Physical Resources	
39.	Square Feet of Indoor Space	
40.	Square Feet of Outdoor Space	
41.	Overall Appearance of Classroom	
42.	Student Attendance	
43.	Pre-Post Interval	

2. Outcome Variables

Seven cognitive measures were originally proposed: the four subscores and the total score on the PSI, the Stanford-Binet IQ, and Animal-House Scale of the WPPSI. Three cognitive measures were selected: total score for PSI, Stanford-Binet IQ, and Animal House scale of the WPPSI. All outcome measures were based upon raw gains (post-pre) or (post-pre) adjusted for pre for a particular outcome variable. The adjustment is described later in this Chapter.

Justification: There has been much interest concerning the effects of various compensatory education programs on Stanford-Binet performance. There is evidence in the literature that certain Head Start and other compensatory education programs have been successful in bringing about significant Binet performance changes, e.g., Smith, (1968); Karnes et al., (1968); Karnes et al., (1970). Most recently, Sprigle (1971) demonstrated substantial and durable Binet gains for disadvantaged children in a Learning to Learn program.

The Preschool Inventory (PSI) (Caldwell, 1967) was developed to yield a measure of achievement for 3 to 6-year-old children in areas regarded as necessary for success in elementary school. It was also developed for use in demonstrating changes associated with interventions such as Head Start. Although the total raw score of the PSI is correlated with SB performance ($r = .45$), there is still a great deal of unique variation believed to represent acquired information and skills of value in later school performance. Some of these skills may be more modifiable by educational interventions than SB performance; Sontag et al. (1969), for example, conclude that the PSI might be quite susceptible to relatively large position shifts due to intervention programs.

The four subscores of the Preschool Inventory (PSI) were combined into a single total score because scale intercorrelations were high. The intercorrelations among the four PSI pre subscales and the Kuder-Richardson 20 reliabilities are shown in Table III.3.

Table III.3

INTERCORRELATIONS AND RELIABILITIES FOR THE PRE-PSI RAW SUBSCORES (N = 1642)

		1	2	3	4
Personal-Social Responsiveness	1	.78*	.63	.63	.66
Associative Vocabulary	2		.74	.55	.59
Concept Activation - Numerical	3			.71	.61
Concept Activation - Sensory	4				.81

* K-R 20 reliabilities in diagonal.

The intercorrelations and reliabilities of the four post PSI scores were of the same magnitude. The decision to combine the four subscores is in agreement with recommendations contained in the test manual provided by Educational Testing Service (1970).

Since raw scores rather than scaled scores were used, and since the PSI scores increase markedly with age, the absolute magnitude of intercorrelations may be somewhat spurious in that scaled score intercorrelations would probably be lower. Such analyses are not reported by ETS, and current norms are provisional, with age groupings of 6 months. Use of such norms could introduce substantial error of interpretation of gains for a shorter pre-post interval.

The decision to use total raw scores therefore is likely to lose the desirable differentiation of effects provided by the four subscores, while gaining reliability of change estimates since the pre-post interval is not involved indirectly in adding error variance as it would be if the 6 months' interval norms were used.

This is a problem in using many standardized tests. The typical 6 month or 1 year norm interval is probably insensitive to the rapid changes preschoolers may show and it is often longer than the actual pre/post testing interval. Testing usually begins a month after programs begin and ends several weeks before programs close; in a typical 9-10 month preschool project, many of the children may have been tested in month 2 or 3 and retested in month 7 or 8. Norms in 3 or 4 month intervals--if there were sufficiently large numbers of items for reliability within age categories--would increase the sensitivity of intervention studies such as this.

On the basis of analyses such as that shown in Table III.3, the final cognitive outcome variables were:

a) SB Performance

b) Total Raw Score of the Preschool Inventory (PSI)

c) Total Raw Score for the Animal House Scale of the Wechsler Preschool and Primary Scale of Intelligence (AH-WPPSI).

Five Behavior Rating outcome variables were originally proposed. A decision was made not to derive Birch Performance Work Scores and Birch Verbal Work Scores because of time and cost constraints in scoring from the 99,354 card images. The three behavior ratings that were used were derived from the Inventory of Factors Affecting the Stanford-Binet. These three scales, derived by examination as rational scales, were:

a) Behavior Problem Scale (BP), which had a score ranging from 0-7 and was derived by counting the tester's yes descriptions of the child's behavior as (1) easily distracted, (2) impulsive, (3) antagonistic, (4) becomes hostile, (5) belligerent, rebellious, (6) hyperactive, and (7) verbose;

b) Motivation Problem Scale (MP), which had a score ranging from 0 to 10 and was derived by counting the tester's yes descriptions of the child's behavior as (1) vaguely inattentive, (2) lacks concern with competence, (3) slow to respond, (4) apathetic, (5) gives up easily, (6) withdrawn, (7) unresponsive apathetic, (8) indifferent to praise, (9) hypoactive, and (10) tactiturn;

c) Feelings of Inadequacy Scale (FI), which had a score ranging from 0 to 4 was derived by counting the tester's yes descriptions of the child's behavior as (1) distrusts own ability, (2) fearful, guarded, (3) shy, reticent, reserved, and (4) needs constant praise.

Kuder-Richardson 20 reliabilities, intercorrelations, means, and ranges were computed for these three variables on the pre administration and are shown in Table III.4. Standard deviations are not presented because the range is a more appropriate index of variability when scales are extremely skewed.

Table III.4

INTERCORRELATIONS, RELIABILITIES, MEANS AND RANGES FOR THE THREE BEHAVIOR SCALES (N = 1651)

Scale		1	2	3	Mean	Score Range
BP	1	.70*	.09	-.05	.68	0- 7
MP	2		.78	.53	1.58	0-10
FI	3			.74	.72	0- 4

* K-R 20 reliabilities in diagonal.

As can be seen from Table III.4, all three scales had satisfactory reliabilities for short instruments. The item intercorrelations within each scale were in the .30s, .40s, and .50s. BP was relatively independent of the other two scales while MP and FI had a moderate correlation of .53. All three scales were skewed in that most children did not exhibit any (0) undesirable behaviors at time of initial testing (See Table VI.3). These three scales were retained as outcome measures reflecting Behavior Ratings.

Justification: The inclusion of a Behavior Problem Scale (BP) is supported by the work of Kagan et al. (1966), which indicates that impulsivity-reflectivity as a "cognitive style" variable is strongly related to achievement and perhaps other process variables. A reduction in the extent of non-goal-directed hyperactive behavior is one objective of typical preschool programs.

The Motivation Problem Scale (MP) attempts to assess the motivation of the child. White's (1965) research indicates that the ability to pay attention is a major component of task performance.

The Feelings of Inadequacy Scale (FI) attempts to measure feelings of inadequacy or self-concept. The development of a healthy self-concept and a feeling of competence is crucial for success and happiness in the educational and occupational world. Cegelka and Thomas (1968) found that disadvantaged children have an unfavorable self-concept, and Passow (1970) gives a profile of various characteristics of disadvantaged children. He characterizes the disadvantaged child as having an orientation towards life that seeks immediate gratification rather than ability to delay for future advantage, an unfavorable self image, modest aspiration, low motivation to achieve academic success, restricted attention span, and a general inability to cope with demands and expectations of school programs. Deficits in the above characteristics would be expected to depress cognitive performance and success in school but more importantly reflect a child that is likely to be unable to cope effectively with life situations.

Others (e.g., Shipman, 1970) have reported that disadvantaged children have a positive self-concept and display fewer personal-social problems than have previously been identified before entering school.

The experimental measure of achievement motivation, the Gumpgookies (GUMP), was retained as a final variable. Although the original pre-GUMP was comprised of 100 items, only the 55 items that were retained for the post test were used in scoring the pre measure. The K-R 20 reliability for the pre-GUMP was .76 while the post-GUMP reliability was .86. The lower reliability of the pretest might be due to its being too lengthy for preschool children. The pre-GUMP had a mean of 34.40 and a standard deviation of 7.0; the total possible score was 55, and scores could range from 0 to 55.

3. Explanatory Variables

a. Child Variables

The following initial child demographic variables were retained as final variables: (1) ethnicity (Black, White, Mexican-American, or Polynesian), (2) sex, and (3) age at time of initial testing,

which was subdivided into the following four categories: 31-47 months, 48-53 months, 54-59 months, and 60-72 months. The variable first born versus other, which was originally proposed, would have been difficult to derive from the available data and was dropped from the final list of variables.

b. Family Variables

(1) Socioeconomic Influences

The inclusion of an SES index and other indicators of the child's quality of life is supported by the consistent relationships reported in the literature between these variables and child development. For example, Mumbauer and Miller (1970) found a significant association between socioeconomic background and cognitive functioning in preschool children. Variables such as family structure, number of moves, etc., are indices of family stability. Shipman (1970) hypothesizes that such family conditions serve to constrict the child's psychological environment and thus create a stressful situation. Green, Hofmann, and Morgan (1967) present evidence that the family variables selected for this study are consistently related to children's intelligence and achievement. Ruben and Womack (1968) report significant relations among ethnic and family variables such as family size, parent's education, parent's occupation and IQ in a sample of disadvantaged school children. This is an especially interesting finding when one considers the relatively homogeneous population in which those relationships were obtained.

There were some major changes in the original list of measures of family influences and current family condition. An index of socioeconomic status (SES) was derived from the following four original variables: (a) presence of telephone in the home, (b) family income, (c) mother's education, and (d) child having own bed. In order to construct this index, the marginal distribution for each of these variables was examined. Enough variation was present for each of the four indicators to contribute to the total variation of the index. The index was constructed by giving a child's family a score of 1 upon satisfying each of the following conditions: telephone in home, family income equal to or above $4,000, mother's education--high school graduate or better, and child has own bed. This resulted in five categories for SES ranging from families satisfying none of the four conditions to families satisfying all four conditions. The number of families falling into each of the five categories on the 1968-69 Parent Interview Pre is shown in Table III.5 below.

Table III.5

DISTRIBUTION OF 1968-69 SAMPLE FAMILIES ACROSS FIVE SES CATEGORIES

Number of Socioeconomic Conditions Satisfied		Number of Families*	
0	Low-Low SES	70	259
1		189	
2	Middle-Low SES	293	
3	Upper-Low SES	210	301
4		91	

* There were 527 families for which an index could not be computed because of missing data. The problem of missing data will be discussed in a later section of this report.

The distribution listed in Table III.5 was fairly normal, but it was decided to categorize this distribution into three SES categories: Low-Low SES (0 and 1 combined), Middle-Low SES (2), and Upper-Low SES (3 and 4 combined). This categorization resulted in N's of 259, 293, and 301 respectively. The utility of the constructed SES Scale is demonstrated in a later section of the report which shows that the means for the cognitive measures on the children's pretests varied significantly and in the expected direction across the three SES categories.

It should be emphasized that the scale was constructed to differentiate within families who are among the most economically depressed, those whose life circumstances are extremely constricted in terms of material possessions and those whose circumstances permit such necessities for our national standard of living as a child having his own bed to sleep in. The correlation of child performance with socioeconomic status within the families eligible for Head Start confirms often-reported debilitating effects of severe privation on the child's development, and indirectly validates the scale as a meaningful index of SES as it affects development.

Information concerning the husband's job and education was not used to reflect family background because only 54% of the families had fathers. On the basis of tabulations, three categories of family structure were defined and retained for analysis: (a) mother only, (b) mother plus father, and (c) mother plus father plus adult relative.

Four categories of mother's employment were considered: (a) not working, not looking; (b) not working, but looking; (c) working, part time; and (d) working, full time.

There were three major categories for primary source of income: (a) mother's earnings, (b) father's earnings, and (c) welfare. The majority of the observations fell into these three categories.

Ratio of rooms to people was retained as an additional quality of life indicator. Accessibility of adults, which was not on the original list, was added to the final list. Accessibility of adults was measured by the ratio of adults to children in the family. The variable, the number of children in the home from 0 to 18, was dropped in the final list of variables because it was found to be redundant to a large degree with whether the child has own bed, ratio of rooms to people, and accessibility of adults. Frequency of getting newspaper was dropped because it was judged to be redundant.

The number of moves in the last three years was retained as a final quality-of-life variable. The information regarding the migration of the family and a rating of the cleanliness and neatness of the home were dropped due to restricted marginals.

(2) Psychological Influences

The manner in which the mother regulates and controls the child's behavior has been extensively investigated by Hess and Shipmann (1965, 1968a). Parental control techniques also have been found by Moustakas et al. (1956) to relate to a variety of cognitive behaviors in the child, and they have been found to predict whether a child exhibits impulsive or reflective behavior in a problem-solving situation.

The variable defined as extent of parent's reading to the child and its relationship to the various outcome measures were of great interest since amount of parental reading has been found to be associated with children's verbal and academic achievement (e.g., Bing, 1963). As will be seen in a subsequent section of this report, the frequency distribution of time spent reading to children may suggest that socially desirable responses may be influencing parent's responses to this question as well as questions similar to this. A relatively large number of parents indicated that they spent much time reading to their child.

The extent to which the mother physically controlled the child showed enough variation among mothers to consider for the final list of variables. The mothers fell primarily into three categories: (a) no physical punishment, (b) scolding, and (c) mild physical punishment. Only a few mothers reported using severe physical punishment. The extent to which the mother reported using rejection and guilt as a means of psychological control showed little variation among mothers. Most mothers were rated as using no psychological control through rejection or guilt. Reaction of the mother to the child's minor infractions demonstrated enough variability to be considered in the final list of variables. Many mothers indicated that they would respond in a constructive manner.

Adult/child influence techniques did not show enough variation to be considered in the final list of explanatory variables. Most of the mothers fell into various categories of unqualified power assertion.

Within the income category "poor," there are individual differences in adult personal styles and attitudes which cannot help affect the child's development. With some oversimplification, internal locus of control seems broadly to account for the family whose life seems to family members most meaningful and worthwhile, and whose children are among the many who do very well. Internal locus-of-control refers to the sense of being able to make a difference, to affect what one's life is like, of internal self-determination. In contrast to internal locus of control is external locus of control, the sense of powerlessness, futility, that what one does really has little effect on one's own life or anything else. Internal locus-of-control may have a great deal in common with the sense of command that is associated with black power and other movements asserting the dignity, worth, and power of a group or individual.

A number of aspects thought related to locus-of-control were assessed in the 1968–69 parent interviews: maternal aspirations and expectations for their children's educational and job attainment, attitudes toward the value of education and educational institutions, and the Srole alienation (Parent Opinion) scale. Parental participation in Head Start is also considered in this area.

Mother's expectations and aspirations for child's education and occupation are important determinants of the child's behavior. Gervasi (1969) and Hess and Shipman (1968b) found that the discrepancy between aspirations and expectations tended to increase as social status decreased. These discrepancies indirectly measure the parents' feeling of influence in the child's life.

Three categories for mother's aspirations for her children were constructed by combining information on mother's aspiration of child's eventual job and education. The "low" category was comprised of mothers who had <u>both</u> low job and educational aspirations for their children (less

than post high school education and less than an administrator of a medium sized business), a middle category included mothers who scored high in either job or educational aspirations, and a high category included mothers who scored high on both job and educational aspirations. Similarly, three categories of mother's expectations for her child's jobs and education were constructed.

Three originally proposed scores from the First Day of School items were not derived because of the computational difficulties and because of the restricted range of raw scores.

The Hess-Shipman Educational Attitude Survey, composed of 23 items, was included in both the pre and post Parent Interview. The items did not appear to measure a unidimensional attitude, but seemed to reflect the following dimensions: (a) attitude toward value of education in general (six items), (b) attitude toward teachers (four items), (c) attitude toward school (five items), and (d) feelings of control over school (two items). Each of the four scales was constructed by giving 1 point for each attitude statement that was responded to in a favorable direction. The intercorrelations among these four scales, K-R 20 reliabilities, means, and standard deviations are presented below in Table III.6.

The reliabilities of these ad hoc attitude scales were low. They did not correlate significantly with other continuously scaled outcome or

Table III.6

INTERCORRELATIONS, RELIABILITIES, MEANS, AND STANDARD DEVIATIONS (SD) FOR THE FOUR EDUCATIONAL ATTITUDE SCALES (N = 1433)

Scale		1	2	3	4	Mean	SD	Score Range
Attitude Toward Value of Education in General	1	.38*	.24	.08	.02	3.49	1.34	0-6
Attitude Toward Teachers	2		.04	.21	.00	3.15	.98	0-4
Attitude Toward School	3			.32	-.05	2.41	1.03	0-5
Feelings of Control Over School	4				.37	1.16	.76	0-2

* K-R 20 reliabilities in diagonal.

explanatory variables. Consequently, these four scales were not considered for further analyses. It should be kept in mind that factor-analytically derived scores from the same or similar items have shown to be correlates of children's behavior in previous studies (Hess, 1971). If the items themselves had been scaled differently, and if the subscores had been based on factor analysis, the information available in these items might have been more useful.

With regard to Head Start participation, there was enough variation in the number of times that the mother went to the Head Start Center and the extent to which the mother volunteered for classroom duties for them to be considered as explanatory variables. The variable concerning whether or not the mother was a member of a parent council was also retained, although about 80% of the mothers were not members of a parent council. Information regarding attendance at parent meetings was not considered further because the dimensions of parent participation seemed to be covered fairly well by the first three parent participation variables. Parent participation is hypothesized to relate to the performance of children in Head Start. Wilmon (1969) found that parental involvement in Head Start influences academic achievement through the parents being able to transfer educational aspirations and the need for achievement to the child.

The last family variable to be considered was a Parent Opinion Scale. This scale (Srole) was a linear combination of five Likert type items and had a range of 5 to 25. In the interview, the high end of the scale reflected extreme pessimism. An example of an item from this scale is: "It's hardly fair to bring children into the world with the way things look for the future." This scale has been reflected for ease of interpretation so that a high score indicates optimism. This scale can be viewed as a measure of the mother's alienation. Shipman (1970) hypothesizes that parental alienation will arrest child development and lead to inconsistency of child-rearing practices. Slaughter (1968) found that the degree of social isolation of the mother was a significant correlate of the child's school achievement. Radin and Karni (1965) report that disadvantaged Negro mothers differed in feelings of alienation as well as in other important attitudes from middle class mothers.

c. Program Variables

(1) Social-Psychological Influences

The social-psychological classroom variables are composed of the six OSCI Class Factor Scores discussed earlier in the report. For descriptive and analytic reasons, each of these six factor score distributions was divided into three categories reflecting low, medium, and high classroom emphasis on the particular factor. Classes were considered low or high on a dimension if they fell respectively

into the lower 25% or upper 25% of the total factor score distribution. Medium emphasis was defined as the middle 50% of the factor score distribution.

The social-psychological teacher variables were measured by the six OSCI Teacher Factor Scores and two rating scales derived from the Post Observation Teacher Rating Scales. The six OSCI Teacher Factor Scores were handled in the same manner as the six Class Factor Scores i.e., scores for each distribution were grouped into low, medium, or high emphasis categories.

The six OSCI Class Factor Scores and six OSCI Teacher Factor Scores are derived from a promising new approach and seem to be measuring program inputs that would be expected to relate to various program outputs. Stern (1968) supports the need for an instrument such as the OSCI by summarizing much evidence that some type of classroom observation must be used to make curricular comparisons because the correspondence between self-reports of program strategies and actual teacher behavior is often quite low.

A question that might be asked of the data is whether emphasis on class structured lessons is associated with gains in the cognitive areas (i.e., Stanford-Binet and Preschool Inventory). Also, the relationship between Socio-Emotional Interaction in the classroom and changes on various behavioral ratings would be of interest. The variation in each of the six OSCI Class Factor Score distributions should represent some of the major differences between various Head Start classroom strategies.

Karnes et al. (1970) using standardized tests compared a rote learning highly structured program, an individually planned diagnostic curriculum, an SES mixed preschool, and a "regular" preschool. For the disadvantaged preschoolers, the evaluation favored the more highly structured and diagnostic programs. Studies by Sprigle (1971) suggest that a different kind of "structure," one directed to learning-how-to-learn is more effective than traditional nursery school oriented programs. The literature suggests that programs are differentially effective in bringing about various cognitive and non-cognitive changes in disadvantaged children. A variety of approaches which share an emphasis on planning, on language development and on cognitive stimulation is likely to be more effective. There are important exceptions that show the need to think carefully before ascribing greater effectiveness to content factors per se: Weikart et al. (1970) for example, found the "traditional" curriculum implemented under his direction was quite effective, comparing favorably with his own curriculum model.

As described previously, each Post Observation Teacher Rating Scale was completed up to five times for the head teacher or teacher in control of the classroom on the five days when the OSCI observations were obtained. If more than one teacher was rated throughout the five ratings, then the

teacher who was present for the majority of the time was selected for analysis. Most of the ratings were based upon three, four, or five rating occasions. An examination of the 33 items resulted in the development of two rationally constructed scales. The Teacher Stimulation Scale was comprised of 13 items which reflected the extent to which the children received teacher attention, advice, and support. The remaining scale, Teacher Sensitivity to Individual Differences, included seven items measuring the teacher's awareness of the children's individual needs.

A score for each scale was derived by averaging the available ratings across occasions for each teacher: some teacher's ratings were based on an average of five ratings and other teacher's ratings were based on three or four ratings.

For descriptive and analytical purposes the score distribution for each scale was divided into three categories. The low category comprised the lower 25% of the distribution, the middle category comprised the middle 50% of the distribution and the high category comprised the upper 25% of the distribution.

Three originally proposed OSCI scores (teacher involvement, locus of control (child or teacher) and teacher aide involvement) involved scanning a large OSCI raw score data file of approximately 29,000 card images. Since 12 OSCI factor scores were already represented in the data base, it was decided that the three variables could be dropped from the final analyses. It should be re-emphasized at this point that the decision not to consider a variable for analysis does not mean that a variable was not regarded as important. Future research based on the available data base could well involve many of the variables that were not analyzed in the present report.

Whether or not the teacher would be willing to participate in similar evaluation programs the following year (a measure of teacher morale) was dropped as an explanatory variable since most teachers responded in the affirmative.

Teachers' experience and preparation have been found to be positively associated with various school output measures in many elementary school studies as reviewed in the Office of Education's "Do Teachers Make a Difference?" (1970). Goodman (1959) found in an earlier study of older students that teacher experience and classroom atmosphere were significantly linked in a positive manner to pupil performance.

(2) Teacher Demographic Influences

A number of head teacher demographic variables were retained on the final list. These were teacher's age, ethnicity, formal

education, formal training in child development, and whether the teacher came from the Head Start neighborhood. Since all but one or two teachers were female and most teachers had little or no experience with preschool or disadvantaged preschool children prior to employment with Head Start, sex and non-Head Start experience were dropped from further consideration. Since the ratio of the number of professional plus paraprofessional staff to the number of children in the class was expected to be associated with other derived indexes measuring the richness of the educational resources and physical facilities of the class, staff ratios were not selected for more intensive analysis.

(3) Physical Influences

The last category of variables involves the physical resources of the Head Start classes. Information from eleven resource variables was combined to yield an index of educational and physical resources. The index was constructed by assigning to each class a point for the presence of each of the following items: slides, swings, sandboxes, dramatic play clothing, puppets, pets, waterplay equipment, learning games, movie slides, books, and science equipment. Scores ranged from 0 to 11. The scale, reflecting substantial variability in resources across the classes, was composed of items that were originally designed to yield two resource indexes.

Three other resource items were retained from the original list. They were square feet of indoor floor space, square feet of outdoor space, and overall appearance. An index earlier proposed to assess classroom physical characteristics was dropped because of the extensive coverage of the preceding resource information.

Administrative factors in the classroom were considered. One variable not on the original list was included in the final list to reflect the amount of exposure that the child had to the Head Start program. This variable was attendance as reported by the teacher. Three levels of attendance (low, medium, and high) were considered. The low category was defined as 90 days or less, the middle category was defined as 91-159 days, and the high attendance category was defined as 160 or more days. Most of the children fell into the middle attendance category. This replaced the originally considered variable, length of class days.

The Administrative Checklist was reduced to a tally of the number of program events that should have occurred but did not. Analyses by program organization (e.g., when someone was responsible for an event, how frequently and in what areas did teachers, parents, Head Start directors, and others take responsibility) were contemplated but not undertaken. This variable was not analyzed as a final variable although the marginal distribution indicated variability across classes on this measure.

The preceding discussion has summarized the considerations and decisions leading to the final list of the 68-69 variables shown in Table III.2. The next section of this report will discuss the variables selected from the more limited 1967-68 data file to be considered for analysis.

B. Variables Considered for 1967-68 Analyses

Some of the same considerations and decision rules used to select the 1968-69 variables were used to select the 1967-68 variables. An additional consideration was that the variables selected should measure constructs similar to the 1968-69 variables so that certain 1968-69 results could be checked on the 1967-68 data. Because of time and cost considerations, only a few explanatory and outcome variables were considered for analysis. The Kansas Social Interaction Observation Procedures (SIO) described earlier was an interesting instrument that was unique to the 1967-68 data; consequently, a few measures were derived from this instrument.

The variables were categorized according to the framework of Table III.1 used in conceptualizing the 1968-69 variables. The 1967-68 variables are presented in Table III.7.

Table III.7

THE 1967-68 EXPLANATORY AND OUTCOME VARIABLES

I. Outcome Variables (Raw Gains)

A. Cognitive Behavior

1. Stanford-Binet Intelligence (same as 1968-69)

B. Behavior Ratings

2. Behavior Problem Scale (same as 1968-69)
3. Motivation Problem Scale (same as 1968-69)
4. Feelings of Inadequacy Scale (same as 1968-69)
5. Number of initiations of interactions from child to adult from the SIO (unique to 1967-68)
6. Percentage of adult-initiated interactions responded to by the child (unique to 1967-68)

C. Achievement Motivation--no measures available for 1967-68.

II. Explanatory Variables

A.1. Initial Child Status--Demographic Variables

(continued)

Table III.7 (continued)

7. Ethnicity (primarily blacks and whites)
8. Sex
9. Age (same four categories as 1968-69)

B.1 Family Influences--Socioeconomic Influences

10. Total number of people in home (four categories--2,3, or 4; 5 or 6; 7 or 8; and 9 or more)

11. Education of mother (two categories--less than 12th grade versus 12th grade or better)

B.2 Family Influences--Psychological Influences.

12. Pre-mother's aspirations for child's education (two categories--go to college versus not go to college)

13. Pre-mother's expectations for child's education (two categories--go to college versus not go to college)

14. Joint aspiration-expectation mix (three categories--high aspirations-high expectations; high aspirations-moderate expectations; moderate aspirations-moderate expectations)

15. Parent's reactions to child's misbehaving through jumping and screaming from the Adult-Child Influence Techniques Instrument (four broad categories of influence--nonintervention, intervention, qualified power assertion, and unqualified power assertion--same as 1968-69 instrument and categories)

16. Attendance at parent meetings for those centers having parent meetings (two categories--yes or no)

C.1.a. Head Start Program Variables--Social Psychological Classroom

17. OSCI Class Factor Score I - Cognitive--Low Structure as defined previously.

18. OSCI Class Factor Score II - Routine and Rules as defined previously

19. OSCI Class Factor Score III - Cognitive--High Structure as defined previously

20. OSCI Class Factor Score IV - Child Centered--Unstructured as defined previously.

(continued)

Table III.7 (continued)

C.1.b. Head Start Programs--Social Psycholgical--Teacher Variables

21. The percentage of times that an adult in the classroom responded to the children's initiation of interactions from the SIO (NI)

C.2 Head Start Programs--Teacher Demographic Variables

22. Level of teacher's education (elementary, high school, bachelor's degree, master's degree, doctorate)

23. Teacher's paid experience with preschool children (under 6 months, 6 months-1 year, 1-3 years, 4-5 years, over 5 years)

24. Teacher's paid experience with disadvantaged children (under 6 months, 6 months-1 year, 1-3 years, 4-5 years, over 5 years)

C.3.b Head Start Programs--Physical--School Plant Resources

25. Inside square feet per child (less than 14, 15-19, 20-24)

26. Outside square feet per child (less than 39, 40-49)

C. Some Analytical and Methodological Considerations

The research designs for the national evaluations are (1) quasi-experimental for 1968-69 and (2) non-experimental for 1967-68. This usually means that inferences regarding the causes of differences, if any, among classes must be drawn cautiously. The ability to replicate findings across years greatly strengthens inferences, however. Replication across sites for the two language development/preacademic readiness approaches in 1968-69 further strengthens inferences regarding these program effects. Replication of program effects across classes in 1968-69 and across years in 1967-68 versus 1968-69, therefore, would present a reasonably firm finding. Where the findings are similar across these Head Start evaluations, Planned Variation and other experimental or quasi-experimental projects, the reliability of the effect would be substantial.

In the absence of control groups, inferences regarding the causes of whatever changes are observed from pre to post are usually made cautiously if at all. Comparison of the magnitude of Head Start study changes with those reported for control groups in other studies (test/retest of non-Head Start controls); for social adjustment and motivation control groups in other studies (e.g., Zigler and Butterfield, 1968);

and for "traditional" and experimental samples in other studies can provide some basis for estimating the extent to which observed change is a Head Start effect.

In addition to limitations on direct inference imposed by the research design,[1/] there are (1) limitations associated with possible non-uniformity of data collected in varying field conditions and under varying systems of quality control, (2) limitations in the sensitivity, reliability and validity of the measures; and (3) limitations in statistical models for estimating main and interaction effects when (a) initial levels of performance differ and (b) some cells in N-factorial ANOVAs are small or empty.

With regard to the quality and uniformity of the data, while they undoubtedly vary to some extent across classes and across E & R Centers, substantial effort went in to ensuring carefully, uniformly collected data.

Among the sources of non-random variation are:

1) Testing conditions, which varied from converted broom closets and front porches to specially equipped air conditioned test vans.

2) Tester skill, which varied from novice to highly experienced.

3) Tester selection, training and quality control procedures. However, all sites had experienced full-time evaluation coordinators who developed and carried out plans for controlling as far as possible testing quality and data accuracy.

4) Delays in communicating decisions on data collection and coding made during the study to all E & R coordinators added "noise" to the data.

1/ Head Start E & R directors and staff did not design a simple random sample study with H/S vs non-Head Start control groups for several reasons. First, true controls are not possible in many real world situations. In rural areas there may be too few eligible children even in a "neighborhood" of several hundred square miles. In urban areas, there is no way to prevent "controls" from participating in the variety of other preschool experiences available. Second, it was felt that more would be learned about what kinds of programs have what effects from selected variations or interventions and that this question, rather than assessment of average program effects, would be more valuable for program planning.

In 1967-68 the Head Start National Evaluation Coordinator was responsible for communications among the 14 Centers;[2] in 1968-69, the Educational Research and Evaluation Unit (EREU) of the Washington School of Psychiatry [3] was responsible for coordinating test and coding manual development and use, for training site coordinators and staff on parent interview procedures, and for monitoring both the intervention studies and data quality. Reports prepared by EREU are on file at the national Head Start Office. For example, reports on the post testing conditions in 1968-69 indicated that accessibility to the classroom, working space, working surface, ventilation, lighting, temperature, and cleanness were optimal or good for about 90 percent of the children. Noise was the greatest problem; the majority of situations were judged to be fair but acceptable. On an overall rating scale, about 45 percent of the conditions were rated as optimal, about 45 percent as good, about 6 percent as fair but adequate and the remainder were judged to be poor.[4] On the basis of these reports, the monitoring efforts of the experienced[5] EREU staff and the equally experienced E & R evaluation coordinators and staff, many of whom were with the project for all three years, the quality of the 1968-69 data is likely to be high for a project of this size.

Coding, keypunching and verifying were done at the E & R Centers: tapes were reverified for code legality and completeness by the OEO Information Center (Mrs. Jane Lee). The complete 1966-69 data bank was edited, checked, reformatted where necessary and prepared for processing by World Systems, Inc. (Contract HEW-OS-69-113.)

2/ The University of Chicago (Dr. V. Shipman) and Teachers' College (Dr. R. Thorndike) participated in the 1966-67 and 1967-68 studies only.

3/ Dr. Russell H. Cort, Dr. Ann O'Keefe, Mrs. Naomi Henderson and Miss Margaret Mathis were responsible for virtually all manuals, codebooks, training and monitoring of the 1968-69 study.

4/ Centers responding were South Carolina, Kansas, Southern, Hawaii, Texas and Tulane.

5/ The EREU researchers were responsible for the 1965 and 1966 Planning Research Corporation National Head Start evaluations as well as other similar projects.

This summary reflects only a portion of data available from the individual site reports prepared by each tester, the evaluation coordinator reports, and the files at EREU, at the Head Start National Office, and at the E & R Centers. Analyses of relationships among tester age, sex, and ethnic characteristics, tester experience and training, testing conditions and performance could form a large work in itself. For this report, the material was scanned and some analyses run by E & R Centers and sites. In general, small-scale studies with very tightly controlled testing would have less variance associated with factors described above; however, there is no reason to expect that major effects or signals will be lost in testing "noise" in these E & R data and for 1968-69, there is considerable reason to expect little variation due to the way in which data were collected and coded.[6/]

The characteristics of the measures have already been discussed. It is, however, appropriate to emphasize that no measures currently available are considered really satisfactory for assessing the full capabilities and characteristics of low-income preschool children, particularly from such diverse cultural backgrounds. Evidence of test-retest reliability is sparse, and of construct or predictive validity, even sparser. While the data are labeled in this report as constructs (e.g., "IQ," "parental optimism," "emotional-social emphasis") it is important to remember these labels represent approximate statements. Carefully as the tests were selected and new measures developed, they capture only fragments of the child, his family, and his experiences, and fragments that do not represent the child as seen by his teachers, parents and friends. The frustration of researchers who are sensitive to the tremendous complexity of the child's experience has been eloquently stated by many. Perhaps few aspects of this report are as sobering as this. The E & R evaluation data are among the most complex, rich and varied ever collected in a national evaluation of an intervention program--and they are about as adequate to describe a child as height, mass, average daily temperature and wind velocity are to describe Mt. Everest in contrast to Mt. Fujii.

6/ All open-end questions were coded under the supervision of Dr. Virginia Shipman, Dr. Robert Hess, Dr. Irving Sigel or Dr. Bela Fehrer. Rater rates reliabilities are .85 or higher, and some measures were entirely doubled coded by independent raters.

With regard to limitations in the statistical models, there are many problems associated with analyzing data of this type. The data are so highly confounded that in many cases it is difficult to estimate the effects of independent variables. In analyzing highly and complexly confounded non-experimental data such as these, the lack of control on other relevant variables means causal inferences should be made cautiously. The problems of pre and post measurement designs and the associated analysis problems of the various types of gain scores also add further to the methodological difficulties. Of particular concern in this regard is the problem of separating regression phenomena, usually attributed to random error in initial measurement which reduces the reliability of the initial score as an index of the child's status, from true changes. Regression as a phenomenon means that scores at both high and low extremes tend to move toward the average score when the child is retested. A child with a score above the mean for the test-as-a-whole is likely to score lower on retesting while a child with a low score is likely to do better.

The psychological dynamics of regression are not well known. There is no really satisfactory way of distinguishing "true" regression and "true" change in a study such as the present one. Covariance adjustments tend to reduce "true" gains or losses; simple gain scores are unreliable since the highest scores come from the extremes where initial measurement is most suspect. Where groups differ on initial score, covariance adjustment, simple gain, and ANOVA often are used to present the consequences of different assumptions about regression and "true" change. In the literature, raw gain and covariance adjustments are both used when groups differ initially. If groups do not differ initially, a simple comparison of final levels is usually considered satisfactory.

With respect to the sequence of analyses, the general strategy involved pointing out the highlights primarily for the 1968-69 data and secondarily for the 1967-68 data. The strategy involved the following steps:

1) The definition and derivation of relevant input and outcome variables--The variance of each derived variable was compared with the largest latent root obtained through a principal components analysis of the item covariance matrix to ensure that the derived variables accounted for a reasonable proportion of variance in the data. These matrices are on file at RTI.

2) The development of information retrieval systems and associated computer software--This was necessary to link child, family, teacher, and program information and create various working tapes for the data analyses from the by-instrument OCD data bank tape prepared by World Systems, Inc.

3) Deleting input and outcome variables on the basis of their univariate frequency distributions--Those variables with a substantial percent of "no data" were omitted along with variables which showed little or no variability. Continuous independent variables were redefined as categorical variables. This procedure eliminated the issue of nonlinear relationships between the original continuous variables and the dependent variables.

4) Conducting univariate analyses relating child-family background and program variables to the prescores for the outcome measures--The analytical method used for this purpose was analysis of variance (ANOVA). ANOVA was performed on each outcome measure for the different levels of the categorical independent variables considered as treatments. The mean square error between levels of the factor was then compared to the mean square error within levels of the factor. The significance of the resulting F statistic was then tested.

5) Examining the confounding among child-family variables and program factors--Interrelationships among the input factors were examined by cross tabulating various child-family background variables against various program variables.

6) Examining the univariate relationships of child-family and program input variables to raw gains on the output measures and raw gains corrected for the prescore--The raw gain for an outcome measure for a given child is the difference between the post and pre score for the outcome measure for that child. The univariate relationships between the raw gain scores and the child-family and program input variables were studied using analysis of variance techniques. The adjusted gain score was defined as the difference between raw gain score and the regression estimate of the gain score using the prescore as the regressor. In other words, the adjusted gain score is that part of the gain which is not accounted for by the prescore. The analytic strategy used for studying the relationships between adjusted gain scores and input factors was the same as that used for the analysis of gain scores.

7) Study of models which examined the joint influence of child-family background variables, program variables and their interactions on the outcome measures--Two approaches have been followed here. The first one involved analyses of variance on nonorthogonal unbalanced designs with four main effects and some of the interesting first order interactions between the factors. The second approach was to examine the multiple regressions of gain scores on the twelve OSCI factors and the pre measure for the outcome measures.

Considering the time and cost constraints and the amount of data available, it was felt that more refined analyses were not warranted. Strong effects in the data should be isolated on the basis of any reasonable analytic technique.

D. Methodological Note on Missing Data

The problem of missing data occurs in most research efforts which involve collection of data on several variables related to an observational unit. There are computational and statistical problems in handling such data. Simply dropping all subjects for whom any datum is missing is often unacceptable. Replacing missing data with the mean value of the overall group on that variable reduces the sensitivity due to diluting the possible between group differences with the weight of the overall mean. In addition to computational problems, the reasons why data are missing could show a biased sample, so that conclusions drawn on the basis of available data are not representative of the original sample and possibly therefore not to the generalization population.

The Head Start data involves missing instances within units and also missing units as a whole for some analyses. Among the reasons for missing data within units are:

1) the child was untestable on some or all measures;

2) the child was not in the class even after the "make-up" day most E & R centers scheduled;

3) tester, keypunch, or other processing error;

4) the child withdrew before the testing was complete;

5) the mother was not available for an interview (or refused) even after three call-backs;

6) the mother lived in a difficult or unsafe area and could not come to the center for interviews;

7) language problems;

8) refusal to answer some items;

9) inadequate records or incomplete returns;

10) accurate reflection of family structure (e.g., no data for father in some households).

The primary reason for missing data from initial to final periods was that 18 percent of the children who had a pre SB did not have a post SB. Table III.8 summarizes what we know about some of the reasons for missing SB data in this study for the 1968-69 data.

Table III.8

SOME REASONS FOR MISSING DATA

Situation	Percent Missing
Child untestable (SB pre)	2
Child untestable (SB post)	1
Child enrolled but unavailable (SB pre)	5
Child enrolled but unavailable (SB post)	6
Child withdrew before testing completed (SB pre)	2
Children who had a pre SB but did not have a post SB	18

In Tables III.9 and III.10 information on missing data is presented on outcome measures and family background variables for all children originally included in the samples. This includes data missing for any reason, and these are data missing for one variable only. The available sample is further reduced for any analysis involving more than one variable. For example, in 1968-69, there were 1756 children with pre and 1518 children with post SB scores; but there were only 1443 children with pre and post SB scores. Some children were tested in the Spring only, if they were enrolled but unavailable in the Fall or if they had enrolled after the Fall testing period but still early in the year. Each analysis was computed on the largest possible N with all data required for that comparison. No attempt was made to impute missing values.

The effects of the biases noted were not tested except for losses due to withdrawal or late entry of the child. Using the SB as a criterion for effect of the bias (Table III.11), no effects were noted for initial scores (pre only vs. pre and post) or for final scores (post only vs. pre and post) for the 1968-69 data. We concluded that whatever bias may be introduced due to withdrawal from the class, the effects are not discernible on the SB.

Table III.9

MISSING DATA ON OUTCOME MEASURES AND CHILD-FAMILY BACKGROUND VARIABLES

(1968-69)

Variable Name	Percentage of Children With Missing Data
1. Outcome Measures	
SB - Pre	10.8
SB - Post	22.9
AH-WPPSI - Pre	12.4
AH-WPPSI - Post	22.9
Personal Social Responsiveness - Pre	16.4
Personal Social Responsiveness - Post	23.8
Associative Vocabulary - Pre	16.4
Associative Vocabulary - Post	24.0
Concept Activation Numerical - Pre	16.5
Concept Activation Numerical - Post	24.1
Concept Activation Sensory - Pre	16.6
Concept Activation Sensory - Post	24.1
Behavior Problem Scale - Pre	16.2
Behavior Problem Scale - Post	23.1
Motivation Scale - Pre	16.2
Motivation Scale - Post	23.1
Feeling of Inadequacy - Pre	16.0
Feeling of Inadequacy - Post	23.1
Gumpgookies - Pre	17.6
Gumpgookies - Post	23.8
Psychological Control - Rejection	31.2
Psychological Control - Guilt	31.2
Reaction to Mild Infraction	32.1
Family Structure	24.2
Accessibility of Adults	24.3
Education of Mother	24.7
Education of Father[1/]	49.9
Mother Working	26.1
Husband Employed[1/]	53.0
Husband's Job[1/]	54.1
Family Income	30.8
Primary Source of Income	26.8
Number of Moves in Last Three Years	34.2
Ratio of Rooms to People	24.4
Does Child Have Own Room	24.3
Does Child Have Own Bed	24.3
Parent Opinion Scale	24.4
Child Influence Techniques - Bad Behavior 1	29.6
Child Influence Techniques - Bad Behavior 2	33.4

(continued)

1/ The percentage of missing data for these variables includes father-absent families.

Table III.9 (continued)

Variable Name	Percentage of Children With Missing Data
2. Child-Family Background Variables	
Ethnicity	0.0
Age	0.0
Language Spoken at Home	0.3
Child Eligibility	21.3
Sex	0.0
Previous Head Start Experience	0.0
Attendance	2.1
Pre-Post Interval - S.B.	23.2
Pre-Post Interval - P.S.I.	22.3
Pre-Post Interval - Inventory of Factors Affecting SB	21.6
Pre-Post Interval - Gumpgookies	23.6
Telephone at Home - Pre	33.5
How Often is Child Read to - Pre	29.5
How Far in School You Would Like Child to Go - Pre	24.2
How Far in School You Think He Will Go - Pre	24.4
What Job you Like Child to Get - Pre	24.8
What Job You Think Child Will Get - Pre	25.3
Mother-Child Interaction - Pre	30.9

Table III.10

MISSING DATA ON OUTCOME MEASURES
AND CHILD-FAMILY BACKGROUND VARIABLES

(1967-68)

Variable Name	Percentage of Children With Missing Data
1. Outcome Measures	
S.B. - Pre	4.3
S.B. - Post	16.7
Behavior Problem Scale - Pre	5.2
Behavior Problem Scale - Post	17.0
Motivation Scale - Pre	5.2
Motivation Scale - Post	17.0
Feeling of Inadequacy - Pre	5.2
Feeling of Inadequacy - Post	17.0
2. Child-Family Background Variables	
Ethnicity	1.8
Age	1.8
Language Spoken at Home	3.4
Sex	1.8
Previous Head Start Experience	25.8
Attendance	10.8
Influence Techniques - Pre	15.9
Influence Techniques - Post	38.4
Aspirations - Pre	13.0
Aspirations - Post	25.7
Expectations - Pre	18.5
Expectations - Post	31.3
Aspirations and Expectations - Pre	18.6
Aspirations and Expectations - Post	31.5
Education of Mother - Pre	13.2
Mother's Job With Head Start - Post	36.0
Total Number of People in Home - Pre	11.5
Ratio of Rooms to People - Pre	12.1

Table III.11

MEAN INITIAL AND FINAL SCORES ON SB
(1968-69)

Category	Mean Score	
	Initial	Final
Those with Initial Score	87.82	
Those with Final Score		92.87
Those With Both Initial and Final Score	88.30	93.12

CHAPTER IV

DEMOGRAPHIC CHARACTERISTICS OF CHILDREN AND THEIR FAMILIES

According to the Head Start manual, 90 percent of the children participating in Head Start should come from families below OEO poverty guidelines, which in 1967-68 was $3,800 per year for a non-rural family of four. Children 3 to 6 years of age may be enrolled, although in most communities most children are a year younger than public school enrollment age. Where there are public school kindergartens, Head Start serves 3-4 year olds; where the public school begins in the first grade, most Head Start children are 5-6 years of age. Some younger children and some older children who are in special need may participate.

Examination of the initial demographic characteristics of the children and their families gives some idea of Head Start's diversity, and also in comparison to Head Start random sample census surveys, of the representativeness of the sample. Neither the 1967-68 nor the 1968-69 sample was designed to be representative; it is, however, of interest to identify areas in which the samples are similar to or different from the national Head Start population. One divergence is already known. Because of testing difficulties, centers with predominantly Spanish-surnamed children or Native American children were not included in the study.

When possible, demographic characteristics will be presented in a comparative framework. Some 1967-1968 demographic data were not retrieved during this study.

A. Demographic Characteristics of Children

Most of the children in both 1967-68 and 1968-69 samples were black (50% and 68% respectively). The next largest subpopulation was white (32% and 18%). The remaining children were predominantly Mexican-American or Polynesian. The ethnic distributions of the samples are presented in Table IV.1. In comparison with the Head Start census survey, black children in 1968-69 are over-represented.

The samples were divided evenly by sex; 50.7% were boys and 49.3% girls, and 50.5% were boys and 49.5% girls in 1967-68 and 1968-69, respectively. This is similar to national ratios.

With regard to age at time of initial testing, the range of ages was from 2 1/2 to 6 years of age. The distributions are shown in Table IV.2. In comparison to national samples, older children

Table IV.1

ETHNIC DISTRIBUTION OF THE SAMPLE

Ethnic Group	1967-68 E&R	1968-69 E&R	Head Start Census
Black	50%	68%	51%
White	32%	18%	38%
Other	18%	14%	11%
N	1889	1998	

Table VI.2

AGE DISTRIBUTIONS OF SAMPLES

Age at time of Initial Testing Chronological Age in Months	Distributions 1967-68	1968-69	Head Start Census
Less than 48	28%	21%	21%
4 59	61%	57%	43%
60+	11%	22%	36%

(i.e., those for whom Head Start serves as a kindergarten year) are under-represented.

The children represented all geographic areas and sizes and types of cities; however, by far the largest proportion (78%) in the 1968-69 sample came from cities with populations over 25,000. Of the remainder, 11% were from the urban fringe areas and 11% were from rural areas. Similar data from the 1967-68 data file were not analyzed.

With regard to prior Head Start experience, 82% of the 1968-69 sample had no previous school experience as required by the E&R sampling guidelines. Neither 1967-68 data nor Head Start census data are available for comparison; recently (1970-71; 1971-72) about 46% of the children have previously attended full-year Head Start.

B. Demographic Characteristics of Families

For the 1968-69 data, 1400 of the 1523 completed family interviews involved the mother. Similar data on the 1967-68 sample were not retrieved but it is a safe assumption to presume the same condition exists.

The distribution of these families on the derived socioeconomic status index was given in Table III.5. Table IV.3 shows the actual income distribution for available data. In comparison with the Head Start Census sample, the E&R respondents seemed to report slightly higher incomes. However, in comparison with the U.S. National Census. the median per capita income for the Head Start E & R sample is substantially lower than the U. S. median of $7,974. The Head Start Census sample survey reported a median Head Start income of $3,210.

Table IV.3

DISTRIBUTION OF FAMILY INCOME

Yearly Income	1968-69 E&R Centers	1968 Head Start Census
Less than $2,000	11%	27%
2,000 to 3,999	41%	40%
4,000 to 5,999	32%	23%
6,000 to 7,999	12%	7%
Over 8,000	4%	3%
N	1383	

Head Start families are larger than the average family size. Head Start survey figures show a median family size of 6.7 persons with over two-thirds of the Head Start children having older siblings living in the home.* Similar data for the 1967-68 and 1968-69 samples were not analyzed.

Parents of E&R sample Head Start families were similar in educational and occupational status to the Head Start census sample; they were less well educated and more likely to be unemployed or employed in low status occupations than the U. S. national averages.

In 1967-68, only 5% of the mothers had attended college while 63% had less than a high school education; 24% had less than a 9th grade education. In 1968-69 only 6% of the mothers had attended college while 63% had less than a high school education; 19% had less than a 9th grade education. Since the quality of education often is lower in low-income areas (Coleman et al., 1966) the educational competency of these mothers is likely to be lower at each category than the U.S. averages. That is, to "compensate" for differences in school quality, these parents would have to receive more years of schooling to be equivalent to the national averages; for example, two years of college for a low-income person may equal a high school degree for a person from a high-income family because of inequities in the educational system. Table IV.4 shows these distributions. The Head Start census also reports two-thirds of the mothers were not high school graduates.

Table IV.4

MOTHER'S EDUCATION

Educational Level	1967-68 E&R Centers	1968-69 E&R Centers	1968 Head Start Census
Less than 9th grade	24%	19%	28%
9th to 11th grade	39%	44%	38%
12th grade (H.S. Graduate)	32%	31%	27%
Some college or more	5%	6%	7%
N	1889	1513	

Considering only families where both parents were present, in 1968-69 only 8% of the fathers had attended college while 64% had less than a high school education; 32% had less than a ninth-grade education. Table IV.5 presents this information. Similar data for 1967-68 were analyzed; the Head Start census data reports 71% of the fathers had not graduated from high school.

Table IV.5

FATHER'S EDUCATION

Educational Level	1968-69 E&R Centers		1968 Head Start Census	
Less than 9th grade	32%	64%	42%	71%
9th to 11th grade	32%		29%	
12th grade (H.S. Graduate)	28%		21%	
Some college or more	8%		8%	

Regarding employment, Table IV.6 indicates that almost one-half of the mothers were not working and were not looking for work while about 20% of the mothers were working full time. U. S. census data show about 19% of mothers in this category. For families where both parents were present, about 76% of the fathers were working full time. Head Start census figures show about 83% of fathers and 31% of mothers employed. The 1968-69 E & R sample fathers were more likely than the census fathers to have completed between 9 and 11 years of school; they were not, however, more likely to be employed full or part time. Similar data for 1967-68 were not retrieved.

Table IV.6

PARENTAL EMPLOYMENT

Employment Status	1968-69 Mothers	1968-69 Fathers	Employment Status	1968 Head Start Census Mothers	1968 Head Start Census Fathers
Not working, but looking	17%	6%			
Not Working, not looking	49%	4%	Unemployed	69%	17%
Retired or disabled	1%	4%			

The primary source of income was fathers' earnings for about 45% of the families; about 28% of the families were receiving welfare. The distribution of primary source of income is shown in Table IV.7 for 1968-69.

The fact that 61% of the families lived primarily on parental earnings and were still eligible for Head Start, may help communicate something of the life experience of people who work in the most physically arduous and least attractive jobs in our society and still are so poor that the average per capita income is below national averages.

Table IV.7

PRIMARY SOURCE OF FAMILY INCOME

Source	1968-69	
Mother's earnings	12%	
Father's earnings	45%	61%
Combined parent earnings	4%	
Other adult	3%	
Welfare	28%	
Earnings plus welfare	3%	
Other sources	5%	
N=	1469	

While most (55%) of the children lived with both parents, 29% came from mother-only homes. These data are shown in Table IV.8. Similar data for 1967-68 were not retrieved; Head Start census figures show 67% of the homes having both parents.

Table IV.8

FAMILY STRUCTURE

Family Structure	1968-69 E&R Centers	
Mother only	29%	
Both parents	48%	55%
Parents and other adults	7%	
Other	16%	
N	1523	

C. Summary

In summary, while there is considerable variability in every demographic characteristic, the typical Head Start child in these samples was black, pre-kindergarten age, and came from a family with both mother and father present. The child had older brothers and sisters, and relative to most children in this country, lived in circumstances associated with low educational attainment, with parents who worked but could earn far less than what was needed for a minimally acceptable standard of living, with overcrowding (.8 rooms per person), and with high mobility (52% moved one or more times in the last three years).

In comparison to the Head Start census samples, black children, younger children (pre-kindergarten) and slightly better off families were overrepresented in the E & R samples. If these characteristics are related to initial psychological status or final status, or interact with program experiences to affect changes, the overall findings from the E & R sample may not represent what "typically" happens in Head Start. Even if the sample were wholly representative, such interactions would mean that general statements would be less precise than more specific conclusions; e.g., a more reliable statement would be "younger children gained 10 developmental months in 8 chronological months, while older children gained 8 developmental months in the same time" rather than "on the average, children gained 9 developmental months." Non-representativeness of the E & R sample would not affect the reliability of the first statement; it would affect the reliability of the second statement in comparison with a national representative sample.

The appropriate generalizations will be indicated in the text; the data in the tables should be cited only with due consideration of limitations on generalization.

Where the demographic variables cited are not related to outcomes, less restriction need be placed on generalization. A conservative approach is to be cautious about generalizations from non-randomly selected samples since divergence on possibly relevant factors which were not considered could bias results. A less conservative approach would emphasize the similarity of E & R and census samples on many variables (e.g., child sex, mother's education) and consider the findings as reasonably general.

CHAPTER V

DESCRIPTION OF TEACHERS AND PROGRAMS

Project Head Start is a national program. There are Head Start classes in all of the 50 states and many of the territories, including Puerto Rico, the Virgin Islands, and Guam. Head Start is also a neighborhood program, employing wherever possible people who form the community of which Head Start is a part. In the guidelines, the competencies an individual brings to a program, rather than formal training, are stressed for many positions; the exceptions are primarily in the health area. In part, this policy was dictated by lack of trained personnel; in 1966-69 if a degree in early childhood education were required for a teacher, most full-year programs would never have opened. In a larger part, this policy was chosen in the belief that the talents and competencies neighborhood residents had developed constituted a reservoir of unused talent. If, by providing enough training to release these talents, neighborhood residents could operate the programs, Head Start could make a direct contribution to reducing unemployment and poverty, and the programs themselves might benefit from the understanding and loving acceptance of the needs of the children that people who had "been there" could bring to the classes. It was recognized that this policy meant the development of in-service training programs on a nation-wide scale; universities and colleges responded to this need, not only by expanding their early childhood education departments but also by developing special seminars, workshops and special degree programs and by making their staffs available as consultants and trainers. It was equally recognized that this policy would mean a slower start-up and later achievement of uniformly high quality programs than might be expected if trained, qualified staff were available for most programs.

These comments should not be interpreted to mean that most Head Start staff was incompetent and unqualified. They do suggest that attributes other than formal training were salient and that staff potential and existing competencies had to be further developed, undoubtedly more so than in many experimental programs. In addition, a system of in-service training and supervision had to be put in place--like those required by many experimental programs--before the centers could be expected to uniformly provide high quality service in all aspects of the comprehensive program.

Staff experience and training, and the resources available within the community, therefore are sources of variation in Head Start programs. Local decisions regarding program emphasis and curriculum approach within classes are other sources of variation. Head Start

as stated in Chapter 1 provide a framework, as do the guidelines, within which all programs operate. The guidelines specify the minimum standards and conditions deemed necessary if a program is to meet the national goals. These define a very broad and flexible program; for example, the minimum adult/child ratios for children of different ages are specified, but not what the adults must do. The Rainbow Series of booklets which supplement the guidelines describes what might be called a traditional, child-centered program in that the objectives emphasize the child's personal-social development in a supportive rather than directive atmosphere. This approach is different from the stereotyped "play school" in that the teacher is responsible for planning how the daily activities can specifically help the individual child; it is assumed that she knows what she is doing and why in relation to each child for whom she is responsible. Head Start programs are free to adopt any approach that seems valuable to the staff and parents, or to mix what seems best from several approaches. National and regional staff provides local programs with information on approaches to early childhood education--which have proliferated since 1965--through many technical assistance and research utilization systems.

In 1967-68, programs were selected to represent examples of the variety of approaches adopted by local programs. In 1968-69, for the first time the national office, on the recommendation of most E & R Center Directors, focused the national evaluations on the programs developed by individual E & R Directors with the agreement and cooperation of the local programs. The classroom observation data described below therefore represent selected natural variation in 1967-68 and a mixture of experimental classes and regular Head Start programs in 1968-69. The primary emphasis in this report is on analysis of 1968-69 data. The 1967-68 data were used where comparable analyses were possible. Therefore, very little 1967-68 teacher-program information will be presented.

A. Teacher Demographic Data

Head teacher demographic data were available for 97 teachers in 1968-69. Head teacher ages varied from quite young (21 year of age or less) to late middle age (46 years or more). Approximately two-thirds of the teachers were between 22 and 39 years of age. The distribution of head teacher's age is given in Table V.1. All but five of the 97 1968-69 teachers were female; 50 were black and 43 were white.

Table V.2 indicates the level of teacher preparation. Over half of the teachers held at least a bachelor's degree. In 1968-69, 66% reported taking early childhood education courses or receiving a degree in this field. Twenty-nine percent of the teachers lived in the Head Start community and 25% had no experience with disadvantaged children

Table V.1

DISTRIBUTION OF HEAD TEACHER'S AGE

Teacher's Age	1968-69
21 years or less	4%
22-27 years	29%
28-33 years	14%
34-39 years	21%
40-45 years	15%
46 years or more	17%
Total N	97

Table V.2

LEVEL OF HEAD TEACHER'S PREPARATION

Educational Level	1968-69
High school graduate or less	5%
Some college	24%
A.A. Degree	10%
B.A. or B.S. Degree	37%
Course credit beyond B.A. or B.S.	19%
M.A. and credit above M.A.	5%
Total N	97

prior to joining Head Start. Regarding previous experience in Head Start, 71% in 1968-69 reported a year or more of paid Head Start experience prior to the evaluation study. These data give some information regarding the mobility of classroom staff.

B. Some Aspects of Class Structure

With regard to program auspices in 1968-69, 37% were LEA (local educational agencies) operated and 47% were CAA (community action agencies) operated.

The teacher/child/classmate ethnic structure is shown in Table V.3. Most classes were ethnically homogeneous, as would be expected of neighborhood programs. In 1968-69 about 60% of the non-white children were taught by non-white teachers while about 59% of the white children were taught by white teachers. Fifteen percent of the children were of a different ethnic group than the majority of their classmates.

Table V.3

TEACHER/CHILD/CLASSMATE ETHNIC STRUCTURE
(1968-69)

Teacher	Child	Class	Percent
White	White	Majority White (75% or more)	4.3
White	White	Mixed (74% or less)	4.0
White	White	Majority other	2.0
White	Other	Majority White	0.4
White	Other	Mixed	6.8
White	Other	Majority Other	25.5
Other	White	Majority White	2.3
Other	White	Mixed	3.1
Other	White	Majority Other	1.8
Other	Other	Majority White	0.3
Other	Other	Mixed	7.8
Other	Other	Majority Other	41.7
Total N			1991

C. Class Physical Resources

Class physical resources varied widely. The resource variation reported by classroom observers and shown in Table V.4, agrees with the self reported census findings, in that some centers were having difficulty finding appropriate space and facilities for conducting comprehensive programs while others were able to find very good facilities and obtain excellent equipment. Examination of Table V.4 shows that the majority of centers appear to have found sufficient space, both indoor and outdoor, although the data do not reflect whether the space was truly functional to its purpose. Outdoor equipment such as swings and sandboxes was present only in about half the centers while indoor equipment was, in general, far more plentiful. Musical instruments, ingerpaint sets and puzzle sets were especially prevalent.

D. Classroom Activities as Reported by Trained Observers

Table V.5 presents the correlations among the four 1967-68 factors, and Table V.6, the 1968-69 OSCI Class and Teacher factor score correlations.

The intercorrelations (Table V.5) involving the four 1967-68 OSCI scores considered in the present analyses indicated no relationships of any practical significance.

The 1968-69 factor correlations are based upon individual child data rather than class average data. They are the correlations involving individual child exposure to the 12 factors (i.e., child was the unit of analysis). These correlations should be quite close to the correlations that would be obtained from using class as the unit of analysis. They can be considered as correlations based upon differential weightings of the classes. Since in most cases classes are approximately equal in size, the weights for classes will tend to be similar in magnitude.

It is apparent from the data that the six class factors are relatively independent of each other and that the six teacher factors are also relatively independent of each other. The cross correlations between these two sets of factors are somewhat higher. There was a strong correlation of .67 between Class and Teacher Structured Lessons and a correlation of .66 between Class Language and Discrimination Learning and Teacher Receptive Learning.

The distributions of the factor scores for the 1968-69 data are presented in Table V.7 and V.8 for Teacher and Class factors respectively. A comparative examination of the distributions of the two sets of factors showing high cross correlations, as well as those other factors relating

Table V.4

SCHOOL PLANT RESOURCES

Square Feet Outdoor Space Per Child

	N	%
0 to less than 39 sq. ft.	330	22.0
40-49 sq. ft.	21	1.4
50-59 sq. ft.	27	1.8
60-69 sq. ft.	67	4.5
70-79 sq. ft.	88	5.9
80-89 sq. ft.	10	0.7
90-99 sq. ft.	45	3.0
100 sq. ft or over	915	60.9

Square Feet Indoor Floor Space Per Child

	N	%
0 to less than 14 sq. ft.	92	6.1
15-19 sq. ft.	24	1.6
20-24 sq. ft.	58	3.8
25-29 sq. ft.	147	9.7
30-34 sq. ft.	116	7.6
35-39 sq. ft.	193	12.7
40-44 sq. ft.	40	2.6
45 sq. ft. or more	847	55.8

Number of Swing Sets

	N	%
0	875	57.7
1-2	428	28.2
3-4	136	9.0
5-6	64	4.2
7-8	14	0.9

Number of Sandboxes

	N	%
0	752	49.6
1-2	685	45.2
3-4	30	2.0
5-6	37	2.4
7-8	13	0.9

(continued)

Table V.4 (continued)

Waterplay Equipment

	N	%
0	544	35.9
1-2	721	47.5
3-4	118	7.8
5-6	39	2.6
7-8	25	1.6
9-10	13	0.9
15 or more	57	3.8

Fingerpaint Sets

	N	%
0	62	4.1
1-2	187	12.3
3-4	162	10.7
5-6	169	11.1
7-8	78	5.1
9-10	77	5.1
11-12	116	7.6
13-14	12	0.8
15 or more	654	43.1

Small Block Sets

	N	%
0	168	11.1
1-2	381	25.1
3-4	354	23.3
5-6	189	12.5
7-8	51	3.4
11-12	20	1.3
15 or more	354	23.3

Puzzle Sets

	N	%
0	73	4.8
1-2	85	5.6
3-4	59	3.9
5-6	144	9.5
7-8	33	2.2
9-10	191	12.6
11-12	50	3.3
13-14	32	2.1
15 or more	850	56.0

(continued)

Table V.4 (continued)

Learning Games

	N	%
0	161	10.6
1-2	229	15.1
3-4	361	23.8
5-6	257	16.9
7-8	72	4.7
9-10	133	8.8
11-12	57	3.8
13-14	28	1.8
15 or more	219	14.4

Musical Instruments

	N	%
0	173	11.5
1-2	444	29.6
3-4	26	1.7
5-6	42	2.8
7-8	25	1.7
9-10	127	8.5
11-12	28	1.9
13-14	44	2.9
15 or more	593	39.5

Science Equipment

	N	%
0	519	34.2
1-2	474	31.2
3-4	179	11.8
5-6	98	6.5
7-8	28	1.8
9-10	37	2.4
11-12	77	5.1
15 or more	105	6.9

Class and Teacher Structured Lessons factors (r=.67) are distributed similarly. The Teacher factor has its mode at scale zero with a normal piling up immediately on both sides, but there is a strong compression toward the positive end of the scale. The Class factor has its mode slightly below zero, but 94% of the remaining scores are above this, i.e., strong positive skew. High scores on these factors represent a high degree of structure.

Class Language and Discrimination Learning and Teacher Receptive Learning (r=.66) are also quite similar in distribution. Differences are slight: the Class factor has a plateau mode broadly on the negative side of zero while the Teacher factor has a clear modal point just slightly negative; but the Teacher factor has a few more negative cases than the Class factor. Both these factors are largely defined by language activities and materials in a watching-listening context. The distributions indicate quite similar degrees of presence in classes.

Although not having high cross correlations, two other sets of factors pairs' distributions should be examined because of their analogous aims. The Teacher and Class Social Emotional Interaction factors (r=.15) are shaped quite closely alike with positive skew. The indications here are not of warmth or coolness of the classroom but rather of an even tempo of interaction tending toward the slightly demonstrative. There are two factors concerned with art activities (r=.31). From the teacher and class viewpoints there is slightly more variability toward the positive than toward the negative end of the scale.

Table V.5

CORRELATIONS AMONG OSCI FACTORS
(1967-68)

	1	2	3	4
1.	Cognitive-Low Structure	.02	-.14	.05
2.	Routines and Rules		.05	-.10
3.	Cognitive-High Structure			.05
4.	Child Centered-Unstructured			

Table V.6

CORRELATIONS AMONG OSCI FACTORS
(1968-69)

		1	2	3	4	5	6	7	8	9	10	11	12
1	Class Factor I Structured Lessons	1	-.08	.01	-.21	-.07	-.07	-.40	.67	-.16	-.31	-.06	-.19
2	Class Factor II Group Activities and Routines			.10	-.03	-.06	-.01	-.19	-.27	-.00	-.22	.37	-.11
3	Class Factor III Social Emotional Interaction				.19	.02	.13	.15	-.31	-.10	-.26	-.22	.05
4	Class Factor IV Verbal Communication					.15	.12	.02	-.28	.20	-.04	-.26	.34
5	Class Factor V Instruction in Creative Arts						.06	-.22	-.08	.31	.36	-.35	.04
6	Class Factor VI Language and Discrimination Learning							.25	-.09	-.11	-.07	.05	.66
7	Teacher Factor I Social Emotional Interaction								.04	.02	.10	.11	.24
8	Teacher Factor II Structured Lessons-Large Group									.00	-.01	-.01	-.14
9	Teacher Factor III Art Activities										.07	-.08	.02
10	Teacher Factor IV Creative Instruction-Small Group											.00	.15
11	Teacher Factor V Routines												.08
12	Teacher Factor VI Receptive Learning												

Table V.7

DISTRIBUTION OF VALUES ON OSCI TEACHER FACTORS
(1968-69)

Intervals (Z-Standard)		I. Social Emotional Interactions N	%	II. Structured Lessons-Large Group N	%	V. Routines N	%
4.00	3.51	13	.8	0	0.0	17	1.0
3.50	3.01	0	0.0	37	2.2	20	1.2
3.00	2.51	23	1.4	34	2.1	0	0.0
2.50	2.01	44	2.7	33	2.0	9	.5
2.00	1.51	59	3.6	97	5.9	58	3.5
1.50	1.01	90	5.4	130	7.9	142	8.7
1.00	.51	204	12.3	109	6.6	242	14.7
.50	.01	301	18.2	232	14.0	259	15.8
0.00	- .49	309	18.7	515	31.2	446	27.2
- .50	- .99	437	26.4	255	15.4	204	12.4
-1.00	-1.49	154	9.3	184	11.1	160	9.8
-1.50	-1.99	19	1.1	14	.8	44	2.7
-2.00	-2.49	0	0.0	0	0.0	40	2.4
-2.50	-2.99	0	0.0	0	0.0	0	0.0
-3.00	-3.49	0	0.0	13	.8	0	0.0
-3.50	-3.99	0	0.0	0	0.0	0	0.0
-4.00	-4.49	0	0.0	0	0.0	0	0.0
		1653		1653		1641	

Intervals		III. Art Activities N	%	IV. Creative Instruction-Small Group N	%	VI. Receptive N	%
-4.00	-3.51	0	0.0	0	0.0	0	0.0
-3.50	-3.01	0	0.0	0	0.0	0	0.0
-3.00	-2.51	0	0.0	0	0.0	0	0.0
-2.50	-2.01	0	0.0	14	.9	0	0.0
-2.00	-1.51	116	7.1	73	4.4	91	5.6
-1.50	-1.01	84	5.1	118	7.2	144	8.8
-1.00	- .51	410	25.0	301	18.3	345	21.2
- .50	- .01	252	15.3	434	26.4	402	24.6
0.00	.49	313	19.1	305	18.6	299	18.3
.50	.99	246	15.0	116	7.1	119	7.3
1.00	1.49	131	8.0	185	11.3	141	8.6
1.50	1.99	20	1.2	59	3.6	40	2.5
2.00	2.49	38	2.3	9	.5	48	2.9
2.50	2.99	24	1.5	17	1.0	2	.1
3.00	3.49	9	.5	10	.6	0	0.0
3.50	3.99	0	0.0	0	0.0	0	0.0
4.00	4.49	0	0.0	0	0.0	0	0.0
		1643		1641		1631	

Table V.8

DISTRIBUTION OF VALUES ON OSCI CLASS FACTORS
(1968-69

Intervals (Z-Standard)		I. Structured Lessons		III. Social Emotional Interactions		V. Instruction in Creative Arts	
		N	%	N	%	N	%
4.00	3.51	0	0.0	20	1.2	0	0.0
3.50	3.01	0	0.0	13	.8	19	1.2
3.00	2.51	30	1.8	0	0.0	20	1.2
2.50	2.01	23	1.4	19	1.1	0	0.0
2.00	1.51	150	9.1	40	2.4	54	3.3
1.50	1.01	129	7.8	88	5.3	168	10.3
1.00	.51	114	6.9	287	17.4	227	13.9
.50	.01	180	10.9	374	22.6	292	17.9
0.00	- .49	350	21.2	369	22.3	340	20.8
- .50	- .99	608	36.8	298	18.0	240	14.7
-1.00	-1.49	60	3.6	93	5.6	192	11.7
-1.50	-1.99	9	.5	52	3.1	70	4.3
-2.00	-2.49	0	0.0	0	0.0	13	.8
-2.50	-2.99	0	0.0	0	0.0	0	0.0
-3.00	-3.49	0	0.0	0	0.0	0	0.0
-3.50	-3.99	0	0.0	0	0.0	0	0.0
-4.00	-4.49	0	0.0	0	0.0	0	0.0
		1653		1653		1635	

Intervals		II. Group Activities and Routines		IV. Verbal Communication		VI. Language and Discrimination Learning	
		N	%				
-4.00	-3.51	0	0.0	0	0.0	0	0.0
-3.50	-3.01	0	0.0	0	0.0	0	0.0
-3.00	-2.51	0	0.0	0	0.0	0	0.0
-2.50	-2.01	15	.9	0	0.0	0	0.0
-2.00	-1.51	13	.8	29	1.8	0	0.0
-1.50	-1.01	157	9.5	132	8.0	102	6.2
-1.00	- .51	228	13.8	433	26.2	471	28.8
- .50	- .01	462	27.9	368	22.3	465	28.4
0.00	.49	330	20.0	309	18.7	271	16.5
.50	.99	227	13.7	165	10.0	165	10.1
1.00	1.49	90	5.4	94	5.7	40	2.4
1.50	1.99	40	2.4	38	2.3	50	3.1
2.00	2.49	31	1.9	31	1.9	30	1.8
2.50	2.99	32	1.9	39	2.4	31	1.9
3.00	3.49	14	.8	15	.9	13	.8
3.50	3.99	0	0.0	0	0.0	0	0.0
4.00	4.49	14	.8	0	0.0	0	0.0
		1653		1653		1638	

CHAPTER VI

INITIAL PSYCHOLOGICAL CHARACTERISTICS OF CHILDREN AND THEIR FAMILIES

While the emphasis in this report is on change, consideration of the initial psychological status of children and their families is of interest in its own right and may be relevant to understanding the changes later discussed.

A. Psychological Characteristics of Children

Table VI.1 presents the means, Ns, and standard deviations for the pre score means on the seven outcome measures. The data presented in Table VI.1 for 1968-69 exclude children with previous Head Start experience while the data for 1967-68 include children with previous Head Start experience. The 1967-68 children with previous Head Start experience could not be identified. Standard deviations are not given on the Behavior Problem, Motivation Problem, and Feelings of Inadequacy scales due to their extremely skewed distributions.

Table VI.1

PRESCORE MEANS ($\bar{X}$), Ns AND STANDARD DEVIATIONS (SD) FOR Ss ON THE OUTCOME MEASURES

Pre Score	1967-68 $\bar{X}$	SD	N	1968-69 $\bar{X}$	SD	N
Stanford Binet IQ	91.35	13.75	1779	88.57	14.37	1420
Preschool Inventory	--	--	--	32.46	11.29	1298
Animal House (WPPSI)	--	--	--	20.82	12.81	1378
Gumpgookies	--	--	--	34.10	7.17	1293
Behavior Problem	0.65	0- 7*	1760	0.68	0- 7*	1322
Motivation Problem	1.87	0-10*	1760	1.58	0-10*	1321
Feelings of Inadequacy	1.00	0- 4*	1760	0.74	0- 4*	1322

* Range of scores.

For the one measure (Stanford-Binet) on which normative data are available, the entering performance is substantially lower than that of the more economically heterogeneous norming sample (Table VI.2). The Preschool Inventory was normed on a low-income Head Start sample but the norm total score mean and standard deviation are not comparable with the present data because of scale differences. The children scored about one-third of a standard deviation below the norming sample of the AH-WPPSI: it was not possible to calculate an exact comparison.

Table VI.2

PRESCORES COMPARED TO NORMATIVE DATA

	1967-		1968-69		Normative	
	$\bar{X}$		$\bar{X}$	SD	$\bar{X}$	SD
Stanford-Binet IQ	91.35	13.75	88.57	14.37	100	16

No normative data exist for the three adjustment rating scales (BP, MP, FI). The frequency distributions on which the means are based are highly skewed; few children exhibited a large number of problems and most showed 0 or 1. This can be seen in Table VI.3. The most frequently reported problem was motivation.

Table VI.3

FREQUENCY DISTRIBUTIONS IN PER CENT FOR ADJUSTMENT RATING SCALES

	1967-68			1968-69		
Raw Score (No. of Problems)	BP 0-7*	MP 0-10*	FI 0-4*	BP 0-7*	MP 0-10*	FI 0-4*
0	66.2	44.2	60.2	64.5	40.7	52.6
1	17.8	18.3	18.9	17.1	15.1	19.5
2	7.5	12.1	9.1	11.0	12.7	11.7
3	4.8	6.9	6.8	4.9	8.4	8.7
4	2.5	6.1	5.0	1.7	9.4	7.5
5	0.8	5.2	--	0.7	6.0	--
6	0.1	3.3	--	0.1	4.2	--
7+	0.2	3.9	--	--	3.5	--
N	1163	1168	1171	1770	1790	1790

* Range of scores.

Interpretation of behavior in a testing situation must be made cautiously since tester rapport and skill doubtless substantially affect the child's ease, cooperativeness and interest. In general, the majority of children, most new to Head Start and most tested in week 4 to 12 of the program, did not show behavior problems or feelings of inadequacy and many did not show motivational problems. This may reflect (1) the skill of the testers, (2) the children's adaptation to Head Start, or (3) that the Head Start children enter programs on the whole, cooperative, confident, and well-adjusted in terms of behavior expected of 3 to 6-year-olds. Much has been written on the damaging effects of the disadvantaged home on the child's personal-social development. While true "pre" data are not

available, these descriptions suggest that, except for motivational problems, the responsibility of the schools is to maintain the positive self-concept (FI) and adjustment (BP) with which the child enters. Some evidence in support of this interpretation is found in the ETS longitudinal study of disadvantaged children (Shipman, 1970). On the Brown Self-Concept test, children who were tested in the spring and summer before entering Head Start had a very favorable self-concept, with scores as high as those of more advantaged children. While the Brown may lend itself to a positive response set, the data are at least suggestive of the need to look for maintenance of self-esteem in preschool and primary school rather than a mean gain.

Additional support for this interpretation may be found in Table VI.4, the prescore intercorrelation matrix. The 1968-69 correlations between initial Binet and PSI performance for the Motivation Problem scale are -.29 and -.27 respectively; for Behavior Problems, the correlations are -.10 and -.08, and for Feelings of Inadequacy, -.13 and -.18. However, these correlation coefficients should be viewed with caution due to the highly skewed distributions of the behavior rating scales.

Table VI.4

INTERCORRELATIONS AMONG PRESCORES*

Scores	SB	PSI	WPPSI	N GUMP	BP	MP	FI
Scores							
1. SB		.45	.24	.30	-.10	-.29	-.13
2. PSI			.52	.50	-.08	-.27	-.18
3. WPPSI-AH				.33	-.09	-.13	-.09
4. GUMP					-.08	-.22	-.13
5. BP	-.12					.09	-.03
6. MP	-.30				.11		.52
7. FI	-.17				.10	.54	

* 1967-68 below diagonal; 1968-69 above diagonal.

The data from the intercorrelation matrix indicate that initial performance on the Binet, the Preschool Inventory and the Gumpgookies may be related more to motivational problems than to feelings of inadequacy and behavior problems.

As an inference, this suggests that on all initial measures--be the measure intended as an assessment of cognitive development (Binet), preschool achievement (PSI) or achievement motivation (Gumpgookies)--the child's scores are more affected by his apparent willingness to

perform and interest in the tests (MP) than by social adjustment (BP) or self-concept (FI).

The data also suggest that:

1) Cognitive performance is associated with preacademic readiness (.45), achievement motivation (.30) and motivational problems (-.29).

2) Preacademic readiness is also associated with ability to learn a new task (.52), achievement motivation (.50) and motivation problems (-.27).

3) Ability to learn a new task is also related to achievement motivation (.33), while achievement motivation is additionally associated with motivational problems (-.22).

4) Within the behavior scales, motivational problems had more in common with self concept (feelings of inadequacy) than with adjustment (behavior problems).

The strength of these relationships as pre measures, with age considered and with initial testing conducted between weeks 4 and 12 o

B. Effect of Prior Head Start Experience on Initial Performance

Prior Head Start experience is reliably associated with initial performance. As far as known, none of the children with prior experience had previously been tested with the evaluation battery, and probably not with other measures since testing has typically been used only if a child would be referred for psychological services' assessments. Children with no previous Head Start experience had the lowest scores on the cognitive measures. Table VII.3 in Chapter VII gives more information regarding these relationships. Table VI.5 shows data for 1968-69; data were not analyzed for 1967-68.

Table VI.5

INITIAL PERFORMANCE AND PRIOR HEAD START EXPERIENCE - MEAN SCORES (1968-69)

Measure	None	Summer	More Than Summer
SB	88.57	91.17	91.94
PSI	32.46	40.82	36.05
AH	20.82	25.90	21.42
Gump	34.10	35.50	35.72
BP	0.68	.77	0.59
MP	1.58	2.03	1.22
FI	0.74	0.79	0.56
CA (months)	53.59	53.13	51.74
N for SB	1420	157	179

Children with prior Head Start tend to be younger than children without prior Head Start. Older children without prior Head Start tend to have lower initial scores than do younger children with prior Head Start. These differences therefore should be considered only as consistent with the belief that Head Start experience benefits the child, rather than as conclusive evidence either of an effect or of the magnitude of changes.

C. Effect of Time of Initial Testing on Performance

While most children were initially tested in weeks 4-8 of their Head Start experience, some were tested earlier and some much later. The association of initial performance on the Binet and time of testing for children with no prior Head Start was investigated and no substantial or significant relationship was discovered.

D. Social-Psychological Characteristics of Families

Five major areas will be considered: parental aspirations and expectations for their children; parental techniques for modifying child behavior; parental stimulation of the child's cognitive development; parental "locus of control"--optimism or pessimism; and parental participation in Head Start. In later chapters, the relation of these characteristics to the child's initial performance and gains will be considered.

1. Parental Aspirations and Expectations for the Children (Pre Parent Interview)

Four items were used in 1968-69 to assess parental hopes for their children:

a) How far do you hope (your child) will go in school?

b) How far do you expect (your child) will go in school?

c) What job do you hope (your child) will have?

d) What job do you expect (your child) will have?

As discussed in Chapter III, these items were combined into an aspirations dimension and an expectations dimension. Table VI.6 summarizes available information for both samples. It can be seen from Table VI.6 that while mothers' expectations for their children are relatively low, their aspirations are relatively high. The mothers have great hopes for their children, but also realistically perceive obstacles to the fulfillment of these hopes.

2. Parental Techniques for Modifying the Child's Behavior (Pre Parent Interview)

Four open-ended questions were intended to assess parental techniques for modifying children's behavior:

Table VI.6

MATERNAL ASPIRATIONS AND EXPECTATIONS
(Pre Parent Interview)

Levels		Percent
Aspiration Level (1968-69) (Education and Occupation)		
Low		8
Medium		39
High		53
N	1101	100
Expectation Level (1968-69) (Education and Occupation)		
Low		57
Medium		29
High		14
N	1101	100
Educational Aspirations Level (1967-68)		
Low (High School and below)		20
High (Beyond High School)		80
N	1657	100
Educational Expectations Level (1967-68)		
Low (High School and below)		70
High (Beyond High School)		30
N	1654	100

a) What do you do when (your child) does something you regard as very wrong?

b) What do you do when (your child) does something you regard as somewhat wrong?

c) What do you do when (your child) does something that you like?

d) What do you do when (your child) does something that pleases you a great deal?

The first two items were intended to elicit parental means of modifying disliked behavior; these typically include verbal approaches (reasoning, scolding, shaming), withholding privileges, taking away favored possessions, and physical control. Generally, the mode of control is associated with social class, with verbal approaches more frequently reported by higher income, better educated families and physical control by lower-income families. Cross-cultural studies have shown a further relationship between authoritarian and egalitarian societies and child rearing practices: physical punishment occurs more frequently in the authoritarian and verbal control in the egalitarian societies. The psychological consequences are not necessarily simply related to later development. The child controlled by shaming and guilt may become, for example, more neurotic than the child whose behavior is modified by immediate physical punishment. On the other hand, the child whose parents present a model of reasoning, of thinking about alternatives and possibilities, of verbal mediation of response even to stressful situations is very likely learning a way of coping that is different from the child whose adult models do not appear to delay response or seek alternatives to violence or physical deprivation ("You'll go to bed without dinner.").

While few studies analyze the range of techniques used by parents, parents who modify their control techniques to fit different situations are providing their children with a quite different coping model from parents whose "average" response varies little with the situation, but which is similar to the "average" response of parents with more variable styles. Such variability would, of course, need to be distinguished from inconsistency of response to similar situations.

More recently, attention has been directed to the techniques used by parents to reward behavior of which they approve. Again, responses include verbal (social) and material rewards. Laboratory studies of the extent to which social class is associated with the child's responsiveness to social and material rewards have generally indicated that more advantaged children perform "better" for social rewards. Others indicate that reward per se (attention, information) may be more salient than type of reward.

Table VI.7 shows the distribution on behavior modification items for the 1968-69 sample in response to "What do you do when (your child) does something you regard as somewhat wrong?" More mothers report using punishment (44%) than attempting to shape the child's behavior through constructive responses by the mothers (30%).

Table VI.7

PARENTAL TECHNIQUES FOR MODIFYING BEHAVIOR
(Pre Parent Interview)
1968-69

Method	Percent
Punishment	44
Punishment and constructive response	16
Do nothing	10
Constructive response alone	30
N	1336

3. Parental Stimulation of the Child's Cognitive Development (Pre Parent Interview)

Socioeconomic status as defined by parental education and occupation, by income level, and by such life circumstances as family size and area of residence powerfully influences the choices the family can make. Within the admittedly limited options imposed by social class, there is great variability in individual life styles, in the way people cope with adversity, in how they feel about themselves, and what they offer their children. The amount of individual attention the child receives that may be directed to stimulating his development has been thought to be a highly important variable, particularly during the period from 0 to 6 when children are almost wholly dependent on their families for such stimulation.[1]

Among the questionnaire items related to parental stimulation of child, frequency of reading to the child was seen as being salient. Table VI.8 gives initial frequencies for the 1968-69 sample. It is interesting to note that 71% of the mothers reported that they read to their child at least several times per week.

1/ Almost all families now have television; many reports indicate that preschool children spend a great many of their waking hours watching TV. Through Sesame Street, Round the Bend, Mr. Rogers, Ripples, Chiquitines and other programs, the potential of television for preschool children is being actively explored.

Table VI.8

FREQUENCY OF READING TO THE CHILD
(Pre Parent Interview)
(1968-69)

Frequencies	Percent
Seldom or never	7
Sometime (at least once a week)	22
Often (several times a week)	35
Regularly (at least once a day)	30
Very frequently (much of each day)	6
N	1387

4. Parental Locus of Control--"Optimism" versus "Pessimism" (Pre Parent Interview)

A broad class of attitudes, beliefs and expectancies is considered under the term "locus of control." As noted earlier, the person who believes that he can affect his own life and the lives of others is said to have an internal locus of control. The person who sees his life as being controlled by "luck" or the influence of others is said to have an external locus of control. Associated with external locus of control is often a sense of anomie--of valuelessness and of alienation--of not belonging to or being part of a society or subgroup. Many low-income families belong to cultures with rich, strong traditions, yet do not believe that what they do can make a difference. The development of a sense of control, power and influence--of much that we mean by the phrase "human dignity"--was one aspect of the Community Action Program and of Head Start's policy regarding parent participation as a decisionmaker as well as parent involvement in the child's education.

As discussed in Chapter III, the school attitude items designed to tap this area were not used in further analyses because of very low reliabilities and consequently low correlations with other outcome or explanatory variables. This is an area in which extensive developmental work needs to be done.

However, the Srole alienation scale also previously discussed was used in further analyses. The frequency distribution of parents' responses to the five stimulus questions was approximately normal, with a mean of 16.9 and a standard deviation of 3.8 with a range from 6 to 25. Examination and interpretation of the responses indicate a slight tendency toward pessimism on the part of the parents in the 1968-69 sample.

5. Parental Participation in Head Start (Post Parent Interview)

Information concerning the participation of parents in the ongoing Head Start program was taken from the post measurements and thus reflect to some degree the success that the programs had in this regard. National Head Start guidelines make it clear that parental participation is a much-to-be-desired program element.

Table VI.9 presents the information on this dimension for 1968-69. With respect to simply going to or visiting the Head Start Center, parents went either very few times per year, or quite a bit. Sixty-five percent of the parents who visited their Head Start Center visited more than once or twice a year. Forty-eight percent of the parents volunteered to help out in class. Less than 20% were members of the Head Start governing board, the Policy Advisory Committee.

Table VI.9

PARENTAL PARTICIPATION
(Post Parent Interview)
(1968-69)

Go to Head Start Center?	Percent	
Times per year of those parents who visited the Head Start Center		
1-2	35	
3-4	23	
5-6	11	65%
7-8	6	
9+	25	
N = 970		
Volunteer in Class?		
No	52	
Yes:		
Less than 1/month	23	
1-2/month	9	
1/week	8	
2/week	4	
3/week	1	
4/week	1	
5/week	2	
N = 1265		

CHAPTER VII

RELATIONSHIPS OF CHILD AND FAMILY VARIABLES TO PRE SCORES

A. Considerations in Presenting Single-Factor Results

Several criteria were developed for deciding which results to highlight in the following chapters. To have merely presented the results of all single-factor analyses for the 1968-69 data and replication results for the 1967-68 data would have placed an undue burden on the reader to distill both which results were robust as well as statistically significant, and what patterns of significant and nonsignificant results appeared in the data. Therefore, it was decided to highlight those results which were judged to be "substantial" as well as statistically significant.

The criterion for accepting a statistically significant result as a "substantial" result was defined in terms of whether or not the largest mean difference within any set of outcome mean scores (pre scores, gain scores, or adjusted gain scores) being compared in a given single-factor analysis exceeded 1/3 of one standard deviation for that outcome measure for the total sample. Table VII.1 presents the criterion value that had to be exceeded in comparing the 1968-69 subgroup mean scores from a particular input variable on each of the seven pre score outcome measures. (The criterion values for judging substantial gain and substantial adjusted gain scores are presented in Table IX.3.) No similar criteria were established for 1967-68 because only a limited number of replication analyses on the 1967-68 data was performed.

One reason for establishing a criterion for a "substantial" result was that attention would thereby be focused on statistically reliable results that were nontrivial. With the large Ns involved in the present analysis, many results were statistically significant without reflecting a substantial relationship between the input and outcome variables. An additional reason for establishing a criterion was that this would reduce in some measure the need to discuss significant relationships which could plausibly have resulted from the inextricable confounding of variables which existed in the data.

Nonsignificant and/or nonsubstantial results are also highlighted if they are "consistent" in some respect with other results and reflect meaningful relationships between input and outcome variables. One criterion of consistency was that at least five of the seven outcome measures involving a given input variable yielded a common pattern of relationships with that input variable. A second consistency criterion was that scores from a particular subset of outcome measures (e.g., the cognitive measures) reflected a pattern of results similar to that found with the same subset of outcome measures and another similar input variable which were judged to be "substantial" in their relationship.

Table VII.1

STANDARD DEVIATIONS AND CRITERION VALUES FOR DETERMINING THE "SUBSTANTIALITY" OF EACH SINGLE-FACTOR ANALYSIS RESULT INVOLVING OUTCOME PRE SCORES (1968-69)

Outcome Measure	Standard Deviation 1968-69	Criterion (1/3 SD) 1968-69
SB	14.5	4.8
PSI	11.7	3.9
AH-WPPSI	13.0	4.3
GUMP	7.1	2.4
BP	1.1	.4
MP	2.0	.7
FI	1.1	.4

Finally, it was decided to highlight primarily those consistent results which involved monotonic relationships between input and outcome variables. This criterion was adopted since instances of psychological theorizing which assume non-monotonic relationships among variables are rare, and the possibilities that non-monotonic relationships in the present data resulted from chance or confounding factors were judged to be high.

Primary emphasis is placed upon the 1968-69 data, but replications from the 1967-68 data are presented when possible.

B. Child Demographic Variables and Initial Status

The tables will present the 1967-68 and 1968-69 results in detail, and the text will discuss 1968-69 results and indicate the correspondence between 1968-69 findings and those for the 1967-68 data. The sample considered for 1968-69 contained only children who had no more than a previous summer of Head Start experience and who were in classes where at least three-fourths of the children had no previous Head Start experience. The sample for 1967-68 included all children since only 13 children had codes which indicated more than a summer of Head Start experience.

Five outcome variables will be considered for the 1967-68 data. Four of them are identical to ones considered in the 1968-69 analyses while one is unique to the 1967-68 data. The four common outcome variables are the Stanford-Binet (SB), Behavior Problem Scale (BP), Motivation Problem Scale (MP) and Feelings of Inadequacy Scale (FI). The unique outcome variable is the Number of Initiations (NI) to an adult in the classroom by the child. As discussed previously, this outcome variable

was available from the Kansas Social Interaction Observation Instrument and was selected for analysis because previous research has shown that disadvantaged children tend to be suspicious of "other" adults and consequently are less likely to interact with them. Table VII.2 summarizes the relationships between the 18 independent variables and the pre scores for all five 1967-68 outcome measures. Also, some interesting findings for 1968-69 could not be replicated exactly on the 1967-68 data since somewhat different measures were involved. For example, an SES index was not constructed from the 1967-68 data, but mother's education and number of children in the home were chosen to reflect SES. Another example involves mother's aspirations and expectations where the 1968-69 measures reflected both occupational and education, and data for 1967-68 were available only for educational aspirations and expectations.

In the tables for 1968-69, mean differences statistically reliable at the .05 level of confidence are indicated by a cross (+); differences greater than 1/3 of the standard deviation of the E&R sample distribution of scores are indicated by an asterisk (*).1/

1. Age (Tables VII.3 & VII.4)

Six of the seven outcome measures were substantially related to age (Table VII.4). Older children had lower scores on the Binet, and better scores on the AH-WPPSI, GUMP, MP, and FI. A similar but nonsubstantial pattern was found on the BP scale. The findings for age are in agreement with the 1967-68 data in that older children had lower SBs and exhibited less problems on the three problem scales. The children in the oldest age group initiated significantly more interactions with adults than the three younger age groups.

The SB scores are age standardized; there is no reason to expect an association with age. This points to some other variable, correlated with both age and Binet performance, which is likely also to be accounting for some of the age relationships with other data. The other six measures for 1968-69 are raw scores uncorrected for age, and the findings thus support the common expectation that with increases in age come increases in various problem-solving skills, achievement motivation, and adjustment to testing situations.

Cross-sectional data such as these frequently show a progressive decline with age of normative scores for low-income populations. Such data contributed to the cumulative deficit hypothesis: that at birth, disadvantaged children and advantaged children are similar in developmental status and predicted potential. The ever-widening disparity on measures of academic achievement and cognition is interpreted to include a

1/ The significance of the BP, FI, and MP scales was tested by Chi-squares derived from contingency tables since the distributions were highly skewed.

Table VII.2

RELATIONSHIPS BETWEEN MEAN PRE SCORE ON SB, BP, MP, FI, NI, AND INDEPENDENT VARIABLES (1967-68)

Independent Variables	SB Mean	SB N	BP Mean	BP N	MP Mean	MP N	FI Mean	FI N	NI Mean	NI N
Child's Age	+		+		+		+		+	
30-47 months	94.27	493	.81	482	2.15	482	1.09	482	4.13	316
48-53 months	91.60	622	.70	616	1.89	616	1.02	616	3.19	410
54-59 months	88.95	469	.53	468	1.69	468	0.96	468	3.69	307
60-77 months	89.35	195	.33	194	1.52	194	0.78	194	5.02	115
Pooled Variance	185.16									
Child's Sex			+						+	
Male	91.16	895	.70	889	1.83	889	0.95	889	3.41	575
Female	91.55	884	.59	871	1.91	871	1.05	871	4.13	573
Pooled Variance	189.27									
Child's Ethnicity	+		+		+		+		+	
Black	90.17	1779	.74	868	1.87	868	0.91	868	3.26	511
White	93.70	570	.59	564	1.82	564	1.18	564	4.59	395
Mexican-American	90.54	115	.62	120	2.22	120	1.12	120	2.72	88
Polynesian	89.62	68	.64	66	2.03	66	0.80	66	3.64	44
Pooled Variance	186.57									
Education of Mother	+						+			
No school	78.33	9	.50	8	3.25	8	0.75	8	5.20	5
1-3 grade	85.74	23	.67	24	1.42	24	0.75	24	4.00	18
4-6 grade	82.86	104	.66	104	2.24	104	1.22	104	2.79	70
7-8 grade	88.73	244	.53	242	1.82	242	1.03	242	3.11	162
9-11 grade	91.22	612	.67	611	1.96	611	1.02	611	3.78	421
12-H.S. graduate	94.50	507	.61	493	1.77	493	1.00	493	4.14	326
Some college	96.08	61	.87	60	1.37	60	0.48	60	3.35	43
College graduate	96.29	17	.78	18	1.61	18	0.78	18	4.71	7
Pooled Variance	177.52									

(continued)

Table VII.2 (continued)

Independent Variables	SB Mean	SB N	BP Mean	BP N	MP Mean	MP N	FI Mean	FI N	NI Mean	NI N
Region										
North	91.97	516	.61	509	1.38	509	0.76	509	3.27	371
South	87.74	325	.48	319	1.71	319	0.82	319	3.29	78
Midwest	93.25	590	.76	594	2.20	594	1.33	594	4.00	442
West	90.86	377	.67	368	2.08	368	0.89	368	4.19	264
Pooled Variance	187.38									
Total Number of People at Home	+		+		+				+	
2-4	92.98	369	.81	361	1.71	361	0.93	361	4.44	252
5,6	92.13	559	.61	551	1.78	551	0.99	551	3.96	371
7,8	91.54	389	.57	389	1.87	389	1.03	389	3.61	246
9 or more	87.51	292	.54	291	2.35	291	1.06	291	2.55	204
Pooled Variance	186.11									
Pre-Aspirations										
H.S. or less (including vocational work)	+ 89.57	373	.60	370	2.06	370	+ 1.18	370	3.90	273
College	91.95	1211	.64	1197	1.83	1197	0.95	1197	3.61	778
Pooled Variance	184.35									
Pre-Expectations										
H.S. or less (including vocational work)	+ 90.32	1114	.63	1099	1.95	1099	+ 1.04	1099	3.74	759
College	95.25	375	.64	367	1.71	367	0.88	367	3.61	231
Pooled Variance	180.11									
Attendance	+		+		+		+		+	
5 out of 5	92.53	761	.70	760	1.92	760	1.01	760	4.09	546
4 out of 5	90.55	580	.50	570	1.64	570	0.89	570	3.43	334
3 or less out of 5	89.68	273	.62	264	2.03	264	1.13	264	3.21	154
Pooled Variance	184.31									

(continued)

Table VII.2 (continued)

Independent Variables	SB Mean	SB N	BP Mean	BP N	MP Mean	MP N	FI Mean	FI N	NI Mean	NI N
Parent Participation									+	
Attended	91.76	962	.61	949	1.91	949	1.03	949	3.91	646
Did not	90.04	342	.64	340	1.81	340	1.03	340	2.90	242
Pooled Variance	185.92									
Teacher Credentials	+		+				+		+	
Elementary	96.52	27	2.22	27	2.52	27	1.57	27	5.74	27
High school	92.52	358	0.60	353	1.88	353	1.06	353	2.52	220
Bachelors	91.93	830	0.70	826	1.90	826	1.05	826	4.20	574
Masters	95.22	128	0.42	122	1.56	122	0.75	122	4.15	88
Pooled Variance	185.66									
Teachers Paid Experience with Pre-Schoolers	+				+		+		+	
Less than 6 months	96.02	116	0.57	113	1.89	113	0.87	113	5.18	66
6 months-1 year	91.90	213	0.64	215	1.74	215	0.92	215	2.52	143
1-3 years	91.94	716	0.65	699	1.73	699	1.01	699	4.01	521
4-5 years	91.23	163	0.88	164	2.43	164	1.36	164	5.01	97
Over 6 years	94.83	135	0.75	137	2.18	137	1.10	137	2.55	82
Pooled Variance	186.87									
Teachers Experience with Disadvantaged	+								+	
Less than 6 months	95.55	125	0.47	122	1.77	122	0.91	122	5.97	64
6 months-1 year	92.40	241	0.61	243	1.55	243	0.85	243	3.00	173
1-3 years	91.90	735	0.72	727	1.99	727	1.12	727	3.95	510
4-5 years	90.82	128	0.82	122	1.85	122	1.02	122	4.84	93
Over 6 years	95.02	114	0.69	114	1.99	114	1.04	114	1.78	69
Pooled Variance	185.16									

(continued)

Table VII.2 (continued)

Independent Variables	SB Mean	N	BP Mean	N	MP Mean	N	FI Mean	N	NI Mean	N
Class Factor CI										
Cognitive-Low Structure					+		+			
Low	92.78	415	0.65	406	1.64	406	0.97	406	4.46	288
Medium	91.05	811	0.63	810	1.87	810	0.91	810	3.49	525
High	91.44	407	0.64	398	2.08	398	1.22	398	3.56	251
Pooled Variance	186.26									
Class Factor CII										
Routine and Rules										
Low	92.33	437	0.50	430	1.85	430	1.03	430	3.42	306
Medium	91.32	736	0.72	732	1.87	732	0.97	732	4.01	488
High	91.23	459	0.64	452	1.86	452	1.00	452	3.47	266
Pooled Variance	187.05									
Class Factor CIII										
Cognitive-High Structure			+		+				+	
Low	92.50	402	0.76	404	1.94	404	1.06	404	3.78	287
Medium	90.69	806	0.67	799	1.98	799	1.01	799	3.30	507
High	92.18	396	0.48	385	1.58	385	0.91	385	4.50	261
Pooled Variance	187.04									
Class Factor CIV										
Child Centered-Unstructured									+	
Low	91.08	432	0.74	418	1.71	418	0.99	418	5.15	306
Medium	91.39	792	0.62	785	1.99	785	1.04	785	2.85	485
High	92.33	421	0.57	423	1.79	423	0.91	423	3.78	281
Pooled Variance	186.87									
Square Feet per Child			+		+		+		+	
Less than 14	92.19	473	0.61	473	1.71	473	0.90	473	3.58	309
15-19	90.24	480	0.52	468	1.57	468	0.7	68	5.03	252
20-24	91.67	797	0.72	792	2.05	792	1.17	792	3.38	585
Pooled Variance	190.97									

+ Significant at .05 level.

cumulative deficit in realizing that potential, presumably due to experiential/environmental factors.

Schaefer's (1968) interpretation of such a deficit reported for Southern, black children on the SB in terms of average IQs at different grade levels by Kennedy, Van De Riet and White (1963) emphasized cumulative selection, i.e., the practice of "holding back" less able children. A follow-up study, retesting the same children after several years, confirmed Schaefer's prediction (Kennedy, 1969).

Another factor to be considered in interpreting the apparent decline in initial performance with age is that most older children were from the South. Relative to other regions of the country, southern children typically have lower mean scores. Comparison of the initial IQs by age and region in Table VII.3 suggests that IQ declines with age in all four regions, but that there was a sharp decline for older children in the South.

Table VII.3

CHRONOLOGICAL AGE AND INITIAL MEAN BINET SCORES BY REGION (Ns IN PARENTHESIS)

Region	Chronological Age (CA) in Months				
	31-37	48-53	54-59	60-72	CA Mean
North	96 (72)	93 (162)	87 (75)	--	50 mos.
South	86 (35)	84 (38)	87 (55)	78 (185)	58
Midwest	96 (10)	93 (25)	91 (21)	89 (84)	58
West	94 (68)	90 (122)	88 (45)	--	50

At present, we cannot offer a further interpretation. The gradual change is consistent with a cumulative deficit hypothesis; it is also consistent with the likelihood (a) that children enrolled at an early age are more able, perhaps coming from more upwardly mobile homes, and (b) that where the older child is deemed unready for public school, he may attend Head Start instead. The data from the South are not consistent with a cumulative deficit hypothesis and suggest that where Head Start takes the place of public school kindergarten, children who are enrolled earlier come from a different population within the economically eligible.

In our judgment there are no statistical techniques for overall analyses that can handle data with complexities such as these with full retention of psychological meaning. In many analyses of change in this report, adjustments are made for initial level of performance. Average CAs for group are given, subgroup analyses are made where feasible, and we have attempted to interpret the data in ways that will help remind the reader of these factors. The CAs are presented for each level of each explanatory variable in the gain tables presented in Chapters IX and X. This is particularly critical where raw scores are reported since variables working in opposed directions impinge on them: (a) the strong tendency for raw scores on these factors to improve as the child grows older--since they were intended as developmental measures, and (b) the tendency for older children to be less able than younger children due to probable non-developmental associates such as selection and to (probably) some developmental retardation cumulatively induced by the environment.

2. Sex (Table VII.4)

There were no substantial relationships between child sex and initial status on the outcome measures in either 1967-68 or 1968-69.

3. Ethnicity (Table VII.4)

These data need to be interpreted carefully since region, age, and, within the poverty guidelines, economic and social factors are confounded with ethnicity. The contribution of (a) these variables, and (b) those of any constitutional differences imposed by differentially poor nutrition or health care and possible genetic effects, on Binet performance can in no way be identified at this time. The data are presented primarily because of the interesting adjustment differences.

On the SB, Black and Mexican-American children scored initially lower (about 86; CA-56 months) than White and Polynesian children (about 93; CA-51 months). On the motivation problem scale, Black and Mexican-American children had substantially more difficulty in "getting with" the testing situation (about 1.8 motivation problems) than White or Polynesian children (about 1.2 motivation problems) despite the fact that the latter children were younger. Black, White, and Polynesian children gave relatively little evidence of lack of self-confidence (about .7 problems for Black and White, .6 for Polynesians) while Mexican-American children seemed least self-confident in the testing situation (1.2 problems).

The ethnicity data for 1967-68 do not clearly follow the trends in the 1968-69 data. The samples for the two years differed in that children in the 1967-68 data base with more than a summer of Head Start experience could not be screened out due to missing information and that ethnicity for 1967-68 is confounded in a different manner with other variables.

Table VII.4

RELATIONSHIP BETWEEN CHILD'S MEAN PRE SCORE ON SB, PSI, AH-WPPSI, GUMP, BP, MP, FI AND CHILD'S INDEPENDENT VARIABLES (1968-69)

Independent Variables	SB Mean	SB N	PSI Mean	PSI N	AH-WPPSI Mean	AH-WPPSI N	GUMP Mean	GUMP N	BP Mean	BP N	MP Mean	MP N	FI Mean	FI N
Child's Age	*+		*+		*+		*+		+		*+		*+	
31-47 months	92.81	231	28.81	214	16.56	222	32.35	211	.75	216	2.16	217	.98	217
48-53 months	90.45	426	31.31	390	19.28	416	33.03	387	.74	394	1.57	394	.82	394
54-59 months	87.80	244	35.23	237	21.95	248	34.18	230	.52	228	1.50	227	.54	229
60-72 months	81.01	334	38.62	321	27.70	329	36.75	332	.58	328	1.40	328	.66	329
Pooled Variance	195.92		123.36		164.98		47.78							
Child's Sex			+						+					
Male	87.23	635	32.92	597	20.40	619	33.88	593	.69	604	1.63	605	.72	607
Female	88.45	602	34.51	567	22.86	598	34.51	569	.62	564	1.61	563	.79	564
Pooled Variance	214.52		136.24		167.97		50.50							
Child's Ethnicity	*+		*+		*		+				*+		*+	
Black	86.00	823	34.48	761	21.52	795	34.66	777	.69	777	1.76	777	.74	780
White	92.82	222	35.14	224	22.61	231	33.13	202	.62	210	1.21	210	.70	210
Mexican-American	85.64	61	33.30	56	30.34	58	33.37	60	.43	60	1.98	60	1.20	60
Polynesian	94.31	75	28.62	71	17.37	76	32.79	73	.49	70	1.07	70	.59	70
Pooled Variance	206.25		130.27		163.34		50.25							
Child's Previous Head Start Experience	*+		*+		*+		+				*+		*	
None	88.57	1420	32.46	1298	20.82	1378	34.10	1293	.68	1322	1.58	1321	.74	1322
2-3 months	91.16	157	40.82	162	25.90	165	35.50	155	.77	151	2.03	152	.79	154
9-10 months	90.04	111	35.26	109	20.23	109	35.71	108	.51	110	1.03	110	.48	110
12 months or more	95.04	68	37.32	68	23.24	71	35.74	66	.73	66	1.55	65	.68	66
Pooled Variance	203.30		126.86		165.42		48.79							

(continued)

Table VII.4 (continued)

Independent Variables	SB		PSI		AH-WPPSI		GUMP		BP		MP		FI	
	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
Child from Region	*+		*+		*+		*+							
North	91.30	381	30.54	358	18.32	358	32.98	341	.76	368	1.73	368	.78	368
South	81.59	404	37.57	406	24.01	423	35.18	418	.56	393	1.72	393	.74	396
Midwest	90.38	156	37.68	139	28.79	145	36.32	141	.53	148	1.28	148	.64	148
West	90.50	296	29.85	261	18.57	291	33.04	262	.70	259	1.49	259	.78	259
Pooled Variance	196.24		123.53		156.17		49.04							

\+ Significant at .05 level.

* Substantial Results.

4. Child's Previous Head Start Experience (Table VII.4)

The amount of previous Head Start experience is shown to be related to pre scores on SB, PSI, AH-WPPSI, MP, and FI. Specifically, the pattern of results was as follows: (a) the greater the amount of previous HS experience, the higher SB pre scores; (b) children with no previous HS experience received the lowest scores on the PSI and the next lowest scores on the AH-WPPSI. Children with 2-3 months of HS (summer) obtained the highest scores on these two measures; (c) children with 9-10 months of previous HS experience obtained lower scores (better adjustment) on MP and FI than any of the remaining experience groups (those with more and those with less experience).

The SB difference was 2.6 points between no experience and just the summer program, about the same for 9 to 10 months of prior experience, and 6.5 points for a full-year or more of prior experience. While selection could be responsible for some of these changes (children with prior experience were enrolled at an earlier age), and younger children tend to have higher scores on the SB at age of entry, they are also consistent with identification of a linear relation between amount of preschool experience and achievement. Comparison of the gains of children with and without prior Head Start to be reported in Chapter IX will elucidate the contribution of an additional year of Head Start vs. possible selection effects.

5. Region (Table VII.4)

There were substantial regional differences on all three cognitive measures and the GUMP for 1968-69. Children in the South scored considerably below (9-10 points) children in other regions on the SB, while children in the South and Midwest scored higher on the PSI, AH-WPPSI and GUMP. However, the children in the Midwest and South were older than children in the North and West and hence could be expected to have higher raw scores on these three measures. For 1967-68, children in the South scored lower on the SB but not as low as in 1968-69.

C. Family Demographic Variables

1. Socioeconomic Status (Table VII.5)

Of the seven measures, only pre scores on the SB were substantially related to SES. As expected, the higher the SES, the higher the SB pre scores. The difference in mean pre scores between the low-low group and the upper-low group was 10.3 IQ points. While not substantial, PSI and AH-WPPSI pre scores were also positively associated with level of SES. The data for the two socioeconomic indicators for 1967-68, education of mother and total number of people at home, supported these findings.

Table VII.5

RELATIONSHIP BETWEEN CHILD'S MEAN PRE SCORES AND FAMILY DEMOGRAPHIC VARIABLES

Independent Variables	SB		PSI		AH-WPPSI		GUMP		BP		MP		FI	
	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
Socioeconomic Status	*+		+											
Low-Low	81.68	240	32.00	232	20.47	247	33.75	235	.61	226	1.62	226	.81	226
Middle Low	86.17	270	33.86	264	21.34	276	34.03	268	.66	262	1.60	262	.65	262
Upper Low	91.95	285	35.29	277	23.19	279	33.94	279	.58	283	1.40	283	.62	283
Pooled Variance	201.95		144.45		173.95		57.26							
Family Structure					+				*+					
Mother only	87.73	303	33.58	281	19.34	297	33.84	281	.80	285	1.53	285	.64	285
Mother and Father	87.85	511	33.58	503	22.28	517	33.97	502	.54	496	1.54	496	.75	496
And Adult Relations	86.84	64	32.94	63	23.44	66	33.87	63	.39	62	1.69	62	.94	62
Pooled Variance	209.00		140.29		168.28		52.06							
Mother's Employment Status			*+											
Not working, not looking	87.13	514	31.75	489	20.52	509	33.36	483	.70	490	1.59	490	.78	490
Not working but looking	86.93	150	35.38	144	21.41	150	34.28	144	.77	143	1.46	143	.69	144
Yes, part-time	86.78	129	36.69	121	23.28	125	34.38	125	.51	124	1.63	124	.79	124
Yes, full time	88.09	211	35.15	210	23.49	218	34.99	216	.50	207	1.41	207	.56	207
Pooled Variance	210.51		137.22		167.32		52.29							
Number of Family Moves in Last 3 Years					+									
None	86.80	418	34.68	405	23.47	416	34.36	414	.56	408	1.47	408	.70	409
One	87.05	249	33.16	238	20.81	248	33.70	237	.60	240	1.50	240	.73	240
Two or more	87.36	230	32.41	220	19.58	233	33.29	218	.70	213	1.60	213	.60	213
Pooled Variance	208.61		140.80		169.08		52.11							

\+ Significant at .05 level.

* Substantial Results.

While interpretation of the raw score data is difficult, since low-low children were predominantly older, black and from southern states, the SB data suggest that even within the OEO Guidelines, differences in life circumstances associated with family economic status are associated with the child's performance on measures of cognitive development and achievement.

2. Family Structure (Table VII.5)

Of the seven outcome measures, only BP was substantially related to family structure. Children living only with their mother showed more behavior problems in the testing situation than children living with both their mother and father or with mother, father, and one other adult relative. These data suggest that the problems experienced by children from mother-only families are more related to behavior control (and classroom management) than cognitive deficiencies. The results from AH-WPPSI showed that children living only with their mother tended to achieve lower pre scores on rate of learning a new task than children in either of the other two family situations.

3. Mother's Employment Status (Table VII.5)

The PSI was the only measure substantially related to mother's employment status; the raw score was lower for children whose mothers were not working and not looking. The PSI finding could be partly due to the fact that at-home mothers had younger children (CA-53 months) than mothers looking for work (CA-55 months), employed part-time (CA-55 months) or employed full time (CA-55 months).

4. Number of Family Moves in Last Three Years (Table VII.5)

No substantial relationships were found between the number of family moves and the seven outcome measures. However a significant result involved AH-WPPSI, suggesting that an increase in number of moves was associated with lower pre scores on a rate-of-learning measure.

D. Parent-Child Psychological Interaction Variables

1. Mother's Aspirations Concerning Child's Educational-Occupational Advancement (Table VII.6)

Four of the seven measures were substantially related to level of mother's aspirations, and six of the seven measures were related in a way that is consistent with common expectations. Specifically, SB, PSI, AH-WPPSI, and GUMP pre scores were both substantially and positively associated with mother's level of aspiration for her child.

Table VII.6

RELATIONSHIP BETWEEN CHILD'S MEAN PRE SCORES AND PARENT-CHILD PSYCHOLOGICAL INTERACTION VARIABLES

Independent Variables	SB Mean	N	PSI Mean	N	AH-WPPSI Mean	N	GUMP Mean	N	BP Mean	N	MP Mean	N	FI Mean	N
Mother's Aspirations for Child's Education and Job	*+		*+		*+		*+							
Low	81.76	90	27.54	83	16.88	85	30.54	83	.70	86	1.59	86	.63	86
Medium	86.91	403	32.36	388	21.05	401	33.31	386	.62	391	1.61	391	.70	392
High	88.58	541	35.67	524	22.77	548	34.96	531	.62	518	1.51	518	.76	518
Pooled Variance	206.84		134.33		168.46		50.36							
Mother's Expectations for Child's Education and Job	*+		*+		+		+							
Low	86.17	600	31.97	578	20.52	600	33.44	582	.63	582	1.65	582	.76	583
Medium	88.09	309	35.68	297	23.22	308	34.39	295	.65	296	1.43	296	.63	296
High	91.36	121	37.48	116	23.33	122	35.53	119	.57	113	1.35	113	.73	113
Pooled Variance	206.66		135.88		169.36		51.64							
Parent's Pessimism Towards Life Scale	*+		+											
Low	92.44	293	35.45	286	22.74	298	33.99	278	.65	278	1.47	278	.61	278
Medium	86.32	458	32.88	432	21.72	453	34.00	438	.57	439	1.53	439	.72	440
High	84.06	289	33.09	283	20.40	289	33.96	289	.70	284	1.65	284	.84	284
Pooled Variance	197.58		139.68		169.91		52.16							
Accessibility of Adults	+		+		+						+			
Low	86.11	268	32.43	256	19.60	265	33.63	258	.72	260	1.69	260	.77	261
Medium	86.94	555	33.45	539	21.84	559	34.13	541	.60	533	1.62	533	.75	533
High	89.93	217	35.67	207	23.40	217	34.01	207	.61	208	1.18	208	.57	208
Pooled Variance	207.90		139.29		168.47		51.99							

(continued)

Table VII.6 (continued)

Independent Variables	SB Mean	N	PSI Mean	N	AH-WPPSI Mean	N	GUMP Mean	N	BP Mean	N	MP Mean	N	FI Mean	N
Frequency of Reading to Child			*+				*		*					
Seldom or never	85.87	64	28.54	59	20.29	62	32.42	59	.97	59	1.68	59	.51	59
At least once a week	88.25	219	33.78	210	20.57	222	33.76	212	.64	213	1.46	213	.75	213
Several times a week	87.54	338	33.51	325	21.93	342	34.03	324	.55	320	1.65	320	.70	321
At least once a day	88.05	277	34.83	267	22.70	272	35.13	271	.59	268	1.49	268	.78	268
Very frequently	88.88	69	39.68	68	22.04	68	34.06	68	.64	69	1.46	69	.65	69
Pooled Variance	204.12		130.65		172.24		50.71							
Extent of Physical Control	*+		*+		*+									
No physical	89.00	390	35.20	372	23.40	384	34.32	377	.63	379	1.49	379	.67	379
Mild physcial	86.75	293	33.35	288	21.50	301	34.13	283	.61	283	1.57	283	.82	283
Severe	79.53	70	30.76	59	18.56	63	34.55	64	.77	65	1.42	65	.55	65
Pooled Variance	201.78		139.10		171.19		50.25							
Parent's Disciplinary Reactions to Minor Rule Infractions	+													
Do nothing	87.31	89	34.49	88	21.44	90	34.56	89	.67	87	1.53	87	.74	87
Punish minor infractions	85.96	399	32.74	384	20.80	403	34.19	380	.66	380	1.48	380	.67	380
Punish negative behavior and shape positive behavior	87.71	154	34.09	148	21.94	151	34.21	148	.67	150	1.55	150	.76	150
Shape positive behavior	89.35	285	34.99	274	22.94	284	34.13	279	.58	276	1.48	276	.68	277
Pooled Variance	205.88		139.66		172.85		50.47							

+ Significant at .05 level.

* Substantial Results.

Inspection of the data also suggests that the largest difference among aspiration levels on the four measures yielding substantial results occurred between mothers with low aspirations and those with medium aspiration levels. The 1967-68 aspiration findings supported these findings.

2. Mother's Expectations Concerning Child's Educational-Occupational Advancement (Table VII.6)

The results concerning mother's expectations basically parallel those for mother's aspirations. Specifically, both SB and PSI pre scores were substantially and positively associated with mother's level of expectation. While nonsubstantial, both AH-WPPSI and GUMP were also positively related to mother's expectations and there were fewer problems in all areas of adjustment to the test situation noted for mothers with high expectations for their child. The expectation data for 1967-68 supported these findings.

3. Parent Pessimism Toward Life Scale (Table VII.6)

While a substantial relationship was found with only one measure, SB, consistent findings were also obtained with PSI, AH-WPPSI, MP, and FI. The lower the parent pessimism, the higher the child's pre scores on the SB, PSI, and AH-WPPSI, and the lower the rated incidence of adjustment problems (MP and FI). The difference in mean pre scores on the SB was particularly noteworthy, with parents rated low in pessimism having children who scored 8.4 points above the children of parents who were rated high in pessimism.

4. Accessibility of Adults (Table VII.6)

Despite the fact that none of the obtained relationships of adult accessibility with the outcome measures was substantial, all but GUMP were consistent with common expectations. That is, the more accessible adults who were in the home, the higher were pre scores on SB, PSI, and particularly the AH-WPPSI, and the lower the incidences of adjustment problems on BP, MP, and FI.

5. Frequency of Reading to Child (Table VII.6)

The PSI was significantly and substantially related to how frequently parents stated they read to their children, and all measures yielded findings consistent with common expectations. That is, the children's SB, PSI, and AH-WPPSI pre scores generally were higher and ratings on MP and BP were lower (better) for parents who stated that they frequently read to their children than for parents who said they seldom read to their children. Differences on the PSI were particularly striking, with children who were "seldom or never read to" obtaining PSI pre scores approximately 11 points below the pre scores of children whose parents stated that they "very frequently read" to their children.

6. Extent of Physical Control (Table VII.6)

Substantial relationships between use of physical punishment and each of the three cognitive measures were found. Specifically, the more severe were the methods of control, the lower the mean pre scores on SB, PSI, and AH-WPPSI. The differences with respect to SB pre scores were particularly striking, with children of parents who reported that they do not use physical punishment obtaining pre scores 9.5 points above those obtained by the children of parents who reported using severe physical punishment. Interestingly, none of the noncognitive measures was related to the use of physical punishment.

7. Parent Disciplinary Reactions to Minor Rule Infractions (Table VII.6)

None of the outcome pre scores was substantially related to parent disciplinary reactions, though a pattern of findings consistent with those obtained with extent of physical control was obtained. Specifically, all three of the cognitive measures yielded the lowest pre scores from children whose parents stated that they punish minor infractions. The highest pre scores were obtained from children whose parents stated that they simply reward positive behavior. No patterns were discernible involving the different types of parental disciplinary reactions to minor infractions and pre scores on the four noncognitive measures.

E. Summary and Interpretation

The results of Chapter VII are impressive and readily interpretable. All but a few results reflected a pattern consistent with common expectations.

1. Family Variables

The findings regarding the relationships of family variables to the children's pre scores on the seven measures are particularly consistent and compelling. Basically, children who perform most poorly on the cognitive and noncognitive measures are from homes: (a) that are ranked in the lowest SES; (b) where mothers are not working or not looking for work; (c) where parents (mothers) have low aspirations and expectations for their children and are pessimistic about life; (d) where parents are relatively inaccessible to their children and spend little time in matters such as reading to their children; and (e) where the nature of disciplinary interactions is characterized by severe punishment and little reward.

2. Child Variables

The patterns of findings concerning child demographic variables are somewhat less readily interpretable than those regarding family variables. First, regarding the relationships of age to the various cognitive measures, it is a common finding in cross-sectional studies that the IQs of children from impoverished environments show a small but steady decline with age. In the present study, however, age was confounded with the children's ethnicity, region, and other variables, some of which were found to be related to IQ in a manner consistent with the age-IQ results. Thus, the obtained decrements of IQ on SB pre scores with age seem most plausibly to reflect selection variables, particularly in regard to the Southern nursery school-age children versus kindergarten-age children.

While the finding that PSI and AH-WPPSI pre scores increased with age is entirely consistent with assumptions about the growth in intellectual ability with age, the obtained means actually may be somewhat depressed at the upper ages, since again, ethnicity, region, selection and several other variables confounded with age could reasonably be expected to lower the obtained results with the older samples.

With regard to the association of ethnicity and SB performance, the 1968-69 analyses suggested that the lower initial scores for Black and Mexican-American children probably were associated with the higher incidence of motivation problems in these children and the additional handicap of lack of self-confidence in the testing situation reported for the Mexican-American children. This may be in part attributable quite realistically to language problems as well as to less self-confidence. These data suggest that SB scores within a low income group are complexly influenced by economic factors, child-rearing practices, and motivational factors. At our present state of knowledge of the relevant factors, how to measure them and how to assess their relative contributions, it is premature to interpret SB differences associated with ethnicity simply as a consequence of innate, genetically determined factors.

The finding that the greater the amount of previous Head Start experience, the higher the SB pre scores, is consistent with recent results from a more highly controlled study by Beller (1971). However, the previous experience-SB findings in this Head Start report must be tempered by the apparently inconsistent pattern found with the remaining cognitive measures and by the small Ns involved in the cases of children who had at least some previous experience. While children with no prior experience consistently were lowest, children with prior summer experience had higher scores on the PSI and AH-WPPSI than those with prior full-year experience.

The increments in GUMP pre scores with age and the decrease in adjustment problems (MP, BP, and FI) with age basically substantiate common expectations (and thereby serve as validating evidence for these four scales) that with an increase in age, children's motivation to achieve increases, and motivational and personality factors such as impulsivity and lack of confidence as a problem solver give way to more mature, task-oriented behaviors. Again, however, the relationships of age to at least one of these scales, MP, might have been even more impressive, had not age been confounded with ethnicity and region. Specifically, the Black sample, which was older than the other ethnic subgroups, obtained relatively higher scores on MP. Thus, mean MP pre scores for the older age groups may have been elevated more (achieving substantiality) than would have been the case had the older groups had the same ethnic mix as the younger age groups. It is interesting that Emmerich (1971), with direct classroom observations of the free play behavior of a similar sample of Head Start children, found substantial changes over the fall to mid-winter period, but relatively few initial age differences. Older children tended to be less submissive, withdrawn and distrusting and engaged in more cognitive activity.

As was the case with the family variables, a majority of the substantial findings in the child variable analyses involved the cognitive measures, though the proportion of substantive findings involving the noncognitive measures was relatively greater for child than family variables.

CHAPTER VIII

INTERRELATIONSHIPS AMONG CHILD-FAMILY, TEACHER AND PROGRAM FACTORS (1968-69); AND THE RELATIONSHIPS OF TEACHER AND PROGRAM FACTORS TO PRE SCORES (1967-68 AND 1968-69)

The first section of this chapter summarizes the interrelationships among child-family, teacher, and program factors for 1968-69 since they aid in the interpretation of subsequent analyses by making explicit the confounding of important subsets of variables in the 1968-69 data. For example, it will be seen that the children exposed to a high level of classroom structured lessons were predominantly older, Black children from the South. This must be kept in mind when interpreting the effects of classroom structure lessons on child gains. The tables are presented in Appendix D. These interrelationships were not generated for the 1967-68 data since the analyses on the 1967-68 data were less extensive. The second section of this chapter describes the relationship of teacher and program factors to the pre scores of the outcome variables for both 1967-68 and 1968-69.

A. Child-Family, Teacher and Program Factors for 1968-69

1. Regional Differences

The centers were grouped in four regions as follows: Region 1 - North (418 children); Region 2 - South (452 children); Region 3 - Midwest (165 children); and Region 4 - West (345 children).

a. Ethnicity (Table D-1)

There were significant differences between the regions in the proportion of children in the different ethnic groups. The percentage of black children ranged from 82% in the South to 48% in the West. The percentage of whites varied from 31% in the North to 7% in the West. All the Polynesian children in the sample were from the West and the Mexican-American children were from the Midwest and West.

b. Sex (Table D-2)

There were no significant differences in the proportion of males or females between regions.

c. Age (Table D-3)

There is considerable difference in the age distribution among regions. Children 60 months of age or older accounted for 56% of the children in the South and 59% in the Midwest. In the North and West there were virtually no children 60 months of age and over. In the South and Midwest, Head Start is more likely to take the place of kindergarten.

d. Attendance (Table D-4)

The average number of days attended is greater for the South and Midwest than for the North and West.

e. Head Teacher's Ethnicity (Table D-5)

The percentage of children taught by black teachers varied from 60% in the South to 31% in the North; for white teachers the range was from 62% in the North to 29% in the West.

f. Head Teacher's Preparation (Table D-6)

All of the children from the Midwest and 72% of the children from the North were taught by teachers with at least a bachelor's degree. The corresponding figures were 28% for the South and 42% for the West.

g. Head Teacher's Training in Child Development (Table D-7)

Teachers who had some training in early child development taught 54% of the children in the North, 68% in the South, 59% in the Midwest, and 85% in the West.

h. OSCI Class Factor CI - Structured Lessons (Table D-8)

In the North and the West none of the children were in classes which offered high levels of structured lessons. In the South, all children were in classes with medium or high levels of structured lessons. Ninety-three percent of the children in the Midwest were in classes with medium or high levels of structured lessons.

i. OSCI Class Factor CII - Group Activities and Routines (Table D-9)

Over 77% of the children from the South and 82% of the children from the North were in classes with low or medium level of group activities and routines. The pattern in the Midwest and the West was different; 49% and 69% of the children in the Midwest and West, respectively, attended classes with low or medium level of group activities and routines.

j. OSCI Class Factor CIII - Social-Emotional Interaction (Table D-10)

In the North, 38% of the children were in classes with high levels of social-emotional interaction. The corresponding percentages were 9% in the South, 26% in the Midwest, and 24% in the West. On the other hand, nearly 50% of the children from the South were in

classes which had a low level of social-emotional interaction while only 9% of the children from the North were in such classes.

k. OSCI Class Factor CIV - Verbal Communication (Table D-11)

Among the children from the South and Midwest, 46% were in classes with low levels of verbal communication. Only 16% of the children from the North were in low verbal emphasis classes. On the other hand, 27% of the children from the North and 36% of the children from the West were in classes with a high level of verbal communication.

l. OSCI Class Factor CV - Instruction in Creative Art (Table D-12)

All the children from the West and 94% of the children from the Midwest were in classes with a high or medium level of instruction in creative art. The corresponding percentages were 62% for the North and 60% for the South.

m. OSCI Class Factor CVI - Language and Discrimination Learning (Table D-13)

Classes which were rated as high on language and discrimination learning accounted for 45% of the children from the North and 34% of the Midwest children. Only 16% of the children from the South and 11% of the children from the West were in such classes. Nearly 40% of the children from the South were in classes rated low on language and discrimination learning.

n. OSCI Teacher Factor TI - Social-Emotional Interaction (Table D-14)

Nearly 65% of the children in the North were taught by teachers who scored high on this factor. Teachers who scored high on social-emotional interaction taught 10% of the children from the South and 13% of the children from the West. The Midwest had no teachers scoring high on this factor.

o. OSCI Teacher Factor TII - Structured Lessons - (Large Group) (Table D-15)

Eighty-three percent of the Southern children were taught by teachers classified as high on this characteristic. None of the Northern children was taught by teachers scoring high on this characteristic.

p. OSCI Teacher Factor TIII - Art Activities (Table D-16)

All the children from the West were in classes taught by teachers rated high or medium on this characteristic. Most of the

children from the other regions were in classes taught by teachers rated low or medium on Teacher's Art Activities.

q. OSCI Teacher Factor TIV - Creative Instruction (Small Group) (Table D-17)

Approximately 36% of children in the North were taught by teachers rated low on this factor. This appears to substantiate the ratings on the corresponding class factor according to which 38% of the Northern children were in classes which were rated low on Instruction in Creative Art.

r. OSCI Teacher Factor TVI - Receptive Learning (Table D-18)

Among the children from the South, 52% were taught by teachers who were rated low on this characteristic, while in the West only 10% of the children had teachers rated low on this factor. Children in the North and West were more exposed to teachers rated high on this factor.

2. Differences Among Child's Ethnic Groups for 1968-69

For analyses discussed in this section, the children are classified into three ethnic groups: blacks, whites and others.

a. Sex (Table D-19)

The sex by ethnicity distribution indicates that among White children there was a substantial difference between the percentage of males (61%) and females (39%).

b. Age (Table D-20)

There are significant differences among ethnic groups in the proportion of children in the four age groups: 31-37 months, 48-53 months, 54-59 months, and 60-72 months. Among Black children, 32% are 60-72 months of age. The corresponding figures for Whites and other ethnic groups are only 12% and 19% respectively. This reflects the predominance of Black children in the South where public schools typically begin in first grade so that Head Start takes the place of kindergarten.

c. Head Teacher's Ethnicity (Table D-21)

Among the Black children, 55% were taught by Black teachers and 41% were taught by White teachers. White teachers taught 59% of the White children, while Black teachers taught only 32% of the White children. Among the children from other ethnic groups 43% were taught by White teachers.

d. Head Teacher's Training in Child Development (Table D-22)

The "Others" group of children had more teachers (79%) with training in child development compared to Whites (61%) and Blacks (65%).

e. OSCI Class and Teacher Factors (Tables D-23 to 28)

There are differences by ethnic group on the OSCI class and teacher factors. The direction of differences is related to regional differences (1) in ethnicity, which is a characteristic of Head Start populations and (2) in the nature of the experimental intervention programs directed by the E&R centers, which is unique to this study and not characteristic of Head Start. Some of these relationships are presented in the Tables. Black children were more often exposed to a high level of Class Structured Lessons and Group Activities and Routines, while White children were more often exposed to a high level of Class Language and Discrimination Learning and Teacher Social-Emotional Interaction.

3. Differences According to Child Age

For the purposes of the age analyses, children were grouped into four age groups: (1) 31 to 47 months, (2) 48 to 53 months, (3) 54 to 59 months, and (4) 60 to 72 months.

a. Attendance (Table D-29)

There were significant differences in attendance across the four age groups. Older children had significantly better attendance records.

b. Pre/Post Testing Intervals for PSI, SB, and GUMP (Tables D-30, A, B, and C

Older children were more likely to be exposed to the longest pre/post testing interval than were younger children.

c. Socioeconomic Status (Tables D-31 to 33)

Older children came from lower SES backgrounds as reflected by absence of telephone at home, low family income, and whether the child had his own bed.

d. Head Teacher's Ethnicity (Table D-34)

Older children had a higher probability of being taught by Black teachers. It should be noted that the South had the highest percentage of older children and Black teachers.

e. Teacher Qualifications (Table D-35)

Older children tended to be taught by less academically qualified teachers than the younger age groups.

f. OSCI Class and Teacher Factor Scores (Tables D-36 to 43)

In general, the oldest children, who were most likely to be black and from the South, were exposed to considerably different Head Start programs than were the younger children in the 1968-69 year, due to the E&R center interventions. The confounding of the child's age with the OSCI factors are presented in Tables D-35 to D-42. Older children were more likely to be exposed to high levels of Class and Teacher Structured Lessons and low levels of Class Verbal Communication, Teacher Social-Emotional Interaction, Teacher Creative Instruction, and Teacher Receptive Learning.

4. Differences Among Teacher Age Groups (Tables D-44 to 48)

For the purposes of the analyses in this section, the teachers were classified into three age groups: (a) 16 to 27 years, (b) 28 to 39 years, and (c) 40 years of age and older. Older teachers were more likely to be in classes rated as high on Class Structured Lessons, high on Class Group Activities and Routines, low on Class Social-Emotional Interaction, low on Teacher Social-Emotional Interaction and high on Teacher Structured Lesson.

5. Differences Among Teacher Preparation (Tables D-49 to 54)

There was a tendency for the most highly trained teachers to be found more frequently in the high level of Class Social-Emotional Interaction and Teacher Receptive Learning, and in the middle level of Class Structured Lessons, Class Verbal Communication, Class Language and Discrimination Learning and Teacher Structured Lessons.

B. Relationships of Head Start Program and Teacher Variables to Pre Score Measures

The analyses described in this chapter are primarily methodological, undertaken to identify the degree of initial confounding of pre-test performance with teacher and program variables. Statistically, inference would be easier if program characteristics were unrelated to child performance before Head Start participation and were related after participation.

Ascription of any initial relationship (confounding) is not easy. In 1968-69, most E&R Centers intervened in the programs. Some interventions would affect OSCI scores; some would not. Table VIII.1 summarizes

Table VIII.1

E&R CENTER INTERVENTIONS AND PROBABLE OSCI EFFECT
(1968-69)

Center	Intervention	Predicted OSCI Effects
Boston	Parent participation	None
Syracuse	None; Selected observations	None
Bank Street	None; Selected observations	None
Temple	Facilities and resources enrichment	Increased play
South Carolina	Buchanan reading readiness	Increased structure and language
Tulane	Buchanan reading readiness	Increased structure and language
Texas	Buchanan reading readiness	Increased structure and language
Kansas	Some behavior modification	None
Michigan	Piagetan curriculum	Increased language
UCLA	Teacher goal setting	None
Hawaii	Hawaii language curriculum	Increased language
Southern	Hawaii language curriculum	Increased language

the 1968-69 interventions and their expected OSCI effects. Region (and thus age, prior Head Start experience, ethnicity, and pre scores) would be expected to be confounded with program characteristics in 1968-69.

In 1967-68, programs were selected for representative variation. No predictions are made regarding confounding for this sample.

The problems of confounding, if any, are twofold. First, no current statistical techniques prove entirely satisfactory in adjusting

for comparison of post scores among groups differing initially for non-random reasons on the pre score. Under these circumstances, raw gains are often greatest for the group with the lowest initial score while adjusted gains are often in the opposite direction, greatest for the group with the highest initial score. One could, depending on the extent to which a priori belief in the efficacy of different treatments coincided with the effect of adjustment for initial differences, reach diametrically opposed conclusions. Second, where the initial confounding is known to be an accurate reflection of population characteristics, generalization of findings (once the first problem has been resolved as well as may be) to the population may be made with appropriate confidence. When non-typical bias is suspected (as in the 1968-69 data) but population characteristics are not accurately known, generalization of the findings must be made cautiously if at all.

There are no OSCI data for a random sample of Head Start classes; the 1967-68 data are believed to represent the range of natural variation but the extent to which the shapes of the distributions are typical is not known. The best anchor is the 1967-68 and 1968-69 Head Start census survey teacher self-reports on classroom emphases and characteristics.

1. 1967-68 Relationships (Table VII.2)

a. Parent Participation

This was significantly associated with pre scores of SB and NI (number of initiations). The children of attenders had a mean of 91.76 as compared to a mean IQ of 90.04 for nonattenders. They also averaged higher on the NI scale. In both cases, however, the relationships are not large.

b. Teacher's Credentials

The teachers with a Master's degree had in their classrooms 128 children with SB mean pre scores of approximately 95, compared with a mean of 92 for children who were placed in the classes of teachers with less training. There was also a trend for the more educated teachers to be assigned children with less behavior problems and lower Feelings of Inadequacy. The children assigned to the more educated teachers also had higher initiation rates (pre NI scores).

c. Teacher's Paid Experience with Preschool Children

Teacher's experience was associated by chance with pre scores of SB, FI, MP, and NI. Teachers with less than 6 months or more than 6 years of experience taught children with higher SBs.

d. Teacher's Paid Experience with Disadvantaged Children

Teacher's experience was by chance significantly related to SB and NI in a pattern similar to the previous variable.

e. OSCI Class Factor CI - Cognitive-Low Structure

This factor attempts to measure the extent of cognitive activities in a low structured class situation. The 1967-68 OSCI Factors were derived on the basis of class-centered information that was different from information collected in 1968-69 and cannot be directly compared to the 1968-69 OSCI factor scores. This factor and the other three OSCI factors were not related to the pre scores for SB which is in contradiction to the 1968-69 OSCI findings. The FI and MP scales were significantly related to Factor I. Classes falling into the high level (i.e., upper 25% of the factor score distribution) had children who tended to be higher (more problems) on FI and MP.

f. OSCI Factor CII - Routines and Rules

This factor was significantly associated only with the BP scale. The lowest level of this factor had the lowest mean (fewer problems) on the BP scale.

g. OSCI Factor CIII - Cognitive-High Structure

This factor reflects the amount of cognitive activity in highly structured class situations (e.g., receptive learning) and was significantly associated with BP, FI, and NI. The children to be exposed to the high level had less behavior problems and feelings of inadequacy and initiated more interactions.

h. OSCI Factor CIV - Child-Centered Unstructured

This factor was significantly associated only with NI. The children exposed to the low level of this factor had the highest interaction initiation rate.

i. Square Feet Per Child (Indoor Floor Space)

The amount of space per child was by chance significantly related to pre scores in BP, MP, FI and NI.

2. 1968-69 Relationships of Teacher and Program Factors to Child's Pre Scores

a. Child's Time in Program Variables

(1) Child's Attendance (Table VIII.2) - Two pre measures, PSI and AH-WPPSI, seemed to influence level of attendance. With respect

Table VIII.2

RELATIONSHIP BETWEEN CHILD'S MEAN PRE SCORE ON
SB, PSI, AH-WPPSI, GUMP, BP, MP, FI AND CHILD'S TIME IN PROGRAM VARIABLES
(1968-6?)

Independent Variables	SB Mean	SB N	PSI Mean	PSI N	AH-WPPSI Mean	AH-WPPSI N	GUMP Mean	GUMP N	BP Mean	BP N	MP Mean	MP N	FI Mean	FI N
Attendance														
Low:			*+		*+		+							
90 days or less	89.52	193	32.66	143	19.33	173	33.52	131	.81	144	1.63	144	.75	144
Medium:														
91-159 days	87.79	729	32.26	715	20.49	731	33.78	717	.62	711	1.58	711	.74	713
High:														
160-196 days	86.97	305	37.47	300	25.63	302	35.33	307	.67	308	1.72	308	.77	309
Pooled Variance	215.06		132.15		164.98		49.87							
Pre-Post Testing Interval														
Short:	*+						+		+				+	
22 weeks or less	91.72	172	34.24	270	22.26	281	34.75	335	.80	172	2.09	172	1.02	174
Medium:														
23-26 weeks	90.15	452	33.57	421	22.09	462	33.55	409	.60	449	1.52	449	.77	450
Long:														
27 or more weeks	84.57	370	33.53	243	21.14	261	34.67	237	.60	389	1.54	389	.67	389
Pooled Variance	194.08		136.61		177.06		49.33							

\+ Significant at .05 level.

* Substantial Results.

to both measures, the higher were their pre scores, the better the children's attendance. A similar but nonsignificant pattern was obtained with GUMP, while an opposite pattern was obtained with SB; namely, the lower were the SB pre scores, the longer the attendance. Finally, there was no obvious pattern of relationships between attendance and BP, MP, or FI.

(2) Child's Pre/Post Testing Interval (Table VIII.2) - Only SB pre scores were substantially related to length of the pre/post testing interval, with higher SB pre scores obtained for those children who had shorter pre-post testing intervals.

b. Teacher Demographic Variables

(1) Teacher's Age (Table VIII.3) - The teachers of different ages did not secure, by chance or confounding, children with substantially different pre scores. There was a tendency for the older teachers to secure a group of children with lower pre scores on the SB and AH-WPPSI and better scores on BP and MP.

(2) Teacher's Ethnicity (Table VIII.3) - The children assigned to Black and White teachers did not differ on the pre scores.

(3) Teacher's Years of Education (Table VIII.3) - Only SB pre scores were substantially associated with the years of a teacher's education. Children who were placed in the classes of teachers who had some professional preparation in early childhood education and development had slightly better pre scores on all variables except SB.

(4) Teacher's Preparation in Early Childhood Education and Development (Table VIII.3) - The results involving teacher preparation in early childhood education and development were consistent but not substantial. All but the SB yielded slightly better pre scores for children who were placed in the classes of teachers who had some professional preparation in early childhood education and development compared to those teachers who had none.

(5) Teacher's Residence in Head Start Neighborhood Table VIII.3) - Nonsubstantial but consistent findings were also obtained regarding whether or not the teacher was a resident of the Head Start neighborhood. The teachers who resided in the Head Start neighborhood were assigned children who scored slightly better on all pre scores of variables other than the SB.

Table VIII.3

RELATIONSHIP BETWEEN CHILD'S MEAN PRE SCORE
SB, PSI, AH-WPPSI, GUMP, BP, MP, FI AND TEACHER DEMOGRAPHIC VARIABLES
(1968-69)

Independent Variables	SB Mean	SB N	PSI Mean	PSI N	AH-WPPSI Mean	AH-WPPSI N	GUMP Mean	GUMP N	BP Mean	BP N	MP Mean	MP N	FI Mean	FI N
Teacher's Age	+				+				+		+			
16-27 years	88.16	348	34.33	344	24.00	340	34.24	326	.76	347	1.76	347	.81	349
28-39 years	88.60	379	32.86	386	20.89	386	34.01	401	.69	380	1.73	380	.77	380
40 years & above	85.57	377	33.15	376	20.30	378	34.24	381	.53	379	1.37	379	.65	379
Pooled Variance	211.78		131.81		164.22		51.02							
Teacher's Ethnicity	+		+		+						+			
Black	86.14	478	34.55	489	21.75	491	34.52	497	.61	480	1.60	480	.74	480
White	87.41	506	33.10	491	22.07	486	34.07	492	.72	505	1.75	505	.79	507
Pooled Variance	211.19		126.45		168.93		50.94							
Teachers Years of Education	*+				+						+		+	
H.S. and/or some college	83.64	421	33.75	429	21.32	433	34.21	442	.55	426	1.42	426	.70	426
2-4 years of college	89.70	447	33.42	438	22.39	433	34.36	426	.75	444	1.87	444	.87	446
Work beyond 4 year degree	89.87	236	32.81	239	20.87	238	33.70	240	.67	236	1.53	236	.57	236
Pooled Variance	199.16		132.33		168.25		50.95							
Teacher's Preparation In Early Childhood Education	+				+						+			
None	88.04	383	32.77	372	20.93	367	33.94	366	.68	385	1.77	386	.80	387
Some	87.11	721	33.74	734	22.00	737	34.26	742	.65	721	1.54	720	.71	721
Pooled Variance	211.07		131.88		169.18		50.96							

(continued)

Table VIII.3 (continued)

Independent Variables	SB Mean	SB N	PSI Mean	PSI N	AH-WPPSI Mean	AH-WPPSI N	GUMP Mean	GUMP N	BP Mean	BP N	MP Mean	MP N	FI Mean	FI N
Teacher's Residence in Head Start Neighborhood	+		+		+		+		+		+		+	
Resides in H.S. neighborhood	84.33	316	35.40	322	23.24	326	35.07	330	.53	316	1.42	316	.60	316
Does not reside in H.S. neighborhood	88.44	646	32.57	643	21.22	637	33.87	634	.69	647	1.69	647	.76	649
Pooled Variance	220.79		132.44		171.47		52.02							

\+ Significant at .05 level.

* Substantial Results.

c. Teacher Interaction Variables

(1) Teacher Stimulation (Table VIII.4) - Of the seven measures, only the PSI was substantially related to the level of teacher stimulation, with those judged to provide more stimulation having children who initially scored higher on the PSI.

(2) Teacher Sensitivity to Individual Differences (Table VIII.4) - There was no substantial finding regarding teacher sensitivity to individual differences.

(3) OSCI Teacher Factor TI - Social-Emotional Interaction (Table VIII.4) - The PSI, AH-WPPSI, and GUMP were significantly and negatively associated with teacher social-emotional interaction. Similarly, higher levels of teacher social-emotional interaction were associated with higher (worse) ratings on BP. However, the children's SB pre scores were positively associated with level of teacher social-emotional interaction.

(4) OSCI Teacher Factor TII - Structured Lessons (Large Group) (Table VIII.4) - Children with higher pre scores on the SB and lower pre scores on the PSI and AH-WPPSI were assigned to teachers with a low level of Structured Lessons.

(5) OSCI Teacher Factor TIII - Emphasis on Art Activities (Table VIII.4) - There were no substantial differences in child assignment as regards teacher emphasis on art activities.

(6) OSCI Teacher Factor TIV - Creative Instruction (Small Group) (Table VIII.4) - Three of the seven outcome measures, SB, FI, and BP, were by chance associated with levels of teacher emphasis in small groups and creative materials. The higher the level of OSCI-TIV, the higher were SB pre scores, and the highest level of OSCI-TIV had children whose pre scores on BP and FI were significantly higher (poorer) than teachers rated either low or medium on OSCI-TIV.

(7) OSCI Teacher Factor TV - Routines (Table VIII.4) - None of the measures were substantially associated with teacher routines.

(8) OSCI Teacher Factor TVI - Receptive Learning (Table VIII.4) - Only SB pre scores were substantially related to level of teacher emphasis on receptive learning. Teachers who later emphasized receptive learning in their classes were assigned (by chance) children with high pre scores on SB and low pre socres on PSI, AH-WPPSI and GUMP.

Table VIII.4

RELATIONSHIP BETWEEN CHILD'S MEAN PRE SCORE ON
SB. PSI, AH-WPPSI, GUMP, BP, MP, FI, AND TEACHER INTERACTION VARIABLES
(1968-69)

Independent Variables	SB		PSI		AH-WPPSI		GUMP		BP		MP		FI	
	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
Teacher Stimulation			*+		+				+					
Low	86.65	257	30.52	256	19.36	255	34.23	244	.56	256	1.32	256	.64	256
Medium	88.58	558	34.56	575	22.69	574	34.00	574	.69	561	1.79	561	.77	564
High	86.88	339	34.53	322	21.94	325	34.55	337	.66	339	1.55	339	.80	339
Pooled Variance	213.04		134.11		169.80		50.80							
Teacher's Sensitivity to Individual Differences														
Low	88.25	268	34.82	283	22.66	284	34.92	276	.65	268	1.80	268	.84	270
Medium	87.11	571	33.50	562	22.00	562	34.20	561	.66	571	1.51	571	.69	572
High	88.13	315	32.86	308	20.43	308	33.60	318	.66	317	1.65	317	.78	317
Pooled Variance	213.57		136.40		170.83		50.63							
Teacher Factor-TI Social Emotional Interactions	+		+		+		+		+					
Low	86.91	249	33.56	281	23.52	237	35.19	240	.46	249	1.57	249	.74	249
Medium	87.15	504	33.28	505	21.06	513	34.41	512	.67	503	1.40	503	.69	503
High	89.97	277	30.91	281	19.38	275	32.98	275	.76	283	1.72	283	.84	283
Pooled Variance	212.34		125.33		167.24		51.35							
Teacher Factor-TII Structured Lessons (Large Group)	*+		*+		*+		+							
Low	90.34	234	31.44	233	19.43	231	33.77	225	.61	236	1.50	236	.73	236
Medium	90.28	509	31.67	505	20.11	502	33.62	500	.73	512	1.49	512	.72	512
High	81.51	287	35.51	287	24.38	292	35.51	302	.52	287	1.62	287	.78	287
Pooled Variance	198.44		123.44		165.20		51.29							

(continued)

Table VIII.4 (continued)

Indepdenent Variables	SB Mean	N	PSI Mean	N	AH-WPPSI Mean	N	GUMP Mean	N	BP Mean	N	MP Mean	N	FI Mean	N
Teacher Factor-TIII Art Activities			+		+		+							
Low	88.61	280	34.68	272	22.65	271	35.43	264	.75	281	1.49	281	.76	281
Medium	87.15	470	32.27	462	21.05	459	33.54	467	.60	474	1.54	474	.75	474
High	87.89	274	31.22	286	19.59	290	34.07	290	.61	274	1.56	274	.72	274
Pooled Variance	212.24		123.87		222.89		51.39							
Teacher Factor-TIV Creative Instruction (Small Group)	*+		+		+		+ -		*+				+	
Low	85.19	296	31.19	288	20.36	289	33.68	295	.62	301	1.56	301	.74	301
Medium	87.88	540	33.83	547	22.12	545	34.88	549	.55	538	1.39	538	.66	538
High	91.91	184	31.69	180	19.67	181	33.20	173	.92	186	1.80	186	.97	186
Pooled Variance	209.13		124.10		168.49		50.77							
Teacher Factor-TV Routines			+										+	
Low	87.40	217	33.13	223	22.55	224	33.57	223	.63	218	1.65	218	.81	218
Medium	88.18	555	31.80	554	20.51	554	34.66	549	.64	555	1.44	555	.63	555
High	87.17	250	34.08	239	21.69	238	33.74	246	.68	254	1.63	254	.93	254
Pooled Variance	213.15		125.63		224.06		52.06							
Teacher Factor-TVI Receptive Learning	*+		+		+		+							
Low	83.82	269	34.25	277	23.22	277	35.11	281	.58	269	1.40	269	.68	269
Medium	88.14	504	32.27	501	21.15	502	34.12	494	.66	508	1.55	508	.71	508
High	91.87	238	31.81	226	19.08	224	33.20	231	.68	238	1.55	238	.76	238
Pooled Variance	206.10		125.52		169.97		50.42							

\+ Significant at .05 level.

* Substantial Results.

d. Classroom Interaction Variables

(1) OSCI Class Factor CI - Structured Lessons (Table VIII.5) - High emphasis on classroom structured lessons was given to children with lower SB pre scores and higher PSI, AH-WPPSI, and GUMP scores.

(2) OSCI Class Factor CII - Group Activities and Routines (Table VIII.5) - None of the measures was substantially or consistently associated with classroom activities and routines.

(3) OSCI Class Factor CIII - Social-Emotional Interaction (Table VIII.5) - Classrooms with emphasis on social-emotional interaction received children whose mean pre score on SB was high and mean scores on PSI and AH-WPPSI were low.

(4) OSCI Class Factor CIV - Verbal Communication (Table VIII.5) - Those classrooms scoring high on Verbal Communication were assigned a group of children with the highest mean pre scores on SB but the lowest mean pre scores on PSI, AH-WPPSI and GUMP.

(5) OSCI Class Factor CV - Instruction in Creative Arts (Table VIII.5) - SB was the only pre score measure substantially associated with level of classroom small group activities involving creative arts. In that instance, the higher the level of instruction in creative arts, the higher the SB pre score of the children who happened to be in these groups.

(6) OSCI Class Factor CVI - Language Training and Discrimination Learning (Table VIII.5) - None of the children's pre scores on outcome variables was substantially associated with level of language training and discrimination learning.

e. Parent Involvement Variables

(1) Frequency of Parent Visits to Center-Post (Table VIII.6) - None of the seven pre scores on the children was substantially associated with the frequency of visits by parents to the center.

(2) Parent Volunteering as Class Aide-Post (Table VIII.6) - As with the preceding variable, none of the seven pre scores was substantially associated with the incidence with which parents volunteered in the classroom as an aide.

Table VIII.5

RELATIONSHIP BETWEEN CHILD'S MEAN PRE SCORE ON
SB, PSI, AH-WPPSI, GUMP, BP, MP, FI AND CLASS INTER CTION VARIABLES
(1968-69)

Independent Variables	SB		PSI		AH-WPPSI		GUMP		BP		MP		FI	
	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
Class Factor CI Structured Lessons	*+		*+		*+		*+							
Low	90.90	236	31.88	229	19.38	227	33.26	224	.80	236	1.74	236	.93	236
Medium	90.13	496	30.97	515	19.53	511	33.54	508	.64	501	1.43	501	.65	501
High	81.64	298	36.51	281	25.53	287	36.08	295	.53	298	1.53	298	.75	298
Pooled Variance	198.18		120.90		161.97		50.57							
Class Factor CII Group Activities and Routines							+		+		+			
Low	87.62	229	32.84	249	22.25	246	34.83	246	.53	230	1.39	230	.69	230
Medium	87.45	511	33.00	491	21.24	493	34.48	488	.76	512	1.71	512	.82	512
High	88.74	290	32.03	285	20.13	286	33.25	293	.53	293	1.33	293	.65	293
Pooled Variance	213.69		126.37		225.24		51.61							
Class Factor CIII Social-Emotional Interaction	+		*+		+								+	
Low	85.72	261	35.08	264	23.06	264	34.66	274	.61	261	1.58	261	.66	261
Medium	88.10	505	32.30	500	20.45	501	33.96	496	.63	512	1.49	512	.83	512
High	89.48	264	31.03	261	20.65	260	34.21	257	.72	262	1.56	262	.65	262
Pooled Variance	212.14		124.30		168.14		51.92							
Class Factor CIV Verbal Communication	*+		*+		*+		*+							
Low	85.38	262	35.61	254	23.92	253	35.75	250	.73	264	1.39	264	.82	264
Medium	87.79	549	33.20	549	21.42	551	34.04	564	.62	551	1.60	551	.74	551
High	90.96	219	28.09	222	17.42	221	32.85	213	.59	220	1.52	220	.65	220
Pooled Variance	210.38		119.68		164.45		51.03							

(continued)

Table VIII.5 (continued)

Independent Variables	SB Mean	SB N	PSI Mean	PSI N	AH-WPPSI Mean	AH-WPPSI N	GUMP Mean	GUMP N	BP Mean	BP N	MP Mean	MP N	FI Mean	FI N
Class Factor CV Instruction in Creative Arts	*+								+					
Low	84.42	261	33.46	250	20.84	250	33.94	263	.63	264	1.52	264	.79	264
Medium	88.25	519	32.76	528	21.22	525	34.37	518	.67	522	1.55	522	.70	522
High	90.59	238	32.23	234	21.58	238	34.28	233	.61	237	1.50	237	.80	237
Pooled Variance	209.59		125.18		170.16		52.30							
Class Factor CVI Language and Discrimination Learning			+				+		+					
Low	86.91	258	34.13	256	22.23	253	35.36	254	.50	258	1.66	258	.90	258
Medium	88.05	474	32.10	481	20.52	485	33.92	477	.71	476	1.46	476	.72	476
High	88.72	284	32.17	274	21.20	272	33.60	281	.61	287	1.47	287	.66	287
Pooled Variance	214.59		124.78		222.42		51.58							

\+ Significant at .05 level.

* Substantial Results.

Table VIII.6

RELATIONSHIP BETWEEN CHILD'S MEAN PRE SCORE ON
SB, PSI, AH-WPPSI, GUMP, BP, MP, FI AND PARENT INVOLVEMENT VARIABLES

Independent Variables	SB Mean	SB N	PSI Mean	PSI N	AH-WPPSI Mean	AH-WPPSI N	GUMP Mean	GUMP N	BP Mean	BP N	MP Mean	MP N	FI Mean	FI N
Frequency of Parents Visits to Center	+								+					
Infrequent: (1-2 times)	87.02	226	33.04	221	22.05	225	34.22	228	.51	226	1.49	226	.64	226
Occasionally: (3-6 times)	86.67	224	33.54	222	21.00	228	34.05	220	.67	221	1.76	221	.74	224
Frequently: (7 or more times)	89.90	200	33.26	194	22.56	197	34.39	196	.54	194	1.72	194	.83	194
Pooled Variance	204.16		142.51		177.92		48.85							
Parent Volunteering as Class Aide														
None	86.94	378	32.93	375	21.10	375	33.95	367	.67	380	1.75	380	.75	380
Yes--Occasionally (less than once a month)	87.60	200	34.65	202	23.31	207	35.24	201	.52	198	1.51	198	.60	199
Yes--Frequently (once a week or more)	89.60	250	33.87	245	22.52	253	33.98	251	.59	241	1.51	241	.80	243
Pooled Variance	207.72		139.46		181.02		48.14							

\+ Significant at .05 level.

CHAPTER IX

RELATIONSHIPS OF CHILD AND FAMILY VARIABLES TO GAIN SCORE MEASURES

A. Overview of Changes in Child Performance

The main purpose of the 1967-68 and 1968-69 national evaluations was to find out what kinds of programs have what kinds of effects for what kinds of children. The interactive hypothesis underlying this objective is that some kinds of experiences may be better than others and that what is best for one child may not be best for another. The most appropriate test of this important and popular hypothesis would be a comparison of programs where the curriculum was in fact "matched" to the child's needs and competencies, with programs where all children receive a similar curriculum. If the interactive assumption is valid, all children would show substantial progress in the individualized program (although possibly at different rates in different areas) while in the across-the-board programs, only some children would progress. Furthermore, if the "match" between child and program characteristics identified in the individualized program could be used to predict which children would benefit in what ways from the "relatively uniform curriculum" approaches, the validity of the interactive hypothesis would be strongly supported.

The present evaluative studies do not provide such a direct test of the interactive hypothesis, nor to our knowledge does such a study exist. To a certain extent, almost all early childhood curriculums are individualized. Many involve extensive testing and observation so that the curriculum materials presented are appropriate to the child's current level of academic competence even where the educational objectives and teaching methods are uniform (e.g., Bereiter-Englemann-Becker, Bushell, and Karnes). Others provide equally intensive but wider-scope assessment and individualization of curriculum across areas (e.g., Hodges, Spiker and McCandless, and Sprigle). Still others rely on the teacher's observations and perceptions and her almost clinical skill in responding to the child's initiations (e.g., Bank Street). Perhaps only in the stereotyped "traditional" program where a similar experience is simultaneously available to all children with unsystematic individualization of teacher initiations and responses to children (if any), can a mostly noninteractive situation be said to exist.

The indirect test of the interactive hypothesis in the E & R evaluations is provided by a systematic search:

1) For overall changes,

2) For univariate program characteristics that may be differentially associated with change, and

3) For interactions among child, family, teacher and program characteristics that may be either differentially associated with change on all variables or associated with changes on different variables in different ways.

The overall mean gain scores (post minus pre) for each of the outcome measures for 1967-68 and 1968-69 are shown in Table IX.1 for all children for whom pre and post data were available on that measure. The gains for 1968-69 are based upon children who had no more than a summer of previous Head Start experience and were in classes where at least 70 percent of the children had no previous Head Start experience. The gains for 1967-68 are based upon all children since children with more than a summer of previous Head Start experience could not be identified. The gains for 1968-69 children with various levels of previous Head Start experience are presented in Table IX.6, to be discussed later. Gains for 1967-68 children as a function of independent variables are shown in Table IX.2.

TABLE IX.1

MEAN GAIN SCORES AND STANDARD DEVIATIONS FOR OUTCOME MEASURES FOR THE 1967-68 AND 1968-69 HEAD START SAMPLES

	1967-68 Data			1968-69 Data		
Outcome Measure	Mean Gain	Standard Deviation	N	Mean Gain	Standard Deviation	N
SB	4.69	10.05	1493	4.82	10.19	1000
PSI				9.38	7.73	983
AH-WPPSI				9.45	12.77	997
GUMP				7.55	7.98	982
BP	-.15	1.16	1460	-.09	1.31	1010
MP	-.56	2.26	1460	-.21	2.42	1008
FI	-.31	1.34	1460	-.20	1.36	1013

TABLE IX.2

RELATIONSHIPS BETWEEN MEAN GAIN SCORES ON SB, BP, FI, MP, AND INDEPENDENT VARIABLES (1967-68)

Independent Variables	SB Mean Gain	N	BP Mean Gain	N	FI Mean Gain	N	MP Mean Gain	N	Mean Age in Months
Age									
30-47 mo.	6.19*	405	-.13	382	-.27	382	-.57	382	
48-53 mo.	3.99	531	-.16	520	-.30	520	-.54	520	
54-59 mo.	4.44	390	-.20	390	-.38	390	-.52	390	
60-77 mo.	3.87	167	-.08	168	-.24	168	-.67	168	
Pooled Variance	100.30		1.34		1.81		5.11		
Sex									
Male	4.62	752	-.16	742	-.29	742	-.57	742	51
Female	4.76	741	-.14	718	-.32	718	-.55	718	51
Pooled Variance	101.04		1.34		1.81		5.11		
Ethnicity									
Black	4.51	764	-.19	732	-.23	732	-.54*	732	50
White	5.12	449	-.19	442	-.47	442	-.68	442	52
Mex-Amer.	3.18	102	-.11	109	-.22	109	-.27	109	52
Polynesian	5.05	60	.13	56	-.07	56	-.37	56	50
Pooled Variance	100.96		1.34		1.80		5.06		
Education of Mother									
No School	7.86	7	.00	6	.67	6	-.20	6	52
1-3 grade	5.33	18	-.21	19	-.26	19	-.58	19	50
4-6 grade	7.07	86	-.11	84	-.18	84	-.27	84	53
7-8 grade	4.52	211	-.02	206	-.25	206	-.29	206	52
9-11 grade	5.02	522	-.20	520	-.36	520	-.62	520	51
12-HS Grad.	4.02	442	-.14	422	-.32	422	-.53	422	51
Some College	6.85	55	-.47	53	.00	53	-.57	53	50
College Grad.	6.53	17	-.50	18	-.44	18	-1.17	18	50
Pooled Variance	100.84		1.33		1.72		5.11		

(continued)

TABLE IX.2 (continued)

Independent Variables	SB Mean Gain	SB N	BP Mean Gain	BP N	FI Mean Gain	FI N	MP Mean Gain	MP N	Mean Age in Months
Region									
North	5.22*	422	-.18*	407	-.12*	407	-.14*	407	49
South	2.95	283	-.16	279	-.81	279	-.33	279	53
Midwest	4.92	493	-.24	499	-.85	499	-.51	499	53
West	5.09	304	.05	286	-.40	286	-.14	286	49
Pooled Variance	100.38		1.33		1.77		6.00		
Total # of People at Home									
2-4	4.90	301	-.26	292	-.36	292	-.59	292	50
5,6	4.45	486	-.12	469	-.38	469	-.55	469	51
7,8	5.27	339	-.18	337	-.23	337	-.41	337	52
9 or more	5.08	256	-.08	20	-.21	255	-.71	20	52
Pooled Variance	102.40		1.33		1.75		5.17		
Pre-Aspirations									
High School or Less (including vocational work)	5.50	305	-.15	298	-.42	298	-.48	298	52
College	4.74	1054	-.17	1033	-.28	1033	-.56	1033	51
Pooled Variance	101.19		1.32		1.73		5.14		
Pre-Expectations									
High School or Less (including vocational work)	4.98	946	-.16	923	-.34	923	-.59	923	51
College	5.24	330	-.18	319	-.23	319	-.53	319	51
Pooled Variance	102.22		1.33		1.75		5.18		
Attendance									
5 out of 5	4.84	714	-.22	707	-.30	707	-.55	707	51
4 out of 5	4.12	531	-.12	517	-.26	517	-.57	517	52
3 or less out of 5	5.31	230	-.03	218	-.42	218	-.49	218	50
Pooled Variance	101.24		1.33		1.79		5.08		

(continued)

TABLE IX.2 (continued)

Independent Variables	SB Mean Gain	SB N	BP Mean Gain	BP N	FI Mean Gain	FI N	MP Mean Gain	MP N	Mean Age in Months
Parent Participation									
Attended	4.40	914	-.16	887	-.32	887	-.59	887	51
Did Not Attend	5.36	321	-.21	316	-.41	316	-.60	316	52
Pooled Variance	102.2		1.33		1.76		5.10		
Teacher Credentials									
Elementary	4.30	23	-1.17*	23	-.43	23	-.48	23	50
High School	6.21	299	-.18	290	-.31	290	-.57	290	51
Bachelors	4.78	679	-.12	664	-.31	664	-.47	664	50
Masters	4.18	103	-.01	94	-.29	400	-.73	94	51
Pooled Variance	92.48		1.38		1.90		4.92		
Teachers Paid Experience with Pre-Schoolers									
Less than 6 mo.	4.37*	100	0.06	94	-.29	94	-.48	94	48
6 mo. - 1 yr.	3.98	176	-.09	175	-.41	175	-.64	175	52
1-3 yrs.	5.97	575	-.13	550	-.33	550	-.44	550	51
4-5 yrs.	5.10	136	-.31	137	-.39	137	-.87	137	51
Over 6 yrs.	3.13	117	-.32	115	-.04	115	-.33	115	48
Pooled Variance	92.0		1.40		1.90		4.91		
Teachers Experienced with Disadvantaged									
Less than 6 mo.	3.60	110	0.03	104	-.33	104	-.61	104	49
6 mo. - 1 yr.	4.13	198	-.11	197	-.41	197	-.60	197	53
1-3 yrs.	5.51	588	-.16	575	-.30	575	-.50	575	51
4-5 yrs.	5.60	112	-.21	104	-.33	104	-.55	104	50
Over 6 yrs.	5.69	96	-.33	91	-.14	91	-.38	91	46
Pooled Variance	92.5		1.40		1.90		4.93		

(continued)

TABLE IX.2 (continued)

Independent Variables	SB Mean Gain	SB N	BP Mean Gain	BP N	FI Mean Gain	FI N	MP Mean Gain	MP N	Mean Age in Months
OSCI Factor I									
Low	5.08	336	-.14	320	-.40*	320	-.48*	320	50
Medium	4.03	684	-.14	677	-.19	677	-.46	677	51
High	4.80	328	-.20	322	-.55	322	-.95	322	52
Pooled Variance	99.35		1.23		1.79		4.94		
OSCI Factor II									
Low	4.62*	365	-.13	352	-.40	352	-.62	352	51
Medium	5.62	602	-.17	595	-.27	595	-.55	595	51
High	3.39	382	-.16	375	-.34	375	-.63	375	51
Pooled Variance	98.51		1.24		1.80		4.97		
OSCI Factor III									
Low	2.94*	331	-.29*	330	-.34	330	-.53	330	50
Medium	5.32	662	-.08	648	-.26	648	-.57	648	50
High	4.54	330	-.14	320	-.45	320	-.72	320	55
Pooled Variance	97.98		1.24		1.77		4.96		
OSCI Factor IV									
Low	4.59	342	-.15	322	-.24	522	-.24*	322	50
Medium	4.61	663	-.21	651	-.40	651	-.83	651	51
High	4.26	355	-.28	358	-.25	358	-.48	358	52
Pooled Variance	99.55		1.24		1.79		4.90		
Square Feet per child									
Less than 14	4.40	403	-.10*	400	-.26	400	-.46	400	53
14-19	4.72	399	-.02	380	-.23	380	-.47	380	51
20-24	4.93	649	-.28	640	-.39	640	-.65	640	51
Pooled Variance	102.43		1.29		1.79		4.94		

* Substantial Results.

Changes in the SB, the PSI raw score, the AH-WPPSI, and the GUMP were statistically reliable and substantial. The decreases in adjustment problems (BP, MP, and FI) in the test situation were not substantial overall, probably reflecting the fact that most children had only 0-1 problems on initial testing.

On the SB, initial performance for 44.2 percent of the children in 1967-68 and 50 percent in 1968-69 was at levels usually indicating learning difficulties (89 or lower); at time of post-testing, only 31.2 percent of the 1967-68 and 36 percent of the 1968-69 children scored at or below this level. Of the almost 1500 children tested in each year, only about 38.7 percent either showed small gains (0-1 points) or had lower scores on post-testing.

The SB scores are age standardized, and thus already take into account the improvement expected because the children are six-months older. It may be useful to compare the overall SB changes with those reported for other studies. SB changes for control groups typically are smaller than those observed for Head Start while changes for experimental programs vary from similar to much larger. Data from 6 control groups showed a median gain of 2 SB points. However, some of these control groups had pre means considerably below those of the Head Start E & R samples.

One of the central questions raised by these data is whether virtually all Head Starts show the "motivational enhancement" level of gain, or whether there are substantial variations among programs.

Appendix E gives the gain score distributions as class averages for the 1967-68 and 1968-69 programs. These data, like those of the 1969-70 Head Start Planned Variation Study (SRI, 1971) indicate substantial variation in average gains among classes, some as high as those reported by Karnes and others and some showing losses.

These data indicate average gains greater for the SB, PSI, AH-WPSSI, and GUMP than the changes anticipated for the approximately 6 chronological months separating initial and final scores. This can be seen by referring to Table VII.4, which indicates the relationship of chronological age to the pre scores of the 1968-69 outcome measures, and contrasting the mean pre score for children in the 48-53 month age category with the mean pre score for children in the 54-59 month category. For example, the mean gain on the PSI was 9.45 points while the difference between the pre PSI means of these two age categories was only 3.9 points. These two age categories differ by approximately 6 months. The children's raw scores gains reflect, therefore, both an anticipated change over time and an acceleration over and above these "maturational" effects.

Gain score intercorrelations were calculated but were not reliable except for the AH-WPPSI and the PSI (.17); the SB and reduction in Motivational Problems (.18); and the Motivational Problems and Feelings of Inadequacy scales (.43).

To summarize: Analyses of the outcome measures indicate gains in cognitive development, preacademic readiness, ability to learn a new task, and achievement motivation greater than expected at usual maturational rates. Adjustment to the testing situation increased slightly but not substantially, which may be due to the initially good adjustment shown by many children. In comparison with data from other studies, the average SB gains are greater than those reported in other studies for control children (2-3 SB points) and similar to those associated with "motivational enhancement" and "traditional" programs (4-5 SB points). The range of mean gains was larger than that for experimental programs: some Head Starts on the average apparently were doing little to enhance development as measured by SB gains while others apparently were as effective as the most effective experimental programs. (This can be substantiated by referring to the listings of class mean gains for 1967-68 and 1968-69 in Appendix G.)

B. A Methodological Note on Interpretation of the Raw and Adjusted Gains

As mentioned earlier, there is no entirely satisfactory way to correct statistically for initial differences between a priori groups. For interpreting the gains related to child-family background characteristics, raw gains will be relied upon. Adjusted gains will be used in interpreting changes associated with teacher and program factors. Since the child-family background characteristics define natural subpopulations and the children in each of these subpopulations are regressing toward different subpopulation means, it is felt that examining the ranking of the subpopulations on raw score gains is more meaningful than examining the ranking of the subpopulations on adjusted gains. However, children within a certain level of a program factor can be assumed to be a mixture of various subpopulations and the adjustment for gains on the basis of one overall regression line is somewhat more defensible.

The criterion for substantiality was the same as that used for describing the prescore analyses. For both gain and adjusted gain scores a difference of one-third (1/3) of a standard deviation of the corresponding gain score distribution between any two levels of a variable or factor was considered substantial. The criterion values are present in Table IX.3. The gain criterion was used for child-family background descriptions and the adjusted gain criterion was used for the teacher-program descriptions. This criterion of substantiality for both gains and prescores is arbitrary and the reader is invited to formulate his own criterion in examining the various tables.

Table IX.3

CRITERION VALUES FOR JUDGING THE "SUBSTANTIALITY" OF THE DIFFERENCES IN RAW AND ADJUSTED GAIN SCORES AMONG SUBPOPULATIONS (1968-69)

Outcome Measure	Standard Deviations		Criterion Value (1/3 SD)	
	Raw Gain	Adjusted Gain	Raw Gain	Adjusted Gain
SB	10.19	9.05	3.40	3.02
PSI	7.73	6.82	2.58	2.27
AH-WPPSI	12.77	11.95	4.26	3.98
GUMP	7.98	7.18	2.65	2.39
BP	1.31	1.01	0.44	0.34
MP	2.42	1.89	0.81	0.63
FI	1.36	0.98	0.45	0.33

C. Child Demographic Variables and Outcomes

1. Age (Table IX.4)

Age was not related to gain on the SB, PSI, GUMP, MP and BP. Older children gained more on the AH-WPPSI and made fewer improvements on the FI (probably because of ceiling effects) than did younger children. The 1967-68 data indicated a significant but nonsubstantial effect of age on the SB with youngest children gaining the most. Overall, there is little evidence of a substantial effect of age on gains, a finding similar to Stearns' (1971) conclusion that "as far as amount of immediate change in intellectual performance is concerned, the time in the individual's early life at which this preschool experience occurs, at least between 2 and 6 years of age, does not appear crucial" (p. 23).

The clearest evidence contradicting this finding is Van de Reit and Resnick's (1972) report that children enrolled in the Sprigle program at 4 years of age clearly "out performed" those enrolled at five, both initially and in the follow-up study. However, all of Sprigle's children were from the South and Southern children showed a possible atypical selection effect.

2. Sex (Table F-1)

Positive gains on SB, PSI, AH-WPPSI, PSI, and GUMP and reduction in behavior problems on MP, BP, and FI were similar for boys and girls. No pattern of consistency for nonsubstantial differences was discernible.

Table IX.4

RELATIONSHIPS BETWEEN AGE OF CHILD AND MEAN GAIN SCORES ON SEVEN OUTCOME MEASURES (1968-69)

A. Age of Child and Mean Gain Scores on SB, PSI and AH-WPPSI

Age in Months	SB				PSI				AH-WPPSI			
	Pre Score	R Gain	A Gain+	N	Pre Score	R Gain	A Gain+*	N	Pre Score	R Gain+*	A Gain+*	N
31-47	93.32	5.11	1.91	185	28.76	8.92	-2.18	177	16.57	6.13	-5.14	180
48-53	90.88	4.05	0.07	347	31.67	9.36	-.86	338	19.67	9.04	-1.18	344
54-59	88.02	5.16	0.25	196	34.98	9.97	0.76	200	21.82	9.61	0.12	202
60-72	81.72	5.40	-1.54	270	38.97	9.31	1.31	266	28.40	12.07	4.80	269
Pooled Variance		103.99	80.99			59.92	45.66			159.71	131.56	

B. Age of Child and Mean Gain Scores on GUMP, BP, MP and FI

Age in Months	GUMP				BP				MP				FI			
	Pre Score	R Gain	A Gain+*	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain+	N	Pre Score	R Gain+*	A Gain+	N
31-47	32.12	7.09	-1.62	171	.78	-.07	.05	180	2.15	-.22	.44	181	1.03	-.43	-.01	181
48-53	33.26	7.64	0.55	331	.68	-.08	-.03	344	1.58	-.34	-.10	344	0.85	-.41	-.13	344
54-59	34.03	7.83	0.01	197	.52	-.02	-.07	200	1.50	.31	-.13	199	0.55	-.05	-.01	201
60-72	36.90	7.58	1.07	281	.58	-.17	-.19	284	1.45	.02	.16	284	0.68	.08	.22	285
Pooled Variance		73.70	52.53			1.72	1.03			5.86	3.48			1.79	.96	

* Substantial Result.

\+ Significant at .05 Level.

3. Ethnicity (Table IX.5)

Substantial differences among the four ethnic subgroups appeared on SB, PSI, AH-WPPSI and GUMP gain scores. Black children, who along with the small sample of Mexican-American children, had obtained the lowest pre scores on SB made the largest gains on SB (5.7 points) while the Mexican-American children made the smallest gains (1.6 points). There were no significant differences on the SB for 1967-68. However, the Mexican-American children gained the least.

On the PSI, Polynesians who had scored 5-6 points lower on the pretest than any of the other three ethnic groups, gained 12.8 points which is approximately 4-5 points more than mean gains for the other three ethnic groups. The largest gains on AH-WPPSI were also obtained by the Polynesian children who had the lowest pre scores, while the smallest gains were made by the Mexican-Americans who had obtained the highest pre scores. While blacks and whites had similar pre scores on the AH-WPPSI, the blacks gained 3.7 more points than the whites.

Finally, while pre scores on GUMP were similar for all ethnic subgroups, gains on GUMP were approximately four points greater for the Mexican-American group than for the remaining groups.

4. Previous Head Start Experience (Table IX.6)

Amount of previous Head Start experience was substantially related to gains on all measures but SB and GUMP, through none of these substantial relationships was monotonic. For the PSI, AH-WPPSI, and GUMP, both the group with no previous Head Start experience and the group with 9-10 months of previous experience obtained larger gains than the remaining two experience groups.

On the other hand, the group with 12 or more months of previous Head Start experience showed greater reductions on BP, MP, and FI than the remaining two (intermediate) experience groups. In general, the previous Head Start experience groups with the lowest pre scores on BP, MP, and FI made the lowest raw score improvement from pre to post test.

These data suggest that the association of Head Start program duration and gains is not linear in the cognitive and preacademic areas. That is, the greatest change occurs in the first year, although there is additional improvement in year 2. This is an extrapolation from cross sectional data, however, since we do not have initial, year 1 and year 2 data for the same S's. To a certain extent, it is unexpected to find any year 2 gain, since, as Stearns (1971) points out, "Often, two years of preschool do not yield a higher IQ score than a single year...it is not the amount of exposure time per se, but rather the experiences occurring during that time which stimulate intellectual growth. It may be...the length of time spent with individual children which is related to their

Table IX.5

RELATIONSHIP BETWEEN ETHNICITY AND MEAN GAIN SCORES ON THE SEVEN OUTCOME MEASURES
(1968-69)

A. Ethnicity and Mean Gain Scores on SB, PSI and AH-WPPSI

Ethnicity	CA	SB				PSI				AH-WPPSI			
		Pre Score	R Gain+*	A Gain+*	N	Pre Score	R Gain+*	A Gain*	N	Pre Score	R Gain+*	A Gain+*	N
Black	55	86.59	5.72	.36	663	34.80	9.02	-.25	639	21.86	10.05	.57	647
White	52	93.62	3.36	.26	162	35.49	8.77	-.29	173	22.93	6.32	-2.79	173
Mex-American	59	85.68	1.57	-4.09	56	32.81	8.19	-1.69	54	30.21	6.14	-.51	56
Polynesian	50	94.37	2.60	-.26	70	28.49	12.83	1.64	70	17.75	12.62	1.75	71
Pooled Variance			102.54	81.36			58.44	47.01			160.11	141.56	

B. Ethnicity and Mean Gain Scores on GUMP, BP, MP and FI.

Ethnicity	GUMP				BP				MP				FI			
	Pre Score	R Gain+*	A Gain+*	N	Pre Score	R Gain	A Gain*	N	Pre Score	R Gain	A Gain+*	N	Pre Score	R Gain	A Gain+*	N
Black	34.65	7.47	-.08	656	.68	-.10	-.05	678	1.77	-.19	.19	678	0.77	-.13	.08	681
White	33.72	7.08	-.90	155	.58	-.17	-.18	161	1.24	-.33	-.35	161	0.74	-.33	-.14	161
Mex-American	33.14	10.87	2.63	55	.44	.22	.12	55	1.91	-.13	.36	55	1.16	-.26	.27	55
Polynesian	32.90	6.72	-1.63	71	.47	-.15	-.23	68	1.09	-.31	-.44	68	0.60	-.40	-.32	68
Pooled Variance		63.06	52.86			1.72	1.04			5.92	3.49		1.83	.96		

* Substantial Result.
\+ Significant at .05 Level.

Table IX.6

RELATIONSHIP BETWEEN PREVIOUS HEAD START EXPERIENCE AND MEAN GAIN SCORES ON SEVEN OUTCOME MEASURES

A. Previous Head Start Experience and Mean Gain Scores on SB, PSI, and AH-WPPSI

Previous Head Start Experience	CA	SB				PSI				AH-WPPSI			
		Pre Score	R Gain+	A Gain	N	Pre Score	R Gain+*	A Gain	N	Pre Score	R Gain*	A Gain	N
None	54	89.21	4.92	-	1153	32.84	9.52	-	1124	21.28	9.74	-	1146
2-3 months	53	90.62	2.32	-	132	40.13	5.32	-	133	25.35	6.86	-	137
9-10 months	51	89.74	3.15	-	97	34.94	7.90	-	99	19.64	11.45	-	99
12 or more months	57	96.03	3.34	-	61	37.95	7.32	-	62	24.47	6.06	-	64
Pooled Variance			96.32				55.01				157.78		

B. Previous Head Start Experience and Mean Gain Scores on GUMP, BP, MP, and FI

Previous Head Start Experience	GUMP				BP				MP				FI			
	Pre Score	R Gain+	A Gain	N	Pre Score	R Gain+*	A Gain	N	Pre Score	R Gain+*	A Gain	N	Pre Score	R Gain+*	A Gain	N
None	34.19	7.96	-	1112	.67	-.12	-	1161	1.56	-.26	-	1160	.75	-.23	-	1161
2-3 months	35.52	5.67	-	136	.86	-.02	-	127	2.13	.44	-	128	.79	.14	-	130
9-10 months	35.51	7.47	-	98	.55	-.15	-	96	1.00	.03	-	96	.44	-.04	-	96
12 or more months	35.59	5.90	-	59	.69	-.36	-	61	1.57	-.90	-	61	.70	-.54	-	61
Pooled Variance		61.61				1.87				5.72				1.74		

* Substantial Result.

\+ Significant at .05 Level.

IQ increase" (p. 25). Beller (1969), Klaus and Gray (1968) and Weikart (1971) find the greatest gain in year 1, with leveling off or slight decline in year 2 of nursery (4 year olds) plus kindergarten (5 year olds). Erickson et al. (1969) found the greatest gain in year 1 Head Start Englemann classes (4's) and leveling off or less gain in year 2 public school Englemann classes (5's). For the experimental program taught by Englemann himself (Karnes, 1969), children gained as much in year 2 (first grade, 6 years) as in year 1 (5 years). Van de Reit (1970), testing classes taught by Dr. and Mrs. Sprigle, found continuing gain from year 1 (5's) to year 2 (1st grade, 6's), and an even greater gain from year 1 (4's) to year 2 (5's) in the South.

Finally, Kraft et al. (1968) found that (a) children with initially higher and lower IQs from relatively better off homes and (b) children with initially higher IQs from very poor homes gained most in year 1 (3's) and leveled off in years 2 (4's) and 3 (5's), while children with initially lower IQs from very poor homes took two years to reach the level attained by relatively better off children in one year.

A conservative interpretation of these data would be that further, systematic experimental data are required. A less conservative interpretation is that where programs are dynamic, highly stimulating and keep pace with the child's rapid development, the longer the duration the better for the child--by the cognitive and academic criteria for immediate effects. There is some evidence that a "rote learning" approach will not show sustained effects when children enter regular public school in year 3 (Karnes, 1969) while a learning-to-learn approach will show substantial sustained effects when children enter regular public schools in year 3 (Van de Reit and Resnick, 1972). Also, where programs are more traditional in content, have continuity-by-curriculum label (e.g., Erickson follow-up of Englemann as taught by others), or less intensively stimulate the child's development, two years are not much better than a one-year program--except for slow learners from very disadvantaged circumstances.

5. Region (Table IX.7)

Children in the South had the largest gain on the SB and the AH-WPPSI, and children from the West had the largest gain on the PSI. The results for the noncognitive measures were also inconsistent: children from the North showed the largest decreases in MP; children from the Midwest showed the largest increase in GUMP; and children from the West showed the largest decrease in FI. The 1967-68 results did not replicate those for 1968-69. Thus the results were inconsistent across both years and outcome measures for region.

Table IX.7

RELATIONSHIP BETWEEN REGION AND MEAN GAIN SCORES ON SEVEN OUTCOME MEASURES

A. Region and Mean Gain Scores on SB, PSI, and AH-WPPSI

Region	CA	SB				PSI				AH-WPPSI			
		Pre Score	R Gain+	A Gain+	N	Pre Score	R Gain+*	A Gain	N	Pre Score	R Gain+	A Gain+*	N
North	50	91.97	5.20	1.57	310	31.06	8.76	-1.65	293	18.47	7.65	-2.97	289
South	58	81.56	6.47	0.52	315	37.76	9.23	0.86	331	24.19	11.44	2.74	339
Mid-West	58	90.89	2.57	-1.41	140	37.31	7.21	-1.29	130	28.85	9.57	2.46	133
West	50	90.95	3.43	0.53	235	29.73	11.63	0.82	229	18.89	8.75	-1.73	236
Pooled Variance			102.11	81.01			58.00	45.87			161.12	136.52	

B. Region and Mean Gain Scores on GUMP, BP, MP and FI

Region	GUMP				BP				MP				FI			
	Pre Score	R Gain+	A Gain+*	N	Pre Score	R Gain	A Gain+	N	Pre Score	R Gain+*	A Gain+	N	Pre Score	R Gain+*	A Gain+*	N
North	33.11	7.58	0.67	274	.74	-.22	-.13	309	1.77	-.59	-.21	309	.82	-.29	-.03	309
South	35.28	6.54	0.72	348	.57	-.12	-.14	334	1.81	-.12	.29	334	.77	-.02	.19	337
Mid-West	36.26	8.88	2.08	130	.49	.03	-.04	138	1.24	.27	.26	138	.64	.02	.14	138
West	32.91	8.30	0.05	230	.67	.05	.09	229	1.41	-.06	.04	229	.80	-.48	-.24	229
Pooled Variance		63.09	52.70			1.71	1.03			5.85	3.52			1.81	0.95	

* Substantial Result.
\+ Significant at .05 Level.

D. Family Demographic Variables

1. Socioeconomic Status (Table IX.8)

For children from families within the Head Start economic guidelines, SES was not substantially related to either raw or adjusted gain scores. A consistent trend was obtained, however, on five of the seven measures. On SB, PSI, and AH-WPPSI, children from extremely deprived homes tended to gain more than did children from higher SES levels, and also generally had larger decreases in problems than did the higher SES levels on MP, and FI. This finding was supported by the 1967-68 data where no significant relationships were found for the two SES indicators, mother's education, and number of people in the home.

2. Family Structure (Table F-2)

None of the raw or adjusted gain scores was substantially related to different types of family structure, nor did any consistent pattern of relationships emerge.

3. Mother's Employment Status (Table F-3)

None of the raw or adjusted gain scores was substantially related to mother's employment status.

4. Number of Family Moves in Last Three Years (Table F-4)

None of the raw or adjusted gain scores was substantially related to family mobility nor did any consistent pattern of differential gains appear in the data.

E. Parent-Child Psychological Interaction Variables

1. Mother's Aspirations Concerning Child's Educational-Occupational Advancement (Table F-5)

None of the gain scores was substantially related to this variable, nor did any consistent pattern of differential gains appear in the data for either 1967-68 or 1968-69.

2. Mother's Expectations Concerning Child's Educational-Occupational Advancement (Table F-6)

None of the gain scores was substantially related to this variable nor did any consistent pattern of differential gains appear in the data for either 1967-68 or 1968-69.

3. Parent Pessimism Toward Life Scale (Table F-7)

None of the gain scores was substantially related to this variable (pre Head Start attitudes) nor did a consistent pattern of findings emerge.

Table IX.8

RELATIONSHIP BETWEEN SOCIOECONOMIC STATUS AND MEAN GAIN SCORES ON THE SEVEN OUTCOME MEASURES

A. Socio-Economic Status and Mean Gain Scores on SB, PSI, and AH-WPPSI

SES	CA	SB				PSI				AH-WPPSI			
		Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N
Low Low	56	81.45	6.94	-.08	193	31.92	9.90	-.24	201	20.28	11.11	1.10	208
Medium Low	54	86.70	4.84	-.49	240	33.89	9.23	-.31	225	21.45	10.14	.52	232
Upper Low	53	92.55	4.42	.98	238	35.17	9.03	-.13	235	23.62	9.27	.39	236
Pooled Variance			96.91	75.37			51.62	49.98			158.13	141.96	

B. Socio-Economic Status and Mean Gain Scores on GUMP, BP, MP and FI

SES	GUMP				BP				MP				FI			
	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N
Low Low	33.64	6.92	-1.09	204	.59	-.12	-.13	199	1.70	-.35	-.01	199	.84	-.30	-.03	199
Medium Low	34.25	7.85	0.12	235	.66	-.16	-.12	237	1.64	-.13	.16	237	.67	-.03	.11	237
Upper Low	34.03	8.15	0.32	240	.60	-.00	.00	243	1.41	-.12	-.01	243	.63	-.13	-.03	243
Pooled Variance		65.29	54.05			1.70	1.07			5.96	3.57			1.71	.96	

\+ Significant at .05 Level.

4. Accessibility of Adults (Table F-8)

None of the raw or adjusted gain scores was substantially related to accessibility of adults.

5. Frequency of Reading to Child (Table IX.9)

There were no significant findings in regard to frequency of reading to child.

6. Extent of Physical Control (Table F-9)

Neither raw nor adjusted gain scores were substantially related to this variable.

7. Parent Disciplinary Reactions to Minor Rule Infractions (Table F-10)

None of the gain scores was substantially related to this variable.

8. Frequency of Parent Visits to Head Start Center (Table F-11)

There is no evidence to indicate that frequency of visiting the Head Start Center in itself plays an important role in bringing about change in children's skills and behaviors. However, cautious interpretation is called for in that post family interview information on this variable existed for only 600 children out of the original 1998 child records. The available data could be somewhat biased like the data for many of the present analyses where a relatively large amount of the data is missing due to systematic selection. The 1967-68 data replicated the participation findings. More detailed analyses of parent participation should be conducted to see if parent participation interacts with child characteristics such as age, pre scores, SES, etc. in bringing about gains.

9. Parent Volunteering as Class Aide (Table IX.10)

Only the PSI had a statistically significant monotonic positive relationship with frequency of parent volunteering as a Class Aide for both adjusted and raw gains.

F. Time in Program Variables

1. Attendance (Table IX.11)

Attendance had a substantial impact on raw SB and PSI gains for 1968-69, but not for 1967-68. For 1968-69, there was a tendency for larger gains to occur for children who attended the Head Start program for longer periods.

Table IX.9

RELATIONSHIP BETWEEN FREQUENCY OF READING TO CHILD AND MEAN GAIN SCORES ON THE SEVEN OUTCOME MEASURES

A. Frequency of Reading to Child and Mean Gain Scores on SB, PSI and AH-WPPSI

Frequency of Reading to Child	CA	SB				PSI				AH-WPPSI			
		Pre Score	R Gain	A Gain	N	Pre Score	R Gain*	A Gain*	N	Pre Score	R Gain	A Gain*	N
Seldom or never	52	85.92	5.02	-.56	49	-28.25	9.30	-1.96	49	21.13	6.47	-3.26	47
Sometimes	54	88.37	5.28	.48	186	-34.09	10.14	.66	186	20.46	10.20	.25	190
Often	54	88.42	5.07	.29	276	-34.11	9.59	.12	276	22.41	9.95	.65	285
Regularly	54	88.57	4.78	.05	231	-34.76	9.80	.51	231	22.81	9.98	.82	226
Very Frequently	54	88.00	4.05	-.86	59	-39.66	7.23	-.56	59	22.02	8.93	-.50	54
Pooled Variance			97.23	75.77			60.13	46.91			167.09	145.08	

B. Frequency of Reading to Child and Mean Gain Scores on GUMP, BP, MP and FI

Frequency of Reading to Child	GUMP				BP				MP				FI			
	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain*	A Gain+	N	Pre Score	R Gain	A Gain	N
Seldom or never	32.18	8.04	-.64	50	.68	.06	.11	47	1.70	-.04	.29	47	.64	-.30	-.19	47
Sometimes	34.21	7.51	-.25	182	.61	-.09	-.08	188	1.45	-.26	-.11	188	.74	-.26	-.06	188
Often	34.19	7.54	-.22	280	.56	-.11	-.13	283	1.66	-.44	-.14	283	.71	-.18	-.01	284
Regularly	35.17	7.27	-.04	229	.63	-.01	.01	234	1.51	.10	.28	234	.79	-.15	.09	234
Very Frequently	34.12	7.26	-.53	57	.68	.08	.14	60	1.55	.13	.35	60	.70	-.10	.06	60
Pooled Variance		63.05	53.80			1.72	1.11			5.93	3.51			1.79	.95	

* Substantial Result.
\+ Significant at .05 Level.

Table IX.10

RELATIONSHIP BETWEEN PARENT VOLUNTEERING AS CLASS AIDE AND MEAN GAIN SCORES ON THE SEVEN OUTCOME MEASURES

A. Parent Volunteering as Class Aide and Mean Gain Scores on SB, PSI and AH-WPPSI

Parent Volunteering as Class Aide	CA	SB				PSI				AH-WPPSI			
		Pre Score	R Gain	A Gain	N	Pre Score	R Gain+	A Gain+	N	Pre Score	R Gain	A Gain	N
No	53	86.95	6.02	.77	350	32.98	8.56	-1.26	346	21.13	9.41	-.32	350
Yes - Occasionally	55	88.31	5.07	.25	189	34.69	9.76	.45	198	23.40	10.46	1.50	201
Yes - Frequently	54	89.81	3.82	-.51	245	33.87	10.45	.90	235	22.46	9.01	-.27	241
Pooled Variance			99.59	80.79			61.99	48.41			158.71	137.87	

B. Parent Volunteering as Class Aide and Mean Gain Scores on GUMP, BP, MP and FI

Parent Volunteering as Class Aide	GUMP				BP				MP				FI			
	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N
No	34.05	8.08	.26	346	.65	-.10	-.07	366	1.78	-.27	.18	366	.77	-.13	.08	366
Yes - Occasionally	35.04	7.35	-.03	194	.49	-.07	-.13	194	1.52	-.32	-.13	194	.61	-.09	-.01	195
Yes - Frequently	33.97	7.29	-.57	244	.60	-.04	-.03	237	1.46	-.04	.11	237	.80	-.26	-.02	239
Pooled Variance		62.17	52.50			1.60	1.02			5.73	3.46			1.71	.98	

\+ Significant at .05 Level.

Table IX.11

RELATIONSHIP BETWEEN ATTENDANCE AND MEAN GAIN SCORES ON THE SEVEN OUTCOME MEASURES

A. Attendance and Mean Gain Scores on SB, PSI and AH-WPPSI

Attendance	CA	SB				PSI				AH-WPPSI			
		Pre Score	R Gain+*	A Gain	N	Pre Score	R Gain+*	A Gain+*	N	Pre Score	R Gain	A Gain	N
Low (90 days or less)	50	94.03	1.30	-1.67	56	32.60	5.90	-4.01	57	18.77	6.75	-3.78	56
Medium (91-159 days)	53	88.26	4.67	-.16	655	32.32	10.06	.04	631	20.54	9.17	-.75	644
High (160-196 days)	57	87.28	5.82	.68	289	37.29	8.61	.10	295	25.43	10.58	2.31	297
Pooled Variance			103.08	81.82			58.68	46.31			162.52	140.28	

B. Attendance and Mean Gain Scores on GUMP, BP, MP and FI

Attendance	GUMP				BP				MP				FI			
	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain*	A Gain	N
Low (90 days or more)	34.52	5.79	-1.8	42	.80	-.29	-.17	55	1.98	-.22	.32	55	1.22	-.56	.01	55
Medium (91-159 days)	33.73	7.56	-.41	638	.60	-.04	-.04	652	1.58	-.21	.03	652	0.74	-.20	-.01	654
High (160-196 days)	35.31	7.78	.53	302	.67	-.17	-.12	303	1.68	-.18	.13	303	0.77	-.14	.07	304
Pooled Variance		63.61	53.19			1.72	1.04			5.93	3.55			1.84	.98	

* Substantial Result.

\+ Significant at .05 Level.

2. Length of Pre/Post Testing Interval (Table IX.12)

While there were no substantial relationships between the adjusted gain and length of pre-post interval, every outcome variable except GUMP and AH-WPPSI was significantly related to the pre-post interval. Longer intervals were associated with greater cognitive gains and larger losses in adjustment problems.

G. Discussion and Summary

Participation in Head Start is clearly associated with improved performance on measures of cognitive development, preacademic readiness, ability to learn new tasks and achievement motivation. Smaller changes were found in measures of adjustment to the test situation; they were probably not larger since most of the children showed few problems on initial testing.

Ascription of these changes in performance to Head Start participation is supported by a comparison of the E & R data to control group data reported in the literature. The overall Head Start gain on the SB (about 5 points) is larger than the median gain across 6 control samples (2 points).* It is likely, therefore, that the changes observed are due to Head Start participation, although this cannot be concluded in the absence of true controls.

These findings are consistent with earlier studies of Head Start which showed consistently reliable but modest gains for summer and full-year programs. Di Lorenzo and Salter (1969), interpreting similar data, pointed out that the distance between final scores and the test's "average development expected for a child of this age" was greater than the distance gained in the program. Overall, this is true of the Head Start data. Stearns (1971) has pointed out that there is ample evidence of dramatic changes in performance on measures of general and specific abilities when achievement of such effects is a clear program goal, usually with "...smaller, well-designed, and expertly staffed programs" (p. 23). In this interpretation, overall data may reflect the fact that few Head Start programs have a primary objective of substantially increasing the child's cognitive performance (Bates, 1970). Stearns emphasizes that "... the major problems for both operating preschool programs and evaluating them is inadequate formulation of goals and objectives" (p. 54).

---

* Information in the literature on the typical SB gains for control groups was supplied by Dr. Lois-ellin Datta.

Table IX.12

RELATIONSHIP BETWEEN PRE-POST INTERVAL AND MEAN GAIN SCORE ON THE SEVEN OUTCOME MEASURES

A. Pre-Post Interval and Mean Gain Scores on SB, PSI and AH-WPPSI

Pre-Post Interval	CA	SB				PSI				AH-WPPSI			
		Pre Score	R Gain+*	A Gain+	N	Pre Score	R Gain+	A Gain+	N	Pre Score	R Gain	A Gain	N
Short (22 weeks or less)	55	91.72	2.60	-1.12	172	34.43	8.05	-1.33	275	22.16	9.29	-.09	278
Medium (23-26 weeks)	53	90.14	4.02	-.20	450	33.59	9.89	.25	450	22.15	9.03	-.36	459
Long (27 or more weeks)	55	84.60	6.89	.88	369	33.60	9.92	.29	258	21.19	10.39	.68	260
Pooled Variance			99.96	81.36			59.17	46.72			163.04	142.86	

B. Pre-Post Interval and Mean Gain Scores on GUMP, BP, MP and FI

Pre-Post Interval	GUMP				BP				MP				FI			
	Pre Score	R Gain	A Gain	N	Pre Score	R Gain+	A Gain+	N	Pre Score	R Gain+	A Gain+	N	Pre Score	R Gain+	A Gain+	N
Short (22 weeks or less)	34.75	6.80	-.70	333	.80	-.18	-.06	172	2.09	.22	.84	172	1.02	.02	.43	174
Medium (23-26 weeks)	33.51	8.16	.09	406	.60	.11	.11	449	1.52	-.10	.19	449	0.77	-.23	-.02	450
Long (27 or more weeks)	34.71	7.72	.20	235	.60	-.28	-.27	389	1.54	-.61	-.39	389	0.67	-.26	-.12	389
Pooled Variance		63.60	53.36			1.68	1.01			5.82	3.37			1.83	.94	

* Substantial Result.
\+ Significant at .05 Level.

This point will be discussed later, after the data on the effects of program characteristics have been presented.

Interpretation of these SB changes as true increments in the variable presumed measured by the SB is less clear. On the SB, the magnitude of the gain overall is within the range of gains reported in the literature for a variety of curriculums. The magnitudes of both average Head Start gains and median traditional class gains are about the same as the 4.0 change found by Zigler and Butterfield (1968) for motivationally enhancing conditions.

These data are consistent with the interpretation that overall, most of the SB gains ascribed to Head Start (and traditional programs) are probably due to changes in the child's readiness to perform, his task orientation, desire to please the adult tester, and sense of confidence in his abilities rather than to basic increments in his cognitive ability.

In some classes, however, both in Head Start and experimental programs, the gains are likely to include in addition to motivational effects, true changes in the child's store of information, readiness and cognitive ability.[1/] In this interpretation, motivational changes may occur most generally and readily, followed by changes in acquiring new skills and information, and with more effort, changes in the ability to profit from experience: to learn, to abstract, to transfer.

The analyses discussed in this chapter suggest the following:

1) Gains are greatest for children from extremely impoverished backgrounds, and for children new to the program. Child sex and age were not related to size of gains. Typically, effects were more likely to be found in cognitive, than non-cognitive measures, probably due to the ceiling effects.

2) The pattern of gains was different for each ethnic group. Black children gained most on the SB, Polynesian children gained most on the PSJ, and Mexican-American children gained most on the GUMP.

3) Parental characteristics such as aspirations, expectations, amount of reading to child, number of visits to the classroom, and child rearing practices were not related to gains. In the pre score analyses, these variables were substantially related to initial levels of performance. At least three interpretations are possible: (a) parental gain scores (which were not analyzed) are better predictors of child gain, (b) classroom experiences "broke up," as it were, effects of the home environment

---

1/ Cross program data are, unfortunately, not available on measures other than the SB.

that earlier determined child performance, or (c) if post parental factors were compared to post child performance (which they were not), the same relationship would be found, only slightly attenuated by the classroom experiences.

Data from other studies are sparse. Stearns (1971) reviewing effects of interventions directed at modifying parental behavior, skills and attitudes comments: "It is clear ... that parents' involvement changed their actions, because children did perform differently when their parents participated. [But] If some parental attitudes were signficantly altered, they were not those measured by these instruments. ...we do not yet understand well what effects we have had directly on the parents" (p. 105).

CHAPTER X

TEACHER AND PROGRAM CHARACTERISTICS AS VARIABLES ASSOCIATED WITH CHANGES IN CHILD PERFORMANCE

Central to the efforts of many researchers, program planners, administrators, parents and teachers is the belief that what happens to the child in the classroom makes a difference. Few beliefs have been as hotly debated, and the evidence is by no means unequivocal. Many writers have detailed the association of poverty and indicators of unequal educational experience: older schools, dilapidated and sparse equipment and facilities, overcrowded classes, inadequate provision for children with special needs, and it would appear all too often, cruelty, shaming and profound disrespect for the children and their families. Analyses of the correlations among indicators of school quality, family background and achievement (Coleman, 1966; Jencks, 1969) typically show a greater association of family characteristics and achievement than of school characteristics and achievement. Interpretation of these correlational data is particularly difficult since the range of "good" characteristics available in the Coleman data falls short of what many educators have been advocating (Datta, 1969).

One means of assessing the potential effects of a changed environment is to change it and see what happens. While this way is fraught with difficulties because of the inertia of a system already in place, it has been most frequently applied, often in a piece-meal way.1/

For preschools, the situation more often has been creation of a new, complete project (e.g., Gray, Weikart, Sprigle), and it may be no coincidence that these programs are often more effective than where an established program is to be modified for a year or so (Di Lorenzo and Salter, 1969; Miller and Dyer, 1970). In 1968-69, the on-going Head Start programs were modified in many of the E & R Center samples; the 1967-68 study involved selection of natural variations. The findings from these analyses should be regarded as tentative indicators of whether variations in the child's classroom experience are substantially associated with differential changes in his development.

The adjusted gain analyses will be emphasized in the interpretation of the results concerning the association of gains with teacher and program variables.

1/ The new experimental schools project initiated by the Office of Education and the British open-classroom experiment are significant exceptions in both the scope of the changes and the time given for the changes to "take hold" before a final assessment.

A. Teacher Demographic Variables

Except for age, there were no substantial or consistent relationships between the teacher demographic variables (ethnicity, years of education, training in early childhood, and residence) for either 1967-68 or 1968-69 (see Tables F-12 to F-15 for 1968-69 and Table IX.2 for 1967-68).

1. Teachers Age (Table X.1)

In 1968-69, older teachers tended to be consistently associated with greater improvement in the children's cognitive behavior as reflected by both adjusted and unadjusted gain scores.

B. Teacher Interaction Variables

Chapter II described in detail the teacher-program variables considered in this analysis. In general terms they were based upon a Post Program Teacher Interview, a Post Observation Teacher Rating Scale, and the Observation of Substantive Curricular Input (OSCI) Teacher and Class factors.

1. Teacher Stimulation (Table F-16)

The results for this variable indicated no substantial or consistent results.

2. Teacher Sensitivity to Individual Differences (Table F-17)

Like Teacher Stimulation, there were no substantial or consistent results for this variable.

3. OSCI-TI - Teacher Social-Emotional Interaction (Table X.2)

Substantial results were obtained for AH-WPPSI in that there was an overall 6.6 adjusted mean gain difference between the high and low levels of this factor. The raw gains also indicated a substantial difference between the low and high levels. The gain was in favor of classes rated low on Teacher Social-Emotional Interaction.

4. OSCI-TII - Teacher Structured Lessons (Table X.3)

Once again, substantial results were obtained for AH-WPPSI. The children (older) exposed to a high level of this factor gained the most. PSI gains were also substantially greater at the high level of emphasis. An opposite pattern which was significant but not substantial emerged for the SB where adjusted gains were lowest for classes with high emphasis. There was also a tendency for more improvement in the noncognitive areas to occur at the higher levels of this factor. This factor was included in the more complex analyses to be discussed in Chapter XI.

Table X.1

RELATIONSHIP BETWEEN TEACHER'S AGE AND MEAN GAIN SCORES ON THE SEVEN OUTCOME MEASURES

and Mean Gain Scores on SB, PSI and AH-WPPSI

	SB				PSI				AH-WPPSI			
CA	Pre Score	R Gain+*	A Gain +	N	Pre Score	R Gain+	A Gain +	N	Pre Score	R Gain	A Gain+	N
55	88.19	3.08	-1.77	290	34.52	8.88	-.48	285	24.31	7.46	-1.19	280
53	88.97	5.40	0.80	334	33.09	9.07	-.72	328	21.06	9.30	0.45	329
55	86.36	6.56	1.12	309	33.29	10.70	.97	323	20.54	11.56	1.63	325
		103.13	82.16			58.94	46.66			162.27	142.15	

and Mean Gain Scores on GUMP, BP, MP and FI

GUMP				BP				MP				FI			
Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain+	N	Pre Score	R Gain+	A Gain+	N	Pre Score	R Gain	A Gain	N
34.33	8.10	.40	265	.76	-.09	.01	292	1.78	-.35	.04	292	.88	-.29	.01	29
33.98	7.62	-.24	338	.68	-.12	-.07	333	1.72	-.17	.17	333	.77	-.20	.01	33
34.34	7.27	-.42	330	.49	-.08	-.15	332	1.39	-.30	-.21	332	.67	-.17	-.03	33
	64.49	54.35			1.75	1.05			5.66	3.33			1.79	.95	

l Result.
t at .05 Level.

Table X.2

BETWEEN TEACHER FACTOR TI, SOCIA -EMOTIONAL INTERACTION AND MEAN GAIN SCORES ON THE SEVEN OUTCOME MEASURES

TI, Social-Emotional Interaction and Mean Gain Scores on SB, PSI and AH-WPPSI

	SB				PSI				AH-WPPSI			
CA	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain+*	A Gain+*	N
57	87.87	5.24	.28	217	33.50	10.38	.72	224	23.27	12.11	3.11	224
54	87.61	5.59	.55	429	33.52	9.48	-.18	425	21.40	10.14	0.51	433
50	90.41	4.08	-.06	229	31.42	9.86	-.44	224	19.71	6.76	-3.45	217
		104.12	83.50			52.52	40.93			164.23	141.61	

TI, Social-Emotional Interaction and Mean Gain Scores on GUMP, BP, MP and FI

GUMP				BP				MP				FI			
Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N
35.33	3.17	.85	223	.45	-.11	-.20	235	1.53	-.24	-.04	235	.74	-.16	.03	235
34.41	7.86	.20	430	.66	-.11	-.07	430	1.41	-.23	-.11	430	.73	-.26	-.08	430
33.11	7.50	-.76	218	.74	-.19	-.10	238	1.71	-.54	-.20	238	.86	-.42	-.13	233
	62.15	51.51			1.60	.93			5.19	2.95			1.79	.87	

Result.
at .05 Level.

Table X.3

:P BETWEEN TEACHER FACTOR TII, STRUCTURED LESSONS AND MEAN GAIN SCORES ON THE SEVEN OUTCOME MEASURES

TII, Structured Lessons and Mean Gain Scores on SB, PSI and AH-WPPSI

	SB				PSI				AH-WPPSI			
CA	Pre Score	R Gain+	A Gain+	N	Pre Score	R Gain	A Gain+*	N	Pre Score	R Gain+*	A Gain+*	N
52	90.63	5.70	1.64	213	31.57	9.28	-.97	213	19.39	9.62	-.69	213
52	90.59	4.18	.10	429	32.06	9.68	-.42	418	20.56	8.17	-1.75	415
60	82.36	6.26	-.47	233	35.79	10.49	1.52	242	24.77	12.73	4.24	246
		103.65	82.95			52.46	40.18			164.26	140.53	

II, Structured Lessons and Mean Scores on GUMP, BP, MP and FI

	GUMP				BP				MP				FI			
re ore	R Gain	A Gain	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain+	A Gain	N	Pre Score	R Gain+	A Gain+	N
.90	7.53	-.37	203	.55	-.02	-.05	218	1.44	-.13	.01	218	.71	-.14	.03	218
.59	8.27	.24	411	.71	-.15	-.08	435	1.49	-.26	-.09	435	.76	-.37	-.17	435
65	7.43	.34	257	.54	-.20	-.23	250	1.65	-.58	-.29	250	.83	-.22	.04	250
	62.05	51.67			1.59	.92			5.18	2.95			1.79	.87	

sult.
.05 Level.

5. OSCI-TIII - Teacher Emphasis on Art Activities (Table F-18)

This factor was not substantially or consistently related to adjusted gains, except for PSI gains which increased with more emphasis on Art Activities.

6. OSCI-TIV - Teacher Emphasis in Small Groups and Creative Materials (Table X.4)

The results for the PSI were substantial and the SB and AH-WPPSI almost reached the criterion for being substantial. Children exposed to a medium level of this factor improved the most on these three outcomes. MP and BR also improved the most for children exposed to the middle level of this factor. This variable was included in the more complex analyses to be discussed in Chapter XI.

7. OSCI-TV - Teacher Routines (Table F-19)

This factor yielded no interesting findings.

8. OSCI-TVI - Teacher Emphasis on Receptive Learning (Table X.5)

Children in low emphasis classes gained substantially more on the AH-WPPSI than children in high emphasis classes. SB and PSI showed a non-significant tendency to follow this pattern. Children in low emphasis classes also showed a significantly greater decrease in BP.

C. Classroom Interaction Variables

1. OSCI-CI - Classroom Emphasis on Structured Lessons (Table X.6)

There was an extremely large mean difference between the low and high levels of this factor on the AH-WPPSI. The mean adjusted gain for children exposed to the high level of this factor was 8.6 points higher than the mean adjusted gain for children exposed to the low level of this factor. A substantial pattern in the same direction can also be noted for the PSI. In the noncognitive area, there was a significantly greater decrease in BP at the high levels of this factor, but a significant increase in FI. This factor was selected for inclusion in the Chapter XI analyses.

2. OSCI-CII - Classroom Group Activities and Routines (Table F-20)

Nothing of interest was found for this factor.

Table X.4

:TWEEN TEACHER FACTOR TIV, CREATIVE INSTRUCTION AND MEAN GAIN SCORES ON THE SEVEN OUTCOME MEASURES

, Creative Instruction and Mean Gain Scores on SB, PSI and AH-WPPSI

	SB				PSI				AH-WPPSI			
CA	Pre Score	R Gain+*	A Gain+	N	Pre Score	R Gain+	A Gain+*	N	Pre Score	R Gain	A Gain+	N
55	86.11	4.85	-.67	253	31.06	10.01	-0.40	242	20.71	10.42	0.55	242
55	88.31	6.25	1.43	458	34.22	10.17	0.73	472	22.30	10.29	0.96	473
51	92.05	2.15	-1.46	158	31.99	8.50	-1.63	153	19.87	7.55	-2.60	153
		101.17	80.01			52.26	40.45			166.93	145.60	

, Creative Instruction and Mean Gain Scores on GUMP, BP, MP and FI

GUMP				BP				MP				FI			
e	R Gain	A Gain+	N	Pre Score	R Gain	A Gain+	N	Pre Score	R Gain	A Gain+	N	Pre Score	R Gain+	A Gain+	N
4	7.06	-.90	249	.64	-.08	-.06	258	1.59	-.27	-.02	258	.77	-.11	.10	258
0	7.90	.47	465	.51	-.13	-.18	476	1.40	-.38	-.27	476	.69	-.25	-.10	476
9	8.91	.78	151	.95	-.21	.01	163	1.70	-.17	.16	163	.98	-.56	-.19	163
	61.42	51.49			1.69	.93			5.15	2.88			1.77	.87	

, Level.

Table X.5

ETWEEN TEACHER FACTOR TVI, RECEPTIVE LEARNING AND MEAN GAIN SCORES ON THE SEVEN OUTCOME MEASURES

Receptive Learning and Mean Gain Scores on SB, PSI and AH-WPPSI

	SB				PSI				AH-WPPSI			
CA	Pre Score	R Gain+	A Gain	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain+*	A Gain+*	N
57	84.37	7.16	1.08	234	34.53	9.45	.10	229	23.80	11.46	2.64	231
54	89.00	4.10	-.49	422	32.56	10.07	.12	431	21.51	10.08	0.49	431
50	91.99	4.61	.98	201	32.24	9.64	-.40	193	19.09	7.07	-3.35	191
		103.01	83.52			52.22	41.01			167.47	144.03	

Receptive Learning and Mean Gain Scores on GUMP, BP, MP and FI

GUMP			BP				MP				FI			
R Gain	A Gain	N	Pre Score	R Gain	A Gain+	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain+	A Gain+	N
8.28	.93	231	.57	-.21	-.23	236	1.39	-.05	.05	236	.69	-.02	.13	236
7.96	.17	426	.64	-.17	-.15	447	1.54	-.41	.20	447	.75	-.35	-.19	447
7.31	-.60	195	.67	.03	.08	201	1.54	-.26	-.05	201	.76	-.28	-.08	201
61.10	51.09			1.59	.91			5.16	2.98			1.73	.86	

Level.

Table X.6

SHIP BETWEEN CLASS FACTOR CI, STRUCTURED LESSONS AND MEAN GAIN ON THE SEVEN OUTCOME MEASURES

Structured Lessons and Mean Gain Scores on SB, PSI and AH-WPPSI

	SB				PSI				AH-WPPSI			
CA	Pre Score	R Gain+	A Gain+	N	Pre Score	R Gain+	A Gain+*	N	Pre Score	R Gain+*	A Gain+*	N
51	91.50	3.29	-0.50	201	32.31	8.20	-1.83	188	19.74	6.77	-3.43	187
52	90.32	5.31	1.14	430	31.07	10.43	0.03	446	19.69	9.19	-1.03	443
61	82.48	6.24	-0.45	244	37.06	9.91	1.33	239	25.99	13.27	5.18	244
		103.37	82.91			51.90	39.92			162.43	136.59	

Structured Lessons and Mean Gain Scores on GUMP, BP, MP and FI

	GUMP			BP				MP				FI			
re	R Gain	A Gain	N	Pre Score	R Gain	A Gain+	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain+	A Gain+	N
.21	8.18	0.03	187	.76	-.07	.03	202	1.65	-.19	.10	202	.92	-.51	-.17	202
.50	7.87	0.21	431	.61	-.12	-.11	439	1.45	-.29	-.15	439	.69	-.23	-.09	439
37	7.58	0.82	253	.56	-.20	-.23	262	1.55	-.44	-.23	262	.78	-.16	.06	262
	62.16	51.64			1.59	.92			5.20	2.94			1.78	.87	

ult.
.05 Level.

4. OSCI-CIV - Classroom Verbal Communication (Table X.7)

The children exposed to the low level of this factor gained over five and one-half AH-WPPSI points more than the children exposed to the high level of this factor. Children at the low level also did substantially better on the GUMP. The findings for the noncognitive measures were non-substantial and inconsistent.

5. OSCI-CV - Classroom Small Group Instruction in Creative Arts (Table F-22)

There were no substantial or consistent findings regarding this variable.

6. OSCI-CVI - Classroom Language and Discrimination Learning (Table X.8)

There were no substantial findings regarding this variable. However, SB gains were significantly higher for high emphasis on this factor and AH-WPPSI gains were significantly lower for high emphasis on this factor.

7. Class Factors for 1967-68 (Table IX.2)

The four OSCI class factors for 1967-68 were not as confounded with the pre scores as the 12 OSCI teacher and class factors were for 1968-69. Consequently, the gains for these four factors were not adjusted for pre score and instead, raw gain scores were used in the analysis. There is not a one-to-one correspondence between these factor scores and the 1968-69 OSCI factor scores, but the significantly greater gain in SB for the medium and high emphasis classes on 1967-68 OSCI-III (Cognitive-High Structure) relative to low emphasis classes supports the findings for 1968-69 in respect to greater cognitive gains occurring in the more highly structured programs.

D. Discussion and Summary

Except for age, teacher demographic variables were not clearly related to the outcome measures for either 1967-68 or 1968-69. The cognitive measures tended to be more susceptible to program variation, as measured by the OSCI, than the noncognitive measures. The AH-WPPSI was extremely sensitive to program variation. Substantially higher gains on the AH-WPPSI were found for those children exposed to a low level of Teacher Social-Emotional Interaction, a low level of Teacher Emphasis on Passive-Receptive Learning, a low level of Classroom Verbal Communication, and a high level on both Teacher and Classroom Emphasis on Structured Lessons. Children also did substantially better on the PSI when exposed to a high level of both Teacher and Classroom Emphasis on Structured Lessons. Of the cognitive measures, the SB tended to be the least sensitive to program variation,

Table X.7

BETWEEN CLASS FACTOR CIV, VERBAL COMMUNICATION AND MEAN GAIN SCORES ON THE SEVEN OUTCOME MEASURES

, Verbal Communication and Mean Gain Scores on SB, PSI and AH-WPPSI

	SB				PSI				AH-WPPSI			
CA	Pre Score	R Gain	A Gain	N	Pre Score	R Gain+	A Gain	N	Pre Score	R Gain+*	A Gain+*	N
59	86.12	4.64	-0.88	212	35.94	9.61	0.69	214	24.38	12.80	4.17	214
53	88.26	5.35	0.52	461	33.70	9.16	-0.45	454	21.87	8.51	-.97	456
50	91.14	5.05	1.15	202	28.28	11.45	0.20	205	17.47	9.57	-1.40	204
		104.43	83.04			51.79	40.90			164.87	141.96	

, Communication and Mean Gain Scores on GUMP, BP, MP and FI

GUMP				BP				MP				FI			
re core	R Gain	A Gain+*	N	Pre Score	R Gain+	A Gain+	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain	A Gain+	N
5.94	8.17	1.22	212	.76	-.42	-.32	233	1.43	-.29	-.16	233	.84	-.24	.02	233
4.05	8.03	0.20	464	.58	.00	-.01	465	1.55	-.25	-.03	465	.77	-.23	-.02	465
2.97	7.08	-1.24	195	.57	-.10	-.12	205	1.56	-.49	-.26	205	.69	-.40	-.26	205
	62.03	51.13			1.57	.91			5.20	2.95			1.80	.87	

sult.
.05 Level.

Table X.8

CLASS FACTOR CVI, LANGUAGE AND DISCRIMINATION LEARNING AND MEAN GAIN SCORES ON THE SEVEN OUTCOME MEASURES

, Language and Discrimination Learning

	SB				PSI				AH-WPPSI			
CA	Pre Score	R Gain+	A Gain+	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain+	A Gain+	N
56	87.79	5.03	0.05	214	34.10	10.09	.60	213	22.76	11.39	2.22	211
53	88.57	4.20	-0.53	414	32.51	9.85	-.12	415	20.72	9.72	-.14	422
53	89.09	6.54	1.98	234	32.47	9.51	-.47	232	21.40	8.28	-1.36	227
		103.53	82.59			52.72	41.22			168.34	146.17	

, Language and Discrimination Learning and Mean Gain Scores on GUMP, BP, MP and FI

GUMP				BP				MP				FI			
Pre core	R Gain	A Gain	N	Pre Score	R Gain	A Gain+	N	Pre Score	R Gain	A Gain	N	Pre Score	R Gain+	A Gain+	N
5.48	7.78	.61	217	.50	-.18	-.24	228	1.70	-.50	-.17	228	.93	-.43	-.10	228
.83	8.23	.31	405	.70	-.14	-.07	420	1.42	-.15	-.04	420	.74	-.31	-.12	420
3.80	7.27	-.67	235	.58	-.04	-.05	242	1.47	-.31	-.15	242	.70	-.09	.07	242
	62.70	52.07			1.57	.91			5.17	2.98			1.78	.87	

.05 Level.

and did not clearly follow the patterns of the AH-WPPSI and PSI. It is interesting to note, however, that the raw SB gains did follow the AH-WPPSI and PSI patterns. The adjusted gains for SB tended to "overcorrect" for the pre SB which was highly confounded with the OSCI factors. Furthermore, the pre SB was confounded in such a way that low pre SBs occurred at the OSCI factor levels where program effects would be hypothesized to be greatest (e.g., pre SB mean of 81 for high level of Classroom Structured Lessons). The OSCI factors are also, of course, confounded with other factors such as region, which are related to gains and which could be partly responsible for the SB results in these univariate analyses. The analysis of variance and regression procedures discussed in the next chapter attempt to adjust for some of these other important factors.

In regard to the noncognitive factors, there was a tendency for improvement on the BP scale to follow the AH-WPPSI and PSI patterns.

In summary, cognitive gains were associated with programs having a high proportion of time spent on language, arithmetic and "structured" activities. These findings support the notion that high-structured, focused, and well-implemented compensatory programs can bring about greater immediate cognitive gains than low-structured, diffuse, and less well-implemented programs.

CHAPTER XI

ANALYSES OF VARIANCE AND REGRESSION ANALYSES FOR SELECTED OUTCOME GAINS FOR 1968-69

The previous chapters presented some of the univariate effects of child-background and program variables on gain scores. One problem with these analyses is that the explanatory variables are confounded with each other and the univariate estimates of effects thus may be somewhat spurious. That is, the effect that is observed in the univariate analysis could be a result of the variable of interest, some other variable(s) that c varies with the variable of interest, or a joint function of the variable of interest and other variables that covary with it.

In this chapter we first discuss the results of the analysis of variance (ANOVA) of selected outcome gains with selected combinations of child-family background and Head Start program factors for 1968-69. This analysis amounts to an analysis of variance on a nonorthogonal unbalanced design. Since the factors considered are confounded, certain combinations of the factor levels do not appear in the design; consequently, all interactions cannot be estimated. This analysis is equivalent to a regression analysis in which the independent variables are dummy variables (0, 1 variables) which indicate the level of a factor in which the observation falls.

Among the factors considered for the analysis, there are categorical variables as well as continuous variables. There are two approaches to including continuous variables in the analysis of variance models. The first approach is to treat them as covariates. The second approach is to categorize them and treat the categories as levels of a factor. For the categorical variables, the different categories are treated as levels of that factor in the analysis of variance model. In the analyses discussed here, the continuous variables were categorized.

There were a number of reasons for categorizing the continuous independent variables: it was easier to include and interpret interaction terms in the model; it was easier to talk in terms of categories for programs or children; some of the continuous variables were nonlinearly related to the dependent variables; and the problem of considering the prescore as a continuous covariate was eliminated.

A. Analyses of Variance

interactions were not included in the model because of their difficulty in interpretation and the nature in which the factors were confounded. In some cases even first-order interactions between two factors were not estimable because there were some cells with no observations. An example would be the interaction for 1968-69, between geographical regions and OSCI Classroom Structured Lessons. There were no classrooms in the South that operated under the lowest level of this OSCI program component. Consequently, there were empty cells and not enough information to estimate interaction parameters.

The analyses attempt to estimate and analyze the significance of the main effects of child-family factors and program factors and their interactions. If the main effect of a factor is statistically significant it could be interpreted to mean that the factor explains some variation in gain when adjusted for other factors in the model. The statistical significance of all the parameters for each main effect and interaction is tested in the presence of all the other factors considered in the model.

The child-family background variables to be included in these analyses were either statistically significant in the univariate gain analyses or were hypothesized to interact with program factors. There were numerous possible child-family background factors to consider. The variables selected were child's age, sex, pre score, socioeconomic level, and geographical region in which he or she resided. Geographical region was related to many other child and program characteristics such as age, race, and, in 1968-69, OSCI factor scores, as well as to gain scores. Consequently, it would seem to be an important variable to consider in the model for adjusting the effects of other factors and interactions in the model. Pre score was included to adjust the other parameters for regression effects. Pre score by program interaction parameters were also included in the models. Of primary interest would be any age or SES main effects, age-program interactions, SES-program interactions or pre score-program interactions. The length of the pre/post testing interval was also included in the analyses because it was significantly related to both OSCI scores and gain scores.

The program factors considered for these analyses were four OSCI Class and Teacher factors that were consistently related to the various outcome gains in the univariate analyses. The OSCI factors considered in the analyses were OSCI Class Factor I (Structured Lessons), OSCI Class Factor VI (Language and Discrimination), OSCI Teacher Factor II (Structured Lessons), and OSCI Teacher Factor IV (Creative Instruction - Small Group).

The outcome variables to be considered in these analyses were SB raw gain, PSI raw gain and MP raw gain. The Gumpgookies was not considered because it was relatively insensitive to child and program variation. The other two rating scales from the Inventory of Factors Affecting the Stanford-Binet were not considered because they contained less items and had less desirable distribution properties than the Motivation Problem scale which contained ten items. The AH-WPPSI was extremely sensitive to program factors in the univariate analyses but because of time and cost constraints was not further considered.

The child-family, program, and outcome measures selected for these detailed analyses were judged to be a reasonable subset of variables on the basis of previous analyses and theoretical interest. Many other potentially important child-family and teacher-program variables for both 1967-68 and 1968-69 were not considered in these analyses and should be considered in subsequent analyses of these data.

Another problem was that due to missing data only a relatively small number of factors and relevant interactions could be included in any one analysis. In many cases, the inclusion of four explanatory variables and one outcome measure reduced the original data file of children from 1998 records to 500 or 600 observations. This is not serious if the explanatory variables are not highly confounded and can be subdivided into relatively independent subsets, because the parameter estimates will not change significantly by adding variables that are relatively independent of variables already in the model. This problem was rectified somewhat by considering the effect of the OSCI factors in the context of different sets of explanatory variables. In this way, a check could be made on the robustness of the explanatory variable.

A definition of the explanatory variables follows. The children were divided into two age categories--less than four and one-half years of age and four and one-half years of age and greater. This resulted in a fairly even split in age for all analyses. The children were also classified into one of three SES categories as discussed previously: low-low, medium-low, and upper-low. Each of the four OSCI factors was categorized into the three levels (low, medium, and high) discussed previously. Three levels of SB and PSI prescore were considered. For these two measures, the low, medium, and high levels were comprised of roughly the lower 25%, the middle 50%, and the upper 25% of the distribution, respectively. Only two levels of the pre score for the Motivation Problem scale were considered: children who exhibited no motivational problems, and children who exhibited one or more motivational problems. The definition of the levels for the various factors seemed justifiable from both a conceptual and a data analytic viewpoint.

Sixteen ANOVAs were undertaken for each of the three selected outcome variables. The set of explanatory variables and the models were identical for each of them, and Table XI.1 summarizes the models used for each of the three outcome variables.

Table XI.1 describes the main effect parameters and interaction parameters estimated for each of the 16 models. For example, in model 1 the main effect and interaction parameters were estimated for region, pre score, age, OSCI Class Factor I, Pre by Age Interaction, Pre by OSCI Class Factor I Interaction, and Age by OSCI Class Factor I Interaction for the gains on each of the three outcome variables. The ANOVA model for SB gain can be written as:

$$\text{SB gain } Y_{ijkln} = \mu + R_i + P_j + A_k + O_l + (PA)_{jk} + (PO)_{jl} + (AO)_{kl} + \varepsilon_{ijkln}$$

where SB gain Y_{ijkln} is the gain for the nth child in region i, pre score category j, age category k, and OSCI Factor I level l,

μ is the overall effect

R_i is the effect for Region i

P_j is the effect for pre score level j

A_k is the effect for age category k

O_l is the effect for OSCI Factor I level l

$(PA)_{jk}$ is the effect for the jkth pre score-age cross classification

$(PO)_{jl}$ is the effect for the jlth pre score-OSCI cross classification

$(AO)_{kl}$ is the effect for the klth-age OSCI cross classification

ε_{ijkln} is the error for the nth child in region i, pre score category j, age category k, and OSCI Class Factor I level l.

The constraints on the parameters are

$$\Sigma P_j = 0,\ \Sigma A_k = 0,\ \Sigma O_l = 0,\ (\Sigma_j (PA)_{jk} = 0,\ \Sigma_k (PA)_{jk} = 0),$$

$$(\Sigma_j (PO)_{jl} = 0,\ \Sigma_l (PO)_{jl} = 0),\ \text{and}\ (\Sigma_k (AO)_{kl} = 0,\ \Sigma_l (AO)_{kl} = 0).$$

Table XI.1

THE SIXTEEN ANOVA MODELS USED SINGLY WIT SB, PSI, AND MP RAW GAINS AS DEPENDENT VARIABLE

ANOVA Model	Main Effects				Interaction Effects		
1	Region	Pre-Score for outcome	Age	OSCI-C-I	Pre x Age	Pre x OSCI-C-I	Age x OSCI-C-I
2	Region	Pre-Score for outcome	Age	OSCI-C-VI	Pre x Age	Pre x OSCI-C-VI	Age x OSCI-C-VI
3	Region	Pre-Score for outcome	Age	OSCI-T-II	Pre x Age	Pre x OSCI-T-II	Age x OSCI-T-II
4	Region	Pre-Score for outcome	Age	OSCI-T-IV	Pre x Age	Pre x OSCI-T-IV	Age x OSCI-T-IV
5	Region	Pre-Score for outcome	Sex	OSCI-C-I	Pre x Sex	Pre x OSCI-C-I	Sex x OSCI-C-I
6	Region	Pre-Score for outcome	Sex	OSCI-C-VI	Pre x Sex	Pre x OSCI-C-VI	Sex x OSCI-C-VI
7	Region	Pre-Score for outcome	Sex	OSCI-T-II	Pre x Sex	Pre x OSCI-T -II	Sex x OSCI-T-II
8	Region	Pre-Score for outcome	Sex	OSCI-T-IV	Pre x Sex	Pre x OSCI-T-IV	Sex x OSCI-T-IV
9	Region	Pre-Score for outcome	SES	OSCI-C-I	Pre x SES	Pre x OSCI-C-I	SES x OSCI-C-I
10	Region	Pre-Score for outcome	SES	OSCI-C-VI	Pre x SES	Pre x OSCI-C-VI	SES x OSCI-C-VI
11	Region	Pre-Score for outcome	SES	OSCI-T-II	Pre x SES	Pre x OSCI-T-II	SES x OSCI-T-II
12	Region	Pre-Score for outcome	SES	OSCI-T-IV	Pre x SES	Pre x OSCI-T-IV	SES x OSCI-T-IV
13	Region	Pre-Score for outcome	Pre-Post Interval	OSCI-C-I	Pre x Pre-Post Interval	Pre x OSCI-C-I	Pre-Post Interval x OSCI-C-I
14	Region	Pre-Score for outcome	Pre-Post Interval	OSCI-C-VI	Pre x Pre-Post Interval	Pre x OSCI-C-VI	Pre-Post Interval x OSCI-C-VI
15	Region	Pre-Score for outcome	Pre-Post Interval	OSCI-T-II	Pre x Pre-Post Interval	Pre x OSCI-T-II	Pre-Post Interval x OSCI-T-II
16	Region	Pre-Score for outcome	Pre-Post Interval	OSCI-T-IV	Pre x Pre-Post Interval	Pre x OSCI-T-IV	Pre-Post Interval x OSCI-T-IV

Region has four levels: (1) North, (2) South, (3) Central, and (4) West. Pre-Score and OSCI factors have three levels: (1) low, (2) middle, and (3) high. Age has two levels: (1) young, and (2) old. SES has three levels: (1) low-low, (2) middle-low, and (3) upper-low. Pre-post Interval has three levels: (1) short, (2) medium, and (3) long.

These are the arbitrary constraints used so that the parameters in the model could be estimated. The predicted gain for a child can be found by estimating the parameters (e.g., R_1, etc.) of the above model and then adding the appropriate parameters to yield the predicted gain. For example, if we wanted to predict the SB gain for a child in Region 1, SB pre-score level 1, age level 1, and OSCI Factor T-I - level 1, then we would simply add up the appropriate parameters to obtain the predicted gain from the model. In this case, the predicted gain for any child falling at the specified levels of these 4 factors would be

$$\hat{Y}_{1111n} = \mu + R_1 + P_1 + A_1 + O_1 + (PA)_{11} + (PO)_{11} + (AO)_{11}.$$

If the set of estimated parameters corresponding to a particular main effect or two factor interaction is close to zero, then the set of estimated parameters would not contribute to the prediction equation.

1. SB ANOVA

The significant main effects and interactions for the 16 SB models are shown in Table XI.2. The table follows the same format as Table XI.1 but only the significant effects are listed. In regard to the SB gain analyses, region was highly significant (.001) for each of the 16 models. The models yielded positive effects for the North and South and negative effects for the Midwest and West.

As expected, the pre SB was highly significant (.001) for all 16 models. Earlier analyses indicated that pre SB and raw gains correlated approximately -.45. The lowest SB category yielded a positive effect while the highest SB category yielded a negative effect in all of the analyses. This primarily reflects quite large changes in the extremely low (SB 70 or less) initial scores.

All four models that included age as a factor indicated nonsignificance for the age main effect but the significance of age by pre score interaction in three of the models indicated that age was needed as an "explanatory" variable. The univariate relationship of age to SB gain indicated no monotonic trend of any practical significance. The four models indicate that age is not needed to account for gains when other explanatory variables are included in the model. Both region and pre scores are significantly related to age in these analyses. It should also be recalled that the predominantly Southern oldest children also experienced the most structured programs in 1968-69.

Likewise SES was not a statistically reliable associate of SB gain when adjusted for the effects of region, pre SB, and other factors, probably because both region and pre SB were significantly associated with SES. SES seems to add no unique contribution to pre score, region, and other factors in explaining SB gains.

The univariate analyses indicated that sex did not relate significantly to gain. Child sex was not a significant factor in the four models involving sex.

Table XI.2

SIGNIFICANT FACTORS IN ANOVAS WITH RAW GAINS ON SB

Model Number	Factors Significant at .05 Level
1	Region, and Pre Score
2	Region, Pre Score, OSCI Class Factor VI and (Pre Score by Age Interaction)
3	Region, Pre Score and (Pre Score by Age Interaction)
4	Region, Pre Score, OSCI Teacher Factor IV and (Pre Score by Age Interaction)
5	Region, and Pre Score
6	Region, Pre Score and OSCI Class Factor VI
7	Region, Pre Score and OSCI Teacher Factor II
8	Region, Pre Score and OSCI Teacher Factor IV
9	Region and Pre Score
10	Region, Pre Score and OSCI Class Factor VI
11	Region and Pre Score
12	Region, Pre Score and OSCI Teacher Factor IV
13	Region, Pre Score, Pre-Post Interval, OSCI Class Factor I and (Pre-Post Interval by OSCI Class Factor I Interaction)
14	Region, Pre Score, Pre-Post Interval, OSCI Class Factor VI and (Pre-Post Interval by OSCI Class Factor VI Interaction)
15	Region, and Pre Score
16	Region, Pre Score, Pre-Post Interval, OSCI Teacher Factor IV and (Pre-Post Interval by OSCI Teacher Factor IV Interaction)

The findings regarding the association of child sex and gains are among the most puzzling and inconsistent in the literature. The E & R data are quite clear in showing no discernible effects of child sex on initial or final scores, whether as a main variable or in combination with other variables. Weikart (1971) found that most of the effects of the Perry Preschool program were on girls, both immediately and in the follow-up study. Kraft, et al. (1968) report greater benefits for boys. Di Lorenzo and Salter (1969) found no reliable effects of child sex, both immediately and in the long run. Many studies (e.g., Schaefer, find teachers view girls, even as young as 5, 6 and 7 years of age, as more task-oriented, and extraverted--far more favorably than they do boys. Typically, boys are reported to have lower initial scores than girls, and as Erickson et al. (1969) conclude, the plight of the disadvantaged boy who must enter first grade without a preschool is pitiful indeed, a finding consistent also with Miller and Dyer's (1970) observations: both of these projects studied Head Start children. The question of the conditions under which sex differences do and do not appear is an important one, which clearly needs further investigation.

The length of the pre/post testing interval had a significant monotonic relationship with raw gain scores in a previously discussed univariate analysis. There was also a significant monotonic relationship between the length of the pre-post interval and the gain scores corrected on the basis of the pre scores. In both cases, the longer the pre-post interval, the greater the gain. The parameters in the model for pre-post interval indicated the same trend for each of the four SB analyses which considered the pre-post interval. In three cases, the pre-post interval was significant at the .05 level and in one case the .10 level.

However, it is difficult to interpret the pre/post testing interval as a measure of program exposure since in some instances the pre test was given two or three months after the child was first exposed to the program.

The four OSCI Class and Teacher Factors considered in these analyses were selected because of their power to explain gain scores in the univariate analyses. Of the four analyses that included OSCI Class Factor I (Structured Lessons), one of them indicated significance at the .05 level (model 13), two indicated significance at the .10 level (models 1 and 5) and one analysis (model 9) indicated no significance. The model which indicated the greatest significance for OSCI Class Factor I was the model which included the length of the pre-post interval. In this analysis, the classroom with middle emphasis on classroom structured activities had the largest positive effect and classrooms with the lowest emphasis had the largest negative effect. The other two analyses where OSCI Class Factor I was significant at the .10 level indicated the same phenomena, greatest gains at the middle level of classroom emphasis of structured lessons. The univariate adjusted SB gains analyses also indicated that children at the middle level were gaining significantly more than either the low or high levels.

The results of the four analyses which considered the level of language and discrimination emphasis as reflected in OSCI Class Factor VI yielded more consistency than the analyses involving classroom structured lessons. All four analyses indicated significance at the .05 level. The effects of this factor seemed to be quite strong in that the adjustments on the basis of four subsets of variables did not alter the conclusion that those classes where language and discrimination were highly emphasized gained the most in SB. These results were also consistent with the univariate analyses with both corrected and uncorrected SB gains. All analyses indicated a two to two and one-half points adjusted SB gain difference between classes low and high on this factor.

OSCI Teacher Factor II (Structured Lessons) yielded a pattern of results similar to that of OSCI Class Factor I (Structured Lessons). That is, two analyses indicated no significance, one marginal significance (.10) and one significance at the .05 level. The two analyses that were significant indicated a negative monotonic trend. The low level had a positive parameter and the high level a negative parameter, i.e., low emphasis gained more than high emphasis. The two insignificant analyses indicated trends in the same direction. The univariate analyses of adjusted SB gains also indicated that the lows gained the most and the highs gained the least.

OSCI Teacher Factor IV (Creative Instruction-Small Group), like OSCI Class Factor VI (Language and Discrimination), was highly significant in accounting for gain score variation in the context of four different subsets of explanatory factors. All four of these analyses indicated that the middle level of Teacher Creative Instruction produced the most gains, and the highest level produced the least gains. These results are consistent with the univariate analyses. The adjusted gain differences between the high and middle groups for the analyses averaged around three SB points. It is not unreasonable to expect classrooms where teachers do not either overemphasize or underemphasize creative instruction to produce high SB gains relative to either of the two extremes. Teachers spending too much time on creative instruction activities might be sacrificing time better spent on other activities while teachers spending too little time on creative instruction may not be emphasizing other skills necessary to improve performance on the Stanford-Binet.

It is interesting to note that there were no significant interactions between child-family background variables and OSCI Factors. The lack of any child-program interactions may reflect insensitivity of the statistical model, given the confounding that reduced cell Ns, to interactive effects, or it may mean that child-program interactions are not as powerful, given the variables studied, as some educators would anticipate.

An examination of the significant interaction effects in the SB raw gain analyses of variance shows that older children with high pre scores gained more than would be expected on the basis of age, pre score and other main and interaction effects (models 2,3, 4). Children who had a short pre post testing interval and who were under a low level of Class Structured Lessons gained less than would be expected on the basis of age, pre score and other main and interaction effects (model 13). However, short pre/post testing interval children under a high emphasis on Class Language and Discrimination (model 14), or under a middle level of Teacher Creative Instruction (model 16), gained more than would be expected on the basis of main and other interaction effects.

In summarizing the ANOVAs on SB gains, it can be said that program variation as measured by the OSCI contributed significantly more in explaining SB gains than child variables, and that there were no interactions of child-background factors with the OSCI Factors. Classroom emphasis on language and discrimination, and teacher emphasis on creative instruction (moderate level) seemed to play a major role in explaining SB gains in these data.

2. PSI ANOVA

The significant main effects and interactions for the PSI models are shown in Table XI.3. Region was significant for all 16 PSI models. Children in the South and West had the largest gains on the PSI. Children in these regions were significantly different from children in the North and Midwest in a number of ways as discussed previously. The regional effect could be due to any of these subpopulation difference. As to be expected, the prescores were significant in all 16 PSI analyses. Sex only had marginal significance (.10) in one of the four analyses and consequently will not be further discussed. Socioeconomic level was nonsignificant in all four analyses. Age, which interacted significantly with the pre score and resulted in nonsignificant age main effects in the SB. analyses, was significant at the .05 level in all four PSI analyses. The parameters in the model for age suggested that the older children gained more on the PSI. This is in agreement with the univariate analyses with adjusted PSI scores. It is interesting that age adds explanatory power to the model even when pre score and regional information are included in the model, because both pre score and region are associated with age. Of course, other differences between young and old children (such as ethnic composition) that were not included in the models could account for the finding that older children gain more on the PSI.

OSCI Class Factor I (Structured Lessons) was significant in all four analyses. The low groups gained significantly less than the high group. The adjusted mean difference in gains between low and high was around three PSI points.

Table XI.3

SIGNIFICANT FACTORS IN ANOVAS WITH RAW GAINS ON PSI

Model Number	Factors Significant at .05 Level
1	Region, Pre Score, Age and OSCI Class Factor I
2	Region, Pre Score and Age
3	Region, Pre Score and Age
4	Region, Pre Score, Age and OSCI Teacher Factor IV
5	Region, Pre Score and OSCI Class Factor I, and (Sex by OSCI Class Factor I Interaction)
6	Region, and Pre Score
7	Region, and Pre Score
8	Region, and Pre Score
9	Region, Pre Score, OSCI Class Factor I
10	Region, and Pre Score
11	Region, and Pre Score
12	Region, and Pre Score
13	Region, Pre Score, Pre-Post Interval, OSCI Class Factor I and (Pre-Post Interval by OSCI Class Factor I Interaction)
14	Region, Pre Score and Pre-Post Interval
15	Region, Pre Score, Pre-Post Interval, OSCI Teacher Factor II and (Pre-Post Interval by OSCI Teacher Factor II Interaction)
16	Region, Pre Score, Pre-Post Interval and OSCI Teacher Factor IV

OSCI Class Factor VI (Language and Discrimination), which was highly significant in explaining SB gains, was nonsignificant in all four PSI analyses.

OSCI Teacher Factor IV (Creative Instruction - Small Group) was significant for two analyses, marginally significant for one analysis, and nonsignificant for the remaining analysis. The analysis which was not significant did not consider age in the model. The three significant analyses indicated that the children exposed to the high level of this factor gained significantly less than those children exposed to the low and middle levels of this factor. These results are consistent with the SB findings in that the highest gains for SB were at the middle level and the lowest gains at the high level. It is clear that relatively high emphasis on creative instruction activities is associated with lower gains on both the PSI and Stanford-Binet. OSCI Teacher Factor II (Structured Lessons) only showed up as significant when the length of the pre-post interval was considered in the analysis. In that one instance, the largest gains occurred at the high level of Teacher Structured Lessons.

The length of the pre/post testing interval was significant in all four models where this effect was included. Children under the long interval gained 2 or 3 more PSI points than children under the short interval when adjusting for the other main and interaction effects. Although the interval length had a significant impact for both the univariate and multivariate analyses, it is difficult to interpret this interval as an exposure measure.

On the PSI, examinations of the significant interactions showed that females gained more than males under a high level of Class Structured Lessons than would be expected on the basis of the main effects for sex, OSCI-C-I and other main and interaction effects (model 5). A high level of this factor combined with a medium pre-post testing interval produced less gain than would be expected (model 13). Finally, a short pre/post testing interval combined with a high level of Teacher Structured Lessons, and a long interval combined with a low level of this same factor (model 15), produced more gain than would be expected on the basis of the main and other interaction effects.

3. MP ANOVA

The significant main effects and interactions for the 16 MP models are shown in Table XI.4. The Motivation Problem scale analyses indicated no significant main effects for the four OSCI factors. Region was significant in all but three of the 16 MP analyses. Children in the North and South improved more on MP than children in the Midwest and West. Children who had a long pre/post interval improved the most on MP. SES and Age had no significant main effects; however, in model 4 age interacts with OSCI Teacher Factor IV. Sex had a significant main effect in only one of the

Table XI.4

SIGNIFICANT FACTORS IN ANOVAS WITH RAW GAINS ON MP

Model Number	Factors Significant at .05 Level
1	Region and Pre Score
2	Region and Pre Score
3	Region and Pre Score
4	Region, Pre Score and (Age by OSCI Teacher Factor IV Interaction)
5	Region and Pre Score
6	Region, Pre Score, and Sex
7	Region and Pre Score
8	Region and Pre Score
9	Region and Pre Score
10	Region, and Pre Score
11	Region and Pre Score
12	Region and Pre Score
13	Pre Score, Pre Post Interval, and (Pre Post Interval by Pre Score Interaction)
14	Region, Pre Score Pre-Post Interval, and (Pre-Post Interval by OSCI Class Factor VI Interaction)
15	Pre Score and (Pre-Post Interval by Pre Score Interaction)
16	Pre Score, Pre-Post Interval and (Pre-Post Interval by Pre Score Interaction)

four analyses and consequently will not be discussed further. As expected, pre score was significant in all 16 analyses. The children with one or more problems on the pre measure showed improvement while those children with no problems showed an increase in problems. This was probably artifactual in that a child with no problems could not "improve," but could only remain problem-free (no gain) or show problems. There were no pre MP score by OSCI, SES by OSCI, or sex by OSCI interactions. Age interacted significantly only with OSCI Teacher Factor IV (Creative Instruction); older children showed a greater decrease in MP than did younger children under a high level of Teacher Creative Instruction (model 4). The remaining interactions were not interpretable.

B. Regression Analyses

The previous analyses involved the consideration of four OSCI Class and Teacher Factors singly, in conjunction with child variables (e.g., age) and control variables (e.g., region). In the present section, all 12 OSCI factors are considered jointly with the pre score in predicting gain scores for six of the seven 1968-69 outcome measures. (The AH-WPPSI was omitted.) The original OSCI factor score distributions were used in these analyses. The multiple regression model was used to regress raw gains on the 12 OSCI Factor scores and the appropriate pre-score. The weights are for each of the 12 OSCI factor scores and the appropriate pre-score. A particular regression weight indicates the effect of that variable adjusted for the effects or influences of the remaining variables. The magnitude and sign of these weights can be considerably different than what would be obtained if gain was regressed on each OSCI Factor separately, because the relationships of the independent variables with one another as well as with the dependent variable are considered in determining multiple regression weights.

1. SB Gain (Table XI.5)

In the SB gain analyses, four OSCI scores had regression weights significant at the .01 level and three more OSCI scores had regression weights significant at the .05 level. The pre score for SB was statistically significant at the .01 level. The two most significant positive regression weights were for Class Verbal Communication and Class Language and Discrimination. It should be noted that the Class Language and Discrimination factor was highly significant in the previous SB ANOVA. Class Structured Lessons had a

Table XI.5

MULTIPLE REGRESSION WEIGHTS FOR OSCI FACTOR SCORES WITH SB GAIN AS DEPENDENT VARIABLE

Independent Variables	Multiple Regression Weights
SB Pre Score	-.305**
OSCI Class Factors	
I. Structured Lessons	-1.92*
II. Group Activities and Routines	- .56
III. Social-Emotional Interaction	1.00*
IV. Verbal Communication	1.58**
V. Instruction in Creative Arts	-2.07**
VI. Language and Discrimination	1.94**
OSCI Teacher Factors	
I. Social-Emotional Interaction	-2.02**
II. Structured Lessons	1.43*
III. Art Activities	- .10
IV. Creative Instruction - Small Group	.65
V. Routines	.71
VI. Receptive Learning	- .93
Constant	32.47

R^2 = .190 for pre score alone

R^2 = .252 for full model

* Significant at .05 level

** Significant at .01 level

negative regression weight, while Teacher Structured Lessons had a positive weight. The Class factor scores seemed to be more predictive than the Teacher factors for SB gains. The R^2 for the 12 OSCI scores and pre-SB was .252 while the R^2 for the pre-SB alone was .190. R is the correlation of the predicted gains derived from the multiple regression model with the observed gains, and R^2 indicates the proportion of gain score variation that is explained by the independent variables. The OSCI scores added significant information to the pre-SB score in predicting SB gain scores. Although an increment in R^2 of only .062 is relatively small, it should be noted that there is considerable measurement error in all independent and dependent variables and that only a simple linear regression model was used to explain the variation in the SB gains.

2. PSI Gain (Table XI.6)

The regression analysis of PSI gain on the 12 OSCI Factor Scores and pre-PSI score yielded significant regression weights for only the pre-PSI score and Class Structured Lessons. These findings were in agreement with the previous PSI ANOVAs where Class Structured Lessons had the largest impact on PSI gains for the four OSCI factors considered. The regression weight was positive indicating a tendency for PSI gains to increase as emphasis on Class Structured Lessons increases. This same conclusion was also drawn from the univariate PSI adjusted gains analysis. The OSCI class and program factors were not as predictive of PSI gains as they were of SB gains. These regression analyses support previous findings where SB and PSI were differentially affected by OSCI Class and Teacher factors. The regression weights are presented in Table XI.6.

3. Motivation Problem Scale Gain

In the analysis of motivation gains, only Class Social-Emotional Interaction and Class Verbal Communication had regression weights significant at the .05 level. The weights were +.19 and -.19, respectively. These two factors were not considered in the previous motivation gain ANOVA where the four OSCI factors that were considered turned out to be insignificant. The R^2 for the full model was .466 while the R^2 for the pre score alone was .443. The regression weights indicate that classes with relatively low interactions of an emotional nature and high interactions of a verbal level tended to have larger decreases in motivational problems. Class observations were more predictive of change than teacher observations.

4. Behavior Problem Scale Gains

There was no significant contribution of the OSCI scores in predicting Behavior Problem gains. The R^2 for the pre score alone was .419 while the R^2 for the full model was only .430.

Table XI.6

MULTIPLE REGRESSION WEIGHTS FOR OSCI FACTOR SCORES WITH PSI GAIN AS DEPENDENT VARIABLE

Independent Variables	Multiple Regression Weights
PSI Pre Score	-.31**
OSCI Class Factors	
I. Structured Lessons	1.50*
II. Group Activities and Routines	.34
III. Social-Emotional Interaction	-.46
IV. Verbal Communication	.01
V. Instruction in Creative Arts	.16
VI. Language and Discrimination	-.74
OSCI Teacher Factors	
I. Social-Emotional Interaction	.71
II. Structured Lessons	-.30
III. Art Activities	.44
IV. Creative Instruction - Small Group	-.09
V. Routines	.40
VI. Receptive Learning	.31
Constant	20.00

R^2 = .214 for pre score alone

R^2 = .251 for full model

* Significant at .05 level

** Significant at .01 level

5. Feelings of Inadequacy (FI) Gain (Table XI.7)

Of the three scales from the Inventory of Factors Affecting the Stanford-Binet, the Feelings of Inadequacy gains were most predictable from the OSCI factor scores given the pre score. The R^2 for pre score alone was .477 and the R^2 for the full model was .531. The most significant variables in the regression were Class Social-Emotional Interaction and Class Verbal Communication.

Other regression weights significant at the .05 level were for Class Instruction in Creative Arts, Class Language and Discrimination, and Teacher Routines. The profile of regression weights is similar to the profile of regression weights for the Motivation Scale.

6. Gumpgookies Gain

The Gumpgookies analysis yielded two regression weights significant at the .05 level. The two regression weights were for Class Verbal Communication and Instruction in Creative Arts with weights of -1.02 and -.69, respectively. The R^2 for pre alone was .187 while the R^2 for the full model was .216.

C. Summary

The ANOVAs discussed in this chapter supported the previous univariate results and further substantiated the univariate findings that program factors were more important than child background factors in explaining gain. There was little evidence of any important interactions between child background and program variables in explaining gains.

When the OSCI factors were considered jointly, they were more predictive of SB gains than one would suspect from examining the results from the preceding univariate OSCI analyses and ANOVAs. The SB gain multiple regression analyses resulted in more significant regression weights than the parallel analyses for the remaining outcome variables. These regression analyses must be interpreted cautiously since they are based upon a linear model and we have seen previously that many relationships between the OSCI factors and the outcomes are nonlinear. They do, however, give us a rough indication of the relative predictability of the various outcome gains and the relative influence of teacher-based versus class-based observations.

The univariate adjusted gain analyses, ANOVAs, and multiple regression analyses all indicated that class emphasis on language and discrimination learning (OSCI-C-VI) was consistently and positively related to SB gains, but not to PSI gains. On the other hand, these same three sets of analysis indicated that emphasis on Class Structured Lessons (OSCI-C-I) was consistently and positively related to PSI gains, but not to SB gains. These findings indicate that different program strategies are necessary to enhance different

Table XI.7

MULTIPLE REGRESSION WEIGHTS FOR OSCI FACTOR SCORES WITH FI GAIN AS DEPENDENT VARIABLE

Independent Variables	Multiple Regression Weights
Feelings of Inadequacy Pre Score	-.81**
OSCI Class Factors	
I. Structured Lessons	.03
II. Group Activities and Routines	-.08
III. Social-Emotional Interaction	.13**
IV. Verbal Communciation	-.17**
V. Instruction in Creative Arts	-.09*
VI. Language and Discrimination	.12*
OSCI Teacher Factors	
I. Social-Emotional Interaction	-.11
II. Structured Lessons	.02
III. Art Activities	-.02
IV. Creative Instruction - Small Group	-.06
V. Routines	-.11
VI. Receptive Learning	-.02
Constant	.29

R^2 = .477 for pre score alone

R^2 = .531 for full model

* Significant at .05 level

** Significant at .01 level

cognitive skills. Although the PSI and SB pre scores are moderately correlated, gains on these variables are differentially related to the various OSCI Class and Teacher Factors. These data indicate that in order to simultaneously enhance the development of a large number of cognitive and noncognitive skills, effective program strategies must consider incorporating a large number of well-implemented program components.

CHAPTER VII

SUMMARY AND IMPORTANT FINDINGS

The Head Start data reported here are for 1967-68 (FY68) and 1968-69 (FY69), the first fully operational years for the nine-to-twelve month programs.

This is a report on one aspect of project Head Start: some characteristics of children, their families and the programs they attended, and the relationship of these experiences to their development. There are no control groups of children who have not attended Head Start: the comparison is of development in different kinds of Head Start programs.

The present report is to a certain extent a national statement of accountability on some aspects of the immediate impact of the 1967-68 and 1968-69 Head Start programs. It is also part of an on-going effort begun in 1968-69 and now scheduled for completion in 1974, to find out what kind of classroom experiences can best contribute to the attainment of Head Start's objectives for the psychological development of children from low-income families.

The evaluation data reported here are intermediate stages in this evaluation. In 1967-68, naturally occurring Head Start variations were sampled; in 1968-69, about one-third of the classes were "regular" Head Start programs and two-thirds were experimental programs. There were 1889 and 1989 children on which some data were available for the 1967-68 and 1968-69 analyses, respectively.

For 1968-69, child-family and teacher-program information was collected through the use of 24 forms, questionnaires or test instruments. In 1967-68 the collection of child-family and teacher-program information was less extensive, but still represented a good sampling of relevant measures.

For both 1967-68 and 1968-69, child measures such as the Binet were collected both pre and post. Family measures were also collected pre and post while certain teacher and program factors were measured throughout the duration of the program.

Two broad classes of variables were identified: outcome variables (whose pre scores in some analyses also served as predictor variables) and explanatory variables. For all analyses, the individual child was the focus.

The data suggested that even within the OEO Guidelines, differences in life circumstances associated with family economic status are also

associated with the child's performance on measures of cognitive development and achievement upon entering a Head Start program.

Children who performed most poorly on the cognitive and noncognitive measures were from (1) homes where families are ranked in the lowest SES; (2) where mothers are not working or not looking for work; (3) where parents (mothers) have low aspirations and expectations for their children and are pessimistic about life; (4) where parents are relatively inaccessible to their children and spend little time in matters such as reading to their children; and (5) where the nature of disciplinary interaction is characterized by severe punishment and little reward. Children who had the greater amount of previous Head Start experience had the higher Binet pre scores.

As age increased, achievement motivation (Gumpgookies) increased and adjustment problems (MP, BP, and FI) decreased, confirming the expectation that with an increase in age, children's motivation to achieve increases and motivational and personality factors such as impulsivity and lack of confidence as a problem solver give way to more mature, task-oriented behaviors.

The OSCI factors were more confounded with pre scores in 1968-69 than in 1967-68. In general, the "natural" variation data of 1967-68 were less confounded than the "planned" variation data of 1968-69. However, emphasis in this report was on the 1968-69 data rather than the 1967-68 data because (1) more measures existed for 1968-69, (2) more program variation was present in the data base, and (3) programs had a greater chance of being well implemented in 1968-69 than in 1967-68.

Analyses of the outcome measures indicated significant overall gains on cognitive development (Binet), preacademic readiness (Pre School Inventory), ability to learn a new task (Animal House Test of the Wechsler Pre School and Primary Scale of Intelligence), and achievement motivation (Gumpgookies) greater than expected at usual maturational rates. Adjustment to the testing situation increased slightly but not substantially, which may be due to the initially good adjustment shown by many children.

In comparison with data from other studies, the average Binet gains are greater than those reported in other studies for control children and similar to those associated with "motivational enhancement" and "traditional" programs. Some Head Start programs on the average were doing little to enhance development as measured by Binet gains while others were as effective as the most effective experimental programs. There were also indications that gains were highest for those children with no previous experience, but that there is also additional improvement for those children with previous Head Start experience.

These data are consistent with the interpretation that most of the Binet gains ascribed to Head Start (and traditional programs) are probably due to changes in the child's readiness to perform, his task

orientation, desire to please the adult tester, and sense of confidence in his abilities rather than to basic increments in his cognitive ability. In some classes where extremely high gains have been found, the gains are likely to include in addition to motivational effects, true changes in the child's store of information, readiness, and cognitive ability.

Gains were greatest for children from extremely impoverished backgrounds, and for children new to the program. Child sex and age were not related to gains, nor were parental characteristics such as aspirations, expectations, and child-rearing practices. In general, child and family background characteristics were not related to gains.

Except for age, teacher demographic variables were not related to the outcome measures for either 1967-68 or 1968-69. The cognitive measures tended to be more susceptible to program variation, as measured by the OSCI, than the noncognitive measures. The relationships of the cognitive outcomes with the OSCI supported the hypothesis that higher cognitive gains were associated with programs having a high proportion of time spent on language, arithmetic and "structured" activities. The Binet results followed the PSI and AH-WPPSI patterns for raw gains but not for gains adjusted for the pre score. The 3 adjustment scales (Motivation Problems, Behavior Problems, and Feelings of Inadequacy) and the Gumpgookies were less sensitive to program variation than were the cognitive measures. One reason for their being less sensitive is that the inherent nature of these scales (ceiling effects) makes them less susceptible to change. These scales were rather short and highly skewed and indicated that most children had few problems initially. However, there was a tendency for the Behavior Problem scale to follow the AH-WPPSI and PSI patterns of relating to program and teacher factors.

These findings support the notion that programs do make a difference and the high structured, focused, and well-implemented compensatory programs can bring about greater immediate cognitive gains than low structured, diffuse, and less well implemented programs.

The findings in this study were based upon univariate relationships of one explanatory variable with either raw or adjusted gain scores. Unbalanced analyses of variance and regression analyses were run where a number of explanatory variables could be considered jointly so that the main effects of the program, child and family variables on the Stanford Binet, Pre School Inventory, and Motivation Problem Scale could be adjusted for the influence of other confounding factors. This also permitted the estimations of interactions between program factors as measured by the OSCI and child background factors such as sex, age, pre score and socioeconomic level. The child-family and program variables selected for these analyses were either significant in the univariate analyses or were hypothesized to interact with other factors.

For the Binet, SES and age did not have significant main effects, but age did interact significantly with pre score. Older children with higher SB pre scores gained more than would be expected on the basis of age, pre score level, and other main and interaction effects. Children in the North and South tended to have higher gains and, of course, pre Binet scores (categorized into 3 levels) had a significant main effect. The length of the pre-post interval also had a significant main effect in three out of four analyses. However, it is difficult to interpret the pre/post testing interval as a measure of program exposure since in some instances the pre test was given 2 or 3 months after the child was first exposed to the program.

Children at the lowest level of Class Structured Lessons gained the least on the SB. On the other hand, children exposed to the highest level of Class Language and Discrimination Learning had significantly higher SB gains. Middle emphasis in Teacher Creative Instruction - Small Group resulted in significantly higher Binet gains. Program variation as measured by the OSCI contributed significantly more in explaining SB gains than child variables, and there were virtually no interactions between child and program factors.

For the PSI, pre score and region had significant main effects. Children in the South and West had the largest gains on the PSI. SES was insignificant. Age, which was not significant in the IQ analyses, was significant in all four PSI ANOVAs. In agreement with the univariate analyses, older children gained more on the PSI. Children exposed to the high level of Class Structured Lessons did significantly better than children exposed to the low level of this factor. Like the Binet analyses, children exposed to a high level of Teacher Creative Instruction gained the least on the PSI. Again, similar to the Binet, children under the long pre-post interval exhibited the highest PSI gains. There was some tendency for both sex and length of pre-post interval to interact significantly with both Class and Teacher Structured Lessons. However, these interactions are difficult to interpret. The Motivation Problems (MP) Scale was the least sensitive of the three gain measures considered in the unbalanced analyses of variance. In general, the results of the ANOVAs agreed with the univariate analyses in respect to OSCI Class and Teacher effects.

The univariate adjusted gain analyses, ANOVAs, and multiple regression analyses indicated a consistent and strong positive relationship of SB gains with Class Language and Discrimination Learning while no significant relationship was found for the PSI. On the other hand, children exposed to a high level of Class Structured Lessons gained more on the PSI and gained less on the SB. The effects of the program and teacher factors were fairly uniform for different sub population of children as indicated by the relatively small number of interactions between program factors and child-family background factors.

This report disclosed that there was substantial variation in cognitive gains among classes in the E & R samples for the two years. More important, this report demonstrated that a significant amount of this variation could be predicted from class and teacher-based observations in a conceptually meaningful way.

REFERENCES

Ballif, B. L. and Adkins, D. C. Preliminary Manual for Gumpgookies: A Test of Motivation to Achieve, University of Hawaii Head Start Education and Research, 1968.

Bates, B. Project Head Start, 1968. A Descriptive Report of Programs and Participants. Washington, D. C.: U.S. Department of Health, Education, and Welfare, Office of Child Development, 1970.

Beller, E. K. The Evaluation of Effects of Early Educational Intervention on Intellectual and Social Development of Lower-Class Disadvantaged Children, In E. Grotberg (Ed.), Critical Issues in Research Related to Disadvantaged Children. Princeton, N. J.: Educational Testing Service, 1969.

Beller, E. K. The Evaluation of Effects of Early Intervention on Intellectual and Social Development of Lower Class, Disadvantaged Children, unpublished report, 1971.

Bereiter, C. and Englemann, S. Bereiter-Englemann approach, Education Forum, 1968, 32, 359-362.

Bing, E. Effects of Childrearing Practices on Development of Differential Cognitive Abilities, Child Development, 1963, 34, 631-648.

Bissell, J. S. Review of and Summary of the Stanford Research Institute Interim Report: First Year of Evaluation. Washington, D. C.: U.S. Department of Health, Education, and Welfare, Office of Child Development, 1971.

Bissell, J. S. The Cognitive Effects of Pre-school Programs for Disadvantaged Children. Unpublished Ph.D. thesis, Harvard University, 1970.

Brink, E. J. and Sarason, I. Social Change Evaluation Project, Final Report. Seattle, Wash.: University of Washington Press, 1968.

Butler, A. L. Current Research in Early Childhood Education: A Compilation and Analysis for Program Planners. Washington, D. C.: American Association of Elementary-Kindergarten-Nursery Educators, 1970.

Caldwell, B. M. The Preschool Inventory. Princeton, N. J.: Education Testing Service, 1967.

Cegelka, W. and Thomas, P. A. The Culturally Disadvantaged: An Analysis of the Role of the Environment, the Child and the School, The Training School Bulletin, 1968, 65, 46-56.

Cicrelli, V. G., Cooper, W. H. and Granger, R. L. The Impact of Head Start: An Evaluation of the Effects of Head Start and Children's Cognitive and Affective Development. Westinghouse Learning Corp., June 1969. Contract No. OEO-B89-4536.

Coleman, J. S., Campbell, E. O., Hobson, C. T., McPartland, J., Mood, A. M., Weinfeld, F. D. and York, R. L. Equality of Educational Opportunity. Washington, D. C.: U.S. Department of Health, Education, and Welfare, 1966.

Datta, L., Schaefer, E., and Davis, M. Sex and Scholastic Aptitude as Variables in Teacher's Ratings of the Adjustment and Classroom Behavior of Negro and Other Seventh Grade Students. Journal of Educational Psychology, 1968, 59(2), 94-101.

Datta, L. A Report on Evaluation of Project Head Start. Washington, D. C.: U.S. Department of Health, Education, and Welfare, Office of Child Development, 1969.

Di Lorenzo, L. J. and Salter, Ruth. "Pre-Kindergarten Programs for the Disadvantaged: A Third Year Report on an Evaluation Study." December 1969.

Dittman, Laura et al. Study of Selected Children in Head Start Planned Variation, 1969-70. University of Maryland, Institute for Child Study, March 1971.

"Do Teachers Make a Difference?" Washington, D. C.: U.S. Department of Health, Education, and Welfare, Office of Education, 1970.

Evans, D. Contemporary Influences in Early Childhood Education. New York, N. Y.: Holt, Rinehart and Winston, Inc., 1971.

Freeberg, N. E. and Payne, D. T. Dimensions of Parental Practice Concerned with Cognitive Development in the Preschool Child, Journal of Genetic Psychology, 1967, 3, 245-261.

Gervasi, M. Lower Class Mothers' Aspirations for Their Children, published doctoral dissertation, Boston College, 1969.

Goodman, S. M. The Assessment of School Quality. The State Education Department of New York, 1959.

Green, R., Hofmann, L., and Morgan, R. Some Effects of Deprivation on Intelligence, Achievement and Cognitive Growth, Journal of Negro Education, 1967, 36, 5-14.

Grotberg, E. H. Review of Research: 1965-1969, OEO Pamphlet, 1969.

Guthrie, J. W. A Study of School Effectiveness: A Methodological Middleground, American Educational Research Association, March 1970.

Hallahan, D. P. Cognitive Styles - Preschool Implications for the Disadvantaged, Journal of Learning Disabilities, 1970, 3, 5-9.

Hawridge, D. G., Tallmadge, G. K., and Larsen, J. K. Foundation for Success in Educating Disadvantaged Children. Washington, D. C.: U.S. Department of Health, Education, and Welfare, Office of Education, 1968.

Hertzig, M. E., Birch, H. G., Alexander, T., and Mendez, O. A. Class and Ethnic Differences in the Responsiveness of Preschool Children to Cognitive Demands. Monographs of the Society for Research in Child Development, 1968, 33, 1.

Hess, R. D. Private Communication, 1971.

Hess, R. D. and Shipman, V. C. Early Experience and the Socialization of Cognitive Modes in Children, Child Development, 1965, 36, 869-886.

Hess, R. D. and Shipman, V. C. Cognitive Elements in Maternal Behavior, Minnesota Symposium on Child Psychology, Vol. I. Minneapolis, Minn.: University of Minnesota Press, 1968a.

Hess, R. D. and Shipman, V. C. Maternal Attitudes Toward the School and the Role of the Pupil: Some Social Class Comparisons, In A. H. Passow (Ed.), Teaching and Learning in Depressed Areas. New York, N. Y.: Teachers College Press, 1968b.

Hess, R. D., Shipman, V., Brody, T. E. and Bear, R. M. The Cognitive Environments of Urban Preschool Children. Chicago, Ill.: Graduate School of Education, University of Chicago, 1969.

Jencks, Christopher. "Some Natural Experiments in Compensatory Education." Paper presented at Society for Research in Child Development, April 1969.

Jensen, A. R. "How Much Can We Boost I.Q. and Scholastic Achievement?" Harvard Educational Review, 1969, 39, 1-123.

Kagan, J., Pearson, L. and Welch, L. Conceptual Impulsivity and Inductive Reasoning, Child Development, 1966, 37, 583-594.

Karnes, M. B. Research and Development Program on Preschool Disadvantaged Children. Final Report. Vol. I of III Volumes. Project No. 5-1181. Contract No. OEO-10-235. (1969).

Karnes, M. B., Teska, J. A., and Hodgins, A. S. An Evaluation of Two Preschool Programs for Disadvantaged Children: A Traditional and A Highly Structured Experimental Preschool, Exceptional Children, 1968, 34, 667-676.

Karnes, M. B., Teska, J. A., and Hodgins, A. S. The Effects of Four Programs of Classroom Intervention on the Intellectual, and Language Development of 4-year-old Disadvantaged Children, American Journal of Orthopsychiatry, 1970, 40, 58-76.

Kennedy, W. A. A Followup of A Normative Study of Negro Intelligence and Achievement. Monographs of the Society for Research in Child Development, 1963, 34, 2.

Kennedy, W. A., Van de Reit, V., and White, J. A. A Normative Sample of Intelligence and Achievement of Negro Elementary School Children in the Southeastern United States. Monographs of the Society for Research in Child Development, 1963, 28, 6.

Kirchner Associates, Inc. A National Survey of the Impacts of Head Start Centers on Community Institutions. Contract No. B89-4638. A report prepared for Project Head Start, 1970.

Klaus, R. A. and Gray, S. W. "The Early Training Project for Disadvantaged Children: A Report after Five Years." Monographs of the Society for Research in Child Development, 1968, 33, 4.

Kraft, I., Fuschillo, J., and Herzog, E. Prelude to School: An Evaluation of an Inner-City Preschool Program. Washington, D. C.: U.S. Department of Health, Education, and Welfare, Children's Bureau, 1968.

Kutzman, T. M. Distribution and Production in a Big City Elementary School System, Yale Economic Essays, 1968

Maccoby, E., Dowley, E. M., Hagen, J. W., and Degerman, R. Activity Level and Intellectual Functioning in Normal Preschool Children, Child Development, 1965, 36, 761-770.

McCarthy, J. L. Changing Parent Attitude and Improving Language and Intellectual Abilities of Culturally Disadvantaged Four-year-old Children Through Parent Involvement, Dissertation Abstracts, 1969, 30, 189-190.

McDavid, J. W. Basic and Applied Research in Project Head Start, paper presented at Southeastern Psychological Association, 1967.

McDavid, J. W. Innovations in Education of the Disadvantaged Child: What Have We Learned? Paper presented at the Conference on Compensatory Education, Portsmouth, N. H., 1969.

Miller, J. O. Review of Selective Intervention Research with Young Children. ERIC Occasional Paper, 1968.

Miller, L. B. and Dyer, J. L. Experimental Variation of Head Start Curricula: A Comparison of Current Approaches. Annual Progress Reports, University of Louisville, OEO No. CE 8199, 1970.

Moustakas, C. E., Sigel, I. E., and Shalock, H. D. An Objective Method for the Measurement and Analysis of Child-Adult Interaction, Child Development, 1956, 27, 109-134.

Mumbauer, C. and Miller, J. D. Socioeconomic Background and Cognitive Functioning in Preschool Children, Child Development, 1970, 41, 471-480.

Passow, H. A. Early Childhood and Compensatory Education, In Harry A. Passow (ed.): Teaching the Disadvantaged Learner. New York, N. Y.: Columbia University Teachers College, 1970.

Possner, F. Evaluation of Successful Projects in Compensatory Education, Washington, D. C.: U.S. Department of Health, Education, and Welfare, Office of Education, Office of Planning and Evaluation, Occasional Paper No. 8, 1968.

Radin, N. and Karni, C. The Child-rearing Attitudes of Disadvantaged Negro Mothers and Some Educational Implications, Journal of Negro Education, 1965, 24, 138-146.

Rotter, J. B. Generalized Expectancies for Internal Versus External Control of Reinforcement, Psychological Monographs: General and Applied, 1966.

Ruben, M. and Womack, M. The Intelligence of Preschool Children as Related to Ethnic and Demographic Variables, Expectional Children, 1968, 34, 609-614.

Shipman, V. Disadvantaged Children and Their First School Experiences, Princeton, N. J.: Educational Testing Service, 1970.

Slaughter, D. T. Maternal Antecedents of Academic Achievement Behavior in Negro Head Start Children, unpublished doctoral dissertation, University of Chicago, 1968.

Smith, M. P. Intellectual Differences in 5-year-old Underprivileged Girls and Boys with and without Pre-kindergarten School Experience, The Journal of Educational Research, 1968, 61, 348-350.

Smith, M. S. and Bissell, J. S. Report Analysis: The Impact of Head Start. Harvard Educational Review, 1970, 40(1), 51-104.

Sontag, M., Sella, A. P., and Thorndike, R. The Effects of Head Start Training on the Cognitive Growth of Disadvantaged Children, The Journal of Educational Research, 1969, 62, 387-389.

Sprigle, H. A. How to Make Rich Kids Poor and Poor Kids Rich, paper delivered at the annual meeting of the Tulsa (Oklahoma) County Mental Health Association, May 1971.

Stanford Research Institute. Implementation of Planned Variation in Head Start. Menlo Park, California, 1971.

Stearns, Manan S. Report on Preschool Programs: The Effects of Preschool Programs on Disadvantaged Children and Their Families. Washington, D. C.: U.S. Department of Health, Education, and Welfare, Office of Child Development, 1971. DHEW (OCD) 72-27.

Stern, C. The Observation of Substantive Curricular Interactions, UCLA Head Start Evaluation and Research Center, 1968.

Temp, G. and Anderson, S. B. Project Head Start, Summer 1966. Princeton, N. J.: Educational Testing Service, 1967.

Van de Reit, V. and Resnick, M. B. Sequential Approach to Early Childhood and Elementary Education, Phase II. Gainesville, Fla.: University of Florida, OCD No. H8222 B/H/O, HEW, December 1970.

Van de Reit, V. and Resnick, M. B. Sequential Approach to Early Childhood and Elementary Education, Phase III. Gainesville, Fla: University of Florida, OCD No. H8222 B/H/O, HEW, January 1972.

Van de Reit, V., Van de Reit, H., and Sprigle, H. The Effectiveness of a New Sequential Learning Program with Culturally Disadvantaged Preschool Children, Journal of School Psychology, 1969, 7, 5-15.

Wargo, M. J., Campeau, P. L., and Tallmadge, G. K. Further Examination of Exemplary Programs for Educating Disadvantaged Children. Palo Alto, Calif.: American Institutes for Research, July 1971.

Weikart, D. P. Preschool Programs: Preliminary Findings. Journal of Special Education, 1967, 1, 163-181.

Weikart, D. P. The Ypsilanti Preschool Curriculum Development 1968-1971. Ypsilanti, Mich.: High-Scope Educational Research Foundation, 1971.

Weikart, D. P., Deloria, D. J., Lawser, S. A., and Weigerink, R. Longitudinal Results of the Ypsilanti Perry Preschool Project. Ypsilanti, Mich.: High-Scope Education Research Foundation, 1970.

White, R. W. Motivation Reconsidered: The Concept of Competence, Psychological Review, 1959, 66, 297-333.

White, S. H. Evidence for a Hierarchical Arrangement of Learning Processes, In L. P. Lipsitt & C. C. Spiker (Eds.): Advances in Child Behavior and Development, Vol. 2. New York, N. Y.: Academic Press, 1965.

Wilmon, B. Parent Participation as a Factor in the Effectiveness of Head Start Programs, The Journal of Educational Research, 1969, 62, 406-410.

Yarrow, M. R., Campbell, J. D., and Burton, R. V. Child Rearing: An Inquiry into Research and Methods. San Francisco, Calif.: Jossey-Bass, 1968.

Zigler, E. and Butterfield, E. C. Motivational Aspects of Changes in I.Q. Test Performance of Culturally Deprived Nursery School Children, Child Development, 1968, 39, 1-14.