DOCUMENT RESUME

ED 043 397 PS 004 025

TITLE	From Theory to Operations. Disadvantaged Children and Their First School Experiences, ETS-Head Start Longitudinal Study.
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	Office of Economic Opportunity, Washington, D.C.
REPORT NO	OEO-4206
PUB DATE	Aug 69
NOTE	250p.
EDRS PRICE	EDRS Price MF-$1.25 HC-$13.05
DESCRIPTORS	Behavior Rating Scales, Community Problems, Community Study, Culturally Disadvantaged, Data Analysis, Data Collection, Educational Development, *Longitudinal Studies, Parent Child Relationship, Physical Development, *Preschool Children, *Program Evaluation, *Research Methodology, *Research Tools, Social Development
IDENTIFIERS	*Project Head Start, Westinghouse Report

ABSTRACT

In the summer of 1966, Educational Testing Service (ETS) began identifying the 1,650 3 1/2-year-old children in four sites across the country who would become part of a six-year longitudinal study designed to assess the impact of Head Start. The sites were Lee County, Alabama, St. Louis, Missouri, Trenton, New Jersey, and Portland, Oregon. The children were measured on a large number of relevant variables before any of them had experience with some or no Head Start program. This report (the second of three) describes attempts to design an evaluative program based upon conceptions of the complexity of the human organism and an interaction model of human development. Six chapters present: (1) a short history of the ETS study, (2) impressions of the study communities, (3) measures used in initial assessments, (4) data collection procedures, (5) data storage and retrieval system, (6) plans for data analysis. The seventh and final chapter anticipates the activities of 1969-1970. Four appendixes are included: (a) tryouts of measure, (b) working papers, (c) the ETS-Head Start Longitudinal Study and the Westinghouse Study, (d) project personnel. Another part of this study is available as PS 003 657. (WY)

PR-69-12

DISADVANTAGED CHILDREN AND THEIR FIRST SCHOOL EXPERIENCES

ETS-Head Start Longitudinal Study

From Theory to Operations

Report under

OEO Contract Number 4206

and

OEO Grant Number C6-8256

August 1969

CONTENTS

*The pages of this appendix section are numbered as if they ran consecutively. Gaps occur, however, where exhibits appear. The exhibits are either unnumbered or bear the page numbers of the original documents from which they are reproduced.

1. A SHORT HISTORY OF THE LONGITUDINAL STUDY OF DISADVANTAGED CHILDREN

Spring 1967 - Summer 1969

Scarvia Anderson

A SHORT HISTORY OF THE LONGITUDINAL STUDY OF DISADVANTAGED CHILDREN

SPRING 1967 - SUMMER 1969

The Longitudinal Study of Disadvantaged Children and Their First School Experiences was born in the spring of 1967, in conversations between representatives of the Head Start Research Office (Office of Economic Opportunity) and the Curriculum Studies group of Educational Testing Service. Edmund Gordon was director of the former, Scarvia Anderson of the latter. All those involved in the discussions were suffering from some frustration.

In 1961, ETS had begun a longitudinal study of academic prediction and growth spanning grades 5 through 12 (Anderson & Maier, 1963). The grade 5 starting place was dictated primarily by judgments about the adequacy of existing educational measurement techniques for children over and under ten. While the study was progressing satisfactorily and some interesting answers to questions about student progress were in the offing, there was an uncomfortable feeling that some of those answers might be wrong, or at best incomplete, without compatible information on the influences of the earliest school years. Of course, a substantial national commitment to preschool programs did not exist in 1961.

In 1966, the second summer of Head Start, Joseph Boyd, George Temp, and others at ETS joined the troops of behavioral scientists around the country asked to take a look at this innovation that had defied all predictions about how long it takes to get an educational idea put into operation on a large scale (see, for example, Cocking, 1951). The findings in this Head Start study (Anderson & Temp, 1967) were similar to those of other studies: children showed some gains in intelligence test scores over the 6- to 8- week period; mothers became "involved" in the centers and wanted their next children to have an opportunity to go too; there were wide variations in content, structure, and atmosphere of programs labeled "Head Start" (although with the measures of program characteristics and children's performance used, it was not possible to tie particular program variations to particular aspects of children's progress). But what long-term effects, if any, could one expect from exposure to Head

Start for 3 hours a day over a period of a few weeks? How could Head Start influences counteract or reinforce the possible influences of the home, the community, and--later--the school which had so much more than "equal time"? What would be appropriate criteria of Head Start effectiveness? Certainly the central purpose of the program was not to raise IQ! These were some of the questions plaguing ETS.

They were also the concern of the Head Start Research Office. Head Start was predicated upon the intuitively compelling assumptions that the earliest years are generally the most important in shaping people's lives and that an important segment of the population is likely to have "inadequate" early shaping experiences. However, most attempts to evaluate specific early education programs had foundered, based as they were upon "engineering" models--and frequently inadequate ones at that. The time was right to try to design an evaluative research program at the early years, based upon conceptions of the complexity of the human organism and an interactional model of human development. Not incidentally, by the spring of 1967 sentiment was moving strongly toward establishment of full-year Head Start programs preceding, or in some places acting as, kindergarten.

The first attempts at the more sophisticated research design suffered from Einstellung effects. The investigators felt bound to the 1966 summer Head Start sample and following up those children in some way; they boggled at the logistics and costs that a truly comprehensive study would involve; they could not repress the guilt feelings induced by entertainment of the possibility of a nonrandom, nonfactorial design. However, during the second half of 1967, with the continuing counsel and advice of Urie Bronfenbrenner, Edmund Gordon, and Edward Zigler (the Head Start Research Council), and John McDavid (who had become Director of Head Start Research), the research design slowly began to take the general shape it has today.

ETS senior research psychologists, who during this period had thrown themselves into the effort, included Albert Beaton, Walter Emmerich, and Samuel Messick. Also participating in the overall planning were Marshall Smith, Trenton State College; Silvan Tomkins, Rutgers University; and Melvin Tumin, Princeton University.

On February 26, 1968, a detailed report of plans was submitted to OEO for

> a comprehensive study of the cognitive, personal, and social development of disadvantaged children over the crucial age period 4-9. The principal aims of the study are to
>
> * identify the components of early education (Head Start and other preschool and primary programs) that are associated with children's development
> * determine the environmental and background factors which influence such associations
> * describe how these influences operate
>
> in order to provide information that will contribute to
>
> * educational planning and improvement of early education programs
> * general social planning for the lower socioeconomic groups
> * psychological theories of child development
> * measurement practices in the assessment of young children and their environments.

Design plans to date call for selection of six communities of different types with three to five elementary school districts in each, identification of all children in those districts who will be eligible for first grade in two years (N = 2000), assessment of those children (and their families) prior to any educational intervention, and close monitoring of the children, their families, and the educational programs they are exposed to in the communities during the next five years. The educational tracks which will be available to these children can be summarized as follows:

Year 1 of Study	Year 2 of Study	Year 3-5 of Study
No preschool	No preschool	Grades 1-3 regular
or	or	or
Head Start	Head Start	Grades 1-3
	or	Follow Through*
	Kindergarten	

* "Follow Through" is a United States Office of Education program fostering new and improved curricula for K-3 classes attended by large numbers of Head State "Graduates."

> Other aspects of the design provide for appropriate cross-sectional comparison groups; a comparison group assessed only at the beginning and end of the study; periodic, though less intensive, assessments of children who move away from the study communities; and, with respect to some variables, augmentation of the original sample with children moving into the educational programs under investigation.
>
> (Educational Testing Service, 1968a)

The summer and fall of 1968 were characterized by two major activities: development, adaptation, and pretesting of measures; and efforts to engage the interest and cooperation of people in potential study sites.

The first kind of activity was given strength by the arrival of Virginia Shipman from the University of Chicago, to serve as coordinator of measurement and training efforts for the study. She had achieved national recognition for her work at the Head Start Research Center in Chicago. Meanwhile, other staff members were well along in identifying practical and conceptual pitfalls in planning assessments of disadvantaged adults and young children. George Gordon tried out dozens of perceptual measures, with many variations, to identify a few that could be used validly with 3 1/2-year-old youngsters--and that would serve as the basis for continuing assessment over time. Donald Medley, with years of experience in techniques of observing teacher behavior, faced the problem of chronicling in a valid and reliable way the teacher-child interactions of the preschool class. His laboratory was just abo. every New Jersey Head Start classroom within driving distance. Stanley Zdep worked not only on the content of family questionnaires but also on techniques for getting them administered by indigenous interviewers to not-always-receptive ghetto mothers. Scarvia Anderson, Dolores Ahrens, Masako Tanaka, and Carolyn Massad developed a "taxonomy" of "academic" objectives of early education programs and proceeded to try to fill the cells with appropriate measures. Thomas Barrows, Gerald Bracey, Ann Jungeblut, Diran Dermen, and Marianne Amarel were at the same time adapting and field testing a number of measures of learning, cognitive style, and social perception. Michael Rosenfeld gave his attention to the problems of measuring the administrative side of preschool and school programs, including physical facilities, school

personnel, pupil services, and school-community relationships. George Temp outlined the types of measures that might be taken directly from teachers, in order to help assess the impact of their skills, experience, attitudes, and personalities on children. With the arrival of Michael Lewis, Edward Chittenden, Samuel Ball, and William Ward on the scene, previously sketchy ideas for measurement of attentional processes, behavior related to the theories of Piaget, attitudes toward school experiences, and creativity began to take definite shape. Melvin Tumin, concerned that sociological points of view be adequately represented, recommended Robert Althauser, Princeton University, and Ann Ryan to the study. They addressed themselves primarily to those factors in extra-familial or community life that can be expected to shape the growth and education of small children, directly or through the effects of parents. Not the least of the efforts to wed theory to the practicalities of assessing young, disadvantaged children were those of Walter Emmerich and his colleagues working in the personality area. Paper-and-pencil personality tests, seldom meeting professional standards even for adults, were rejected in favor of comprehensive observational techniques applied to children during free-play situations.

Parallel to the activity in the measurement realm was the stumping of the country to locate suitable study sites. While the February 1968 report had proposed the selection of six communities, four was the number finally decided upon. It was not an easy decision and it followed many heated discussions by the Steering Committee and advisers from OEO--with strong arguments advanced for more than six, for only one, and for most of the numbers in between. Logistics, representation of a variety of preschool and school programs, possibilities of community dropout, and public relations all figured in the arguments. Fortunately, there to present the facts were two new study staff members who, with Daniel Landis and representatives of ETS regional offices, had by the early fall of 1968 visited some dozen possible locations and checked hundreds of pages of school, OEO, and community statistics. Joseph Boyd was not really new to Head Start projects since he had directed the ETS 1966 summer study; however, in between he had coordinated other large testing programs at ETS. Samuel Barnett was not new to dealing with and

understanding the problems of the disadvantaged in cities; he had been a policeman and later a director of several community programs in Philadelphia. When they visited the communities, the Boyd-Barnett team talked to school people, OEO people, and leaders of various community agencies. They also talked to the man on the street (and in the barbershop and café) and to the woman whose children might be eligible to go to Head Start.

The final choice of communities was announced by ETS and OEO in December: Lee County, Alabama (including the city of Auburn); St. Louis, Missouri; Portland, Oregon; and Trenton, New Jersey. All had year-long Head Start available, and relatively comprehensive Head Start programs; three were scheduled to have Follow Through programs in K-3 in some schools; three had OEO Parent-Child Centers; they varied in size and typical life styles of their residents; and leaders in all four communities had indicated that they were prepared to give the study extraordinary cooperation and support.

The month before (November 1968) the study had received its first public announcement to a professional audience at a special symposium sponsored by the National Council on Measurement in Education (Anderson & Doppelt, 1969). The first five papers--and even their titles--reflected some of the major theoretical and practical considerations the project planners had been contending with in the preceding months:

Early Schooling: What Is It All About? - Marshall P. Smith

The Family and Community: What Are Their Roles in the Educational Process? - Melvin Tumin

The Child: His Cognitive, Personal-Social and Physical Development--A False Trichotomy? - Edmund W. Gordon

How Are Measurement Strategies Related to Models of Human Development? - Walter Emmerich

Can You Do Real Research in the Real World? - Samuel Messick

Scarvia Anderson then summarized the study objectives and design, and John McDavid (who had recently left his post at OEO to return to the University of Miami) commented on the scientific and social significance of the study. He concluded, "My perspective is such that I believe firmly that good basic research and good program evaluation can be integrated, and I am extremely pleased to have been a part of planning

and developing this longitudinal project. But it is important that all of us recognize that Rome was not built in a day, and that no one single study, no matter how massive, can ideally provide all critical needs of both the scientific community and the federal bureaucracy."

By the end of 1968, the study staff had built, if not Rome, at least a monumental two-volume report on the "Theoretical Considerations and Measurement Strategies for the Longitudinal Study of Disadvantaged Children and Their First School Experiences" (Educational Testing Service, 1968b). Edward Zigler and Lois-ellin Datta, representing the Head Start Research Office, gave the draft an intensive review, cautioning ETS about the "looseness" of some of the brain-damage talk, suggesting that information be collected about testers as well as the children they tested, arguing for the dropping of some measures and further research on others, requesting the addition of "global" classroom ratings, encouraging a specific statement about the study staff's stand on the use of intelligence measures, and always showing concern for the logistics of field operations.

The report was revised and then subjected to further review by Boyd McCandless, Albert Yankauer, Edward Suchman, Richard Orton, Richard Armstrong, and other OEO consultants and staff members. The principal theoretical statements were prepared by Samuel Messick (Cognitive and Perceptual Development), Walter Emmerich (Personal and Social Development), Edmund Gordon (Physical Health and Nutritional Status), Virginia Shipman and Anne Bussis (The Family), Donald Medley (The Classroom--actually an antitheoretical statement!), Samuel Ball and George Temp (The Teacher), Samuel Ball and Marshall Smith (The School), and Robert Althauser and Ann Ryan (The Community). The outlines of the multivariate analyses to be applied to the data collected in the "multiple time series quasi-experimental design" were described by Albert Beaton.

Action at last! In the early months of 1969, field offices were set up in Lee County, Trenton, St. Louis, and Portland. Joseph Boyd and Samuel Barnett, aided by ETS regional office personnel, interviewed dozens of candidates for the coordinator jobs. The "ideal" coordinator had to have administrative skills; be familiar with the power structure

in the community and on communicative terms with diverse elements in it; show signs of "caring" about disadvantaged children, their families, and opportunities for early education; be able to speak well and influentially; be in good health--and drive a car! Lida Campbell, once a student of Marshall Smith's at Trenton State College and for some years a resident of Auburn, was appointed in Lee County. Then: Verna Shepherd in Portland (she had left the world of the home for the world of work when "community action" was in its infancy); Ronald Greeley in St. Louis (he had devoted the last few years to counseling teenagers and adults about jobs and about life); and Conrad McLean in Trenton (he came directly from the local anti-poverty agency). Two males, two females; three blacks, one white.

They each set about getting an office, an assistant, a $29.95 copying machine, paper clips, and the other paraphernalia that most operations take for granted. They also started recruiting community members who might be interested in temporary employment as testers, drivers (to transport children and mothers), and baby-sitters (to look after children when they weren't with a tester--or even younger brothers or sisters who were brought along). With the help of ETS regional staff members, Robert Lambert and George Temp (Portland), J. A. Davis and Roderick Ironside (Lee County), Daniel Norton and Chandra Mehrotra (St. Louis), and Princeton staff (Trenton) they began to do a lot of talking--to school people, Head Start people, other community leaders, mothers, newspaper reporters, OEO regional offices, and the many others who asked, "What's this thing all about?" or "What's in it for us?"

Teams of zealous psychologists and research assistants went out from the Princeton office to train community "housewives" in the art/science of administering some 30 different testing instruments to 3 1/2-year-old disadvantaged children (and, for some measures, to their mothers). The testing materials were not yet in adequate supply; the manuals were being completed in airports and bus stations; the optimistic estimate of the amount of time needed for training was two weeks. In practice, it took something like four to six weeks to turn out "professional" administrators, and then the tester trainers' hearts were broken if Mrs. Smith or Mrs. Jones still couldn't quite make it.

The testing centers were churches, preferably ones with a number of small Sunday school rooms. They were conveniently located in the neighborhoods where the children lived, their ministers were happy to have them rented during the week (although there was great activity on Friday afternoons stowing away all the testing equipment), and one or two of the ministers took an active role in the project themselves.

Of course, before there could be children at the centers, the children had to be identified. Audits and Surveys, Inc., New York City, had taken on the job of, first, taking a census of all children who would be eligible to enter the first grade of selected elementary schools in each site in the fall of 1971 and, then, interviewing the mothers of those children on such topics as their aspirations for the children and what they thought was the best way to bring them up. A&S, like ETS, was committed to using community people as their local coordinators and interviewers, purchasing many supplies from merchants in the town or city, and, in general, counteracting to the extent possible the image of the clipboarded researcher who cavalierly invades people's lives and then leaves them with nothing.

There were, of course, some local problems that were almost impossible to contend with: a rent strike, restrictions on the number of occupants permitted in public housing units, and blocks (and blacks) that had been oversurveyed in the very recent past. Under these conditions, people knocking on doors asking questions were bound to be met with suspicion, hostility, or, at best, boredom.

However, some operational problems stemmed from the fact that research organizations such as ETS had never really done this kind of community-centered research before--certainly not on such a large scale. Consider, for example, some of the financial crises. Educational Testing Service pays its regional and Princeton personnel twice a month, after the work has been performed. By the time pay checks got to the first tester-trainees, they had been in the training program three to four weeks. Many were in debt to baby-sitters and to their neighborhood grocers who couldn't quite appreciate why the Princeton computer turned so slowly. The ETS office did not anticipate what the "start-up" expenses would be like in testing centers (cookies and juice, extension cords, playroom

blocks, and extra underpants), and had not provided adequate petty cash funds (Woolworths's doesn't usually "charge" an extension cord!). Princeton personnel visiting the sites were using up their "living" allowances to purchase equipment.

The spring "tooling-up" operations for the longitudinal study took place in the context of more published controversy about the education of disadvantaged children than any of the staff could remember. First, there was the Jensen article in the Harvard Educational Review (1969); then, the Westinghouse-Ohio University report on Head Start (1969). The telephoned queries to ETS, like the newspaper accounts, frequently confused the two: "Do they really show that Head Start is no good?" "How is your study different?" "Are you getting the same results?"

We had to reply that so far we had no results--we were just trying to get started, that the longitudinal study was different in intent and procedures from the Westinghouse study (see Samuel Ball's notes in Appendix C), that we were dubious about using an IQ measure as the principal criterion of the effectiveness of compensatory education (in reference to Jensen's article), and that we hoped to be a lot more knowledgeable about the whole business in a few years.

So we continued looking for subjects, trying to improve training techniques and testing facilities, requisitioning Tootsie Rolls and painting toys to the color specifications of the tests, explaining ETS procedures to many new employees who would probably never see Princeton, and constantly adapting our expectations to the exigencies of research in the real world.

. . .

The initial enumeration, interview, testing, and physical examination phases of the longitudinal study are still going on and are expected to continue through August--and, in some cases, even into the fall. This review has only touched upon the present status of things. The report that follows details the decisions and operations that occurred between the December 1968 report and July 31, 1969. The last chapter anticipates the activities of 1969-70: the continuation of efforts to

ensure adequate first year data; the second-year data collection associated with the longitudinal sample of children, some in Head Start and some not; and a cross-sectional look at the kindergarten and primary grade children and programs in the selected school districts, to find out what the current picture is like.

Scarvia Anderson

Princeton, New Jersey
July 31, 1969

References

Anderson, S. B., & Doppelt, J. (Jt. Chm.) Untangling the tangled web of education. Symposium sponsored by the National Council on Measurement in Education and presented in conjunction with the conferences of the Educational Records Bureau and Educational Testing Service, New York, November 1968. Princeton, N. J.: Educational Testing Service, RM-69-6, 1969.

Anderson, S. B., & Maier, M. H. 34,000 pupils and how they grew. Journal of Teacher Education, 1963, 14, 212-216.

Anderson, S. B., & Kemp, G. Project Head Start, Summer 1966. Final Report, Section 3, May 1967, Educational Testing Service, Contract OEO 1359, Office of Economic Opportunity.

Cocking, W. The regional introduction of educational practices in urban school systems of the United States. New York: Bureau of Publications, Teachers College, Columbia University, 1951.

Educational Testing Service. Disadvantaged children and their first school experiences. Interim report, February 1968, Contract OEO 4206, Office of Economic Opportunity. (a)

Educational Testing Service, PR-68-4. Disadvantaged children and their first school experiences: Theoretical considerations and measurement strategies. Report in 2 volumes, December 1968, Contract OEO 4206 and Grant OEO C6-8256, Office of Economic Opportunity. (b)

Jensen, A. R. How much can we boost IQ and scholastic achievement? Harvard Educational Review, 1969, 39, 1-123.

Westinghouse Learning Corporation & Ohio University. The impact of Head Start: An evaluation of the effects of Head Start experience on children's cognitive and affective development. April 1969, Contract OEO B89-4536, Office of Economic Opportunity.

2. IMPRESSIONS OF THE STUDY COMMUNITIES

Lee County, Alabama

Roderick Ironside

. . . . .

St. Louis, Missouri

Daniel Norton

. . . . .

Trenton, New Jersey

Marshall Smith, Gray Sidwell

. . . . .

Portland, Oregon

Robert Lambert

Lee County, Alabama

Roderick Ironside

As you leave Interstate 85 and drive north through the countryside toward Auburn, you see not only the hills and cuts of red clay but also the houses and shopping centers of the more prosperous residents. However, when you leave Auburn and head west, south, or east, you quickly become aware that Lee County is a rural county. The dirt back roads (and most roads are back roads) are dotted with shacks and lined with farms and groves. The villages and crossroads tend to be clusters of dilapidated stores and filling stations, many of them no longer occupied. If you are tempted to generalize, you would probably say that it is a poor county, although many of the occupants of the rundown houses have telephones and automobiles. You would also infer that the people live off the land on small holdings, although a number of them depend upon the two large towns of Auburn and Opelika for their livelihoods. And you would notice that "the church" is everywhere--along the federal highways and down the dirt tracks--sometimes brick, sometimes weatherbeaten, sometimes newly-painted, but always in your view and almost always evangelistic.

Lee County is not a community. There are, of course, two small cities, Auburn and Opelika, and they provide a sense of community for their residents. But beyond town limits, people live far apart--in small enclaves or isolated dwellings. While the churches in the county provide some community feeling (for their own congregations), the schools do not appear to do so and there are no public recreational facilities. Even the schools do not have playgrounds. There are several general stores that serve as gathering places, and this is the sort of county where villages with a population of fifty still have post offices.

In the two towns, on the other hand, you can see that the usual bases for community awareness and participation exist. Auburn, for example, has two picnic parks and a community center (located in a condemned building) as well as a racially integrated pool. Both Auburn and Opelika have local bus service and theaters, and in addition, Opelika serves as the county seat. Everyday needs are provided by the towns' stores, but for "big" shopping, people either drive or take the train to

Montgomery or Atlanta, many miles away. Auburn has its university, too, which is an important employer as well as the center of cultural activities--music, drama, lectures, films, exhibits--almost exclusively for the white community. But even with these kinds of contact with the outside world, you feel that Auburn and Opelika are somewhat isolated and quite unsophisticated. The majority of the college students are Alabamans, not surprisingly, and the local daily and weekly newspapers are principally local in coverage. The one vocational training school is considered inadequate in curriculum and plant, and the attained educational levels are disappointingly low: median grade completed for whites is roughly 9, and for blacks about 6 (1960 census). The school systems of the towns and county are under court order to integrate beyond token freedom-of-choice levels; and, although the faculties are biracial, the children are still thoroughly segregated. You might find, too, that travelers are not expected often, except along the federal highways and the Interstate. Auburn has two boxed-in motels, no hotel, and a surprising number of second-rate restaurants. Finally, there is little new housing construction. In the county, that which is bad continues to deteriorate; in Auburn new homes are being built by university faculty and there are apartments for students--but downtown Auburn is replete with seedy rooming houses; in Opelika the only "urban renewal" effort in the county is proceeding--along two streets.

You have looked around Lee County and have found it to be rather a bleak place. On the other hand, consider that the local school systems have, in fact, accepted and aided the ETS-Head Start Longitudinal Study, as have a number of churches, organizations, and individuals. Consider, too, that there are four or five factories in the county which provide a degree of economic stability, and that Auburn has been the headquarters of the Alabama Council on Human Relations. The advent of Interstate 85 has brought more commerce and travelers to the county, and, of course, the university continues to grow and to extend useful educational and cultural activities in the area.

You might say that Lee County is out of the mainstream, and is unlikely ever to be threatened by ungainly skyscrapers or great wealth. It is quiet, rural, and conservative in many ways. Yet, its gently rolling

hills and lush green growth make it attractive in spite of the summer heat, and the winds of change and equal opportunity blow slowly over those rolling hills.

St. Louis, Missouri

Daniel Norton

St. Louis is a city half bounded by the Mississippi River and half by suburban developments. The population of the city in 1960 was 750,000. Its estimated population in 1965 was 666,000; projected populations are 592,000 for 1970 and 526,000 in 1975.

As the population of St. Louis has declined, so has its racial make-up altered. The white population has decreased and the nonwhite population has grown. In 1960 the nonwhite population was 216,022 or 28.8% of the total city population. The estimated percentage of nonwhites in 1965 was 42.7; projected nonwhite populations for 1970 and 1975 are 49.1 and 56.3, respectively.

As in other urban centers, the nonwhite populations are concentrated in the central city where poverty is highest. In 1960, 47.4% of the families in the central city resided in poverty areas and 19.1% of the families of the central city areas were below the poverty level. Outside the central city only 9.6% of families were below the poverty level.

When this data is arranged by racial groups, 12.4% of white families resided in poverty areas while the corresponding percentage for nonwhites was 86.9. Only 8.9% of white families were below the poverty level, but 42.1% of the nonwhite families were recorded as below.

The dominant structure of St. Louis in 1969 is the Saarinen Arch, which sweeps across Jefferson National Expansion Memorial, near the river, at the eastern edge of the city. One boundary of the study area actually passes beneath the arch; and this arch, its slender lines reaching for the skies, more graceful than any skyscraper, is an ever-visible reminder of man's ingenuity and skill. But there are less satisfying examples in this city.

The study area contains some of the oldest sections of the city. The Henry school attendance district, southern section of the study area, is bounded on the south by downtown St. Louis. It was, in 1874, occupied by neighborhood clusters of immigrants having German Catholic, German Protestant,

Irish, Italian, and Orthodox Jewish origins. They were later replaced by other groups. Since World War II, black migrants have occupied crumbling and decayed brick row structures built by earlier residents.

The Webster school district, center of the study area, sweeps upward from the industrial flatlands along the river. Serious and partially successful efforts have been made to rehabilitate landmark homes in the area but indications of blight are frequently visible.

A mile back from the river, on a line through the Webster school area, stand the Pruitt-Igoe Apartments which constitute the Jefferson school district. (Wendell Oliver Pruitt was a World War II black hero.) The Pruitt section was ready for occupancy in 1954 and consists of 1736 apartment units in 20 buildings, each 11 stories high. Part of an urban renewal project, it provided one- to five-bedroom apartments of low cost housing. The Igoe Apartments were built for people displaced by redevelopment and highway construction. At the time they were opened in 1956 they provided 1,132 new high rise units.

In February 1965, an anthropological team which conducted an investigation in the Pruitt-Igoe area found that about one-fourth of the units were empty. Suspiciousness was, they believed, a notable characteristic of the residents. Conditions were described in part by a resident who said:

> The Reeds are bad, their kids throw things from the ninth floor... The Olins are bad, they break windows... The Pullians are bad, they urinate in the elevator.
>
> (St. Louis Globe-Democrat, February 6, 1965)

Later the same year, the St. Louis Post-Dispatch ran an article titled "Misadventure in Public Housing." The existence of flaws at the Pruitt-Igoe apartments was reported to be the reason for a projected expenditure of $7,000,000 for correction of conditions. The article read, in part:

> Ten years ago, Pruitt and Igoe were the proud names of the two newest public housing projects in St. Louis. Today, the names stand for the worst in public housing. The image of Pruitt-Igoe is one of crime, vandalism and anti-social behavior. The vacancy rate is high--about 26 per cent--because eligible renters are afraid to

> love in or wish to avoid the stigma of "project" living. The buildings themselves are formidable, rising above the surrounding slums like huge fortresses. The lawns are trampled and littered with glass. Inside, the corridor walls are unpainted and scrawl-covered, giving the impression of age and decay. The undersized elevators (which stop only at the fourth, seventh and tenth floors) are brutally defaced, battered and reek of abuse. Unused laundry rooms and stairways, with light fixtures shattered, provide dark havens for illicit activities. The corridors have become crowded playgrounds for unsupervised youngsters and gauntlets to be run by residents entering or leaving their apartments. Breezeways under the buildings form the only convenient loitering places for teenagers and the noises from their impromptu gatherings intrude into nearby apartments. Many ground-floor apartments are empty and many have broken windows covered with plywood sheets.
>
> This is home for 10,736 individuals, including 7523 youths (under 21 years old), 2223 adult females and only 990 adult males. Of the 2100 families, 1100 are receiving welfare payments. The median annual income is $2300.
>
> (St. Louis Post-Dispatch, October 31, 1965)

Pruitt-Igoe is still considered to be an unsuitable place to live. Taxi drivers are reluctant to call there for passengers. Nearby structures such as the Transfiguration Lutheran Church, are commonly protected by chain link fencing and are locked at all possible hourse of the day.

North and northwest of Pruitt-Igoe are the Jackson and Blair school attendance areas. At this time, business streets in these neighborhoods are, in most cases, lined by vacant, plywood or chain-link protected buildings. A reporter for the St. Louis Post-Dispatch wrote recently:

> From his $35-a-month apartment on North Market Street, Floyd West can look out his front window and watch a neighborhood die.
>
> A couple of years ago his block, between Twenty-second and Twenty-third Streets, was a small but busy commercial center. West could buy anything from wax to window shades by walking across the street.

> Today his block contains an appliance store, vacant; a clothing store, vacant; a confectionery, vacant; a tavern, vacant--and a few places still open but possibly soon to close.
>
> Another, in fact, closed yesterday. When tomorrow comes, those who had shopped at the Ben Franklin variety store at 2312 North Market may have to go half a mile for a spool of thread.
>
> To city officials and poverty workers who fought to keep the store in the neighborhood, the closing was the latest bitter chapter in the story of urban decay in St. Louis.
>
> (St. Louis Post-Dispatch, May 25, 1969)

Other school attendance areas lie east and north of those just described. The proportion of white residents can be seen to increase as one travels into these neighborhoods. The homes in which white families live are most frequently 4-flat and 6-flat apartment buildings. The residents are not prosperous, but an effort has been made to keep many buildings painted and in modestly good repair. Blacks continue to replace white residents who move away. One must suspect that the severe blight and decay which have occurred in attendance areas further south is also likely to occur in the currently white neighborhoods.

Today, Pruitt-Igoe is again the center of controversy and bad feeling which have culminated in a rent strike involving hundreds of families. If the strike continues, Mayor Alfonso J. Cervantes says that the projects will be closed within a year--the Housing Authority is near bankruptcy.

In a recent incident, a hydrant on the eighth floor of one building was broken by vandals. For about three hours, water poured down the staircases; soaking through ceilings; flooding apartments knee-deep; flushing dirt and garbage before it; ruining food, furniture, and clothing; and driving some 60 families into the street--where the children played happily in the flood.

The mayor, while working with Teamster leader Harold Gibbons to end the strike, places responsibility for maintenance of the projects squarely upon the federal government, and responsibility for better welfare payments on the state.

Such are the problems of this city, where the climate induced by the rent strike and the suspicious natures of the residents have combined to produce some difficulties for the study.

One must hope that man's ingenuity, which bridged the narrowing gap between the soaring pillars of the Saarinen Arch, skillfully uniting them at a point high in the air, will be as successful in closing the widening rift on the ground beneath them.

Trenton, New Jersey

Marshall Smith and Gray Sidwell

Almost two hundred years ago, from a slight rise above a village, George Washington directed the artillery fire which raked the unsuspecting Hessians below and destroyed them.

Many things have changed: the village is now a city; five busy roads meet near the spot where Washington stood; and the Battle Monument erected in commemoration of that long-ago victory stands at the crossroads. But one thing is unchanged, the city is as beleagured as was that small village in 1776. And by enemies more implacable than a few tattered revolutionaries. Some are modern; the roots of others reach back before Washington and began in another place. Urban blight, lost industries, high taxation, deterioration, discrimination, intolerance--all concentrated in an area too small to contain them easily.

In every direction from the Battle Monument there are spreading slums. As you go down into the heart of the city you cross the Penn-Central spur, which still has grade crossings in the center of town. If luck is with you, you may avoid the delay and inconvenience occasioned by a train parked across the highway. A little further along you cross the abandoned Delaware-Raritan canal, once a cross-state waterway, and now just a water source for industries up on Route 1. In less than a mile from your starting point you reach the Delaware River and the state buildings along West State Street.

The new cultural center, the fine museum, the library, the planetarium, and other buildings erected in the last few years are a handsome sight. But their white starkness, heightened by the sun, forms a dramatic contrast to the shabby, decaying district through which this short journey has taken you. And even here, from the city's point of view there is a dark side, for these buildings in company with many others (the State Prison, the State Hospital, the State Home for Girls, the School for the Deaf, and the Mercer County buildings among them) create vast areas of virtually tax-free real estate. These areas certainly do not generate or contribute the income which industrial use of the properties could.

Around the state complex are acres and acres of parking lots. As a gesture, trees have been planted which stand, regimented and pathetic, in small holes in the barren macadam. Beyond the parking lots and the broad rushing lanes of a modern divided highway lies the Delaware.

They say that the shad are coming back in the river--and other fish, too--as anti-pollution measures begin to take effect. The river is not navigable at Trenton which is located just above the mild rapids marking the head of tidal influence. Swing down the river on that modern highway, the John Fitch Way (named after the Trenton-based inventor of the first steamboat), and you find on your left, a little way inland, a large cleared area which, it is hoped, will become a great shopping mall. But the project is in trouble and with it all plans for a revival of downtown Trenton. Many potential sponsors, dubious about the future of the city, are hesitant to commit their resources.

From the right vantage point about here you can see the proud sign--Trenton Makes; the World Takes--stretched along one of the several bridges to Pennsylvania, a slogan which celebrates the industrial base which once made Trenton a vigorously prosperous city. Famous potteries; the gigantic complex of the Roebling family where wire cable was spun for suspension bridges all over the world and for whom the Brooklyn Bridge was just another job; these, and many others, contributed to that prosperity. Now, industry is pulling out of Trenton, thereby reducing tax ratables and forcing greater burdens on real estate. Now, too, middle class people are moving to the suburbs and across the bridges to Pennsylvania. And, as they go, the poverty areas spread as poor people move in. Moving also are institutions of higher education. Once there were three; soon, with the departure of the Mercer County Community College, there will be none. It would be wrong, of course, to suggest that these institutions exist solely to serve the city. On the contrary, they have a responsibility to the wider county and state communities. Moreover, the moves, none very distant from the city, have undoubtedly resulted in improved facilities and an enhanced capability for service. Nonetheless, it is regrettable that this could not have been achieved within the city.

Certain aspects of education in Trenton are encouraging. The preschool and Head Start programs have been outstanding from the beginning.

Managed and staffed by deeply concerned and able people they have secured wide public support and participation.

This matter of wide participation in community efforts to save the city and relieve the problems of poverty has been characteristic of Trenton for the last several years. It is a development which offers considerable hope.

Against this, the Central High School, which accomodates all the high school students of the city, presents a less happy picture. A huge structure of wide wings and long corridors, it has for some time been beset by racial problems. Securing a black superintendent of schools has not seemed to help the situation much. The junior high schools and elementary schools, in general, have good plants, though by virtue of their location they seem to have a high degree of de facto segregation.

There is developing in Trenton an unfortunate situation where ethnic groups, such as those of Italian and Polish extraction, are lining up against the encroachments of blacks and Puerto Ricans. This is evident not only in bitter political fights, but in incendiary statements and accusations of overreaction on the part of the predominantly white police force, and in the formation of citizens "patrols."

There are large areas of the city where the lawns and shrubbery are well tended and the houses are large and gracious. There are handsome high rise apartments; fine modern office buildings stand along State Street; and a new Penn-Central station to match the Metroliner is in the planning stage over on Clinton Street. Cadwalader Park is beautiful, especially so in May and June. Still, the predominant impression is one of squalor, dirty streets, and deteriorating neighborhoods.

Trenton of today needs some victories as urgently as did Washington those many years ago.

Portland, Oregon

Robert Lambert

The Columbia, mightiest river in the West, is majestic in its beauty and awesome in its economic importance to the Northwest. Portland lies at the confluence of the Columbia and Willamette rivers. According to historical fact, two enterprising Yankees set up their general store in a muddy patch of land on the banks of the Willamette in 1843. They flipped a coin to determine the name of the city. Mr. Pettygrove wished to name the future metropolis for his native city, Portland, Maine, while his partner, Lovejoy, opted for Boston, his birthplace. Pettygrove won the toss.

Eighty miles downstream to the west, on the edge of the Pacific, stands Astoria. A fur-trading post established here in 1811 was the foundation of the Astor fortune--John Jacob died the richest man in the country in 1848. Sixty miles east of the city Mt. Hood rises out of the mists. With heavy rainfall and clouds typical of the climate, it is understandable that being able to see Mt. Hood should be the city's touchstone for good weather. On a splendid day the Portlander can enjoy not only Hood but can also look northward into Washington and see Mt. Adams and Mt. St. Helens. Mt. St. Helens is absolutely symmetrical and with its snowcover resembles a high-peaked dish of ice cream. When the clouds roll away and the grass grows so fast that you can almost watch it happen, the natural beauty of the Portland area brightens the spirits of all but the most insensitive. But, when the sun vanishes--as it does--for 35 days, a certain gloom is forgiveable, perhaps.

Portland is bisected by the Willamette River as it flows northward through the fertile valley that bears its name. The older and wealthier part of the city lies on the hills west of the river, but the flatter expanse that stretches to the east has its share of hillocks which provide topographical relief for the inhabitants. The old business section on the west bank has suffered some urban blight. The grand old Multnomah Hotel, once the proudest hostelry in the Northwest, has been converted into offices. Still, only three blocks away, the unique

Benson, the most solidly elegant hotel in the Northwest, boasts its London Grill room, long famous for limestone lettuce and flaming entrees. The blighted downtown area embraces only a few city blocks around Burnside Street, the skid row of Portland. Dan and Louie's Portland Oyster House, renowned for its Quilcene and Olympia oysters, is located on Ankenny Street in the heart of the urban decay and still does a thriving business even though its well-fed patrons step out to find the panhandlers waiting.

South, along the Willamette, the freeway runs to Salem, the state capital, and so to the valley's principal city, Eugene--home of the University of Oregon--100 miles from Portland. And north, just across the Columbia, lies Vancouver, whose residents seem to prefer Portland's restaurants and substantially cheaper cigarettes.

Portland, the "City of Roses"--from the rose and bulb farms that verge on the eastern suburbs--is considered not a city by many, but just an overgrown town. For, even though it is a deep-water port and boasts a population of 400,000, it lacks the relative sophistication of Seattle, 140 miles to the north. Oregon's "Four Hundred" are regarded as hicks by travelers from more arrived cities such as New York, Chicago and San Francisco. Yet the Portlanders live well, enjoy their boats and summer houses on the Columbia, and, on balance, are proud of their city. Many people live in Oregon because of the distant panorama, the fishing, and the hunting. Just thirty minutes from the center of the city at day's end will put you in a picnic spot by a waterfall with a view of the Columbia sweeping westward. Throughout the state you will find a multitude of professional and managerial people who trade less income for less pressure and a chance to catch steelhead salmon in the evening. Oregon is a beautiful state and Portland a handsome city.

According to the best census figures and school district estimates the black population of Portland is about 6%. About 8% of the children in the Portland Unified School District are black. One reason why Portland was chosen as a study site was because it presents a unique situation in the Urban West: unlike in St. Louis, which has rotted in the center and bloomed on the periphery, the Portland whites have not fled to suburbia. Consequently, Portland is an intact cultural-socio-economic organism. True, racial incidents have occurred and there is some of the

national phenomenon of racial fear between the blacks and whites. Nevertheless, the polarization is not as marked and tensions are not as high as in Oakland, San Francisco, and cities of the Middle West and Eastern Seaboard.

Our operation is in Albina, an area roughly a mile wide and 2 1/2 miles long, bounded by the east-west freeway on the south, a northbound freeway on the west, a light industrial area to the north, and a lower middle class white community on the west. Originally, Albina was inhabited by people of German descent. It has had the reputation of being a tough area for 50 years. There are still several German churches and the parishioners who now live elsewhere come into Albina on Sunday and attend church unmolested. People of all races have free access to any part of Portland, unlike San Francisco where whites are adjudged to be foolish if they travel, particularly at night, to Hunters Point, one of the black ghettos. Those who expect Portland's ghetto to resemble Harlem or downtown St. Louis are doomed to disappointment or will receive a pleasant surprise--depending on their expectations. Except for one street in Albina which has a few blocks of business district with its share of vacant store fronts and small, unprepossessing, "Mom and Dad" business efforts, it is reminiscent of Las Veges. Albina resembles many other small Western towns such as Cody, Wyoming. The houses are modest, but most have tended lawns. The TV antenna is omnipresent and 87% of all homes and apartments in Albina have telephones.

Portland is a city of churches and Albina has its full share. Those whose facilities we have used have permitted us to move furniture and to adapt the quarters into functioning test centers. The church bulletins have contained descriptions of the study and the ministers have made announcements from the pulpit urging or suggesting participation. Although the power of the church is strong, particularly among the older people, some of the younger couples are disaffected. They seem to look upon the church as an enemy, or, at best, as little help in the fight for needed social change. One of our host ministers has a paying congregation of 2000--attendance is rarely over 200. Another of our hosts has demonstrated good rapport with the young and enjoys the respect of many young men and adolescents in Albina. But, all in all, the power of the church is being diminished--or, at least, questioned--here as it is throughout the land.

Although the operation has gone well we have had our problems. Albina has been oversurveyed; during the last five years numerous university and OE supported projects have canvassed and interviewed the residents. Despite the benignity of the longitudinal study, some mothers reject us as being possible lackeys of the relief agencies. Others are just tired of telling their stories. Most rejection has come from members of the middle class black community who do not want their children involved in anything which smacks of Head Start or any other "help us" efforts. With that exception, however, we have not encountered any group or bloc rejections. The white and black communities have cooperated in large part, and the decision to participate or not has been made at the individual family level. The small but important amount of money pumped into the community by our presence there is positive. The $5.00 cooperation fee presented to parents when their children have completed the process is very, very important. It constitutes a pleasant token to the more affluent and more groceries to the poorer families.

The children are as little children always are, terribly winning. Most of them appear clean, brushed, and fed. Some are hungry. In many cases this is because mama couldn't get herself up in the morning before she came to the test center. In Portland, as in any other community, poor people have "poor ways." But, there is a lot of pride in Albina.

At first, there was some fear that the study might be boycotted by certain subpolitical groups. To this time, these fears have been unfounded. The acid test came in early July when a small riot occurred. The incidents which led up to the riot are hard to determine and depend on whom you are talking to. The consensus is that a young black girl who badmouthed the police was beaten by them. Next day, a group of black youths allegedly set upon a white couple without provocation. The man was beaten and the woman kicked and stripped to the waist. In any case, the following night was marked by a number of set fires and two furniture stores were burned out at an estimated loss of $400,000. Fifty blacks were arrested. The following night, the older black leadership pleaded with the Chief of Police to keep the patrols out of Albina and let them try to cool it. At 1:30 in the morning, with the situation out

of control, the senior black leader, tears in his eyes, called for the police to return in force. They did and there have been no difficulties of any consequence since. We expected that a number of mothers, black and white, would withdraw their children, but, to our knowledge, only one mother did so.

This inspiring section of our country may prove to be a specific for integration at its best. However, with the most aspiring hopes, Albina remains much like our other ghettos; it will take work and a little bit of luck. Jobs for youth are not enough. Employment is a necessary but insufficient condition to prevent outbreaks of violence. As the young black males (previously castrated by our society), armed with their new freedom of expression and acting out, grope toward their new identities, all educational means must be employed to reduce tensions and promote understanding.

3. MEASURES USED IN THE INITIAL ASSESSMENTS, 1969

Virginia Shipman

MEASURES USED IN THE INITIAL ASSESSMENTS, 1969

Virginia Shipman

The overall design of the study calls for the following operations in the first year after identifying children for the longitudinal sample: interviewing their mothers, observing the interaction of children and their mothers in standard situations, testing the children in order to obtain initial estimates of their cognitive and personal-social characteristics, and obtaining medical information on each child.

Parent interview. The final parent interview consists of 274 items focusing on the mother's individuation of the target child; her attitudes toward the educational system; the availability of home resources and her knowledge of utilization; and attitudes toward various community resources, in addition to providing necessary demographic information. It typically requires between 1 and 1 1/2 hours.

As part of the questionnaire development stage, it was necessary to conduct an exhaustive pilot test. The pilot test had several goals: In order to keep within budgetary limitations, preliminary testing was necessary to ensure that the questionnaire took no longer than 90 minutes to administer. Moreover, it was feared that the validity of the results would be suspect if a longer questionnaire were used because of respondent fatigue or loss of interest. Another purpose of the pilot test was to test training methods which would be used when the full staff of interviewers was being trained. The pilot test was used by the project staff to evaluate whether or not the interviewer understood the intent of the questions. This information was helpful in preparing final questionnaire instructions. Each question was evaluated to ensure that the interviewer was able to communicate the intent of the question clearly to each respondent. Modifications in question wordings were instituted, where required, based upon pilot test results. Finally, the pilot test was intended to provide insight into whether the question had meaning to the respondent, within her frame of reference.

Following initial piloting in the metropolitan New York area, a full-scale pilot test consisting of about 10 completed interviews was

conducted in each of the four test cities, for a total of approximately 40 interviews. The interviewing procedures paralleled the final design and execution to as great an extent as possible. Audits & Surveys' project director traveled to three sites for the pilot test and ETS staff administered the interview in the fourth site (Trenton). Three interviewers in each city underwent an extensive briefing in order to conduct the pilot test.

Reactions of the interviewers to various parts of the briefing were useful in evaluating the training methods to be used for the main part of the survey. Once the briefing was completed, each interviewer was given an assignment to complete and instructed to return to the training center for a debriefing. The debriefing took place at a meeting with all three interviewers present. On a question-by-question basis, each interviewer was asked about her reactions to the question, how she thought each respondent reacted, and about the types of answers she obtained. The comments of one interviewer frequently served as a springboard for another interviewer to comment on some related experience she had had with that same question. The entire debriefing was taped for further analysis by executives of Audits & Surveys and Educational Testing Service. The debriefing report by the project director, supported by the tape recordings and independent analysis of the pilot test questionnaires were quite useful in the final revision of the questionnaire and the training procedures.

Since changes in the interview involved only deletion of a few ambiguous and/or alternatively worded questions, or modifications in format rather than in the nature of an item, a previously scheduled second pilot test was considered unnecessary.

Measures of children and of mother-child interaction. The December report (Educational Testing Service, 1968) specified a wide variety of measures that we felt would help us describe more adequately the complex interrelationships and structure of children's abilities and characteristics over time, and enable us to tease out their interaction effects with particular preschool and primary school programs. Even a brief scanning of these measures revealed certain inherent assumptions about what we felt was necessary to accomplish the goals of the study. Whenever possible, multiple sources of information about a particular phenomenon were proposed

(e.g., verbal behavior was seen as a function of the stimulus materials, the communicator-communicant relationship, and the purpose of the act, e.g., to inform, seek help, express emotion). We emphasised process rather than static variables, especially those process variables involving parent-child and teacher-child interactions, such as modes of information-processing and reinforcement strategies. Implicit throughout was our belief that only for purposes of analytic discussion could one separate cognitive-perceptual and social-personal domains or study the child without taking into account his environment.

It should be remembered that the measures described in December were chosen only after considerable pruning. In the excitement engendered by the manifold possibilities offered by this study it was difficult not to want to attempt to investigate more than was practicable. Constraints related to available testing time, balance among areas and modes of response, sufficient knowledge of the appropriateness of the task for the intended population, level of administration difficulty, readiness for inclusion by spring, and, of course, expense, led to the exclusion of many candidate measures. The following paragraphs describe the nature of and reasons for the changes that led to the final selection of the measurement procedures to be used with the 3 1/2-year-old sample children and their families.

It became apparent that the budget required reduced testing time. To eliminate one day's testing (from 5 days to 4), several measures were shortened or combined with a closely related one. For example, instead of studying both constancy of physical identity and sex role, the more salient and age-appropriate variable, sex role constancy, was chosen. Of the three measures proposed for assessing vigor (running, crank turning, and broad jumping), the more difficult and possibly hazardous one was dropped--broad jumping. Previous investigations with the Etch-a-Sketch interaction task indicated that comparable data would be obtained by using three designs, rather than five, for the mother and child to copy. By incorporating the nonverbal Uses Test into the Open Field situation, and by reducing the Open Field to one setting rather than two, additional time was saved. Further analysis of the Beery Developmental Test of Visual-Motor

Integration indicated that the 3 1/2-year-old could be expected to reproduce few figures beyond those already included in the Caldwell Preschool Inventory (line, circle, square, triangle). It was decided, therefore, to incorporate these other figures (slant line and X) with the Caldwell and then to score the Caldwell drawings according to both Caldwell and Beery criteria.

Given the necessity to reduce the time taken for interviewing the mother in the home and at the testing center, the proposed internality-externality scales were eliminated from this year's battery. Other indices of powerlessness, however, remain in the interview and it is planned that the omitted scales will be restored next year. In addition, two tasks (Stanford Memory and Form Memory) were dropped because additional pilot data revealed that few disadvantaged youngsters of this age would be able to attain more than a minimal level of performance. Given the lack of clinical expertise expected in the testers to be employed and the time available for testing, this was also the reason for dropping TAMA Tell-a-Story from the battery.

Additional changes or deletions in measures occurred following our pilot experience in using indigenous testers. It soon became evident that the various ratings proposed (Head Start Inventory of Factors Affecting Test Performance and Test Situation Ratings) were too many and too complex in format for the testers to assimilate quickly and well. Since the various demands at the time did not allow adequate revision, these ratings were dropped for this year. Testers, however, were trained to describe the child's test behavior as completely as possible so that modified Hertzig ratings could be obtained. Preliminary perusal of answer sheets suggests that it might be possible to develop codes for obtaining some of the earlier proposed ratings (e.g., child's task enjoyment, cooperation, attention). The only other major modification was in the mode of data collection for the three mother-child interaction tasks. To save transcribing and coding time it had been proposed that testers should code "on-the-spot." Given the relatively brief training period for the number of tasks involved and the difficulties experienced by naive testers in learning how to relate simultaneously to both mother and child, this procedure was abandoned for taping of the entire interaction plus instructions to the tester to note relevant nonverbal behaviors.

The above paragraphs have described major deviations in the measures proposed previously for the 3 1/2-year-old battery. The following paragraphs describe the procedure followed in arriving at the final version and sequencing of the measures selected.

Prior to initial selection, each measure had been administered to children similar in age and socio-economic level. None, however, had been given by indigenous testers under the supervision of an ETS researcher; typically, a research assistant of graduate student level had been used. Although considerable rewriting of test manuals and changes in test format to facilitate handling of the testing materials had taken place both before and during the training of tester trainers, refinement of procedures awaited piloting in the field. The first two sites (Auburn, Alabama and Portland, Oregon) were, therefore, used for continued simplification and clarification of testing and scoring procedures based on trainer experiences and trainee suggestions. Similarly, the pilot batteries for the four days had been arranged to take into consideration the need to balance type of response (active vs. passive, verbal vs. nonverbal), to maintain constancy of certain sequencing (Johns Hopkins Perceptual Test before Matching Familiar Figures, since the former involves practice on the response demanded), to offer a variety of stimuli, and to provide something to take home (a picture, bag of toys, coloring book, Tootsie Roll). Nonetheless, the batteries still had to be representative of the various domains. The first week of dry-run cases in each site piloted the adequacy of the sequencing. Following experience in the first two sites, minor adjustments were made enabling more equivalent testing time and level of difficulty of test administration across batteries. Trainees and trainers were encouraged to discuss the merits of the various modifications, and not until testers were ready to test actual sample children were procedures stabilized for final production of manuals. From such cooperative efforts, of course, are derived not only more adequate measurement procedures but the type of community-based research for which this study stands. (Table 1 shows the final order of the tests in batteries.)

Medical examinations.* As preparations in the medical area got underway it appeared that our initial conceptions had been somewhat naive and

*This section was prepared by George Gordon.

Table 1

THE MEASURES AND TESTING SEQUENCE USED IN THE

INITIAL ASSESSMENTS

Day I

Mother-child Interaction tasks:

Hess & Shipman Toy Sorting Task
Hess & Shipman Eight Block Sorting Task
Hess & Shipman Etch-a-Sketch Interaction Task

Motor Inhibition Test
ETS Matched Pictures Language Comprehension Task I

Battery A	Time in minutes
Cooperative Preschool Inventory (Caldwell)	20
Vigor I	5
Spontaneous Numerical Correspondence	5
Massad Mimicry Test	10
TAMA General Knowledge I	10
Risk Taking 1 and 2	20
Picture Completion (WPSSI)	5
Battery B	
Sigel Conceptual Style Sorting Task	20
Mischel Technique	5
Johns Hopkins Perceptual Test	15
Open Field Test	10
ETS Story Sequence Task	20
Seguin Form Board	5
Matching Familar Figures Test	15
Battery C	
Fixation Time	20
Vigor 2	5
Brown IDS Self-Concept Referents Test	10
Preschool Embedded Figures Test	15
Children's Auditory Discrimination Inventory	10
Peabody Picture Vocabulary Test, Form A	15
Sex Role Constancy Task	5
ETS Enumeration I	5

that modifications of approach would be necessary. A previous report indicates that "In contrast to the other areas of the study, many of the techniques and instruments here (in the medical area) are almost universally familiar to the appropriate practitioners" (Educational Testing Service, 1968). Further consultation and study indicated, however, that of the techniques which are most appropriate to the research questions posed, many are unstandardized, are not generally available, or are still in a strictly experimental state. To a large extent, the information obtained in a typical physical examination utilizing familiar methodology is not highly productive of the type of data required to investigate relevant research questions.

In light of these developments, it was decided that if research were to be done on the relationships of physical variables to educational development, it would be necessary to concentrate as much developmental effort in the medical area as in the other areas of the study. Since ETS does not have research capability in this area, the most desirable course of action seems to be the development of a joint effort with an institution equipped to conduct research on the medical aspect of child development. Possible relationships of this type are being explored.

Since no such relationship existed during the first year, the approach has been to obtain data which for the most part parallels information obtained in the normal Head Start procedures. As a result, a Head Start type of physical examination and medical history will be available for the entire longitudinal sample, whether or not the children attend Head Start. However, some of the appropriate research questions and instruments await the establishment of a liaison with a qualified medical institution.

The form currently being used in the collection of medical data follows page 46.

Reference

Educational Testing Service, PR-68-4. Disadvantaged children and their first school experiences: Theoretical considerations and measurement strategies. Report in 2 volumes, December 1968, Contract OEO 4206 and Grant OEO C6-8256, Office of Economic Opportunity.

A LONGITUDINAL STUDY

CHILD HEALTH RECORD

NAME OF CHILD (LAST, FIRST, MIDDLE)

HOME ADDRESS

Child I.D. Number

ILLNESS HISTORY

HAS CHILD HAD OR DOES HE HAVE:	YES	NO	DATE	DESCRIBE DETAILS OF ANY ITEM CHECKED "YES"
MEASLES (RUBEOLA)				
MUMPS				
CHICKEN POX				
RUBELLA (3-DAY OR GERMAN MEASLES)				
WHOOPING COUGH				
SEIZURES, FITS, OR SPELLS				
TONSILLECTOMY				
ANY HOSPITALIZATION				
EXPOSURE TO TUBERCULOSIS OR PERSON WITH CHRONIC COUGH				
FREQUENT BEDWETTING NOW				
ANY KNOWN CHRONIC DISEASE OR HANDICAPPING CONDITION				
OTHER SERIOUS ILLNESS				

IMMUNIZATION RECORD

1. DPT
   a. ___ Has never been immunized for DPT
   b. ___ Has received at least one dose, but not fully immunized
   c. ___ Was fully immunized for DPT (had at least 3 doses of vaccine, the most recent within the past 2 years)
   d. ___ Has unknown DPT immunization status

2. Polio-myelitis
   a. ___ Has never been immunized for polio
   b. ___ Has received at least one dose of polio vaccine, but was not fully immunized
   c. ___ Was fully immunized for polio (has received at least 3 doses of trivalent oral polio vaccine, the most recent within two years; or had received polio vaccine, the most recent within two years)
   d. ___ Has unknown polio immunization status

3. Small-pox
   a. ___ Was never vaccinated or vaccination was not successful
   b. ___ Has received a successful smallpox vaccination DATE_____

4. Measles
   a. ___ Was never immunized against measles
   b. ___ Has been immunized against measles DATE_____

PREGNANCY AND BIRTH HISTORY

PLACE OF DELIVERY (NAME OF HOSPITAL)

PREVIOUS PREGNANCIES			BABY'S BIRTHWEIGHT	BABY'S HEIGHT
TOTAL NO.	MISCARRIAGES	STILL BIRTHS		

MOTHER'S HEALTH DURING THIS PREGNANCY ☐ EXCELLENT ☐ OTHER (DESCRIBE)

DELIVERY ☐ NORMAL SPONTANEOUS VERTEX ☐ OTHER (DESCRIBE)

DID BABY ARRIVE
☐ ON TIME ☐ EARLY BY _____ WEEKS ☐ LATE BY _____ WEEKS

ILLNESS OR COMPLICATION IN NEWBORN PERIOD ☐ NONE ☐ OTHER (DESCRIBE)

FAMILY AND HOUSEHOLD

NAME	DATE OF BIRTH	LIVES WITH child		HEALTH PROBLEMS AND SCHOOL PROGRESS
		YES	NO	
CHILDREN IN ORDER OF BIRTH (LIST ALL PREGNANCIES INCLUDING PATIENT) 1.				
2.				
3.				
4.				
5.				
6.				
7.				
8.				

ARE THERE ANY DISEASES WHICH "RUN IN THE FAMILY"?
☐ NO ☐ YES (DESCRIBE)

DEVELOPMENTAL HISTORY

COMPARED WITH HIS BROTHERS AND SISTERS AND WITH OTHER CHILDREN HIS AGE, HAS THIS CHILD BEEN PARTICULARLY FAST OR SLOW IN:	FAST	ABOUT AVERAGE	SLOW	COMMENTS
WALKING, RUNNING, CLIMBING				
TALKING				
PLAYING WITH TOYS, COLORING, DRAWING				
UNDERSTANDING WHAT IS SAID TO HIM				
GETTING ALONG WITH CHILDREN HIS OWN AGE				

IS THIS CHILD CONSIDERED BY HIS MOTHER OR BY OTHERS TO BE PARTICULARLY:	YES	NO	COMMENTS
"DIFFICULT" OR "DIFFERENT"			
HYPERACTIVE			
CLUMSY			

I give my permission for _____
to have all necessary medical examinations and laboratory tests from the physicians and other health personnel of the ETS-OEO study of young children. I further give my permission for personnel of this study to obtain copies of all birth and other hospital records of my child.

_____

Date Signature of parent or guardian

SCREENING TESTS RECORD

	CHARACTERISTIC	TYPE OF TEST	RESULTS		COMMENTS
			NORMAL	ABNORMAL	
VISION SCREENING	ACUITY (FAR)				
	ACUITY (NEAR)				
	OCULAR MOTILITY				
	FUSION				
AUDITORY SCREENING	ACUITY	AIR CONDUCTION			

Blood:
Hemoglobin, in grams/100ml _____

PHYSICAL EXAMINATION

HEIGHT	WEIGHT	AGE	
Inches	Lbs.-Ozs.	YEARS	MONTHS

DOES THE EXAMINATION REVEAL ANY ABNORMALITY IN:	ABNORMAL	NORMAL	NOT EXAMINED	DESCRIBE FULLY ANY ABNORMAL FINDINGS
GENERAL APPEARANCE, POSTURE, GAIT				
SPEECH				
BEHAVIOR DURING EXAMINATION				
SKIN				
EYES: EXTERNALS				
OPTIC FUNDI				
EARS: EXTERNAL AND CANALS				
TYMPANIC MEMBRANES				
NOSE, MOUTH, PHARYNX				
TEETH				
HEART				
LUNGS				
ABDOMEN (INCLUDE HERNIAS)				
GENITALIA				
BONES, JOINTS, MUSCLES				
NEUROLOGICAL EXAMINATION				
OTHER				

DEVELOPMENTAL SCREENING EXAMINATION

	NORMAL FOR AGE	OTHER (EXPLAIN)	REMARKS
GROSS MOTOR FUNCTION			
FINE MOTOR AND MANIPULATIVE FUNCTIONS			
ADAPTIVE FUNCTION			
LANGUAGE FUNCTION			
PERSONAL-SOCIAL FUNCTION			

SUMMARY OF FINDINGS, TREATMENTS, AND RECOMMENDATIONS

ABNORMAL FINDINGS	ADVICE AND TREATMENT GIVEN	RECOMMENDATIONS OR FURTHER EVALUATION, TREATMENT OR SOCIAL OR EDUCATIONAL SERVICES.

SIGNATURE OF PHYSICIAN

DATE

4. DATA COLLECTION PROCEDURES

Establishing Local Coordination Offices

Joseph Boyd

. . . . .

Securing Community Support

Samuel Barnett

. . . . .

Enumeration and Interview Activities

David Schwartz

. . . . .

Training Testers

Virginia Shipman

. . . . .

Operating Testing Centers

Virginia Shipman, William Ward

. . . . .

Medical Examinations

Joseph Boyd

Establishing Local Coordination Offices

Joseph Boyd

Earliest plans for conducting this study had provided for the direction of field operations from ETS regional offices, but it soon became apparent that full-time, on-site staff would be needed. Even before site selection had been completed, interviewing began of candidates for local coordinator positions. Criteria for coordinators were community rather than academically oriented. A coordinator had to be a person with strong ties and high acceptance in the disadvantaged community, who was, at the same time, acceptable to the establishment--especially the educational hierarchy. Other criteria included knowledge of community resources and ability to communicate.

A number of qualified individuals were interviewed in each of six potential study sites. As selection of sites became firm, the selected coordinators were employed as regular ETS staff. Concurrent with interviewing, Princeton staff members sought suitable office space.

Of the coordinators who were selected two are men, two women. Three are black, one white. One has completed college and some graduate work; one completed high school through GED equivalency tests. The two others have had some college work. Two coordinators had been working in community action agencies in the area of job training and placement, one had been a community agent with the school system, and the fourth had worked in the development of an adult basic literacy program. The coordinators comprise a particularly able, articulate, energetic, and involved group. They demonstrated their qualities to many of the Princeton ETS staff during a week-long orientation meeting in early January 1969.

The January meeting combined orientation to ETS employment, discussion of the project objectives, and discussion of details of field operations. During the discussion of operations it became clear that each local coordinator would need an assistant to handle office work and communication during the testing cycle. Since then, it has become clear that the assistant will be needed year-round. Steps are being taken to convert these positions from ETS temporary employees to ETS regular employees.

Local coordinators were project directors in their areas, with a broad range of responsibilities. As personnel directors they recruited, interviewed and nominated candidates for positions as testers. They processed personnel records, and were responsible for certifying payroll records. As purchasing agents, coordinators bought items that ranged from office desks to paper cups. As public relations officers coordinators spoke to community groups and interested individuals, appeared on television and radio programs, and quashed rumors. They established and maintained communication with community action agencies, school systems, and the Head Start organizations. Basic news releases were prepared and distributed from the Princeton office, but coordinators fielded inquiries from the news media as they emerged. They located sites for test centers and space for training the testers.

A major role of the coordinators was making arrangements to test the children. This involved communicating with parents, scheduling transportation of each mother and child to the test center, and assuring that center staff and facilities were adequate for the assigned load. Local coordinators learned to cope with the unpredictable--equipment breakdowns, sick testers, unwilling parents, untestable children. They knocked on doors to locate children, borrowed child-size furniture from the school, signed certificates of training, and balanced petty cash funds. In a word, coordinators were jacks-of-all-trades and masters of most.

In seeking office space we were looking for about 500 square feet, in two rooms, with the idea of using one room for an office and the other for training and, later, for storage of test material. In all four cities adequate space was found near the study areas. In all, except Trenton, the offices are on the second floor. The desirability of second story offices was reinforced by the only break-in experienced so far: In Trenton an air conditioner was removed from the back wall of the building and taken, along with a typewriter.

The offices were furnished with only essential furniture, mostly used, which was purchased locally. The policy of purchasing from local businessmen was adhered to whenever possible. For example, certain test materials were specified by make and model number for local purchase. For some of these,

the local coordinators were actually quoted lower prices than the ETS purchasing agent (in Princeton) was able to get for material to use in Trenton! Even when prices were comparable, many items could still be purchased locally more economically than the combination of central purchasing followed by shipping to the three remote sites.

A second round of real estate operations occurred in each study site as the time for testing approached: the location of testing centers. Ideal space requirements for a testing unit include six private testing rooms, a room for "Open Field" equipment, a playroom, and rest rooms. The only buildings with such facilities, standing empty most of the time, are churches. So, with one exception, all test centers were located in churches. Early plans called for simultaneous establishment and operation of four testing units per site, each near a school used to define the study population. Several factors prevented this, including number of qualified testers, number of children identified, and availability of test centers and testing equipment. Two unit staffs were organized in Portland and Lee County, three in St. Louis, and two and a half in Trenton. A testing unit includes nine people: a supervisor, six testers, a playroom supervisor, and a driver. Operation of test centers is described later in this chapter.

Locating suitable churches for use as testing centers was not always easy. Many small buildings, and the older ones in the inner city, did not have enough church school rooms. A number of the best choices had already been occupied by Head Start centers. Eventually suitable premises were located, and church boards granted permission for us to use their facilities. In Trenton, two and a half testing units are operating in the buildings of one church. In St. Louis, originally, two units occupied one church, and one unit another; now all three are located in one church. In Portland, two churches were used initially, with the shift of a unit from one location to another midway in the year's operations. A church-related university student center was used for testing in Auburn, along with several rural churches stretching from one end of Lee County to the other. It is indicative of the dedication of the testers that they endured the shifting of the center from place to place to be near where the children were. Although we were able to find church schools with enough rooms, we also found many rooms

to be unfurnished, or equipped with adult-size furniture. Child-size tables and chairs were purchased where necessary. In St. Louis the school system loaned the needed number of low tables to ETS.

Inadequate or nonexistent public transportation, and the need to make it easy for mothers to cooperate by allowing children to be tested during four days dictated that the project assume responsibility for transporting the children between home and test center. To that end, a car or station wagon has been leased on a monthly basis for each test center. Most of the drivers are men, and they are needed around each center to move furniture, pick up and deliver supplies, and so forth.

The logistics of this study are unprecedented in ETS experience. Other projects involve more subjects, or more test administrators, or more locations, or more test books and answer sheets. But no project has required the procurement of so many different and diverse items. Quantities required vary from a low of 4 (typewriters) to 7000 (tiny umbrellas and cars for the Risk Taking Task). Each of the more than 30 tests or tasks required a printed answer sheet and a manual. Most tasks call for material or equipment for the child to manipulate. A number of tasks had been developed from toys rather casually selected by the test developer--without too much thought to the requirements of production of test kits in modest quantities (20 or 60 kits). The Sigel Conceptual Style Sorting Task includes a red, white, and blue metal toy top; we had to import a supply from Detroit. The same task requires a blue rubber ball; blue is no longer popular, only red, orange, or multicolor. So, 24 red rubber balls were spray painted blue in the ETS paint shop, and they took weeks to dry. For other tasks, the paint shop also sprayed plastic cars, cups, spoons, chairs, and wood blocks in specified colors, because "You can't buy it that color any more."

The woodworking shop made blocks, three different kinds of sorting boards, and crank devices for the Vigor Task.

A random sampling of requisitions sent to the ETS purchasing department shows: marbles, tow trucks, rulers, Etch-a-Sketch, alphabet blocks, rag dolls (white and black), modeling clay, tape recorders, Polaroid film, and foot lockers (for equipment storage over the weekend). And the list could go on to more than 200 items new to the ETS inventory.

Securing Community Support

Samuel Barnett

The history of poor minority groups in urban areas has not been a pleasant one. Each group has had to fight to free itself from the mire of poverty and the lack of economic opportunity, and has had to develop the ability to protect itself from those agents of the power structure that would exploit it for whatever reasons.

Each group, in turn, was relegated to the poorest housing and the poorest sanitation, to the least-protected and the least-cared-for neighborhoods. In the area of education, the schools have been old, ill-equipped, and not prepared to offer the kind of education that was needed for the inhabitants to break into the mainstream of life. In order for members of minority groups to be able to deal with these environmental limitations, they developed a society closed to outsiders. All members of the group contribute to the overall effort in varying degrees. When children are asked questions, they respond by staring with blank or confused expressions. Adults, when queried for names or addresses, do not know any such person or address, even if they happen to be standing in front of it and have lived in the community all their lives.

For too long have poor urban communities (usually black) been the subjects of high-sounding idealistic surveys and studies. Some have been surveyed to such an extent that many residents feel they can answer the questions before they are asked. It is very common for them to state "I don't want to be part of a test tube operation," or, as one "oriented" radio announcer wanted to know, "Isn't this an experiment where you know the answer already?" Going into communities like these and securing support for a project requires the development of communication not only with the traditional power structure, but also with the emerging groups or leaders in the immediate black or white communities. The people in these areas identify the traditional leadership throughout the black community as having served the white power structure and its agents for years. These spokesmen often state for publication what the black communities or poor white communities will or will not accept. But, almost never do those in traditional leadership roles seek the counsel of the neighborhood residents.

together, community people were located, hired, and trained, and the census operation was begun. This first phase, when completed, was followed by interview sessions with the mothers of eligible children. Again the ETS local coordinators aided in the recruitment of community people to do this most important job. Our coordinators met with community groups, released information to the press, and generally made themselves and information about the study readily available. It was assumed that by the time the survey staff was ready to interview the mothers, they would be quite aware of the project.

Rumors were beginning to fly, and the importance of the initial contact for the mothers was most important, as it would set the pattern of cooperation so badly needed. A bad impression left by enumerators and interviewers would not only kill the project immediately, but might slam the door shut for any future research in these locations. To guard against this, ETS insisted upon having a major part in the orientation and instruction of those involved in the interviewing process. Many of the interviewers were unprepared to deal with attitudes such as that demonstrated by a father in one of our study areas. Exasperated after having been prevailed upon to allow his son to be in the project, he finally blurted out "I don't want him in the study--he is dumb and stupid (3 years old) like his brother and he will always be dumb. Why should I waste time on him?" Or to have a small child come to the door and say, "My mother says she is not at home," and slam the door.

Even widespread coverage in many of the large daily newspapers did not reach the people with whom we needed to communicate. There had to be person-to-person contact where community residents could see and assess the sincerity of project staff members. They asked questions like: What will the project do? Who is sponsoring it? Will it get my kid into Head Start? It sounds great, but will the present administration allow it to run its course? Another source for disseminating information was through use of the black media. Newspapers and radio stations cooperated. Unfortunately some areas like Trenton and Lee County do not have a black newspaper serving the area.

Community people showed their sensitivity to outsiders doing research by the questions they asked. They wanted to know about the authors of the tests. "Were any of them black?" "Were their ideas about the black learning experience being included?" "What about some of the questions in the family interview? How important is it to know where I shop, or how I spend my free time, or how much time I spend with my kid?" It was interesting to note the reaction to questions concerning time spent with the children. The assertion was that they spent more time with their children than middle class parents because they could not afford to have maids, house cleaners, or baby-sitters. And those fathers residing with their families claimed they spent more time at home than the ambitious middle class fathers who were often away on business.

The local coordinators became quite adept at answering most of these questions. Their greatest asset in separating this project from the many others carried out in the area was the project's use of community people.

A second major test of the project was the appeal to the community for tester applicants. If there were any doubts in some people's minds about our ability to locate people with at least some of the skills we were looking for, this was quickly dispelled when in each area there were more applicants meeting our criteria than were needed. It became necessary for coordinators to screen applicants, and notify those less qualified that they could perhaps fit into another segment of the study, but for this part they seemed unsuited and would find it most trying. For the most part, coordinators have maintained control of public relations and community goodwill even though not all applicants could be employed, and some who were hired were later released.

When training of testers started they were informed of the short duration of the training program (4 to 6 weeks), and the fact that only the best people would be chosen to do the actual testing. They were also informed of working hours, given an orientation to ETS and the study, and informed that they were in the vanguard of new research techniques.

As we continued to build support for the program through meetings with groups, discussion with community leaders, and response to questions

about the study, people began to question the measures. How were they being administered and who was administering them? It is quite strange that there were no criticisms in any location about the academic backgrounds of the testers. The fact that they were community people seemed to satisfy most questions on credibility and reliability. As tester training continued, the community people began to build a certain comradeship that crossed racial and neighborhood lines. By the time we had reached the final selection time, the trainees were talking about "our" study, and discussing technical areas of cognition and learning styles. These people were becoming great supporters of the project, and some questioned why the study was so long in coming.

To date, our latest efforts to maintain community support have centered on the following:

1. A graduation ceremony for those who successfully completed tester training. This has been enthusiastically received in all locations and trainees have in some instances prepared themselves several weeks in advance to attend and bring husbands or wives. Community leaders have usually been invited.

2. Periodic luncheons in each location, bringing together clergy, school people, community leaders and other influential members of the community with the members of the study team to discuss present problems and future goals.

3. Development of a study newsletter which will include notes of interest from each of the four locations.

No matter what we do in the area of maintaining community support, in the last analysis much of the success will depend on that personal "selling" ability which we hope our coordinators and others in the project will continue to use. As long as the public is convinced that the project offers tangible results and intends to follow through on its promises, the public will continue to buy it.

Reference

Educational Testing Service, PR-68-4. Disadvantaged children and their first school experiences: Theoretical considerations and measurement strategies. Report in 2 volumes, December 1968, Contract OEO 4206 and Grant OEO C6-8256, Office of Economic Opportunity.

Enumeration and Interview Activities

David Schwartz

Audits & Surveys, Inc.

Audits & Surveys has played a critical role in the Longitudinal Study of Disadvantaged Children. Its task has been to locate all eligible households within the geographic areas being studied, based on a definition of eligibility supplied by Educational Testing Service, and then to complete a 90-minute interview with the mother or mother surrogate of that household. When completed, the work will have involved screening approximately 40,000 households.

The purpose of this paper is to indicate the nature of the assignment, the kinds of unique problems encountered, and the solutions to these problems. In meeting this purpose, this paper is divided into several sections:

A. Overview of the Assignment
B. Preliminary Field Set-Up
C. Prelisting Phase
D. Administration of the Interviews

A. Overview of the Assignment

Audits & Surveys was assigned to collect data in four areas by Educational Testing Service. These areas consist of parts of Lee County, Alabama; Portland, Oregon; St. Louis, Missouri; and Trenton, New Jersey. Within each area, several school districts were selected by ETS. Audits & Surveys then screened every household in these districts to determine if any eligible children resided there. An eligible child was defined by birthdate, in terms of eligibility to enter first grade in fall 1971. As noted below, the admissible birthdates differed slightly from area to area, paralleling differences in first grade admission qualifications.

Eligible Children

Area	Date of Birth
Lee County, Alabama	October 3, 1964 - October 2, 1965
Portland, Oregon	November 14, 1964 - November 15, 1965
St. Louis, Missouri	October 1, 1964 - September 30, 1965
Trenton, New Jersey	December 1, 1964 - November 30, 1965

B. Preliminary Field Set-Up

Prior to the beginning of any screening in the four areas, a series of steps were taken to organize and set up the field operation. These fell into several categories:

1. Community Relations
2. Interviewer Recruitment
3. Interviewer Supervision

1. Community relations. Previous experience with similar surveys had demonstrated that a key issue is the generation of community support. Without community understanding and support, the survey could easily be plagued by problems due to the community's uncertainty and/or fear of the project, and of the project personnel. Conversely, good community relations could help the study in terms of increased respondent willingness to be interviewed and to give valid and reliable answers.

Attempts were made to secure community understanding, acceptance and cooperation through comprehensive press releases for the use of the local broadcast and print media. This achieved a high level of cooperation with local media, but resulted in little or no improvement in community attitudes. Subsequent investigation suggested that many community residents have little contact with any of the media. Therefore, the problem of generating community support was additionally handled by going directly to community leaders. These did not just include the local pastor, elected officials, or other "visible" leaders. Project personnel also sought out social club members, civil rights leaders, and Head Start and VISTA workers.

2. Interviewer recruitment. Contact was made directly with local community leaders. They have contacts throughout the community and are

sometimes able to supply a list of already trained local interviewers, or of local residents who wish to interview and who have the necessary educational qualifications. Educational Testing Service, as part of its own field set-up, had already made contact with key community leaders in each test area. Audits & Surveys was able to take advantage of these contacts.

3. Interviewer supervision. Past experience has shown that an effective supervisory system should have these elements, which are discussed in the following paragraphs:

a. A local field office
b. A full-time local supervisor
c. Continuous interviewer contact
d. On-the-spot supervision by the home office staff

A small, vacant store or office was rented in each test area as close as possible to the center of the study school districts. This office was equipped with desks, chairs, telephones, and other materials. One advantage of having a local field headquarters was that the office helped to identify Audits & Surveys to the local community and gave interested community residents an opportunity to get to know supervisory personnel and program goals from the beginning of the study. However, an even more important reason for setting up a local office was to provide a centralized location from which new interviewers could be hired and trained. As well, it allowed Audits & Surveys management to exert on-the-spot supervision over all field work. An extensive validation of each interviewer's work could be executed directly from the local office.

Audits & Surveys hired a full-time local supervisor to be on duty during all interviewing hours and exercise control over the quantity and quality of each interviewer's work. The supervisor was also part of the supervisory team responsible for the initial hiring and training of interviewers. Later, the supervisor was responsible for training replacements for those interviewers who were not performing satisfactorily or who resigned for personal reasons.

Each interviewer was regarded as a full-time member of an office staff rather than a temporary worker. She reported to work at regular hours and worked closely with the supervisor. Periodically, each interviewer reported in to receive her next assignment and materials. The assignment completed, she returned to headquarters, left all materials and completed work, and "signed out," recording the hours she had worked. The time allowed to complete assignments varied from one area to another. In rural Lee County, where some interviewers had to make a round trip of up to 60 miles, the period was about one week. In Portland and Trenton, on the other hand, the assignment period was generally about two days.

Members of the home office staff frequently visited the study sites. Thus, problems that local staff were unable to solve were quickly referred to senior staff members.

C. Prelisting Phase

Several major elements were involved in the prelisting phase, including:

1. Defining school district boundaries
2. Training interviewers
3. Handling unanticipated prelisting problems

1. Defining school district boundaries. What was thought to be a relatively uncomplicated task of defining school district boundaries in order to begin prelisting proved to have special problems.

For example, in Lee County, Alabama, Audits & Surveys representatives went first to the County Board of Education, only to find that no official could define accurately the boundaries of each school district. There was open enrollment and any student could enroll in any school throughout the county. Further investigation did show, however, that most students lived near the school they attended, and that only a few traveled long distances.

So, to define the school boundaries, the interviewers rode the school buses on those routes that traveled to the perimeter of each district. In effect, the bus route was made to serve as the district boundary. Interviewers noted where children got on the bus and asked each child how far he had

walked to get to the bus, and from which direction. These households were then plotted on a map of that school district, along the perimeter.

2. Training interviewers. As noted earlier, contacts were made with local community leaders, both by Audits & Surveys and by ETS staff members. Interviewers recruited in this manner were encouraged to have their friends apply. In some cases, advertisements were placed in local newspapers. In order to qualify as an interviewer for the prelisting phase the candidate had to be a resident of the community and be able to read and write. Each training session--or briefing--was generally limited to between six and nine trainees. When the background briefing was completed, a one-page questionnaire was given to all interviewers to look over.

The first section of this questionnaire provided room to record the interviewer's name, and to record the respondent's name and address.

The second section provided space to record the results of each listing attempt. Interviewers were instructed to make as many as three attempts to list a household. In some situations (such as a refusal), the interviewer might be instructed, after the very first call, to make no further attempts to get information from a household. In other situations, the interviewer could make a second or even a third attempt to complete the listing. Because of the various possibilities, a flow chart was incorporated in the questionnaire which gave instructions for various situations, while also providing space to record the result of each interviewing attempt.

The third section of the questionnaire consisted of several questions to be read verbatim to a responsible adult in each household. The respondent was asked if one or more children resided in the household who were born after December 1, 1961. If so, the name and date of birth of each child was obtained. It was also determined if any of the children were enrolled in the first grade of public school.

As noted earlier, the purpose of the prelisting phase was to locate households with 3- to 4-year-old children. Audits & Surveys included the above-mentioned series of questions primarily for control purposes. Some mothers forget the exact age of a child; others sometimes forget about one of their young children. By asking a respondent to think of all young children in the family and by structuring the questions so that the respondent is literally forced to think about each child, a subtle yet powerful form of probing is introduced.

Since there was no information available about the number of 3-year-olds living within each school district, Audits & Surveys was concerned that it would have no way to evaluate the accuracy of the listing. However, U. S. census records show that the number of 3-year-olds in a given area is generally similar to the number of 6-year-olds; hence first grade enrollment was obtained as a check.

Once questionnaire training was completed, each interviewer was given a sample location map which covered only a small part of the school district. The interviewer had to cover this small area completely by locating and visiting every household.

Interviewers were given the following instructions once they had studied the location map and been given a general introduction on its use.

Urban Areas

You are to cover each household within the blocks shown on the map. Do not go beyond the area assigned nor are you to omit any households on the blocks. Generally speaking you should start interviewing at the household that is in the top left hand corner of your map. Then walk around the block in a clockwise manner contacting every single household. Do not cross any streets. Just go around the block. When that block is finished move on to the next block on your map and go around that one in the same clockwise manner.

Rural Areas

You are to cover every single household within the area shown on your map. Do not go beyond the area assigned nor are you to omit any households within the area. You must use your judgment to see that every single household in your area is located. For example, you may find little paths or unmarked roads not indicated on this map. People may be living along these roads. Track down all unmarked roads within the land area on your map. Don't be afraid to use available clues such as a cluster of mail boxes on an unmarked road which suggests that several households are nearby. Similarly, several children playing in a deserted field suggest nearby residences. As you complete an interview, ask the respondent if he knows of any other families in the area.

Independent checks were made by the supervisor to confirm the number of households in the area. Interviewers who completed this first small assignment satisfactorily were then given additional assignments.

3. Handling unanticipated prelisting problems. During the prelisting phase, there were a series of problems.

One immediately encountered concerned the development of individual location maps to control interviewer assignments. This was particularly difficult in rural areas of Lee County. Frequently there were no named streets or official county roads. Usually, available landmarks are used when preparing location maps for rural areas. This was not possible in Lee County because there was no up-to-date official map of the county on which such landmarks could be identified. The most recent county map was prepared in the late 1930's with some irregular and scattered updating in 1948.

The problem was severe. Without a detailed map it would be difficult to control interviewer assignments. So, Audits & Surveys turned to the County Tax Assessor, only to find that even he had no detailed map of the county, but was using maps inherited from several decades earlier.

The problem was resolved by hiring several local residents who had lived in the area for many years. These people traveled through the county making detailed maps of each school district. This appeared to be an effective solution and these maps were used to prepare the needed location maps.

Other steps take in Lee County to locate potential subjects included the following:

a. Local grocers throughout Smith's Station were interviewed. Many were able to provide help on where to find some residents.

b. The U. S. Post Office in Smith's Station was approached and the local postmaster was interviewed.

c. Audits & Surveys local area supervisor met with officials of the Alabama Power Company. They assisted in locating additional households.

d. The local Head Start office was visited. Their wide contacts with underprivileged families throughout the county aided the location of still more households.

Even then, the problem of locating all families residing in Smith's Station was not completely resolved. Again several area residents with detailed knowledge of the district were able to locate some households that had been missed.

This problem of locating the expected number of households was not unique to the rural areas of Lee County. In St. Louis it was found that many of the neighborhoods being studied have houses with entrances in alleyways which do not appear on official maps. Here, too, the solution was based mainly upon the use of local residents.

As noted earlier, a question about first grade enrollment was used as a cross-check to ensure that as few as possible eligible households were missed during prelisting since there were no official figures to check against. However, unexpected variations in local enrollment practices caused problems here, too. For example:

a. In several school districts, students are admitted who reside outside the district.

b. In Portland several school districts suddenly switched grade coverage among schools in the district so that some have students at the third grade level or above and others provide first and second grade coverage.

c. In St. Louis the grading system proved confusing. School authorities assign each child a combination alpha-numeric code which refers to the child's level of academic proficiency rather than to the number of years he has been attending school.

d. Certain schools are thought to be "better" than others. Some parents falsify their addresses in order to get their children into the "better" schools.

e. One final problem was that a few children are enrolled in the first grade who were born before December 1, 1961. These children were over 7 years old at the time of the interview. (The questionnaire was concerned only with children born after December 1, 1961.) Although this is relatively rare on a national level, within the specific populations being studied there is a greater likelihood of encountering the situation. As a result, the number of first graders found in the prelisting is likely to be lower than the number actually enrolled.

In view of these kinds of problems a replacement cross-checking procedure was used. Supervisors in each of the test cities were instructed to have a different interviewer recheck each completed location map by counting the number of dwelling units. This helped to find several location maps which had too few households. Once these areas were found, supervisors sent interviewers back into the area to prelist again.

Some additional problems were encountered during the prelisting phase. One major problem was that many ghetto area residents were afraid to be interviewed. These refusals were related to the level of community cooperation. Audits & Surveys, working with Educational Testing Service, attempted to stimulate favorable community attitudes through articles in local newspapers, circulars given to school children to take home to their parents, as well as paid advertisements.

D. Administration of the Interviews

Once the prelisting was completed, work proceeded with the implementation of the main interview phase. This consisted of four steps:

1. Supervisor training
2. Interviewer training
3. Administration of the interviews
4. Check-in of survey results

1. Supervisor training. As the first step in the process of interviewing, an intensive supervisor training program was conducted in Audits & Surveys' offices in New York City. All four supervisors came to New York for a 4-day period.

Perhaps the most important aspect of training related to the questionnaire. One objective was to teach supervisors to train interviewers on how to administer the questionnaire. A second objective was to teach supervisors how to evaluate the work of each interviewer. To help a supervisor develop skills in these areas she was first made proficient in interviewing with the questionnaire.

Then the supervisors traveled to Trenton, New Jersey to conduct several interviews. They were told that each interview would be with a

"real" respondent who had been located in the prelisting phase, and that the questionnaire would actually be used in the analysis of survey findings. Unknown to the supervisors, Audits & Surveys hired several women accomplices who posed as respondents. These "respondents" deliberately gave each supervisor some kind of trouble, such as exhibiting a lack of desire to be interviewed, giving an unclear answer, or showing a general lack of interest midway through the survey. After the interviewing was completed, the supervisors again met as a group.

Once supervisors developed familiarity and skills with the questionnaire, the training moved to a discussion on how to evaluate interviewer work. The following is an extract of an outline regarding interviewer evaluation:

> When the interviewer brings her work into your office, be sure to find out what happened at each interview. Review her questionnaire carefully and point out all apparent mistakes to her and give her any hints you can think of which would solve her problems. Only if you are completely satisfied with her work up to this point are you to give her any further assignments.
>
> While the interviewer is still in the office check the following:
>
> 1. Completed questionnaires
>
>   a. Are all questions answered?
>   b. Are answers written legibly? (If "translation" of answers is needed have interviewer take care of this immediately.)
>   c. Is a properly completed classification form inserted in the questionnaire?
>   d. Is the time the interviews took reasonable? (If less than 1 hour or more than 1 1/2 hours, make a special note of it, and find out why.)
>
> 2. Completed contacts--no interview
>
>   a. Is the reason the interview was not completed a reasonable one?
>   b. Is the information on the classification form complete?
>
> 3. Any problems the interviewer had or comments she made.

> If, at the time you examine the questionnaire, you find certain "factual" questions that have not been asked, return the questionnaire to the interviewer and have her recontact the respondent and get the necessary information. If the question is about an "attitude" do not recontact the respondent. When the respondent must be contacted, almost all of these corrections can be made by telephone. If possible, have the interviewer do this right in your office.

Another objective of the training program was to give supervisors guidance on how to hire interviewers in the study sites. Audits & Surveys instructed supervisors that interviewers must be female and matched in race to the respondents. They should also be:

1. Over 21 (or at least close to 21) years of age
2. Neat and clean in appearance
3. The kind of person who would be patient and understanding with respondents
4. Cheerful and friendly ("out-going")
5. Able to follow directions exactly as given to them
6. Conscientious and accurate workers
7. Able to write legibly

This training program appeared to be successful, and the supervisors had the opportunity to meet each other, to compare problems, and to stimulate each other. Based upon subsequent on-the-spot observations by the project staff, and on an examination of completed questionnaires, the interviewers that were selected were satisfactory.

2. Interviewer training. The interviewer training program in many respects paralleled that for supervisors. Training lasted for two days and each program was limited to a maximum of five interviewers.

The first day was devoted to introducing interviewer candidates to the project. The orientation was similar to that given in the prelisting phase. Once the background orientation was completed, training began on how to use the questionnaire. The reasons for asking each question were pointed out. Then, each interviewer interviewed a partner while the others observed. After all interviewers had completed an interview, the efforts of each one were discussed. Later that day each interviewer completed a test covering the instruction booklet.

The following day each interviewer returned to the test center to have her "test" graded. The "test" served as a point of departure to review several key elements of the questionnaire.

3. Administration of the interviews. Once interviewer training was completed, the interviewing of qualified mothers or mother substitutes began. The administration of these interviews went smoothly. Only a few unanticipated problems arose.

One problem related to locating respondents who had qualified at the prelisting stage. This problem was most severe in the rural areas of Lee County, Alabama, for a variety of reasons:

a. Some interviewers who worked on the prelisting were either not interested or not qualified to work on the main interview phase. Consequently, difficulties arose when attempts were made by new interviewers to locate the prelisted households.

b. Similarities in family name caused problems. Several areas had many families with the same last name, thus making it difficult to find the prelisted family.

c. Another problem arose because the U. S. Post Office had recently switched route numbers in part of the county, while elsewhere the old route numbers were still being used. In several instances there were two different locations with the same route designation.

The standard policy throughout the survey was that all staff members (including interviewers) were to be completely briefed on all aspects of the survey. One experience suggests, however, that there is a limit to how much information should be given. In Portland, interviewers were told that in order to optimize the validity of the survey, black interviewers would be assigned to black respondents and white interviewers would be assigned to white respondents, wherever possible. Reaction was fierce; several interviewers stated that such a policy was "racist" and that "it perpetuated segregation." They insisted on being treated as "interviewers" and not as "black interviewers" or "white interviewers." Rather than jeopardize the study, this policy was withdrawn. Instead, the supervisor attempted to assign interviewers to areas where they would be most likely to find respondents of the same race, without specifically calling this to the attention of the interviewers.

4. Check-in of survey results. Each interview was reviewed on a question-by-question basis. Answers were checked for consistency, clarity, and completeness. If any answer was insufficient, for whatever reason, a decision was made as to whether the question was "factual" or "attitudinal." Where the question was factual (that is, the answers were not thought to vary over a short period of time), the respondent was recontacted to obtain clarification. If the question was of an attitudinal nature (that is, answers to it could vary over a short period of time), the respondent was not recontacted.

Regardless of how generally satisfactory an interviewer's work seemed, the supervisor recontacted at least 15% of every interviewer's respondents by either telephone or personal visit to the home. Since the program is being conducted over several years, it was important not to antagonize any respondents during this validation phase. Therefore, recontacts were explained as a request for additional clarification, rather than as a check on whether the interview was actually done. The respondent was told that the questionnaire was not clear at this point and asked if she could please just help to clarify "the record."

Training Testers

Virginia Shipman

From the beginning of the study it had been argued that using local testers would facilitate community cooperation, increase the validity of the data obtained and, in addition, provide training that would contribute to future employment possibilities for community residents. Operational details, however, necessarily remained vague until sites had been selected and information could be gathered on the particular communities involved. The following paragraphs describe what we did. Having no previous study on which to rely, and given limited time and resources, our procedures should be recognized for what they were--a pilot attempt. We are gratified by the results obtained and look forward next year to a more efficient and thorough approach that will capitalize on our gains and reduce, if not eliminate, the losses.

One month prior to the date when we were scheduled to begin training, eleven ETS staff members were approached as to their willingness and availability for training testers in distant cities. These staff members were chosen primarily for their experience in testing preschool children and their ease in relating to others. To this small group, all of whom agreed to participate, were added two other staff members who had related experience (such as kindergarten teaching), but required more extensive training in child testing. The first two weeks were spent familiarizing the group with the variety of tasks chosen for the 3 1/2-year-old battery. Each test was first demonstrated by the researcher who had proposed the measure. Next, a videotape of the test being administered was observed, followed by a discussion of the various questions raised. The group then practiced giving the test to one another, then to children of the appropriate age. For our two less-experienced trainers, we arranged that they first administer the tasks to children somewhat older than the study subjects and, preferably, known to them, and then to lower socioeconomic preschool children. Those previously unfamiliar with a particular task were observed administering the test before being checked out as able to train someone else on it. Each one of the original group of eleven had had prior experience administering at least one of the tests.

It should be noted that during this training period materials had to be shared or improvised since they were very scarce, and manuals were being revised continually to increase their uniformity, clarity, and comprehensiveness. Fortunately, trainers revealed the patience, flexibility and humor required.

Due to the limited number of trainers and the limited time they could afford to be away from their responsibilities at home and in the office, original schedules for data collection had to be revised. Staff resources did not allow us to begin training simultaneously in the four sites. Training commenced at two-week intervals, starting March 17 in Auburn (March 31 in Portland, April 14 in Trenton, and April 28 in St. Louis). It was also necessary to limit the number of trainees. Consequently, we staffed only two testing units in the first two sites and three units in the last two. If necessary, these staffs were then transferred to new locations to accomodate the children in other sample school districts within a community. This schedule also enabled us to use our more experienced trainers in the field first and to provide a more prolonged training period for those requiring it. An additional benefit was that we could transfer any knowledge gained in our first field experiences to our other sites before actual testing began. For example, following the suggestions of trainees in Auburn and Portland, we revised further the various test manuals.

The general procedures in the field were the same in each site. Prior to the arrival of the training team, the local coordinator preselected the tester trainees, choosing approximately 30% more than the number that would eventually be hired. Depending on a variety of factors, such as the resources in the community, the local coordinator's preferences, publicity concerning the project, and intra-community relations, trainees varied both within and between sites. As specified, all trainees were female. The usual educational credentials were not required, but experience working with young children was considered essential, as was the ability to read and speak easily. Trainees varied considerably in age and educational background. Most were housewives and had limited work experience. Most were black, but considerable effort was expended to recruit white trainees wherever possible.

The first two training weeks were held in the local coordinator's office. Following a general orientation to the project and to testing young children, trainees began on the very first day to practice one of the simpler tasks. It was felt that modes for handling the variety of problems that a tester is likely to encounter are best discussed in the context of a particular test. These general procedures were then repeated more meaningfully in the context of other tasks. As in training trainers, the task was first demonstrated and then the trainees were assigned to pairs and practiced administering it to each other.

The first tasks demonstrated were those in the Day I battery (see Table 1 in the chapter, Measures Used in the Initial Assessments) since all testers were required to administer that sequence. To reduce the number of tasks that each trainee would be required to learn, each one was then assigned to learn one of the three remaining batteries. The three groups were formed by counting by threes in seating order, a method which though appearing random helped to separate friends and to increase the racial mix. Each group (ranging from 6-10 members) then met with one trainer for the remainder of the two weeks. Following demonstration, trainees practiced administering each of the tasks to each other and then to children volunteered by other trainees, their friends, and neighbors. In most cases, these children were verbal, cooperative, and easily tested, enabling the new trainee not to be overwhelmed by her first testing experiences. Additional teaching methods used in the last two sites were the use of videotapes of trainees administering tests and brief tests to assess the trainee's knowledge of the tests in the battery.

In the third week, the trainees moved to the actual testing centers. A trainer was assigned to each testing center to ensure the adequacy of the physical arrangements, to arrange for the necessary testing supplies, and to function as a center supervisor so that trainees could concentrate on improving their testing skills. The local coordinator arranged for practice subjects who would be comparable to our sample subjects. This was not an easy matter, since it usually involved bringing children from considerable distances to the testing center. The fourth (and sometimes fifth) week of testing practice subjects was observed by Mrs. Shipman and Mr. Ward to select those trainees best prepared to be center supervisor,

tester, or play area supervisor. (The respective functions are described in the next section.) In addition to observing each trainee test different young children, the evaluators role-played 3-year-old subjects in order to test the trainee's understanding of particular procedural details which may not have occurred in the practice sessions. Following evaluation, each trainee not selected was seen individually, and every attempt was made to structure the situation as a growth instead of a failure experience and to maintain the person's interest and involvement in the study. The most common reasons for non-selection involved difficulties in relaxing with and relating easily to the young children, and being willing and/or able to administer the tasks in a standard manner. Those trainees selected as center supervisors and play area supervisors were also seen individually to explain more fully their respective responsibilities. Once evaluations were completed, each center operated one to two weeks more for a dry run. A Princeton office trainer continued to be assigned to each center to provide general assistance and additional test training while the center staff practiced their new roles. Monitoring of center operations was assumed by ETS regional office staff with the assistance of Princeton office staff once actual testing began.

Following completion of tester training in the four sites (mid-June in St. Louis), new training situations arose. For a variety of reasons, some testers resigned in each of our sites. Thus, from mid-June to the present, trainers have had to resume most of the activities described above. Training time has been shortened, however, for both trainee and trainer since the trainee obtains more individual attention and the trainer can share training duties with regional office and local center staff and does not have to manage center functioning as well. As with the original trainees, however, evaluations are made in the same way by the Princeton office trainers.

We should note the discrepancy between plan and practice that was evident here as in almost every other applied research endeavor. From the beginning the number of available trainers was small and no arrangements for back-up support were made. Thus, when illnesses and personal emergencies arose, increased demands were made on the remaining trainers. Although the training period was originally scheduled for 4 weeks, training

took 6-7 weeks in each site, and we soon used up the time for which trainers had agreed to participate in this phase of the study. Trainers were more than generous in their willingness to spend additional time in the field, often at considerable personal cost, but there were obvious limits to which this could be done. We did recruit two additional trainers at a later date, but for the past 2 months we have worked with a maximum of six trainers.

Our original plans provided for trainers to remain at least 1 week in a site, with one trainer, familiar to the trainees and the local situation, being present to introduce new trainers when they arrived. However, various factors, such as those mentioned above, interfered with the consistent application of this practice. It would have been much less confusing to trainees if they could have worked continually with the same staff. In addition, trainee and trainer would have become more knowledgeable about each other, and those frictions that did occur might have been avoided or at least would have had time to be worked through.

The last chapter of this report describes some planned changes in procedure for next year. The use of local trainers will help alleviate all the above problems. Moreover, testers will have always available to them the necessary, continued, on-the-job training. In addition, there will be more opportunity to pay attention to personal as well as task demands. We will, of course, also take cognizance of these factors when training trainers next year. Our experiences have emphasized for us the need to train all field workers in the special sensitivities required to handle possible conflicts such as those which arise when trainers are younger or of a different race, or when mothers are asked to manage their child subjects in a manner quite different from what they consider "right" behavior for young children.

The above is only a very dry, brief account of our training operations this past year. It is hoped that at a future date a more complete account of the many exciting, exasperating, funny, and tender episodes can be related. Such a document is needed as a guide to the many other social scientists who are only now beginning to carry out research in the "real world."

As was pointed out in the beginning, this was a pilot year. One result, however, is certain. Individual child testing is not the prerogative of the educational elite; the local women in our study can learn--and have learned--a wide variety of demanding tasks. They have managed well under many difficult local situations. We are tapping a large pool of unused resources. Given a better training program profiting from our experiences this year, these women will do even better. An even more important result, however, is the strengthening of our belief that traditional training models must be questioned, that effective training must involve mutual learning and cross-socialization.

Operating Testing Centers

Virginia Shipman and William Ward

Testing centers are located in churches or community recreation facilities in or near the districts where the children live. Each center provides, at a minimum, six individual testing rooms (or spaces which can be partitioned off from larger areas) and a larger play and rest area; most also include some kitchen facilities.

Each testing unit is staffed by nine persons, whose roles are:

a. Testers: Six testers are assigned to each unit. Each tester was trained to administer the "Day 1" battery, and one of the three subsequent test batteries.

b. Supervisor: The center supervisor is responsible for all aspects of the day-to-day administration of the testing center. Her major duties include scheduling of children and mothers for testing, after initial appointments have been made by the local coordinator; assignment of children to testers; checking that answer sheets are properly completed and collected for transmittal to Princeton; and arranging with the local coordinator all matters of needed supplies and equipment replacements. When necessary, she serves as a substitute tester. When time permits, she also monitors testing in her center, relaying to the Princeton measurement staff or to regional office monitors any needs for consultation on measurement problems.

c. Play area supervisor: The play area supervisor is responsible for children during periods in which they are not being tested. She provides snacks for children and mothers during testing breaks; comforts shy and frightened children; and provides general assistance to the center supervisor and testers in maintaining a smooth testing operation. Sometimes, she also has to baby-sit with younger brothers and sisters.

d. Driver: The driver is responsible for transportation of mothers and children to and from the testing center, and for providing such other assistance to the center and to the local coordinator's office as is needed.

After an initial training period of approximately two weeks, the centers were opened so that the trainees could become accustomed to working in the actual study setting. During the next phase of training (from 2 to 4 weeks), members of the ETS Princeton office staff were responsible for the administration of the centers, while the trainees directed their major efforts to administration of the test batteries to one another and to practice subjects (no real subjects were in the centers during this period). Trainees were also given some opportunity to practice supervisory and play area roles. After evaluation of their performance, the trainees were assigned to tester, supervisor, and play area positions. Testing of practice subjects continued for another one to two weeks, while the training staff remained available to complete training and to provide assistance as needed.

The first longitudinal sample children were tested seven to eight weeks after the beginning of training. During actual testing, the center staffs work independently except for periodic visits by a monitoring staff. This staff consists, in three testing sites, of personnel from ETS regional offices, and, in the fourth (Trenton), of several members of the original training group. The monitor is responsible for providing general advice on both testing and administrative problems to the center staff and to the local coordinator, and for observing testing to determine that standard procedures are indeed being followed.

In planning for the testing operation, it was expected that testing centers would generally run four days each week. Each Monday, 24 children and their mothers would be tested, half in the morning and half in the afternoon, on the "Day 1" battery. The children would return in groups of 12 for the next three mornings or afternoons to complete the testing battery. Each tester would be assigned two children in each half-day session, and would alternate in testing these children throughout the session; thus, a child would be engaged in testing for only about half of a 3-hour period, and could rest and play for the remainder of this time. For children who missed a testing session, who took unusually long to complete a battery, or who would not cooperate in testing on a given day, it was expected that make-up tests would be administered on Fridays.

It was planned, however, for the child to spend as many days in the testing center as were needed to complete all the tests; particularly difficult children might have to return for several weeks before they finished.

In practice, there were several changes in these plans. During training, it was discovered that most children could complete an entire battery in 1 1/2 to 2 hours, with only short rest periods; and that working with one child at a time was more convenient for both testers and those who had to schedule and transport the children. The scheduling procedure was therefore altered so that groups of no more than six children arrived at the testing center at a time. Scheduling details differed across sites and centers, but an "ideal" was established which required groups of children to begin testing at 9 a.m., at 10:30, at 1 p.m., and at 2:30. Saturday and early evening appointments were also arranged to accomodate the working mother. Second, it was found that there were continual needs for make-up testing, and centers routinely operated for 5 days each week. Finally, instead of assigning children to testers randomly for the three test batteries following the first day's testing, children were routinely tested on the second day by the same tester who had worked with them on Day 1. This procedure provided some continuity between Day 1, when the mother interacted with the child during most of testing, and the remaining days, when she was not present.

These structural and operational procedures appear to have been satisfactory ones; no major shift in center organization for future testing operations is needed. There have been, however, three persistent problems in the testing operations. First, the supply of children to testing centers has been less than desired. Most centers do appear capable of testing 24 children per week, but have worked with many fewer than this. Many factors have been responsible, including in some cases a failure to obtain names of children interviewed at an adequate rate, and in several sites a reluctance of parents to allow their children to participate. Steps taken to combat this problem have included an increase in publicity for the project, and on several occasions having testers make personal contact with the families.

A second problem concerns the need for more consistent monitoring of the testing operation. Regional office personnel, supplemented by Princeton staff, have devoted 1 to 2 man-days per week to each testing site. However, this arrangement has led to inconveniences for all concerned and, in some cases, to ambiguity in instructions given to local centers. Emphasis on the role of local professionals, rather than out-of-town personnel, will be attempted in the future to eliminate these difficulties.

Finally, there has been greater turnover in the testing staff than was anticipated, requiring that members of the Princeton office training staff continue to make field visits to train replacement testers. Again, there are several reasons for this turnover. The tester position is a temporary one, so some of the most capable testers have resigned for more permanent employment. More important, perhaps, has been the effect of the inadequate supply of children for testing. Centers have had to continue testing throughout the summer in an attempt to test the desired number of children, despite initial predictions that all testing would be completed around July 1; thus, some individuals with other commitments for the summer have been lost. Various personal emergencies have been responsible for some additional loss of personnel. It is a hard fact of life for our testers that they lack resources and, in consequence, such emergencies have arisen more frequently and have been more incapacitating than might otherwise have been the case.

It is anticipated that these problems can be alleviated in future operations. The children to be tested will be ones who have already been involved in the project, or ones who become classmates of our subjects, so that obtaining their names and securing parental cooperation should present fewer problems. The presence of a local professional serving in both training and monitoring roles should provide more thorough quality control and more immediate response to operational difficulties that do arise. With these improvements, the testing center operation should continue to provide an adequate way of combining community participation in the study with the gathering of scientifically useful information.

Medical Examinations

Joseph Boyd

The type of medical information collected is described in Chapter 3. As is true for many other aspects of the study, there are regional variations in the procedures for conducting the medical examinations.

In St. Louis, for example, where a Neighborhood Health Center has contracted to do the examinations, the letter of agreement specified:

> Grace Hill Settlement House (GHSH) will assume responsibility for providing health examinations to the children selected by the Educational Testing Service. The exams will be provided over a thirteen-week period....at the rate of 25-37 exams per week. GHSH will be responsible for the provision of the services described; specifically, GHSH will provide the management functions, and the Neighborhood Health Center (NHC), 2500 Hadley, will conduct the actual examinations.
>
> The Neighborhood Health Center is a non-profit institution which has been providing health services to North St. Louis for over 60 years....
>
> Each child will be given a thorough examination by the Pediatrician based upon the outline on pages 3 and 4 of the Child Health Record. These two pages will be completed for each child examined. Furthermore, the parent of each child will be advised of any problems or recommended treatment by the NHC staff and the NHC offered as <u>one</u> resource for such treatment. The OEO Health Specialists will also be notified in cases where treatment is advised so that they may work with the parent in securing such treatment. Basic immunizations will be provided to any children needing them.
>
> It is the policy of the NHC that children must be accompanied by a parent or guardian. This policy will also apply to the examinations for the ETS except in special cases....
>
> In addition, the following services will be provided by the Block Captain System of Grace Hill Settlement House:

> GHSH will assume responsibility for setting up health examination appointments and for getting the selected children and parents to the NHC as scheduled. This service will utilize, as required, such techniques as letters, home visits, follow-up visits, Block Captain reminders, transportation, and special appointments in unusual circumstances.*

In Portland, one physician has contracted to do the medical examinations. The ETS local coordinator will schedule children for examinations as mutually agreeable to parents and the physician. When required, ETS will transport the parent and child to the doctor's office, which is centrally located in the study area. One or more colleagues may assist the contracting physician to conduct the examinations. In a small number of cases, the parents prefer that the child be examined by another physician whom the family has consulted previously. In these cases the contracting physician coordinates the data collection with the other physician.

The arrangements in Trenton parallel those in Portland--one physician is responsible for getting the work done.

Distances in Lee County have made it impossible to concentrate the medical examinations in one location. Service in the eastern quarter of the county, the Loachapoka area, was provided by a physician in Tuskegee, in adjoining Macon county. Another physician in Auburn examined the children from that city and the unincorporated area south of Auburn. Still a third physician is conducting the physical examinations of the children in the southeastern portion of the county. At the suggestion of the physician first engaged, the Tine test for tuberculosis was added to the routine in Lee County because of the relatively high incidence of tuberculosis in the area. Physicians in the other three study sites felt the test was not needed in their areas.

The medical examinations have been purposely scheduled after the testing of the child has been completed so that the possibly unpleasant taking of the blood sample would not be associated with the testing sessions. Ideally, children who were tested one week would receive physical examinations the following week. Scheduling problems make

---

*Letter from Richard Gram, Director of Neighborhood Development, Grace Hill House, May 6, 1969.

this ideal impossible to achieve, but children have been examined in approximately the same sequence as they have been tested.

5. DATA STORAGE AND RETRIEVAL SYSTEM

Jesse Webb

DATA STORAGE AND RETRIEVAL SYSTEM

Jesse Webb

Figure 1 shows the Data Storage, Retrieval and Analysis system design as it now stands. The system is expected to be composed of eight computer programs. The function of these programs and their present status are discussed briefly below.

1. Student Identification Number Generator. This program is a one-time use program for generating unique identification numbers for subjects in the study. The identification numbers are composed of seven digits of which the seventh is a self-checking digit, computed from the first six, which serves as a means for checking the correctness of identification numbers on data from the field.

This program has been completed and was used to produce a set of number lists that were distributed to the field staffs. These lists are used to assign an identification number to each child and to record the child's name and address. The local coordinator keeps one copy for his own records and returns the remaining completed copies to ETS so that Data Processing can create a name and address file. To date, all the study subjects have not been selected so that the lists have not yet been received.

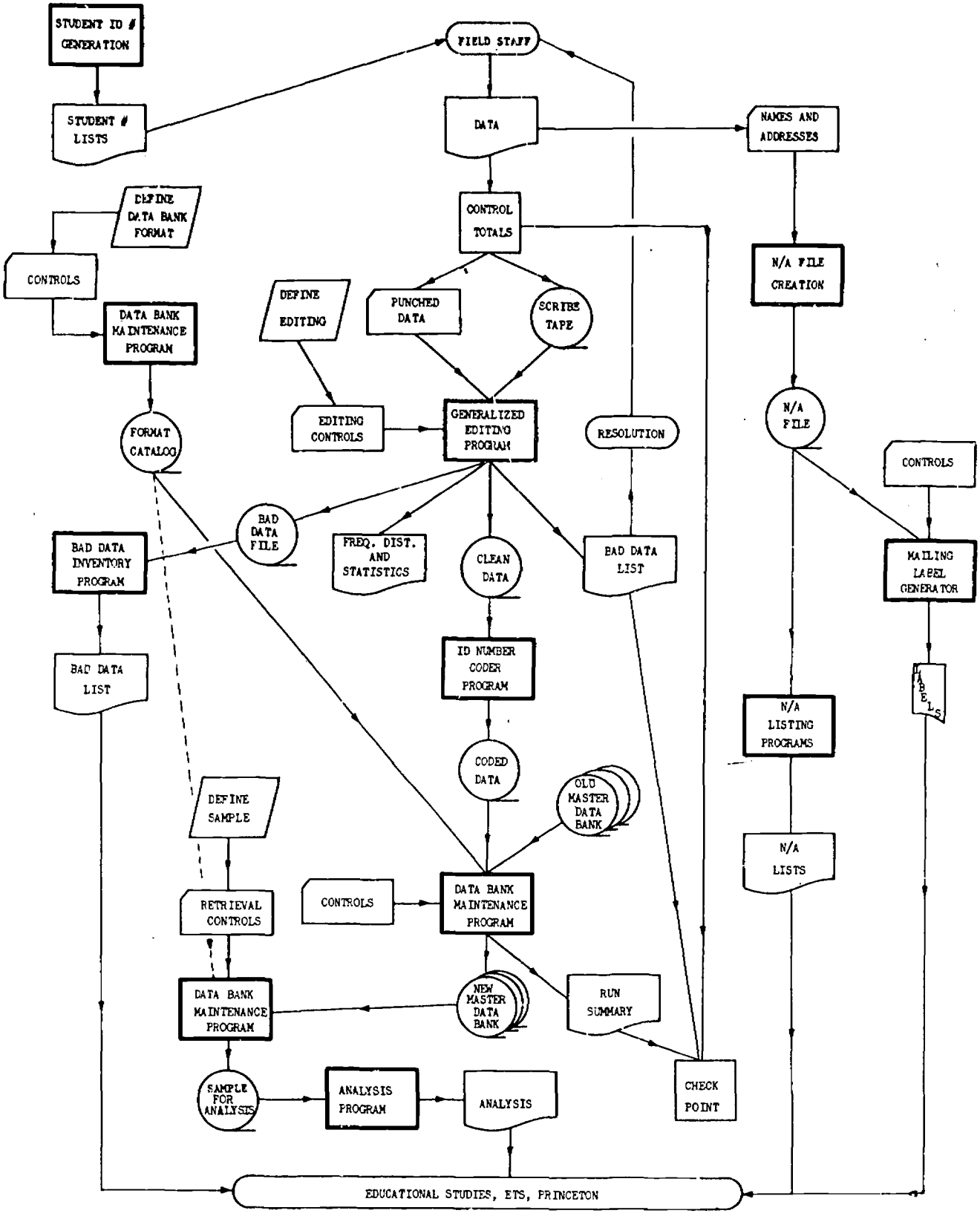
2. Name and Address File Creation Program. This program has been written and tested and is ready for use. It will create and maintain the name and address file using the number lists from Program #1 as source input documents.

3. Name and Address Listing Program. This program can be used to produce any one of several possible name and address lists. Lists of study subjects by school or by city can be produced. Lists of field staff can be produced. The program has been written and tested and is ready for use when all number lists have been returned from the field and the name and address file has been created.

4. Mailing Label Generator. This a program for producing gummed labels printed with names and addresses from the Name and Address file for mailing purposes. Labels can be produced for study subjects, field staff, study schools, or any combination thereof.

Figure 1

DATA STORAGE, RETRIEVAL, AND ANALYSIS SYSTEM



The program has been coded and tested and can be used as soon as the name and address file has been created.

5. Generalized Editing Program. No data will be deposited in the data bank until it has passed through this editing program. The program will check the validity of various data fields and the correctness of the subject identification number. The program can also produce a limited variety of statistics such as frequency distributions, means, standard deviations, and counts.

The program has been designed and coded and is now being tested. The program is generalized, and the type and degree of editing must be defined for each new type of data that is processed. Data that fails to meet all the editing requirements will be separated from "good" data and held on a separate file. Only clean data will be allowed to pass through to the data bank.

6. Bad Data Inventory. This program is not yet written. It will produce listings of data that is being held pending resolution, as a means of ensuring that initially rejected data finds its way to the data bank after resolution of errors.

7. Student Identification Coder Program. Since the security of the data stored in the data bank is of utmost importance, all data destined for the data bank will pass through the Coder program at which time new identification numbers will be coded from those originally assigned. The student numbers in the bank will, therefore, never be the same as those used in the field, thus making it impossible for an individual researcher to associate data output with any study subject. The privacy of the subjects is thereby protected. This program will also sort the input data into the order required for entry to the bank.

This program has not yet been written, but it will be a relatively simple program, and will probably be a "job step" of the maintenance program, which is described next.

8. Data Bank Maintenance Program. This program is the core of the system. It is a generalized maintenance program which allows for depositing data in the data bank and retrieving data for analysis. The file structure definition for the bank is stored in a format catalog which the maintenance program enters for updating the bank or for retrieving samples

for analysis. The format catalog itself is maintained and updated by this program. Before each new data bank file can be serviced by this program, it is necessary to define the file structure for inclusion in the format catalog. This definition will be handled by the programmer responsible for the maintenance program in conjunction with the researchers for whom the data is to be stored.

The program has been written and is now being tested. If it is not completely operational by the time data begins coming in, the data will be temporarily held on tape until the data bank is created.

The design of the data bank is such that three generations of it will exist at all times in a cyclic fashion. That is to say, when the fourth generation is created, it will be created on the first generation tapes. This will keep the bank self-contained and will prevent obsolete generations of the bank from making their way back to the general tape library.

Information Flow

The main stream of information flow shown in Figure 1 begins with the collection of data in the study sites. When the data is received at ETS, control counts are made so that at the check point, shown at the bottom of the diagram, it can be determined whether or not all received data has passed through the system. Sometime before the first data of a given type is received, the researchers responsible for the data should arrange to meet with the data bank programmer to define the file formats for that data. When this has been done, the programmer will schedule a maintenance run to insert the new formats into the data bank format catalog. The maintenance program will then be able to process the data as it is received. At the same time as the file formats are being defined, the data editing parameters should be defined so that all data can be checked before entry into the data bank. The researchers responsible for the data should arrange to define editing parameters with the programmer.

After the control totals have been made, the data is either SCRIBED,* or punched into cards for entry into the system. Again, before the first

*SCRIBE is the name given to the scoring machine used by Educational Testing Service.

data is received, the responsible researchers should arrange to meet with the data bank programmer to decide on punched card formats. It is not recommended that they attempt to define card layouts themselves. This is better left to a Data Processing representative.

The punched cards or SCRIBE tape plus the editing controls, resulting from the editing defining phase, are input to the Generalized Editing program. The program makes all the required editing checks and separates the "clean" data from the "unclean" data. The "unclean" data is held on a temporary data file while resolution is being attempted. This temporary data file will probably have new data added to it on each run, and should constantly have data removed from it as the data errors are resolved. The data file can be interrogated by means of the Bad Data Inventory program. The Generalized Editing program can also produce limited statistics on the input data. The statistics desired must be defined at the same time as the editing parameters are defined.

Clean data from the editing phase is then passed through the coding program where identification numbers are coded for residence in the data bank. The coded data along with maintenance controls and the master data bank are finally processed with the Generalized File Maintenance program. The updated data bank is created at this time, and a summary of the run produced. The run summary, the listing of bad data from the editing phase, and the control totals generated when the data was received, are used to check that all data was processed.

The Generalized File Maintenance program is also used to create sample data files for analysis purposes. The analysis programs will be specially written as need arises.

Data Received to Date

The only data that has been received so far is the pilot data on the Classroom Observation Ratings (see Appendix A). The multiple observations on the subjects are being placed on magnetic tape and some simple analysis programs are being written to produce distributions, variances, correlations, etc. The results of these analyses will determine whether or not the instrument will require modification before it is used for collecting real data. The pilot data will probably not be stored in the data bank.

6. PLANS FOR DATA ANALYSIS

Albert Beaton

PLANS FOR DATA ANALYSIS

Albert Beaton

This longitudinal study will collect large masses of data on individual children, their school environments, families, communities, and many other things that might affect their attitudes and behavior in their environment, according to the design in Table 1. The major task in data analysis is to reduce the information for both the purpose of identification and the study of psychologically important factors. The organization of this study suggests the data organization. Basically, the data is organized in:

1. Time Periods. Time periods as defined here are the years of the study, each year running from September to August. Occasionally, we will need to differentiate testing sessions within a year.

2. Domains. Within each time period we have measured domains of children's activities and environment which are theoretically distinct, though usually highly correlated. These domains are cognitive/perceptual, personal/social, physical health and nutrition, family, classroom, teacher, school, and community. Each domain will be studied separately before interdomain analyses are performed.

3. Instruments. The data is collected by administering a series of tests, physical examinations, or some such, each of which is a logical entity called an instrument. One or more instruments cover the important aspects of a domain.

4. Variables. Each instrument generates one or more variables which are the units studied. We use "variable" very generally here to describe, for example, continuous and discrete variables, and variables with a normal or other distribution.

5. Items. The item is the basic unit of information. Items may be discrete or continuous, depending on the instrument.

Table 1

OVERALL STUDY PLAN

Sample	Sp. 69	69-70	70-71	71-72	72-73	73-74	74-75	Comments
LD - Longitudinal disadvantaged LDE - LD who leave site LDI - LD who move into site	Pre	HS	K	Gr. 1	Gr. 2	Gr. 3		In 69-70, 70-71, some children will not attend Head Start, Kindergarten
C_{HS} Cross-sectional (same feeding areas, elementary schools)						HS		73-74 cross-sectional samples optional but desirable to assess change. HS may bear a different name in 73-74.
C_K "		K				K		
C_1 "		Gr. 1				Gr. 1		
C_2 "		Gr. 2				Gr. 2		
C_3 "		Gr. 3						
SC - Special comparison groups (different but comparable feeding areas, elementary schools)			HS				Gr. 3	74-75 Gr. 3 sample a re-testing of 70-71 HS sample as a comparison group to assess impact of intervening testing.

The aim of this section is to discuss the analytical techniques to be used in each section of this hierarchy. We will work in reverse order, starting at the basic item level and working up to the analyses of time periods. We note that the organizational data is not rigid, for it will often be useful to examine items against variables in other domains. We will try to maintain as much flexibility as possible, while assuring that each level is fully and carefully covered.

Data reduction and analyses entail a risk of losing critical information; this is, of course, always a problem to research people as well as data analysts. Thus, the first task is to examine items for reasonableness, and combine them into scales or variables by instruments. We feel strongly that this process should be under the personal control and consideration of the psychologist responsible for the instrument. Each instrument has been recommended by a psychologist who has been responsible for overseeing the development or adaptation of the measure, the training of testers, and general quality control. We wish to take advantage of his expertise and interest in the area for judgments as to the validity and reasonableness of responses as well as for the basic control of the internal analysis of each instrument.

However, certain common elements must underlie the analysis of each instrument. In all cases where possible, reliability estimates, relationships among scales within an instrument, and basic descriptive statistics will be generated to record the properties of the variables. In many cases, we will want to look immediately at breakdowns by community, sex, and race.

Items

Each instrument is constructed of various items which must be aggregated to form more global measures. At these early years there are a large number of different types of items since the child is not yet old enough to respond to standard pencil-and-paper multiple-choice test items. Many items are actually multiple-choice items which are given orally to

the children and whose response is recorded by a tester. There are also many others which are measured on continuous scales; e.g., height, time in responding, or angle of rotation of a cross. Each item must be examined individually before aggregation.

The analysis of a single item is relatively straightforward. For multiple-choice-type items, we will compute the proportion responding to each choice. For continuously measured items, a frequency distribution will be computed as well as means and standard deviations. We are, in most cases, interested in whether or not items vary, although in this study a nonvarying item at one time period might vary in another and is not necessarily statistically useless. We will also examine the various distracters for each item for unforeseen biases to one or another subgroup of the children. We will have the facility to examine the distribution of item responses by community, sex, race, and so forth, as needed.

Variables (Scaling)

The second step is the aggregation of test items into scales or variables which are then used in inter-instrument analyses. In some cases the scaling will be a priori; e.g., the score on a measure is a weighted sum of the number right or a single item is considered a scale by itself. However, in both cases we will usually wish to check the intercorrelations of a priori variables with the items which compose it, with other items in the battery, and with the one item scales.

The basic inter-item analysis technique will be correlational and factor analytic. First, a correlation matrix of items and derived scores will be computed for each instrument. This matrix will contain the intercorrelation among items as well as between items and a priori scores. Factor analyses of the items will be computed to find systems of organizing items which form more or less cohesive scales. A priori scales will be treated as extension factors so that the correlations of factors and these scores can be compared. Factor analysis will be computed with or without rotation as seems appropriate for individual instruments.

For each scale, some descriptive information will be computed:

1. Frequency distribution;
2. Mean and standard deviation;
3. Reliability coefficient (where possible);
4. Analysis of variance by community, sex, race, etc.

Where there is an appropriate external criterion, we will use the criterion scaling method (Beaton, 1969). This procedure maximizes the validity of an item with respect to an external variable. Such scaling avoids problems of nonlinearity and nonresponse and bridges the gap between scaling and analysis of variance methodology.

Instrument

An instrument is really an administrative grouping of scales and, as such, will not usually result in a single overall score. However, since these tests may be used by other research people in different settings, the information usually associated with a published test will be prepared for dissemination, including distributions, descriptive statistics, intercorrelations, and reliability coefficients.

Domain

Each of the instruments was carefully selected to generate information about children's behavior in various domains of their activities. The analysis of a domain is the joint endeavor of the several psychologists responsible for its instruments.

Correlational, discriminant and factor analysis will be the major tools used in exploring the several instruments relevant to a domain. The purpose is to understand the interrelationships among the variables and, if possible, reduce the number of scales through factor scores. The variables will be analyzed for coverage of relevant domain, redundancy, and simple underlying structure.

In addition, other basic analyses such as multivariate analysis of mean vectors by various groupings and discriminant analyses which maximize or minimize inter-group differences will be computed so as to assess the various differences of the children in the sample.

Time Periods

The most general level of analysis is the longitudinal study of variables over the years of the study. The primary purpose is, of course, to study the growth of children and to identify critical conditions of maturation. To contribute to the area of child education, broadly conceived, we hope to isolate factors in a child's world which might be manipulated to understand and enhance his growth. We would like construct a reasonably accurate model for theoretical manipulation, a simulation perhaps, of a child's development and for study of theoretically viable alternative educational methods.

The idea of changing patterns and amounts of learning is troublesome for it leads us to the age-old dilemma of attempting to find cause and effect through empirical methods. In the final analysis, such cannot be done. This study is weakened somewhat because it is not an experimental design in the proper sense; that is, the subjects were not randomly assigned to experimental treatments nor are the treatments defined tightly enough for the isolation of pure interactive effects. In addition, we are seldom able to reach the variables of actual concern in the behavioral sciences for we must measure variables like performance on intellectual acts, not capacity for action; for we must use family background, not family influence, and so forth; in a way, we are using the shadows of causal variables to predict the shadows of effects.

However, we can and must try to approach reasonable proxies for cause and effect because our end is the proposing of educational treatments that will promote the attainment of personal and social goals for children. The general framework over time periods requires psychologists to postulate causal models explaining the relationships among variables over time. The models are fit with the data for support or for

rejection. If a model has strong theoretical basis and reasonable fit, then it may be used with caution, of course, for examination of the consequences of change in the variables in the model.

The modification of variables leads to simulation. That is, given a reasonable theory, we may consider what would happen to the interactions among the variables by changing a variable by a given amount, or by changing a set of variables, a treatment say, in a certain manner, then following these changes to their logical conclusion. The power of such a model is obvious, for it would lead to the simulation of educational experiments and perhaps the pinpointing of treatments to be tested and encouraged in the real world.

There is no intention to adopt this procedure uncritically, for such simulation is only as good as the model and the model, in turn, is only as strong as the theory and data from which it was generated. Although this study is not a proper experimental design it does conform to the Campbell and Stanley (1963) Quasi-Experimental Design 14 which is relatively strong, though it leaves us unable to be sure that some uncontrolled variable does not overwhelm the educational treatment in a "causal" sense. We have guarded against unmeasured variables by the vast number of variables collected for each child. A second problem is the ever-changing sample; each year children will leave our target areas as well as immigrate. The model as described in the December 1968 report (Educational Testing Service, 1968) permits us to use all children available for any two consecutive years in estimation of the relationship between the years. Another problem may be the errors generated from less than perfect measuring instruments; this will be monitored and carefully controlled through the internal analyses.

Operationally, the procedure is as follows: First, the team of researchers encode one or more theories to represent the transition of a set of variables at one point in time to the set at the next point in time. The theory of transition is represented by a matrix. The variables in the two sets are examined in all possible pairs and their relationship is postulated. Ordinarily, only the existence of an important relation or no relation will be codified, although if one wanted to postulate a "strong" theory, an actual degree of relationship could be assigned.

Although relationships will ordinarily be postulated from one time period to the next, effects within the same time period are permissible.

The next step is to fit the model. The procedure is essentially a least squares procedure which leans heavily on the work in path analysis by Wright (1960) and Duncan and Blau (1967), causal analysis by Simon (1957) and Blalock (1964), and the work done in structural analysis by various economists. The goodness of fit for a model is examined by computing the residual of the scores at the later time from the predicted values of the model, and examining their correlations and covariances. A model's fit can be improved if necessary by modification of the model. The steps of postulating a model, fitting, and examining residuals can be repeated as often as necessary until an acceptable model is found or, perhaps, until several viable models are identified.

A model is fit between each pair of adjacent time periods. The procedure is recursive; thus third year results may be expressed as a function of first year and variables that come into relevance during the second year; the fourth year is also expressed as a function of the first with orthogonal variables entering in the second and third years, etc. Changes in relationship over time can be carefully monitored.

The simulation phase is contingent upon the goodness of fit of the model and its correspondence with the real world. If the model truly represents the growth of children, then changes of the input to the model should generate changes in the output similar to the effect of such changes on the behavior of children. We come, of course, to the problem that certain variables cannot be changed, thus we must be careful to manipulate only the malleable variables within control of those responsible for a child's development. The model might suggest the effect of changes in educational treatment; that is, the effect of giving educational treatment to those children who do not receive it. The model might also suggest new treatments which might be tried to answer special needs of less well adapting children. Making broad assumptions, the model might also be used to explore the possible effects of major changes in the child's living conditions or environment. Without exaggerating the

possible import of such a simulation model, we know of no similar attempt to construct such a broad-base model of change in the child.

References

Beaton, A. E. Scaling Criterion of Questionnaire Items. Socio-Economic Planning Science, Vol. 2, pp. 355-362, Pergamon Press, Great Britain, 1969.

Blalock, H. M. Jr. Causal inferences in nonexperimental research, Chapel Hill, University of North Carolina Press, 1964.

Campbell, D. T. & Stanley, J. C. Experimental and quasi-experimental design for research. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 171-246.

Duncan, O. D. & Blau, P. M. The American occupational structure. New York, John Wiley & Sons, Inc., 1967.

Educational Testing Service, PR-68-4. Disadvantaged children and their first school experiences: Theoretical considerations and measurement strategies. Report in 2 volumes, December 1968, Contract OEO 4206 and Grant OEO C6-8256, Office of Economic Opportunity.

Simon, H. A. Spurious correlation: A causal interpretation. Models of Man. New York, John Wiley & Sons, Inc., 1957.

Wright, S. Path coefficients and path regressions: Alternative or complimentary concepts? Biometrics, 1960, 16, 189-202.

7. OPERATIONS PLANNED FOR THE SECOND YEAR OF THE STUDY (1969-70)

Scarvia Anderson, Samuel Ball

OPERATIONS PLANNED FOR THE SECOND YEAR OF THE STUDY
1969-70

Scarvia Anderson and Samuel Ball

Whether it is said

politically

> A state without the means of some change is without the means of its conservation.
>
> (Burke)

poetically

> Since 'tis Nature's law to change,
> Constancy alone is strange.
>
> (Wilmot)
>
> Fate, Time, Occasion, Chance, and Change? To these
> All things are subject...
>
> (Shelley)

or psychologically

> The world of the ghetto...is not a textbook world... .The study will have to make constant adaptations to this world...
>
> (Anderson, Beaton, Emmerich, and Messick)

it seems prudent to begin this review of study plans for 1969-70 with a reminder that some changes are inevitable. Eventually, the study may have to be judged more in terms of the scientific and practical merits of its adaptations than in terms of its strict adherence to preconceived procedures. Two months ago, we could (and did) write the following:

The overall study design, as described in the December report (Educational Testing Service, 1968b), and summarized in Table 1 (with revised estimates of N's), calls for two major lines of investigation in 1969-70:

1. Follow-up of the longitudinal sample that was selected and measured during the first year of the study. These children will be between four and five years of age. As before, data will be obtained on

Table 1

OVERALL STUDY DESIGN (ESTIMATED N'S)

	Summer 1969	1969-70	1970-71	1971-72	1972-73	1973-74	1974-75
LD - Longitudinal Disadvantaged	Pre (1650)	HS (575) not HS (1075)	K (1240) not K (185)	Gr. 1 (1230)	Gr. 2 (1060)	Gr. 3 (915)	
LDE - LD who leave site		(225)	(195)	(170)	(145)		
LDI - LD who move into site		(225)	(195)	(170)	(145)		
C_{HS}						HS	
C_K Cross-sectional		K[a]				K	
C_1 (Same feeding		Gr. 1				Gr. 1	
C_2 areas, elementary		Gr. 2				Gr. 2	
C_3 schools)[b]		Gr. 3				Gr. 3	
SC - Special comparison group (different but comparable feeding areas, elementary schools)			HS (480)			Gr. 3 (250)	

[a] HS in Lee County.

[b] All classes in target elementary schools in grades indicated and (in 1973-74) HS centers in tne same districts.

the children's cognitive, social, emotional, and physical development. In addition, their families, communities, and the Head Start classes and centers they attend will be studied. Thus, by the end of the second year of the study, there should be sufficient data for making a provisional evaluation of the effects of Head Start over a one-year period, for comparing the individual and familial characteristics of those children who go to Head Start classes and those who do not, for describing in some detail the Head Start programs themselves, and for specifying key interactions among home, community, program, and child variables.

2. Study of appropriate cross-sectional groups (K through grade 3). Children will be tested, and teachers, programs, and administrators will be surveyed. The information obtained on the cross-sectional groups will provide answers to questions about the general characteristics of disadvantaged children in the study sites and about the educational programs they presently attend. This data will be valuable in its own right because there is a dearth of objective knowledge in this area. However, as discussed in the Interim Report (Educational Testing Service, 1968a, pp. 5-6) and the December 1968 report (Educational Testing Service, 1968b, Chapter M), the cross-sectional measurements will also allow comparisons with the main longitudinal sample. Of equal importance, the cross-sectional data on preschool and early school programs will serve as a baseline for study of the possible effects on school programs, themselves, of such educational-social interventions as Head Start.

However, before we can measure the longitudinal subjects a second time--or give much attention to the cross-sectional study--we must complete the assessments planned for the first year. Earlier sections of this report have alluded to a number of explanations for the present state of being considerably behind schedule in St. Louis and Trenton: delays in final choice of study sites, difficulties of training testers, problems in local coordination, community conditions leading to wariness toward enumeration and interview, and so on. More important than a rehash of these explanations is a statement of what we propose to do about the situation.

Table 2 gives estimates of the number of children in the study districts within sites who probably will go to Head Start. It also presents

Table 2

THE SUBJECT SITUATION

	Lee County	Portland	Trenton	St. Louis
No. of elementary schools involved[a]	8	7	6	7
Est. no. of 3 1/2-year-olds (based on current Gr. 1 enrollment)	530	525	590	680
Est. no. who will attend Head Start	300	260	150	300
Est. no. of possible study participants[b]	460	450	425	450
No. interviewed by 8/8	510	498	285	269
No. tested by 8/29	417	454	226	244
Est. no. 10/1/69	460	450	380	350[c]
Est. no. in HS, fall '69	---	225	125	225

[a]Lee County - Beauregard, Beulah, Boykin, Cary Woods, Dean Road Elem., Loachapoka Junior High, Smith Station High, Smith Station Elem.

Portland - Boise, Eliott, Humboldt, Irvington, King, Sabin, Vernon

Trenton - Grant, Jefferson, Jr. 5, Jr. 2, Jr. 3, Robbins

St. Louis - Ames, Blair, Clay, Henry, Jackson, Jefferson, Webster

[b]Some children will not be "findable"; some parents will not agree to participation; these conditions are especially serious in the highly urban areas.

[c]We hope that this is a conservative estimate.

information on the number of children tested by August 29. The major consideration here is to test all (or, realistically, nearly all) children who will attend Head Start before they are actually exposed to the Head Start program. Previous studies have indicated that just a few weeks or even days of Head Start may be related to a difference in test performance. Thus, the following order of operations is planned:

1. Secure Head Start advance registration lists for all centers within the districts of the study. Check these lists against the names of children already tested and schedule for testing as quickly as possible all other children destined for Head Start. If necessary, test the children first; interview their mothers later (see phase 3).

2. At the opening of Head Start, provide each Head Start teacher (through the Center Director) with a list of all children who have been tested in the district. Ask her to send any untested children directly to the nearest testing center before they begin the Head Start routine. Of course, appropriate provisions will be made for securing parental permission and for transporting children. The influence of the national Head Start Research Office is very important for this phase.

3. While phases 1 and 2 are going on, continue with parent interviews. Since most of the Head Start decisions have been made by this time, the possible influence of the interview on parental decisions about Head Start can probably be discounted. Schedule interviews to precede testing when possible; schedule interviews of mothers of Head Start children before interviews of mothers of non-Head Start children when possible; complete interviews in September if necessary.

4. Complete the testing of any "left over" non-Head Start children in September after the last Head Start children have been tested.

We should stress again that these extraordinary efforts are most relevant to the Trenton and St. Louis locations, though we shall want to get as complete a sample as possible in Portland and Lee County.

The activities scheduled for the remainder of 1969-70 can be summarized under eleven main headings: Test Development and try-out, Training, Interviewing mothers, Observations in classes, Testing children, Collecting information from testers, Collecting information from educational personnel, Collecting community information, Keeping track of

Table 3

MEASURES* TO BE USED IN TESTING CENTERS WITH LONGITUDINAL SUBJECTS (4-5 YEARS OLD)

Spring 1970

	Min.
Reasoning and Analytic Styles	
Hess & Shipman Eight Block Sorting Task	35
Preschool Embedded Figures Test	15
Attention, Learning, Memory	
Fixation Time	20
Relevant Redundant Cue Concept Task I	20
Stanford Memory Test	10
Attitudes, Interests	
Brown IDS Self-Concept Referents Test	15
Social Schemata	15
Controlling Mechanisms	
Matching Familiar Figures Test	15
Mischel Technique	5
Motor Inhibition Test	5
Risk Taking, Task 2	5
Sigel Conceptual Style Sorting Task	20
Creativity	
Naming Category Instances	15
General Knowledge	
Cooperative Preschool Inventory (Caldwell)	20
TAMA General Knowledge II	10
Perception	
Children's Auditory Discrimination Inventory	10
Johns Hopkins Perceptual Test	15
Seguin Form Board	5
Physical	
Vigor 1 & 2	5
Piagetian	
ETS Enumeration II	5
ETS Spatial Egocentrism Task	10
Sex Role Constancy Task	10
Spontaneous Numerical Correspondence	5
Social Motives	
Gumpgookies	25
Open Field Test	(10)
Verbal	
ETS Matched Pictures Language Comprehension Task II	10
ETS Story Sequence Task, Part I	10
ETS Story Sequence Task, Part II	10
ITPA: Auditory-Vocal Automatic Subtest	10
Massad Mimicry Test	10
Peabody Picture Vocabulary Test and ETS Adaptation	20

*Many measures cut across more than one general area; a measure is listed under the single general area that is most relevant.

children who move, Data preparation and analysis, and Community relations. In addition, explorations will be conducted into possible study extensions.

Test development and try-out. Work will proceed during the fall on further development and piloting of tests to be used in spring 1970. Examples of such work include the following: (a) Modifying the Peabody Picture Vocabulary Test to make the items more meaningful for the population under study--in terms of racial characteristics and activities pictured; equating scores on the modified form to scores on the standardized form, to enable direct comparison between scores obtained in the first and second years of this study and between scores obtained in this study and other studies. (b) Pretesting the questionnaires designed for Head Start center supervisors; these questionnaires have been adapted from materials used in earlier Head Start facilities inventories and in questionnaires designed for elementary school principals. (c) Trying out simple and relatively inexpensive means of getting clearer recordings of children's speech. (d) Developing and trying out questionnaire items suitable for the mother of a 4 1/2-year-old, as opposed to the 3 1/2-year-old of the first year; in the family interviews some items will remain fixed across the years of the study, some will be introduced only periodically, and others will be specific to the year of the study and the age of the study child.

A list of the measures scheduled for use with the 4 1/2-year-olds in spring 1970 is given in Table 3. These test materials must go into production in November, so as to be ready for the training cycle beginning in January.

In addition, throughout the year, work will proceed on instrumentation required for the third and later years of the study. Examples include measures of general knowledge, interests, ideational fluency, conception of natural events, and writing abilities.

As a crucial part of the development activities, a special advisory panel of educators and citizens from the four study communities has been appointed and will be asked periodically to review proposed measures and the items in them. The focus of the reviews will be on the appropriateness of the content to the experiences of disadvantaged children,

many of them black. All the items used in the first-year battery were pretested with subjects judged to be comparable to those of the study and many of the instruments were reviewed by panels especially concerned with the reactions of disadvantaged children. However, the formation of this advisory committee represents an additional step to ensure that the measures do not appear to present any particular problems for the subjects in Lee County, Portland, Trenton, and St. Louis. The first meeting of the panel was held on August 26-28, 1969.

Training. In the spring of 1969, training was conducted by ETS Princeton staff members who went to the four study sites, with three to six people each spending from one day to two weeks at a site. For future years, it is proposed to appoint a professionally qualified technical director at each site who will operate in parallel with the local coordinator. The major responsibilities of the technical director will be for directing the training of indigenous testers and observers and for monitoring the collection of data in the field.

Six types of training will be involved: training in administration of tests to children (including administering the mother-child interaction task); training in the use of PROSE for observation of classrooms; training in (a) observation of children in free-play situations and making personality ratings with the Emmerich instrument, and (b) use of the appropriate forms to make global ratings of classrooms; training in administration of questionnaires and other instruments to teachers, administrators, and other educational staff members; training in collection of the appropriate data from children, teachers, and administrators in the cross-sectional study; and training in administration of the parent interview.

The technical director will designate himself or one other person (the "head trainer") to be in charge of each type of training locally. In the case of most of the types, the technical director and his "head trainer" will then be responsible for training any additional trainers he may need "back home" and, with their help, training the local trainees. However, if no more than two trainers are required locally for a type of data collection, they will both be brought to Princeton for training and

one level of the process will be eliminated. Furthermore--and this may be the case with PROSE for Head Start classes--if no more than two trainees per site are required for a type of training, they will be brought to Princeton and trained directly, eliminating all intermediate training stages. All training in the field will be monitored not only by the local technical director but also by regional and Princeton office personnel.

Technical directors will be selected on the basis of professional training and experience relevant to the requirements of the project and comparable to that of senior professional staff on the project in Princeton and in ETS regional offices. The local academic community is seen as the source of technical directors, and it is expected that they will choose their training staffs largely from faculty members, faculty wives, and graduate students.

Interviewing mothers. During the spring of 1970, the second set of interviews will be conducted with the mothers of the children in the longitudinal sample. This time, however, the interviews will be given in testing centers on the day that the mother comes in for the mother-child interaction task. Such a plan should increase the probability of getting complete data. It is anticipated that interviewing will again take place in homes in spring 1971, because of the added information that periodic direct observation of the child's living conditions may provide.

Special provisions, as described below, will have to be made for interviews of parents who have moved away from the study sites.

Observations in classes. The study will involve four types of classroom observations in 1969-70: observations in Head Start classes with the PROSE time- and child-sampling instrument (about 20 half-days per class, spread over the year); observations of children in Head Start free-play situations in order to make personality ratings (about 60 minutes per child in the fall and again in the spring); global observations of Head Start classrooms (about 10 per class during the year); global observations of the cross-sectional classes, kindergarten through grade three (about 2 per class).

Experiences in pretesting PROSE and the Emmerich personality rating instrument have pointed to the exceptional logistical problems of observing Head Start classes. Even when, as they must be, schedules are

Table 4

TENTATIVE SELECTION OF MEASURES FOR CHILDREN IN CROSS-SECTIONAL GROUPS, 1969-70

Test	K	Gr.1	Gr.2	Gr.3
Metropolitan Readiness Tests	35[1]			
Cooperative Primary: Pilot		10	10[4]	
Cooperative Primary: Listening		35	35	35
Cooperative Primary: Reading		35	35	35
Mathematics[2]		20	25	25
Preschool Embedded Figures Test	15			
Children's Embedded Figures Test		15	15	15
Brown IDS Self-Concept Referents Test[3]	15	15	15	15
Gumpgookies[3]	25	25	25	25

1 Administration time in minutes.

2 Selected items from longer tests proposed for the longitudinal sample.

3 Group version proposed for Gr. 1, 2, 3.

4 Optional.

carefully worked out in advance, they are frequently wrecked by suddenly competing events, illnesses, and even the weather. Classroom observers and those charged with monitoring observations need special measures of good will, humor, and flexibility.

Testing children. Testing of the longitudinal subjects will be done during the spring by about three testing units per site. As in 1968-69, centers will be located in churches or other suitable facilities, and each unit will be staffed by a supervisor, six testers, a play area supervisor, and a driver (for transport of children). Units will probably be operated on a three-shift basis: 6 children scheduled for each of three 2 1/4-hour shifts a day, working nearly the entire time with a tester. Thus a unit could handle a maximum of 90 children in 3 weeks; however, experience has shown that maximums are not likely to be achieved. Arrangements will be made with parents and teachers (if a child is in Head Start) for a child to be sent to a testing center for about three days to receive about 6 1/2 hours of assessment. (It is important to note that while we think of hours of assessment or testing, the children just seem to think of "going to play games with nice ladies.")

Cross-sectional subjects will be tested at about the same time, but most measures will be group measures administered in their regular classrooms by local project staff or the classroom teacher (depending on how the particular instruments were standardized). The cross-sectional batteries will be shorter than the batteries proposed for the longitudinal subjects as they progress from kindergarten through grade 3; however, the longitudinal batteries will include all measures given to the cross-sectional subjects. A tentative selection of measures is given in Table 4.

All testing operations will be monitored by the technical director, the nearest ETS regional office, and occasionally by staff from Princeton.

Collecting information from testers. As in 1968-69, testers will be asked to supply the project staff with certain background information about themselves and to answer questions that permit some assessment of general ability levels and attitudes judged relevant to the project. Some of this information is obtained at the time of hiring; the rest is obtained later, after they have been on the job. The purpose, of course, is to allow study of the degree to which tester characteristics are associated

with differential performance on the part of the children with whom they interact so intimately. Protecting the privacy of testers is as important as protecting the privacy of the other subjects of the study (children, mothers, teachers, etc.). The fact that all information they provide is treated confidentially and, for the purpose of any reporting, anonymously, is carefully explained to them. Even so, in 1968-69 there were some refusals by testers to supply the requested data.

Collecting information from teachers, school administrators, and other educational personnel. During the spring of 1970, Head Start teachers will be asked to perform two major kinds of tasks: fill out questionnaires relating to their own backgrounds, attitudes, classroom programs, etc., and provide ratings of certain traits for each of the pupils in their classes (these ratings parallel some of those used in the Emmerich personality observation procedure). Aides will be asked to provide some of the same information. Procedures will be fully explained to teachers personally by local project staff, and teachers will be compensated for the extra time the tasks involve. Head Start administrators will also be asked to provide information about themselves and the center as a whole.

Shorter questionnaires will be used with teachers and administrators in the cross-sectional study.

Collecting community information. Much information about the communities and perceptions of them is obtained directly from interviews with mothers of the longitudinal subjects. However, since the parents may not constitute a broadly representative sample of adults in the study districts, some questions will be asked of other community people and agencies in order to obtain a more balanced picture of what it is like to live in the study areas in Trenton, Lee County, St. Louis, or Portland. In addition, certain demographic data will be sought from public and community agency records.

Keeping track of children who move. The study strategy calls for trying to keep track of the study children even if they move--from the block, neighborhood, city, or state. Keeping track of them as long as they live in the study districts and finding out where they go if they leave are the responsibility of the local coordinator. Personal contacts,

community agencies, the post office, postcards left with parents ("If you move, please fill this out and send it to... We'll send your child a surprise gift."), and other sources should all prove to be helpful. However, the central project staff is investigating the possibility of using professional agencies to whom such tracking, on a per-family basis, is more routine.

Plans for follow-up assessment of "movers" have not been formulated in detail. However, in general terms, the assessment battery will include the parent interview, plus medical information as reported by the mother; about two hours of testing of the child (key instruments selected from the regular battery for the age group, with practical restrictions on complicated equipment); and an interview of about 45 minutes with the child's current teacher if the child is in school or preschool (the teacher interview will consist of questions about the program the child has experienced during the year, paralleling some of the types of information obtained through observation in the study sites, and items of background and attitudinal information).

Present plans also call for the training of a cadre of professional personnel affiliated with the seven ETS regional offices, each of whom would be prepared in a one-day visit to handle the parent, child, and teacher assessments in the child's new location. Each spring during a defined period, these tester-interviewers would be notified by the Princeton office of the new addresses of study children who have presumably moved into their regions. They would then do as much checking as possible and try to arrange appointments in advance. Again, professional investigating agencies could be used where it was felt they might be helpful in locating children.

Data preparation and analysis. It is anticipated that the preparation (checking, scoring, coding) of the data collected through September 1969 will continue through most of the fall and winter. As preparation is completed for each measure, internal analyses, as specified by those responsible for the individual instruments, will be carried out. These analyses will provide information about the characteristics of the measuring instruments for this population; e.g., internal consistency, inter-rater

or scorer agreement, difficulty, factorial structure, and intercorrelations with other measures. As soon as possible, of course, the full data analysis, as described by Albert Beaton in an earlier chapter of this report, will be initiated.

As data is collected during 1969-70, it too will begin the progression through the preparation-analysis sequence. Needless to say, at that point the job will become increasingly complex since it will be necessary to match not only all the different scores and descriptive statements obtained from many sources about a child in one year but also all the information obtained across two years.

Community relations. So far the study has been able to attract a great deal of interest and cooperation of people in the four locations. However, as time goes by interest could diminish, and a continuing program of community relations is considered a necessary part of the study operations. The four local coordinators take major responsibility for relations with educators, anti-poverty agencies, families of the study children, and other local groups; and they have used radio, television, newspapers, informal luncheons, speeches before groups, and church bulletins to get their messages across and to receive feedback from the communities. During the coming year, they expect to continue these activities. The new technical directors, as well as Princeton and regional office personnel, will help, especially with respect to contacts with the educational-professional community. The Princeton office will also be responsible for assistance on publicity, and plans are underway for an informal newsletter that can go to parents of study children and others to keep them abreast of the progress of the project--nationally, as well as in Lee County, Portland, Trenton, and St. Louis.

Possible study extensions. As has been stated before, the longitudinal study may be thought of as the nucleus for a comprehensive program of research on early childhood education and development. The first year of the study pointed to an immediate need to propose and seek separate funding for two auxiliary investigations: a parallel, longitudinal study of middle-class children; and an intensive study of the medical and physiological characteristics of the disadvantaged children of the present study. Plans for preparing and presenting these proposals are being made.

In addition, continuing efforts are being made to engage the interest of other investigators in launching parallel studies focusing on the development and early schooling of American Indian, Mexican American, Puerto Rican, and other special groups of children.

References

Educational Testing Service. Disadvantaged children and their first school experiences. Interim report, February 1968, Contract OEO 4206, Office of Economic Opportunity. (a)

Educational Testing Service, PR-68-4. Disadvantaged children and their first school experiences: Theoretical considerations and measurement strategies. Report in 2 volumes, December 1968, Contract OEO 4206 and Grant OEO C6-8256, Office of Economic Opportunity. (b)

APPENDIX A

TRYOUTS OF MEASURES

Classroom Observation Rating Scale (Personality)

Walter Emmerich

. . . . .

Personal Record of School Experience (PROSE)

Donald Medley

. . . . .

Summary of Pretest Populations for Measures of 3 1/2-Year-Olds

Classroom Observation Rating Scale (Personality)

Walter Emmerich and Gita Wilder

The classroom observation ratings were designed to assess a broad array of children's personal, social, and cognitive behaviors in the "free play" setting of the classroom during the first two years of the longitudinal study (1969-71). This report is the third in a series that includes Emmerich's Working Paper (Educational Testing Service, 1968a), and Appendix b-1 (Educational Testing Service, 1968b).

The present report consists of revised procedural recommendations for application of the instrument in the longitudinal study together with initial results of a feasibility study conducted under field conditions in Trenton during winter and spring, 1969.

Our major conclusions can be summarized briefly. First, despite the obstacles noted in this report, we were able to accomplish the major aims of the tryout study. These aims were to estimate interjudge reliabilities on all scales (simultaneous independent ratings) and the short-term child stabilities of all scales (repeated independent ratings) under actual field conditions. Second, we believe that many of the obstacles faced in the tryout study can be reduced greatly by implementing the revised procedural recommendations spelled out in this report. Third, we find that we cannot reach the very highest standards of measurement that are possible for classroom ratings under optimal circumstances. Nevertheless, and we emphasize this point, our experience also convinces us that our revised procedures will result in sufficiently reliable appraisals of individual differences in behavior to provide information that will greatly enhance the value of the longitudinal study.

We recommend that provision be made in the longitudinal study itself for continuous monitoring of interjudge and short-term child stabilities. We believe strongly that this procedure is well worth the added cost because it will allow us to monitor the instrument's reliability throughout the course of the study, and because it will provide better estimates of children's behaviors within each age period. The latter point is analogous to the maxim, commonly accepted among test constructors, that several

items tapping the same variable will result in more accurate measurement than a single item.

Moreover, we recommend that during the first year in which a child is observed in a classroom that he be observed and rated both within the first half and the second half of the school year. The basis for this recommendation derives from theoretical and methodological considerations spelled out in the accompanying Working Paper, "Models of Continuity and Change in Development" (in Appendix B). In brief, such a procedure will allow for the derivation of change scores on each subject within the presumably critical first year of school experience.

Current Status of the Instrument

Content and Format

The following minor changes have been made in the instrument since its initial development:

1. The four-point scale for the unipolar attributes has been modified slightly, as follows:

   0 - Attribute totally absent during period of observation.

   1 - Attribute occasionally present during period of observation (attribute occurs once during period of observation).

   2 - Attribute frequently present during period of observation (attribute occurs more than once but is not continual during period of observation).

   3 - Attribute continually present during period of observation.

2. A number of unipolar attributes have been redefined or combined, resulting in 127 unipolar attributes.

3. The sequence of the scales has been reordered, and some scales have been bifurcated to differentiate behaviors directed toward adults from behaviors directed toward children.

4. One bipolar scale has been eliminated (oriented toward adults vs. oriented toward children) because of the above bifurcations.

A four-page rating sheet was used and is found at the end of this paper. Identifying information is recorded on the first page, followed

by the 127 unipolar scales, which spread over most of the rating sheet. The 21 bipolar scales appear on the last page (Exhibit A).

Procedures

1. Raters are trained intensively throughout a two-week period. Training consists of (a) learning the manual of scale definitions, (b) learning the observation procedure, (c) observing in selected classrooms, making ratings, and discussing them in groups, and (d) observing as pairs in classrooms under full field conditions, making ratings, and discussing them with the trainer. The manual of scale definitions, with behavioral illustrations, and qualifications appears at the end of this section (Exhibit B).

2. A rater is considered to be "trained" when, in the trainer's judgment, she exhibits few disagreements with an independent rater on all scales on several children.

3. Ratings are made after 30 minutes of continual observation of a child during a free play period. The observer must be reasonably free to follow the child around during the observation period.

4. Simultaneous observations of the same child are made by pairs of observers who independently rate the child on all scales.

5. Observers independently rate a particular child immediately after observing him. Whenever possible, the rater goes to a place free from distraction to make his ratings.

6. Each child is rated a second time by a different pair of raters, ideally within a period of two weeks.

Tryout Study

Problems in Meeting Aims

Since the tryout study was initiated in Trenton almost immediately after that city was selected as a study site, many practical difficulties were encountered. The tasks of recruiting potential raters, setting up the field office appropriately, establishing work hours and payment

schedules, informing Head Start personnel about this phase of the study, preparing teachers for the presence of observers in the classroom, and so on, had to be compressed into a short period. Moreover, schools were shut down on several occasions by two major snowstorms and a teacher strike. There were additional delays because of faculty meetings, dental and eye examinations, class trips, and child absences. Finally, the attrition rate among observers was high. Eight observers (four pairs) were trained, but by the end of the four-month period of the tryout study, only two observers (one pair) were still available.

Procedures of the Tryout Study

1. Initial recruiting. Fourteen applicants were considered, with the intention of recruiting about eight. Three documents were completed for each applicant--Background information; Judgments on potential raters; Child behavior (copies are attached, Exhibits C)--and used in the rater selection process. Eight raters were chosen on the basis of their ability to recognize and describe appropriate behaviors of children, their verbal facility, and their apparent motivation to do the task.

2. Trainee characteristics. Seven of the eight women were black and lived in or close to a neighborhood in which a Head Start Center was located. Six were between 19 and 29 years of age, one was 35, and one was 51. One rater had completed 10th grade, six had completed high school, and one had a master's degree. Seven of the women had at least one child.

3. Training. Initially, two weeks were allotted to the training procedure. Using the Trenton field office as headquarters and various Head Start centers for practice, two researchers and the eight rater trainees spent each day studying and discussing the manual, making practice observations and ratings, discussing the practice ratings, and returning to do more ratings in light of the discussions. Sometimes the entire group of eight met, while at other times there were two groups of four. After the first week, observers worked in pairs and the researchers met individually with pairs of raters to discuss their paired observations. Progress varied considerably among the trainees. Some were ready to start on the tryout study at the end of two weeks, while others needed

three weeks. Absenteeism and inability to get into classrooms for practice observations accounted for most of the additional time. While trainees differed in their motivation to stick to the task, it was our impression that all eight had the essential skills necessary to make satisfactory ratings.

4. Sample. It was our hope that about 120 children would be seen in four Head Start centers (8 classes). One pair of raters was to start in each center, and after a week of first observations, the pairs were to switch centers. This very neat design broke down after the first week, but a total of 92 children were seen, and there were double paired observations (i.e., four observations per child) for 64 of the children.

5. Time schedule. Initial plans called for eight observers, but the actual number, averaged over the period of the tryout study, was about four. As a result, the study took almost twice as long as expected. Original plans also called for six weeks of observations (excluding the two weeks allotted to training) with an additional two weeks for make-up observations. The actual period was 12 weeks.

6. Monitoring. For two weeks after the training period, one of the researchers remained at the Trenton office to monitor the rating procedure. Monitoring involved collecting the rating sheets at the end of each day's ratings, discussing discrepant ratings with pairs of raters, answering questions about the scales, and making estimates of interjudge agreement. After this initial period of close monitoring, the researcher returned to the Princeton office and supervised from there. Rating sheets were delivered to Princeton at the end of each day, and pairs of raters were instructed to discuss their ratings after the ratings had been entered on the sheets. Schedule changes were made by the local coordinator.

Criteria Used to Evaluate Results

Earlier pilot work on the present instrument, as well as rather extensive previous research on child ratings, indicated that moderate to high reliabilities can be achieved under optimal training and observation conditions. However, the basic question posed in the tryout study was

whether the actual field conditions of the longitudinal study would attenuate reliability to the point where the resulting information would have little scientific value. Our aim was to capture enough "true" variance in children to allow for the emergence of functional relations among the scales and between the scales and other variables of the study.

It is important to note that each scale represents a sample of behavior that eventually will be used in a composite score. Composite scores will be derived empirically after the longitudinal data is collected, as outlined (in part) in the accompanying Working Paper (in Appendix B). In effect, each rating on each scale represents an "item" in search of an appropriate composite. Composite scores will be generated in three distinct ways: (1) by summing independent simultaneous ratings (analogous to split-half reliability using the same form); (2) by summing repeated ratings made within a short time interval (analogous to test-retest reliability using the same form); and (3) by summing across correlated scales having similar meanings (analogous to internal consistency reliability). Such composites will parlay scales which meet only minimal reliability standards into total scores which should contain ample "true" child variance for research purposes.

It follows that the proper criterion for evaluating the reliability of a scale ("item") is whether it captures a significant portion of child variance. Ideally, all scales will meet all three reliability criteria, although it is not necessary for a scale to meet all criteria in order to be useful.

The present report provides initial estimates of these reliabilities. Final estimates will be computed on the longitudinal sample itself. Moreover, the presentation that follows takes an overall look at the viability of the instrument as a whole, rather than attempting to appraise the potential utility of each scale.

Interjudge Reliabilities

Three initial estimates of interjudge agreement were computed (Pearson correlations), based upon different pairs of raters. For Type 1,

there were 124 paired ratings made by the one pair who functioned throughout the course of the tryout study. For Type 2, there were 84 paired ratings made by another pair who remained for about half of the study. Type 3 includes the several (least stable) pairs combined into a single correlation, with random assignment of pair members to the two columns (N = 108). Working on the assumption that actual field conditions for the longitudinal study will involve some combination of all three of these types, we have taken as an overall (and probably lower-bound) estimate of interjudge reliability the correlations for all three cases combined (N = 316). These correlations are reported in Table 1. The mean reliability for all scales was .50.

These results are encouraging as it appears that most scales will meet at least one of the three criteria discussed earlier. Moreover, for reasons noted below, we are optimistic that many of the scales which had (relatively) low rater reliabilities in the tryout study will be found to have more satisfactory rater reliabilities in the longitudinal study itself.

In the first place, our most stable pair of raters (Type 1) generally were more reliable than the less stable pairs (Types 2 and 3). For 74% of the scales, Type 1 was more reliable than Type 2, and for 77% of the scales, Type 1 was more reliable than Type 3. <u>These findings indicate the great importance of implementing field conditions which enhance the stability of rater pairs throughout extended periods of the longitudinal study</u>.

Inspection of the 148 standard deviations for all subjects on whom at least one observation was made (N = 92) revealed marked variation among scales in variability. (Unipolar scales also differed considerably with respect to mean scale scores, and, as expected, the means and standard deviations were correlated.) An important factor influencing the magnitude of a reliability coefficient is amount of variability. Since it is reasonable to expect the larger samplings of children, classrooms, and age periods of the longitudinal study to produce greater scale variabilities on scales having low variability in the tryout study, we can expect higher rater reliabilities on these scales to occur in the longitudinal study itself.

Table 1

INTERJUDGE RELIABILITIES (J) AND SHORT-TERM STABILITIES (S) FOR ALL SCALES[a]

Scale Number	J	S	Scale Number	J	S	Scale Number	J	S	Scale Number	J	S	Scale Number	J	S
Bipolar:			8	.61	.21	38	.01	-.03	68	.66	.27	98	.53	-.01
1	.74	.44	9	.81	-.05	39	.47	-.10	69			99	.33	-.18
2	.68	.57	10	.45	-.04	40			70	.64	.31	100	.34	-.04
3	.22	.03	11	.54	.05	41	.49	-.06	71	-.01	-.01	101	.18	.01
4	.22	.16	12	.44	.03	42	.64	.11	72	.71	.15	102	.31	.40
5	.62	.34	13	.67	.04	43	.58	.23	73	.60	.30	103	.42	.04
6	.49	.12	14	.53	.05	44	.71	-.07	74	.45	.06	104	.85	-.04
7	.64	.23	15	.52	.04	45	.60	.09	75	.28	.15	105	.56	-.06
8	.38	.26	16	.42	-.08	46	.58	.38	76	.40	-.03	106	.68	.28
9	.66	.42	17	.19	.18	47	.66	.39	77	.54	-.04	107	.68	.20
10	.55	.36	18	.66	-.03	48	.41	.01	78	.53	-.07	108	.61	-.03
11	.15	.13	19	.68	-.06	49	.33	-.08	79	.55	-.01	109	.28	.04
12	.58	.44	20	.55	-.06	50	.55	.32	80	-.01	-.03	110	.67	-.04
13	.56	.37	21	.53	-.05	51	.78	.25	81	.46	-.03	111	.57	.27
14	.39	.05	22	.55	.14	52	.61	.30	82	.57	.00	112	.40	-.02
15	.36	-.09	23	.51	-.02	53	.57	-.02	83	.61	.05	113	.45	.20
16	.52	.15	24	.36	-.05	54	.70	.13	84	.50	-.07	114	.62	.18
17	.10	.18	25	.71	.10	55	.70	.18	85	.61	.09	115	-.01	-.02
18	.50	.06	26	.28	-.11	56	.47	.05	86	.28	.15	116		
19	.61	.39	27	.31	.02	57	.42	.16	87	.23	.03	117	.95	
20	.42	.31	28	.35	.08	58	.54	.23	88	.59	.43	118	.86	-.02
21	.54	.17	29	.65	.08	59	.32	-.09	89	.55	.36	119	.46	-.03
Unipolar:			30	.45	.00	60	.50	.04	90	.30	.22	120	.66	-.05
1	.99	-.04	31	.58	-.04	61	.40	.17	91	.52	.11	121	1.00	
2	-.01		32	.43	.18	62	.49	.16	92	.26	.03	122	.70	-.01
3	.74	.14	33	.45	-.01	63	.52	.10	93	-.04	-.06	123		
4	.38	.14	34			64	.57		94	.26	-.07	124	.42	.03
5	.66	.30	35	.55	.05	65	.52	.00	95	.36	.28	125	.66	-.02
6	.69	-.09	36	.65	.18	66	.71	.06	96	.50	.15	126	.54	.04
7	.74	.10	37	.62	.21	67			97	.22	-.08	127	.60	-.02

[a]The listing of scale numbers corresponds to scale titles found in the appended rating sheet. For J, N = 316; for S, N = 83. Blank cells occur when SD = 0.

Short-Term Stability Over Time

Short-term stability coefficients were available on 83 subjects, with an average of 11 days between observations. These correlations are also reported in Table 1. It will be noted that these correlations generally are low, especially in the case of the unipolar scales. Inspection of the stability coefficients for (relatively) short vs. long time intervals between the first and second ratings revealed no systematic tendency for shorter periods to produce higher stability coefficients than longer periods.

These findings on short-term stability are not very encouraging. Of course, the factors noted earlier, which attenuated interjudge reliabilities in the tryout study, probably attenuated these stability estimates to an even greater extent, especially for the unipolar scales. Also, it should be noted that since a different observer rated a child on the first and second occasion, stability coefficients were further attenuated by lack of perfect interjudge agreement. Nevertheless, it is clear from these results that some scales may not be sufficiently stable to warrant summing repeated observations.

We think this situation is less serious than it might appear. As noted earlier, we plan to sum across scales as well as across observers and occasions when deriving composite scores. It seems likely that several behaviors will serve as functionally equivalent ways of expressing the same underlying trait. Composite scores derived from scales (having similar meaning) which are positively intercorrelated within both the first and second sets of observations should produce considerably higher stability coefficients than the separate scales considered alone.

Indeed, this situation suggests a strategy for building composite scores. Step 1 would be to determine the interjudge reliabilities of the scales. For scales meeting the criterion for Step 1, which we anticipate will include almost all scales, the second step would be to intercorrelate ratings within each of the observations, and to extract clusters of functionally equivalent behaviors in each observation period which match across periods. Composite scores (within observation periods)

Table 2

INTERCORRELATIONS AMONG BIPOLAR SCALES (N = 92)

Scale		Scale 1	2	3	4	5	6	7	8	9	10
1.	Withdrawn vs. Involved										
2.	Masculine vs. Feminine	-.13									
3.	Tolerates vs. Vulnerable to Frustration	.03	-.09								
4.	Rebellious vs. Compliant	.02	.15	-.09							
5.	Expressive vs. Restrained	-.77	.04	-.02	-.20						
6.	Tense vs. Relaxed	.72	-.11	.03	.22	-.74					
7.	Other- vs. Self-centered	-.65	-.02	.13	-.18	.70	-.70				
8.	Submissive vs. Dominant	.64	-.23	.06	-.24	-.49	.56	-.50			
9.	Active vs. Passive	-.89	.22	-.06	.12	.77	-.70	.65	-.59		
10.	Apathetic vs. Energetic	.88	-.27	.07	.03	-.74	.74	-.62	.58	-.91	
11.	Stable vs. Unstable	-.28	-.14	-.14	-.19	.30	-.20	.16	-.02	.25	-.20
12.	Solitary vs. Social	.89	-.10	.04	.09	-.76	.70	-.65	.61	-.84	.86
13.	Assertive vs. Timid	-.74	.21	-.06	.07	.70	-.64	.62	-.62	.78	-.69
14.	Dependent vs. Independent	.45	-.17	-.07	-.11	-.36	.50	-.44	.57	-.42	.48
15.	Constructive vs. Destructive	-.09	-.17	.19	-.20	.15	-.04	.14	.05	.03	.06
16.	Aimless vs. Purposeful	.29	.08	-.06	.11	-.16	.24	-.14	.17	-.21	.21
17.	Academic vs. Other Motivation	-.03	.08	-.03	.13	.02	.10	.01	-.03	.08	.06
18.	Aggressive vs. Affectionate	-.14	.26	-.15	.43	.16	-.01	.02	-.34	.25	-.15
19.	Secure vs. Insecure	-.81	.17	.03	-.10	.80	-.76	.70	-.65	.76	-.71
20.	Rigid vs. Flexible	.69	-.20	-.09	.28	-.76	.79	-.62	.51	-.66	.74
21.	Happy vs. Unhappy	-.80	.07	.10	-.20	.73	-.70	.72	-.58	.71	-.71

Scale		Scale 11	12	13	14	15	16	17	18	19	20
12.	Solitary vs. Social	-.28									
13.	Assertive vs. Timid	.28	-.67								
14.	Dependent vs. Independent	-.08	.39	-.41							
15.	Constructive vs. Destructive	.40	-.07	.06	-.05						
16.	Aimless vs. Purposeful	-.39	.29	-.25	.31	-.29					
17.	Academic vs. Other Motivation	.20	.06	.17	.02	.23	-.03				
18.	Aggressive vs. Affectionate	-.05	-.08	.30	-.12	-.13	.14	.27			
19.	Secure vs. Insecure	.31	-.74	.78	-.46	.23	-.33	.12	.11		
20.	Rigid vs. Flexible	-.15	.69	-.64	.40	-.10	.22	.15	.00	-.76	
21.	Happy vs. Unhappy	.25	-.80	.58	-.47	.18	-.22	.09	.01	.77	-.61

could then be derived from each cluster. Step 3 would be to compute the stability coefficients of the composite scores found in Step 2. If these are satisfactory, then the final step would be to sum the (comparable) composites across observation periods; the resulting total would become the actual score used in subsequent analyses.

Interrelations Among Scales

There were 92 subjects having one to four sets of ratings. Using each subject's average rating on each scale, the 148 scales were intercorrelated. Inspection of the total matrix revealed promising clusters of interrelations within the set of bipolar scales, within the set of unipolar scales, and between these two sets of scales. For the sake of brevity, we report only the intercorrelations among bipolar scales, found in Table 2.

Two features of Table 2 are noteworthy. First, there are a number of high correlations, even though most subjects' scores are based upon sums across repeated observations. This fact is encouraging with respect to the strategy suggested above for dealing with the problem of short-term stability.

Second, it is clear from inspection of Table 2 that much of the common variance is attributable to a single global dimension. This dimension includes (at one pole) involved, expressive, relaxed, other-oriented, dominant, active, energetic, social, assertive, secure, flexible, and happy behaviors, and, at the other pole, withdrawn, restrained, tense, self-centered, submissive, passive, apathetic, solitary, timid, insecure, rigid, and unhappy behaviors.

Although similar in some respects to the dimension of introversion-extraversion often found in rating studies of young middle class children, we think this dimension has different and much more profound implications for a disadvantaged population. In the first place, it is more global in the disadvantaged, suggesting that basic processes of personality differentiation occur later in this population, an hypothesis that the longitudinal study will explore in detail. Secondly, this dimension seems to implicate so many of the child's psychological systems (energy,

social behavior, affect) that it is difficult to apply a defining label less strong than that of "tuned in" vs. "turned off" in the classroom environment. Moreover, it is likely that this global dimension will be a crucial moderating variable at this age (and perhaps also at later ages) which influences functional relations between specific aspects of the child's environment and specific educational outcomes.

Of course, the emergence of this one pervasive dimension could signify merely that our raters were more subject to halo effects than raters of previous studies in which more dimensions have been found. We cannot rule out this explanation at this time, although the fact that several of the bipolar scales did not enter into this dimension suggests that a single halo effect could not have operated indiscriminantly across all scales. Moreover, the longitudinal study will provide at least two internal checks on the halo interpretation. First, inspection of the unipolar scale correlates of each bipolar scale provides information on this (and other) issues. If the patterning of unipolar correlates varies among bipolar scales, then we will have direct evidence that judges responded to different behavioral cues when arriving at each bipolar rating, rather than responding merely to generalized halo impressions. Second, rater characteristics will probably be quite constant throughout the longitudinal study, so that evidence of increasing child differentiation with age would be inconsistent with the halo interpretation.

Recommended Procedures for 1969-70 Longitudinal Sample

Recruiting

The longitudinal sample for 1969-70 may yield 250 children in each of the three study sites having pre-kindergarten Head Start programs (Portland, St. Louis, and Trenton). It is recommended that the technical director (henceforth referred to as "director") in each of these sites appoint an assistant for classroom ratings, who will share with the director responsibility for activities concerned with the observer ratings of children. We anticipate that four pairs of observers (eight persons) will be needed in each of the three sites during 1969-70.

Using the same selection procedures employed in the tryout study, candidates will be recruited by the classroom rating assistant in each site in consultation with the director and the Princeton office. Sites will maintain some source of back-up personnel to be called upon in case of dismissal or illness of regular observers.

Training

We recommend a two-phase training process. First, the classroom rating assistants and directors from each site will meet in Princeton for 1-2 weeks of intensive training. Then, upon returning to their sites and after recruiting twelve trainees, the classroom assistants will train and select a final group of eight raters during a 2-week period. During this period, Princeton personnel will travel to each site to monitor the training and to judge the readiness of trainees.

Role of the Assistant for Classroom Observations

The assistant for classroom observations will have a number of duties in addition to those of training. During the time she is training in the field, she will also arrange for nursery schools (not involved in the longitudinal study) to provide practice observations. Once the study is under way, she will monitor the progress of ratings very closely, keep records, maintain schedules, coordinate rater activities with Head Start programs, substitute for absent raters whenever possible, and hold regular weekly group and individual meetings with raters to monitor their use of the scales. She will provide daily assignments for pairs of observers, and reassign pairs appropriately when schools are unavailable or observers are absent. She will also keep in close contact with the Princeton office, where a master chart of all field operations will be kept up-to-date.

Observation Schedule

There will be double paired observations on each of the study children in both the first and second halves of the first year in which the child is enrolled in a school program. Each child will be observed initially in the fall by a pair of observers, then rated again by a second pair of observers within two weeks of the first observation. This entire procedure will be repeated in the winter-spring, for a total of eight observations per child during the school year.

Study children will be divided into three groups, to be observed initially in October, November, and December, respectively. The group of children observed in October will be observed again during February; the group observed in November will be seen again during March; and the group observed in December will be seen again in April. In this way, the interval between the first and second sets of double observations (about 4 months) can be kept fairly constant.

References

Educational Testing Service. Disadvantaged children and their first school experiences. Interim report, February 1968, Contract OEO 4206, Office of Economic Opportunity. (a)

Educational Testing Service. Disadvantaged children and their first school experiences: Theoretical considerations and measurement strategies. Report in 2 volumes, December 1968, Contract OEO 4206 and Grant OEO C6-8256, Office of Economic Opportunity. (b)

EXHIBIT A

0176

OBSERVER RATINGS OF CHILDREN

CHILD _____ M F CODE _____

TEACHER _____ CODE _____

OBSERVER _____ CODE _____

CO-OBSERVER _____ CODE _____

DATE _____ TIME OBSERVATION BEGUN _____ COMPLETED _____
Year Month Day

Scale: 0 = Attribute totally absent during period of observation
1 = Attribute occasionally present during period of observation
2 = Attribute frequently present during period of observation
3 = Attribute continually present during period of observation

				COMMENTS
____ 1.	Seeks physical affection from adult			
____ 2.	Seeks physical affection from other child			
____ 3.	Seeks help or guidance from adult			
____ 4.	Seeks help or guidance from other child			
____ 5.	Seeks physical proximity of adult			
____ 6.	Seeks physical proximity of other child			
____ 7.	Seeks attention from adult through positive bid			
____ 8.	Seeks attention from other child through positive bid			
____ 9.	Seeks attention from adult through deliberate negative bid			
____ 10.	Seeks attention from other child through deliberate negative bid			
____ 11.	Seeks attention from adult through weak bid			
____ 12.	Seeks attention from other child through weak bid			
____ 13.	Seeks praise or approval from adult			
____ 14.	Seeks praise or approval from other child			
____ 15.	Seeks evaluation from adult			
____ 16.	Seeks evaluation from other child			
____ 17.	Seeks or makes a comparative evaluation			
____ 18.	Demanding of adult			
____ 19.	Demanding of other child			
____ 20.	Tries to get adult to do what self is expected to do			
____ 21.	Tries to get other child to do what self is expected to do			
____ 22.	Exhibits helplessness			
____ 23.	Rejects positive bid from adult			
____ 24.	Rejects positive bid from other child			
____ 25.	Seeks adult's permission to do something			
____ 26.	Seeks permission of other child to do something			
____ 27.	Conforms to routine or routine request of adult			
____ 28.	Conforms to routine or routine request of other child			
____ 29.	Rejects reasonable request of adult			
____ 30.	Rejects reasonable request of other child			
____ 31.	Engages in complementary behavior			

____32. Engages in parallel activity
____33. Exhibits interest in or concern for other in distress
____34. Praises or expresses approval toward adult
____35. Praises or expresses approval toward other child
____36. Expresses criticism of adult
____37. Expresses criticism of other child
____38. Reciprocates with adult
____39. Reciprocates with other child
____40. Tries to "make up" with adult
____41. Tries to "make up" with other child
____42. Friendly toward adult
____43. Friendly toward other child
____44. Nurturant toward adult
____45. Nurturant toward other child
____46. Exhibits leadership
____47. Behaves competitively
____48. Seeks leadership of adult
____49. Seeks leadership of other child
____50. Smiles and/or laughs
____51. Engages in gross motor activity
____52. Engages in fine manipulative activity
____53. Engages in cognitive activity
____54. Engages in fantasy activity
____55. Engages in artistic activity
____56. Concerned with satisfaction of physical need
____57. Takes initiative in carrying out own activity
____58. Tries to pursue difficult task
____59. Attempts to overcome obstacles by himself
____60. Exhibits persistence
____61. Completes activity by himself
____62. Gets intrinsic satisfaction from activity or task
____63. Praises self
____64. Threatens to act aggressively toward adult
____65. Threatens to act aggressively toward other child
____66. Possessive
____67. Verbally aggressive toward adult
____68. Verbally aggressive toward other child
____69. Bosses adult
____70. Bosses other child
____71. Physically aggressive toward adult
____72. Physically aggressive toward other child
____73. Deliberately aggressive toward property
____74. Expresses negative feeling about self, possession, or own product
____75. Exhibits visual curiosity

Scale:

0 = Attribute totally absent
1 = Attribute occasionally present
2 = Attribute frequently present
3 = Attribute continually present

____76. Exhibits active curiosity
____77. Seeks information from adult
____78. Seeks information from other child
____79. Responsive to teaching by adult
____80. Responsive to teaching by other child
____81. Imitates behavior of adult
____82. Imitates behavior of other child
____83. Instructs or demonstrates
____84. Attempts to communicate verbally to adult
____85. Attempts to communicate verbally to other child
____86. Communicates meaningful complex idea to other child
____87. Communicates meaningful complex idea to adult
____88. Verbally loud
____89. Talks to self
____90. Difficult to understand
____91. Does not concentrate on activity
____92. Inattentive when adult communicates to him
____93. Inattentive when other child communicates to him
____94. Incomplete communicative act
____95. Exhibits goal-directed activity
____96. Shows planning in pursuing activity
____97. Flexible in substituting goal
____98. Corrects or modifies performance to meet own standard
____99. Products or activities have common theme
____100. Perseverates on activity or task
____101. Perseverates verbally
____102. Preoccupied with own thoughts
____103. Unable to tolerate delay
____104. Concerned about physical discomfort or physical danger
____105. Seeks verbal reassurance
____106. Hesitant in relating to adult
____107. Hesitant in relating to child
____108. Hesitant to try things on his own
____109. Unusually good physical coordination
____110. Poor physical coordination
____111. Restlessness
____112. Easily frustrated or threatened by adults
____113. Easily frustrated or threatened by other children
____114. Recovers quickly from frustration or threat
____115. Response to frustration or threat: becomes stubborn
____116. Response to frustration or threat: becomes fearful
____117. Response to frustration or threat: cries
____118. Response to frustration or threat: becomes dejected

_____119. Response to frustration or threat: becomes defiant, rebellious

_____120. Response to frustration or threat: increased quietness

_____121. Response to frustration or threat: increased activity that seems aimless

_____122. Response to frustration or threat: seeks comfort from adult

_____123. Response to frustration or threat: seeks comfort from other child

_____124. Response to frustration or threat: retaliates against person who caused frustration

_____125. Response to frustration or threat: ignores the frustration or threat

_____126. Response to frustration or threat: effectively defends self

_____127. Response to frustration or threat: becomes angry

BIPOLAR SCALES

	1 Extremely X	2 Considerably X	3 Slightly More X Than Y	4 X No More Than Y	5 Slightly Y	6 Considerably Y	7 Extremely Y	
	X							Y
1.	Withdrawn		1	2	3	4	5	6	7	Involved
2.	Masculine		1	2	3	4	5	6	7	Feminine
3.	Tolerates Frustration		1	2	3	4	5	6	7	Vulnerable to Frustration
4.	Rebellious		1	2	3	4	5	6	7	Compliant
5.	Expressive		1	2	3	4	5	6	7	Restrained
6.	Tense		1	2	3	4	5	6	7	Relaxed
7.	Sensitive to Others		1	2	3	4	5	6	7	Self-centered
8.	Submissive		1	2	3	4	5	6	7	Dominant
9.	Active		1	2	3	4	5	6	7	Passive
10.	Apathetic		1	2	3	4	5	6	7	Energetic
11.	Stable		1	2	3	4	5	6	7	Unstable
12.	Solitary		1	2	3	4	5	6	7	Social
13.	Assertive, Bold		1	2	3	4	5	6	7	Timid, Fearful
14.	Dependent		1	2	3	4	5	6	7	Independent
15.	Constructive		1	2	3	4	5	6	7	Destructive
16.	Aimless		1	2	3	4	5	6	7	Purposeful
17.	Academically Motivated		1	2	3	4	5	6	7	Otherwise Motivated
18.	Aggressive Toward Others		1	2	3	4	5	6	7	Affectionate Toward Others
19.	Socially Secure		1	2	3	4	5	6	7	Socially Insecure
20.	Rigid		1	2	3	4	5	6	7	Flexible
21.	Happy		1	2	3	4	5	6	7	Unhappy

EXHIBIT B

Observer Ratings of Children

MANUAL: SCALE DEFINITIONS AND EXAMPLES

Walter Emmerich
Gita Wilder

January 20, 1969

Many of the scales and definitions contained herein are modifications of those found in investigations cited in the report under OEO Grant Number C6-8256 submitted early in 1969. However, this particular instrument was developed pursuant to this grant, and is limited in its use solely for this research study. This use does not constitute publication. No distribution of this material is authorized outside of Educational Testing Service or the Office of Economic Opportunity.

Attribute: 1. Seeks physical affection from adult

2. Seeks physical affection from other child

Definition: Actively seeks physical affection from another.

Examples: (a) Child hangs onto teacher.

(b) Child goes to teacher and clearly wants to be picked up or hugged.

(c) Child seeks to hold teacher's hand.

(d) Target child hangs onto another child.

Qualifications: Merely being receptive to affection initiated by other is not included here.

Attribute: 3. Seeks help or guidance from adult

4. Seeks help or guidance from other child

Definition: Seeks help, guidance, or assistance from other.

Examples: (a) Asks teacher to get play materials off high shelf.

(b) Seeks teacher's help in protecting self from other child's aggression.

(c) Asks another child what he should do next.

Qualifications: Merely accepting help initiated by other is not included here.

Attribute: 5. Seeks physical proximity of adult

6. Seeks physical proximity of other child

Definition: Active attempt to be near another child or teacher.

Examples: (a) Target child follows another child from activity to activity.

(b) Stays near teacher, following her when she moves.

(c) "I want to sit next to Eric."

Qualifications:

Attribute: 7. Seeks attention from adult through positive bid

8. Seeks attention from other child through positive bid

Definition: Active attention-seeking expressed in a positive manner. Attempt is to attract the attention of other toward self, possession, product, or act.

Examples: (a) "Look at what I made, teacher."

(b) Shows something to other child, and expects a response.

(c) "Watch me, teacher."

Qualifications: Unsuccessful as well as successful bids for attention are rated.

Attribute: 9. Seeks attention from adult through deliberate negative bid

10. Seeks attention from other child through deliberate negative bid

Definition: Acts negatively to draw attention of other to self.

Examples: (a) Deliberately runs tricycle into a group to gain their attention.

(b) Child stamps feet loudly during a time of relative quiet in classroom.

(c) Child repeatedly does something he knows annoys other.

Qualifications: Negative attention-seeking is likely to result in scolding, censure, or discipline from another, but a rating is made even when the other does not respond, or responds positively to the child.

Attribute: 11. Seeks attention from adult through weak bid

12. Seeks attention from other child through weak bid

Definition: Mild or incomplete attempts to secure the attention of other. These can be positive, negative, or neutral bids. These attempts will often be unsuccessful in drawing attention because they are too weak.

Examples: (a) Child holds picture up, and waits to be noticed by teacher or other child.

(b) Child's attention-seeking comment is made too softly to be noticed by teacher, and no attempt is made to make a stronger bid.

(c) A negative attention-seeking bid is started but stopped before it is carried out to completion.

Qualifications:

Attribute: 13. Seeks praise or approval from adult

14. Seeks praise or approval from other child

Definition: Child actively seeks indication that other likes him, his possession, his product, or his act.

Examples: (a) Child asks, "Do you like this design?", pointing to a block design he has just made.

(b) Child makes a special effort to do something that other has indicated or implied will receive approval.

(c) Child asks another child, "Do you like this picture?", indicating a picture he has just made.

Qualifications: More than mere attention of the other is sought. Make this rating only if praise, approval, or expression of liking or admiration is clearly sought.

Attribute: 15. Seeks evaluation from adult

16. Seeks evaluation from other child

Definition: Child seeks an evaluation of self, possession, product, or act in relation to a standard.

Examples: (a) Child shows painting to teacher and asks, "Is this right?"

(b) Child trying puzzle asks teacher, "Is that the way it goes?"

(c) Child shows painting to another child and says, "Is this pretty?"

Qualifications: More than attention or approval is sought. Make this rating only if the child wants the other to apply a standard of evaluation. One sign of this behavior is that the child seems receptive to criticism.

Attribute: 17. Seeks or makes a comparative evaluation

Definition: Child seeks or makes a comparative judgment about himself, possession, product, or act. He may compare himself with another, compare two others, or compare his own earlier performance with a more recent performance.

Examples: (a) "Is my tower taller than Craig's?"

(b) "I'm stronger today."

(c) "Patricia is prettier than Louise."

Qualifications:.

Attribute: 18. Demanding of adult

19. Demanding of other child

Definition: Child insists that other meet his request.

Examples: (a) Child insists loudly and repeatedly that teacher meet his request.

(b) Child asks teacher for special privileges.

(c) Target child orders another child to do something that benefits target child.

Qualifications:

Attribute: 20. Tries to get adult to do what self is expected to do.

21. Tries to get other child to do what self is expected to do.

Definition: When other suggests or requests something of target child, he tries to get other to carry out the suggestion.

Examples: (a) Told by teacher to put doll carriage in doll corner, target child tells teacher to do it.

(b) Target child asks another child to put away crayons that target child has been using.

(c) When asked by the teacher to put away blocks she has been using, target child says, "John took them out."

Qualifications:

Attribute: 22. Exhibits helplessness

Definition: Child is passive or ineffective in making needs known to other. Child's goal may be to be cared for, comforted, helped, or given attention, but he makes few or ineffective efforts to communicate these needs.

Examples: (a) Child stands in front of activity shelf for some time without choosing activity or making desire known to teacher.

(b) Child seems unable to start an activity suggested by teacher.

(c) Child seems thwarted in some activity, but neither tries to overcome difficulty by self nor seeks other's help.

Qualifications:

Attribute: 23. Rejects positive bid from adult

24. Rejects positive bid from other child

Definition: Active rejection of positive attention or help from another.

Examples: (a) Turns away from teacher who puts arm around child and tries to say something to child.

(b) Child refuses assistance of teacher.

(c) Crying child rejects comforting attempts by other child.

Qualifications:

Attribute: 25. Seeks adult's permission to do something

26. Seeks permission of other child to do something

Definition: Child asks teacher or other child for permission to do something or to engage in activity.

Examples: (a) "Can I paint now?"

(b) Child asks teacher if he can take out the blocks.

(c) "Is it my turn now to throw the ball?"

(d) Child asks another child if he can use one of his blocks.

Qualifications:

Attribute: 27. Conforms to routine or routine request of adult

28. Conforms to routine or routine request of other child

Definition: Child responds in accordance with accepted classroom routine or routine requests made by other.

Examples: (a) Asked to put his painting on the shelf, child does so.

(b) Spontaneously follows a classroom routine; e.g., putting away materials after using them.

(c) Asked by another child to back tricycle out of way, child does so.

Qualifications: What is "routine" may vary among classrooms. This rating applies only when a routine procedure or request is implied.

Attribute: 29. Rejects reasonable request of adult

30. Rejects reasonable request of other child

Definition: Child refuses to act in accordance with reasonable request made by teacher or other child.

Examples:

(a) Asked by teacher to return crayons to the shelf, target child says "No".

(b) Asked by teacher to stop painting and wash his hands, target child continues to paint.

(c) Asked by teacher to move his wagon away from the direction of the children, target child moves wagon towards the children.

(d) Asked by another child to move out of the way, target child stays where he is.

Qualifications:

Attribute: 31. Engages in complementary behavior

Definition: Child coordinates his own activity to supplement and facilitate a common activity shared by one or more others. Genuinely cooperative activity.

Examples:

(a) Target child builds one part of block structure while another child builds other part.

(b) Several children (including target child) participate in "playing house" or other dramatic games in which each child plays a different role.

(c) Coordination of activity occurs, such as in see-sawing, pushing and receiving a ball, etc.

Qualifications:

Attribute: 32. Engages in parallel activity

Definition: Child engages in same activity as other who is nearby, but their activity is independent, with no mutual coordination.

Examples: (a) Target child builds blocks by himself, with another child using blocks nearby.

(b) Child rides a tricycle in tandem with other.

(c) Child plays with dolls, but no effort is made to coordinate role playing with other children who also are playing with dolls.

Qualifications:

Attribute: 33. Exhibits interest in or concern for other in distress

Definition: When another child exhibits distress of some sort, target child shows interest, concern, or sympathy.

Examples: (a) Child looks at crying child.

(b) Child approaches a crying child and asks, "What's the matter?"

(c) Child approaches a crying child and offers him a toy.

Qualifications:

Attribute: 34. Praises or expresses approval toward adult

35. Praises or expresses approval toward other child

Definition: Child expresses praise or approval of other, either verbally or through gestures.

Examples: (a) "You're a good boy, Eric."

(b) "Jane made a nice road."

(c) Child pats other on back to indicate approval.

Qualifications:

Attribute: 36. Expresses criticism of adult

37. Expresses criticism of other child

Definition: Child critically evaluates other. The criticism is more self-protective or constructive than aggressive.

Examples: (a) "You should use the green paint in your picture."

(b) "Don't run into me."

(c) "You're not supposed to do that."

Qualifications:

Attribute: 38. Reciprocates with adult

39. Reciprocates with other child

Definition: An exchange of favors with another.

Examples: (a) Child pulls another in wagon after other pulls target child in wagon.

(b) "If you help me pick up blocks, I'll help you put away dishes."

(c) Child helps other button smock after other has helped target child.

Qualifications:

Attribute: 40. Tries to "make up" with adult

41. Tries to "make up" with other child

Definition: Child attempts to "make up" with other after behaving in disapproved manner.

Examples: (a) Child bumps table, knocking another child's puzzle to floor. Target child then helps other child pick up pieces.

(b) Child refuses to meet teacher's request. Minutes later, he makes a special effort to please teacher.

(c) Child tries to comfort another whom he has made cry.

Qualifications: The rater must see both the initially disapproved behavior and the "make up" behavior.

Attribute: 42. Friendly toward adult

43. Friendly toward other child

Definition: Affiliative, mutually affectionate, facilitating behavior.

Examples: (a) Two children play together, with exchanges of positive communications and feelings.

(b) Target child makes an outgoing gesture to another by smiling or by saying, "Let's play with the beads."

(c) Child says to other: "You can play with me."

Qualifications: Unlike nurturance, for which benefit of other is primary concern, friendly behavior is affiliative.

Attribute: 44. Nurturant toward adult

45. Nurturant toward other child

Definition: Spontaneous efforts to help, give to, or reassure another.

Examples: (a) Child does something for teacher that is spontaneous and not ordinarily expected.

(b) Child spontaneously gives object to other child.

(c) Child helps other child do something.

Qualifications:

Attribute: 46. Exhibits leadership

Definition: A positive attempt to influence or control the behavior of another.

Examples: (a) Child initiates a game of ball by throwing ball to another child.

(b) Child says, "Let's play house. You be the father and I'll be the baby."

(c) Child oversees the construction of a block tower by several other children.

Qualifications:

Attribute: 47. Behaves competitively

Definition: Attempts to "outdo" another.

Examples: (a) Child rides tricycle faster in response to being passed by another child.

(b) Child tries to build block tower higher than another child's.

(c) Competes for teacher's attention.

Qualifications: Competitiveness for resources should be rated as "possessiveness" rather than "competitiveness."

Attribute: 48. Seeks leadership of adult

49. Seeks leadership of other child

Definition: Actively seeks leadership from other.

Examples: (a) Asks teacher to help him find a new activity.

(b) Follows other child and does what other child does or suggests.

(c)

Qualifications:

Attribute: 50. Smiles and/or laughs

Definition:

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 51. Engages in gross motor activity

Definition:

Examples: (a) Plays with large objects, such as large blocks, trucks.

(b) Engages in activities requiring physical strength, such as pulling heavy objects.

(c) Runs hard.

Qualifications:

Attribute: 52. Engages in fine manipulative activity

Definition: Interested in activities requiring finger-and-hand dexterity and eye-hand coordination. Includes peg boards, puzzles, beads, small block designs, and stacking and nesting toys.

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 53. Engages in cognitive activity

Definition: Includes working or playing with numbers, letters, words; "reading a book"; writing one's name; exploring nature; talking about an experience or asking questions for the purpose of better understanding experience.

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 54. Engages in fantasy activity

Definition: Engages in "make believe" behavior.

Examples: (a) Plays role of parent in doll play.

(b) Dresses up and/or acts like an animal.

(c) "Look, I'm Batman."

Qualifications:

Attribute: 55. Engages in artistic activity

Definition: Use of expressive media, such as crayons, finger or brush painting, work with clay or paper mache, musical instruments, etc.

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 56. Concerned with satisfaction of physical need

Definition: Child's attention directed toward satisfaction of physical need.

Examples: (a) Seeks food during periods other than snack time.

(b) Excessively tired; falls asleep.

(c) Sucks thumb.

Qualifications:

Attribute: 57. Takes initiative in carrying out own activity

Definition: Child knows what he wants to do; a "self-starter".

Examples: (a) A child works on several puzzles, returning each when he is finished, and choosing another.

(b) A child becomes engaged in a series of activities without any apparent direction from the teacher.

(c) After completing an activity, child easily finds another activity.

Qualifications:

Attribute: 58. Tries to pursue difficult task.

Definition: Child attempts to do something that is difficult. Signs of difficulty include (a) inability to perform the task quickly, (b) need to mobilize effort, (c) need to figure something out in order to accomplish task.

Examples: (a)

(b)

(c)

Qualifications: Child need not be successful with task in order to receive this rating.

Attribute: 59. Attempts to overcome obstacle by himself

Definition: When in the course of an activity the child faces some obstacle, he tries to overcome the obstacle by himself. An "obstacle" here refers to something blocking the activity that is not caused deliberately by another.

Examples: (a) When a tool is missing that he needs, he tries to find it without asking or disrupting others.

(b) When a piece of a puzzle is missing, the child searches for it.

(c) When a piece of furniture is in his way, the child moves the furniture.

Qualifications:

Attribute: 60. Exhibits persistence

Definition: Sticks to a task or course of action despite distractions or interference.

Examples: (a) Child works at a puzzle despite interesting distractions that might command his attention or interest.

(b) Child returns to building with blocks repeatedly after interruptions.

(c) Child returns to painting a picture after noticing another interesting activity occurring next to him.

Qualifications: In order to rate here, a distraction or interference must occur with subsequent persistence by child.

Attribute: 61. Completes activity by himself

Definition: Once a task is begun, the child carries it out to completion without requiring encouragement or help from other.

Examples: (a) Child works on a puzzle and completes the puzzle without seeking help.

(b) At the teacher's suggestion, the child starts to paint a picture. He works at the picture without seeking further encouragement from the teacher.

(c)

Qualifications:

Attribute: 62. Gets intrinsic satisfaction from activity or task

Definition: Child appears to enjoy activity or task for its own sake. Signs of intrinsic satisfaction are (a) wholehearted involvement and concentration on activity; (b) child expresses positive feelings while engaging in activity.

Examples: (a) Child is completely absorbed in painting a picture.

(b) Girl sits in corner rocking a doll, singing to self.

(c) Boy kicks ball along whistling.

Qualifications:

Attribute: 63. Praises self

Definition: Child expresses positive self regard.

Examples: (a) "I'm a good boy."

(b) "I'm strong."

(c) "I look pretty today."

Qualifications: This rating is an affirmation of self regard, not an attempt to elicit praise or approval from other.

Attribute: 64. Threatens to act aggressively toward adult

65. Threatens to act aggressively toward other child

Definition: Child threatens other with physical aggression. May be a verbal or gestural threat.

Examples: (a) "Stop doing that or I'll hit you."

(b) Child shakes fist at other, but stops short of physical contact.

(c)

Qualifications:

Attribute: 66. Possessive

Definition: Possessive attitude toward an object or resource.

Examples: (a) Child is unwilling to share or give up something in his possession.

(b) Child attempts forcefully to secure an object that is in another's possession.

(c) Child protects own block structure from all comers.

Qualifications:

Attribute: 67. Verbally aggressive toward adult

68. Verbally aggressive toward other child

Definition: A remark that expresses hostility, derogation, dislike, or rejection of other.

Examples: (a) "You're stupid."

(b) "No. I won't play with you. I don't like you."

(c) "Your picture is ugly."

Qualifications:

Attribute: 69. Bosses adult

70. Bosses other child

Definition: A negative (verbal) attempt to influence or control other.

Examples: (a) Child says to another, "You can't swing now; I have to go first."

(b) Child says to another, "If you won't be the baby, you can't play house with me."

(c) Child says to teacher, "Take these dishes back to the sink right now."

Qualifications:

Attribute: 71. Physically aggressive toward adult

72. Physically aggressive toward other child

Definition: Child actually makes physical contact with other in expressing aggression. Includes hitting, wrestling, kicking, pinching, pushing, biting, spitting, throwing object at another.

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 73. Deliberately aggressive agains. property.

Definition: Aggression directed toward objects and property. Includes disruptive throwing of objects, deliberate breaking of things, tearing up things, destroying products that other children have made or are working on.

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 74. Expresses negative feeling about self, possession, or own product

Definition: A negative, self-depreciating remark.

Examples: (a) "I can't do it."

(b) "Your dress is prettier than mine."

(c) "My painting is no good."

Qualifications:

Attribute: 75. Exhibits visual curiosity

Definition:

Examples: (a) Child looks at ongoing activities in classroom.

(b) Child looks at toys and games on a shelf.

(c) Child is attentive to a new or unusual event in classroom.

Qualifications: Visual curiosity is directed rather than aimless. It need not be accompanied by action, although a child may move from place to place in order to get a clearer view of object of attention.

Attribute: 76. Exhibits active curiosity

Definition: Active interest in a variety of ongoing activities or objects, including manual or verbal exploration.

Examples: (a) Child moves around room and "tries out" a variety of activities.

(b) Child goes to toy shelf and manipulates different toys.

(c) Child asks many questions on a topic.

Qualifications:

Attribute: 77. Seeks information from adult

78. Seeks information from other child

Definition: Asks a question of another child for the purpose of gaining information.

Examples: (a) To another child: "How did you make that house?"

(b) To the teacher: "When will it be Christmas?"

(c) To another child: "How old are you?"

Qualifications:

Attribute: 79. Responsive to teaching by adult

80. Responsive to teaching by other child

Definition: Child attempts to follow another's instructions, to master a skill being taught, or to modify a mistake pointed out by another.

Examples: (a) Child carries out teacher's instructions on how to hold paint brush.

(b) Child practices making a circle with crayon after another child has shown him how.

(c) Teacher shows child how a piece of puzzle is put in the wrong place, and child searches for correct place.

Qualifications. Note here the child's response to teaching of "subject matter" rather than classroom routines.

Attribute: 81. Imitates behavior of adult

82. Imitates behavior of other child

Definition: The child clearly tries to imitate or copy adult's behavior. By "imitation" is meant behavior that is clearly (a) stimulated by the behavior of the adult, (b) very similar to the behavior of the adult, and (c) occurs soon after the behavior of the adult.

Examples: (a) Child imitates clapping of teacher in game or song.

(b) Child copies a design made by the teacher in a demonstration of finger painting.

(c) Other child says, "I'm a horse." Target child says, "I'm a horse."

Qualifications: Merely following instructions does not qualify as imitation unless these instructions call for imitation of the instructor's behavior.

Attribute: 83. Instructs or demonstrates

Definition: The child tells or shows another how to do something.

Examples: (a) Child shows other how to do puzzle.

(b) "This is how you button your coat" (demonstrates).

(c) Child shows another child where the peg board is.

Qualifications:

Attribute: 84. Attempts to communicate verbally to adult

85. Attempts to communicate verbally to other child

Definition: Uses words and sentences in an effort to communicate with other.

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 86. Communicates meaningful complex idea to other child

87. Communicates meaningful complex idea to adult

Definition: Child communicates a train of thought, or a complicated idea.

Examples: (a) Reports an experience at home in some detail.

(b) Tries to tell other that blocks must be put in a certain sequence in order to build tower.

(c) Engages in fantasy play which incorporates a variety of different roles.

Qualifications: The train of thought or complicated idea may refer to something real or fantasied, but it should have some coherence.

Attribute: 88. Verbally loud

Definition: Makes sounds vocally and perhaps noisily to imitate some sound in nature or to express internal state.

Examples: (a) Shouts "zoom" as he runs around room.

(b) Pretends to shoot a gun, saying, "bang, bang."

(c) Shouts something to get other's attention.

Qualifications: Whether child vocalizes words or sounds, this rating refers to vocalization in the service of expression rather than communication.

Attribute: 89. Talks to self

Definition: Child delivers monologue or addresses remarks to non-human objects.

Examples: (a) "Here's a green wheel and here's another green wheel. I think I'll take another one."

(b) Child says "Get off me" to a piece of string that is clinging to his fingers.

(c) Child asks and answers his own questions. "What color shall I make this house? Red."

Qualifications: In order to make this rating, it should be clear that the child is attempting to communicate to himself rather than to other.

Attribute: 90. Difficult to understand

Definition: Child has difficulty making proper sounds for words, leading to difficulty in understanding.

Examples: (a)

(b)

(c)

Qualifications: Do not give this rating if child's speech is clear and rater doesn't understand child's vocabulary.

Attribute: 91. Does not concentrate on activity

Definition: An activity or task fails to sustain the child's attention, interest, and effort.

Examples: (a) Child begins to do puzzle, but leaves it after putting in one piece.

(b) Child flits from one activity to another without becoming involved in any one.

(c) Activity of another child turns child's attention away from picture he is painting, and child never returns to paint picture.

Qualifications:

Attribute: 92. Inattentive when adult communicates to him

93. Inattentive when other child communicates to him

Definition: Does not sustain attention toward other who is attempting to communicate with target child.

Examples: (a) Target child doesn't seem to listen when teacher instructs or communicates with him.

(b) Attention wanders when teacher is instructing a group of which target child is a member.

(c) Doesn't seem to listen to another child's attempt to communicate with him.

Qualifications:

Attribute: 94. Incomplete communicative act

Definition: Does not bring full attention to bear on other when communicating to other.

Examples: (a) Child says something presumably directed to teacher, but does not look at teacher while he is saying it.

(b) Answers teacher's question so softly that teacher cannot hear.

(c) Tries to communicate with other without first capturing other's attention.

Qualifications:

Attribute: 95. Exhibits goal-directed activity

Definition: Purposeful activity directed toward specific goal.

Examples: (a) Begins and completes a drawing.

(b) Makes a structure out of building blocks.

(c) Attempts to put beads on a string.

Qualifications:

Attribute: 96. Shows planning in pursuing activity

Definition: Child approaches activity or task in a careful, orderly, thoughtful manner indicative of planning. Evidence for planning may be seen in

(a) preparatory behavior which makes a task easier, such as putting all pieces of a puzzle face up before putting in the pieces, and

(b) doing things in an orderly sequence.

Qualifications:

Attribute: 97. Flexible in substituting goal

Definition: When a goal is blocked, child readily seeks or accepts a substitute.

Examples: (a) Child approaching tricycle gets there after another child has taken it. Target child turns to another activity.

(b) Child accepts an alternative task in response to teacher's suggestion.

(c)

Qualifications:

Attribute: 98. Corrects or modifies performance to meet own standard

Definition: Child modifies his behavior, apparently in accordance with his own standard, and without external pressure to do so.

Examples: (a) After examining his block structure from several angles, child rearranges several blocks.

(b) Saying, "This is a goof," child throws away drawing he has made and starts another.

(c) Child tries several doll dresses on a doll before deciding which she is satisfied with.

Qualifications:

Attribute: 99. Products or activities have common theme

Definition: Despite variation in child's specific activities, his actions and/or products contain a common theme or idea.

Examples: (a) Child makes engine sounds ("vroom") in moving crayons, truck, and his own body.

(b) Child paints several paintings that differ, but have similar content or form.

(c) Child plays "mother" in a variety of contexts.

Qualifications:

Attribute: 100. Perseverates on activity or task

Definition: Repeated performance of an activity or task beyond the point where the behavior appears to serve any goal beyond repetition itself.

Examples: (a) Child hammers aimlessly on a peg that is already in the hole as far as it can go.

(b) Stacks and unstacks dishes over and over again.

(c) Paces back and forth.

Qualifications:

Attribute: 101. Perseverates verbally

Definition: Repeated performance of a word or phrase or sentence beyond the point where the behavior appears to serve any goal beyond repetition itself.

Examples: (a) Repeats a phrase over and over again, without any effort to communicate with other.

(b) Sings part of song over and over again to self.

(c)

Qualifications:

Attribute: 102. Preoccupied with own thoughts.

Definition: Child appears to be responding more to his own thoughts than to external events; daydreaming.

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 103. Unable to tolerate delay

Definition: Impatience in getting or doing something.

Examples: (a) Told that he must wait in line and take his turn, child pushes in front of other.

(b) Impulsively does something too quickly to be successful.

(c) Tries to skip steps in doing an activity that calls for an orderly sequence.

Qualifications:

Attribute: 104. Concerned about physical discomfort or physical danger

Definition: Child exhibits concern, fear, or anxiety with regard to physical comfort, pain, or danger.

Examples: (a) Child is concerned about and "favors" a past injury.

(b) Child expresses concern about his physical safety.

(c) Child complains about a physical discomfort.

Qualifications: The child's concern may be either realistic or unrealistic, and should be rated here in either case.

Attribute: 105. Seeks verbal reassurance

Definition: Seeks reassuring remark or comment from other.

Examples: (a) "Do you like my picture?"

(b) "Am I doing this right?"

(c) Child starts an activity and hesitates, looking at the teacher as if to seek an expression of reassurance.

Qualifications:

Attribute: 106. Hesitant in relating to adult

107. Hesitant in relating to child

Definition: A tendency to hesitate or to avoid relating to an adult, or to a child or group of children. Hesitancy may occur with shyness or fear, or the child may vacillate between approaching and avoiding the adult.

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 108. Hesitant to try things on his own

Definition: A tendency to hesitate or to avoid doing things by himself. Hesitancy may occur with excessive cautiousness and fear, or the child may vacillate between approaching and avoiding an activity or task.

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 109. Unusually good physical coordination

Definition: Child is unusually well coordinated in use of large muscles, in eye-hand coordination, sense of balance, or rhythm.

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 110. Poor physical coordination

Definition: Child exhibits difficulty in physical coordination, either in use of large muscles or in eye-hand coordination.

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 111. Restlessness

Definition: Does not sit still, fidgets, paces.

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 112. Easily frustrated or threatened by adults

113. Easily frustrated or threatened by other children

Definition: Frustrates easily in response to actual or potential injury, blockage of activity, thwarting by other, or social threat initiated by another.

Examples: (a)

(b)

(c)

Qualifications:

Attribute: 114. Recovers quickly from frustration or threat

Definition: Response to frustration is neither prolonged nor severe. Frustration includes actual or potential injury, blockage of activity, thwarting by other, or social threat initiated by other.

Examples: (a)

(b)

(c)

Qualifications: Quickness of recovery from frustration should be rated independently of how easily the child is frustrated.

Attribute: 115 - 127. Response to frustration or threat

Definition: Frustration includes actual or potential injury, blockage of activity, thwarting by other, or social threat initiated by other.

Examples: (a)

(b)

(c)

Qualifications: Make a rating for all of the following kinds of response to frustration:

115. Becomes stubborn

116. Becomes fearful

117. Cries

118. Becomes dejected

119. Becomes defiant, rebellious

120. Increased quietness

121. Increased activity that seems aimless

122. Seeks comfort from adult

123. Seeks comfort from other child

124. Retaliates against person who caused frustration

125. Ignores the frustration or threat

126. Effectively defends self

127. Becomes angry

RATER BACKGROUND INFORMATION

(To be filled in by the local coordinator)

Check site:

____ Lee County
____ Portland
____ St. Louis
____ Trenton

Name _____

Sex _____ Age _____ Marital status: S__M__ Separated/Divorced _____

Highest grade attained _____

Ages of children (if any) _____

Plans for caring for children during work periods _____

_____

_____

General work experience _____

_____

_____

Special education, training or experience with young children _____

_____

_____

_____

Other relevant information _____

_____

_____

_____

JUDGMENTS ON POTENTIAL RATERS

(To be filled in by the local coordinator)

Check site:

____ Lee County
____ Portland
____ St. Louis
____ Trenton

Name of candidate rated _____

Rate the candidate on the characteristics listed below, using the following scale:

1. Very low or very weak
2. Moderately low
3. Moderate
4. Moderately high
5. Very high or very strong

___ 1. Conscientious motivation to carry out rating task

___ 2. Ability to arrange time to carry out rating task

___ 3. Experience with young children

___ 4. Potential rapport with teachers and school administrators

___ 5. Verbal skills

___ 6. Ability to work independently

Rank order this candidate in relation to all other candidates:

This candidate is ranked _____ out of a total of _____ candidates.

Is the candidate available from February 1 through May 16?

_____ Yes

_____ No (Explain)

Name _____

CHILD BEHAVIOR EXAMPLES

Some of the ways that children behave are listed below. Drawing on your own experience, write several examples of each behavior as it might occur in a young child. (Please take no longer than twenty minutes.)

1. Rebellious behavior:

2. Compliant behavior:

3. Dependent behavior:

4. Independent behavior:

5. Academically motivated behavior:

6. Aggressive behavior:

7. Affectionate behavior:

8. Rigid behavior:

9. Flexible behavior:

10. Purposeful behavior:

This instrument was developed pursuant to a grant funded by the Office of Economic Opportunity under the provisions of the Economic Opportunity Act of 1964 and is limited in its use solely for this research study. This use does not constitute publication. No distribution of the material is authorized outside of Educational Testing Service or the Office of Economic Opportunity.

Try-Outs of Personal Record of School Experience (PROSE)

Donald Medley

The principal focus of this report is on a pilot study conducted in Trenton in the spring of 1969, with the purpose of (1) determining the feasibility of using community housewives as PROSE recorders, and (2) gaining an experiential base for designing procedures for recruiting and and training recorders for the study proper. The bulk of the report is a fairly detailed account of what was done--how the recorders were recruited and trained, how the data was collected, and the plans for data analysis (programming difficulties have delayed completion of the analysis). For readers not interested in the details, the following brief summary, emphasizing what was learned, is presented.

Recruiting procedures need to be somewhat more selective than those used in the pilot study (which were quite lax); experience indicated that some housewives can learn to record behaviors on PROSE, and some cannot. Training procedures need to be revised so that the system is learned in smaller increments, with the easier portions studied first. This will involve re-editing the training materials, particularly the videotapes used.

The data collection plan adopted involved a principal recorder with two major responsibilities: (1) collecting and auditing all completed PROSE records, and (2) making a scheduled round of visits during which only those pupils missed by a regular recorder were observed, so that missing data could be acquired. The first part of this plan worked well enough to be continued; the second should probably be abandoned. Instead, each recorder should be scheduled to obtain the missing data herself.

No evaluation of the scoring procedures and analytical design can be made at this point.

Training Procedures

The pilot study for PROSE took place in Trenton in the spring of 1969. Recorders were community women from urban Trenton hired by the local

coordinator. Three women were hired to work half-days observing children in Head Start centers, and one other, the "principal recorder," was engaged to work full time. In addition to collecting and checking all completed records before turning them in, the latter was responsible for making any observations missed by other recorders.

The recorders were ready to begin collecting data after eleven half-days of training. A shorter training period was planned initially, but the results obtained indicated a need for more practice than was originally thought necessary.

Day 1. Training began with a description of the intent and purpose of the longitudinal study and the role of the PROSE instrument. The rationale of the instrument was then discussed briefly.

Some of the trainees were not familiar with Head Start classrooms, so a brief introduction to the types of behavior to be encountered and the structure of the classroom was necessary before actual training could commence.

Training began with study of the context side of the PROSE answer sheet which describes generally the classroom setting and climate. Each word and its alternatives were described and defined. At the same time, the answer sheets were presented and the observers were shown the mechanics of coding. Examples of some of the behaviors described on the context side were then shown for practice in coding. The training tapes were made in an actual Head Start center and were presented audio-visually. The trainees each received a copy of Definitions of Prose Items (see Exhibit A) for further study and reference.

Day 2. The second day began with a review of the context side with more classroom examples. Special emphasis was placed on increasing speed of coding as well as reaching a higher level of competency.

Following this review, the statement side of the instrument, which deals with the specific interaction in which the target child is involved at the moment of observation, was introduced. First, the three possible types of interactions were defined: child-adult, child-child, and child-material. Then definitions for the alternatives to all the words were presented.

In retrospect, this method of instruction does not seem to be the most efficient way of introducing the statement side. The large number of choices to be learned simultaneously confused the trainees, and their impression of the complexity of the statement side was difficult to dispel. A better method of presentation might be to first introduce the recorders to interactions involving child-material contacts. Each word would be defined separately and examples would be given until a child-material interaction could be coded in its entirety. Child-adult and child-child interactions would then be presented similarly, thus adding gradually to the complexity of the coding task.

Day 3. The third day began with a review of the statement side, followed by examples of longer classroom sequences in which the recorders observed the behavior of one child for a total of five observations, as they would do later in the field. When the recorders seemed reasonably skilled in using the statement side of PROSE, more of the longer sequences were shown and they were instructed to code both back and front as they would in the classroom.

Days 4-7. The next four days consisted of two trips to the Carolyn Stokes Day Nursery in Trenton for live practice in the classroom. Each trip was followed by a day of questions and discussion.

Day 8. On the eighth day the trainees were given a test consisting of several timed sequences in which several individual children were observed, each one for five successive intervals. The recorders' answer sheets were checked for errors in coding, after which each trainee had a private conference with the instructor to discuss individual coding problems.

Day 9. On the ninth day, the Recorder's Guide was distributed (see Exhibit B). Recorders were instructed in procedures for preparing for classroom visits, and keeping account of answer sheets afterward. These procedures included sequencing and coding control cards and answer sheets; seeing to equipment needs for the classroom; observing tasks and expected conduct in the classroom; and finally "cleaning up," or the gridding and checking of answer sheets to identify the recorder and the day on which the record was made.

Day 10. On the tenth day, the trainees visited a regular Head Start classroom for a "dress rehearsal," during which they carried out the duties they would be expected to perform in the actual study.

Day 11. The eleventh day began with more discussion and questions followed by distribution of schedules and prepunched subject control cards. A brief review of the Recorder's Guide ended this final day of training.

Data Collection

Schedules for visiting the six Head Start classrooms were set up so that no two recorders would be in a classroom at the same time, and so that, after twelve days, each recorder would have visited each classroom twice. At the end of several days of observation it became evident that the recorders were still making a number of errors in coding. A brush-up session was held at ETS at which major errors were discussed. Data collected up to this point was discarded and twelve additional days of observation were scheduled (see Exhibit C).

The Principal Recorder

The principal recorder followed a schedule similar to that of the regular recorders, but she observed only those children who had been absent when the classes were visited by one or more of the other recorders. In addition to her work in the classroom, the principal recorder was responsible for collecting completed records from the other recorders and thoroughly checking all data. This included making sure that each child's answer sheet was accounted for, as well as checking for coding errors and errors in identification data. A check list (see Exhibit D) was used to simplify the process. She was also responsible for communication between ETS and the other recorders in the case of any questions or problems.

Once a week, completed and checked data was brought to ETS. Spot checks of the records revealed many mistakes still occurring. All data had to be rechecked and corrected after which the answer sheets were sent to SCRIBE.

Again, in retrospect, it might have been better for recorders to secure their own missing data. This would make for purer and, hopefully, more complete data collection and also leave the principal recorder freer for her other duties.

Following SCRIBE processing, the PROSE data was sorted and checked once more for error. Then, a final error-free tape was created for scoring and analysis.

Scoring

A multi-purpose scoring program is being designed for use with this and future PROSE data. At its simplest level, the program will compute the frequency per record of each item alternative on the statement side of the instrument. The frequencies are then reported and stored on tape or cards in this manner: RECORD IDENTIFICATION, SCORE 1 (INIT); SCORE 2 (STAR): etc., up to SCORE 50 (NEG).

For more elaborate scoring, a second program alternative will enable the user to specify scoring keys based on any combination of items from either side of the instrument. The user may choose how many and which statements to include in each key, as well as the weight to be assigned to each statement. The following sample keys illustrate some of the flexibilities of the scoring program (see Exhibit A for an explanation of the abbreviations used):

Key 1: Hostility

Statement	Weight	Side 1	Side 2
1	+1	4-AGG 5-AGG	
2	+1		6-Lost Temper
3	-2		6-Shaff p

Key 2: Seeks physical contact

Statement	Weight	Side 1	Side 2
1	+1	1-INIT 3-POS	6-CNTC 3-NEXT

Table 1

DESIGN FOR PROSE ANALYSIS

Let: N = total number of pupils

O_u = number of different observers seeing same pupil

$$O = \sum_{u=1}^{N} O_u$$

V = total number of observations of all pupils by all observers

Source	d.f.	MS
Teachers	2	a
Classes (within teachers)	3	b
Pupils (within classes & teachers)	N-6	c
Observers (within pupils)	O-N	d
Visits (within O, P, C, & T)	V-O	e
Total	V-1	

Key 3: Traditional teaching role

Statement	Weight	Side 1	Side 2	
1	+1		2-a	4-e
2	+1		2-b	4-e
3	+1		2-c	4-e

Analysis

Scores on each item, and on a number of keys like the examples cited above, will be submitted to an analysis of variance in an hierarchical design as in Table 1. Rough estimates of two reliability coefficients will be calculated from each analysis:

$$r_1 = (c-d)/c$$

$$r_2 = (c-e)/c$$

The first coefficient estimates the correlation between records made by different coders observing the same pupils at different times; the second estimates the correlation between records made by the same coder observing the same pupils at different times. The discrepancy between the two would indicate the amount of observer agreement.

These formulas produce reliability estimates for items or keys used to discriminate pupils from one another. Similar formulas involving the "Teachers" mean square, a, in place of the "Pupils" mean square, c could be used but the number of teachers observed was too small to work with.

EXHIBIT A

DEFINITIONS OF PROSE ITEMS

CONTEXT SIDE

1. INSTRUCTIONAL CONTENT

The choices which describe what the pupil is supposed to be learning are marked here.

ARITH (arithmetic) is marked when number skills are being studied.

ARTCRF (art or craft) is marked if the pupils are learning how to draw, paint, model, or do any other art or craft.

EXGM (exercise, game) is marked when the teacher has the pupils exercising, or playing a game.

FRPL (free play) is marked when the pupils are free to choose their own activities.

HLTHYG (health, hygiene) is used when the teacher is teaching the pupils habits of good health or hygiene.

LANG (language) is checked when any language skill is being taught--such as learning the letters of the alphabet or connecting words with objects.

MUSRTH (music, rhythm) is marked when the teacher is trying to teach pupils something about music--a new song or how to play an instrument, for instance. If singing is being used to teach vocabulary, health habits, etc., MUSRTH is not marked.

RSTSNK (rest, snack) is marked when the children are eating, drinking, or resting.

RTNTRN (routine, transition) is used when the teacher is trying to organize or reorganize the class, as when she moves from one activity to another.

SCI (science) is marked when the teacher is trying to help pupils learn about nature--as when she teaches them about plants, trees, animals, etc.

SENS (sensation) is marked when the teacher is trying to teach students to recognize different colors, shapes, sounds, feelings, etc.

SOSK (social skills) is marked when the teacher tries to help students learn rules of behavior, manners, ethics, etc.

SOST (social studies) is marked when the pupils are learning about people--their roles, their relationships with others--and the society in which they live.

2. ADULT ROLE

Item 2 is used to record the roles played by adults in the classroom. Boxes T and O are used to indicate whether the teacher (T) or another adult (O) is playing a particular role.

SHOW is marked when the adult is showing the class something.

TELL is marked when the adult is telling or reading the class something.

LEAD is marked when adult and pupils are doing the same thing under the teacher's direction, as when she leads them in games or songs.

DSCS (discuss) is marked when both adult and pupils speak to one another.

PEER is marked when the adult acts like a pupil--is a member of a group but not its leader.

MNG (manage) is marked when the adult is giving directions or orders and the pupils are carrying them out.

SPVS (supervise) is marked when the pupils are working and the adult moves around, watching and offering help or suggestions whether asked to do so or not.

RSRCE (resource) is marked when the pupils are working and the adult is available for help when asked, but does not go to the pupil and offer it.

HSKP (housekeeping) is marked when the adult is not interacting with pupils but works with materials on such public tasks as posting items on the bulletin board, straightening up, or preparing a snack.

INDATT (individual attention) is marked when the adult spends a long time with one pupil while she is not available to others.

LSWT (listening, watching) is marked when the adult is not taking part in what the pupils are doing, but is listening or watching them.

NCNT (no contact) is marked when the adult is not paying attention to the pupils at all.

3. DISTANCE

Psychological distance is used to record how near or far from others the pupil being observed was during the cycle. This item may be left blank when no choice is appropriate.

NEXT is marked when the pupil being observed is closer to the adult than any of the other pupils and the adult is aware of his closeness.

NEAR is marked when the pupil is closer to the adult than any of the others, but the adult does not seem to notice his nearness.

MID (middle) is marked when the pupil is in the middle of the group.

FRNG (fringe) is marked when the pupil is on the edge of the group.

OUT (outside) is marked when the pupil has left the group he was in.

4. GROUP STRUCTURE

This item is used to indicate the size of the group in which the child spent the cycle.

Group size is indicated by marking whichever choice fits the number of pupils in the group: ONE, TWO, 3-5, 6+ (i.e., more than six but not all of the pupils in the room), and ALL (when the whole class is organized as one group). The mark is made at AD if an adult is a member of the group; at NA if no adult is present. For example, if a child is talking to the teacher with no other child nearby, ONE would be marked at AD.

5. CLASSROOM CLIMATE

This item is used to record the behavior of the group as a whole during the cycle. Use this item only when one of the choices is clearly appropriate.

ATT EXC (attentive, excited). The class is listening closely to the teacher or concentrating on the work it is doing, but there is a feeling of eagerness or suspense.

ATT TNS (attentive, tense). Again the class is listening closely or concentrating on work, but there is a feeling of anxiety or fear.

NSY EXC (noisy, excited). The class is in somewhat of an uproar, many or most of the pupils are not attentive or at work, but are talking, shouting, and probably moving about in a disorderly fashion.

NSY BSY (noisy, busy). The class may appear in a slight uproar, but pupils are busy with assigned tasks which need considerable talking.

QU BSY (quiet, busy). The class is orderly and quiet; pupils are working.

QU IDL (quiet, idle). The class is orderly and quiet but few or none of the pupils are working.

6. PUPILS SIGNS (marked only if they happen to the pupil you are watching.)

USED NUMBERS	used numbers
USED WORDS	spoke or wrote
SNG, TKD TO SLF	sang or talked to himself
HELPED OTHER P	helped another pupil
COMFORTED P	comforted, sympathized with another pupil
SHAFF P	showed affection for another pupil
ASKED P FOR HELP	asked another pupil for help
REC'D HLP, AFF	received affection or help from another pupil
REJECTED BY GP	was rejected by a group (excluded from it)
ACCIDENT, HURT	had an accident or was hurt
SHOWED FEAR	showed fear
CRIED	cried
LOST TEMPER	lost temper
TATTLED	tattled
LED OTHER PUPIL	led another pupil
BOSSED OTHER P	bossed another pupil
WRECKED SOMETHING	wrecked something
REFUSED HELP	refused to be helped
RESISTED AD	resisted adult (teacher, aide, etc.)
DISOBEYED	disobeyed adult
SHOST TO AD	showed hostility toward adult
RID	was held up to ridicule by adult

BAD EX	was held up to others as a bad example
GOOD EX	was held up to others as a good example
CHORE, ERRAND	was asked by adult to do chore or run errand
SHAFF AD	showed affection for adult
WAITED	pupil waited for a turn in line or for an adult's attention

7. ADULT SIGNS (if the pupil you are watching is directly involved, mark if he was not involved but watched, mark at W).

CLD 4 QT	adult called for quiet
THRT	adult threatened pupil(s)
LSTMP	adult lost temper
LFTW	adult laughed with pupil(s)
PHRSTR	adult restrained child physically

8. MATERIALS AND LOCATIONS Mark each type of material used by the child at M. Mark the place where he uses materials at L. If the materials and the location are the same, make only one mark at M.

ART (PAINT, CLAY, ETC.)	paint, clay, easel, crayons, etc.
CRAFT	scissors, paste, beads, construction paper, etc.
MUSICAL INSTR	musical instrument
BOOKS, WRITING	books and writing materials
PUZZLES, QU GAMES	puzzles, "quiet" games
BLOCKS	blocks, construction toys
WHEEL TOYS	trucks, cars, trains
TOOLS, WORKBENCH	tools, workbench
SAND	sandbox, sandpile, sandtable
GYM, EXERCISE	jungle gym, ball, bat, etc.
PET	guinea pig, fish, mouse
CLOTHES, JEWELRY	clothing, jewelry, shoes
FOOD, WATER	dishes, food, water
DOLLS	dolls, stuffed animals, doll house, etc.
HOMEMAKING	stove, table and chairs, dishes
CLEANUP TOOLS	broom, dustcloth, dustpan, etc.
AUDIO-VISUAL DEV	audio-visual device--record player, projector, etc.

SP INST DEVICE	special instructional device--teaching machine, etc.
SCIENCE EQUIP	science or laboratory equipment
ASSIGNED SEAT	place apparently assigned to child as his own
TEACHER'S DESK	teacher's desk
OPEN AREA	open area (usually in center of room)
TOILET	toilet
HALL	hall outside classroom
PLAYGROUND	playground

STATEMENT SIDE

1. LEVEL OF PARTICIPATION

One of the four choices to the first item is marked when the pupil is observed in contact with an adult when the timer sounds.

INIT (initiating) is marked if the child asks the adult for more attention than he has been receiving.

STAR is marked if the child being observed is receiving more attention from the adult than any other child.

PART (part of a group) is chosen if the adult is paying equal attention to the child and other children in a group.

LSWT (listening and watching) is marked if the child is not trying to get the adult's attention, but is watching or listening to her while her attention is somewhere else.

2. ADULT IDENTIFICATION

This item describes the adult to whom the child was paying attention. If more than one adult is involved, the mark should be placed in the first appropriate category.

TCHR is checked when the teacher holds the pupil's attention.

AA (adult aide) is used if the child pays attention to an adult aide, assistant teacher, or teacher's helper.

TAA (teenage aide) is marked when a teenage aide or older child has the pupil's attention.

OBS (observer) is checked when the child's attention is on the observer.

OTH (other) is marked when the child's attention is on any other adult.

3. ADULT BEHAVIOR

The third item is marked only if the first and second words are used to describe what the adult is doing. It is left blank if the adult ignores a pupil initiation.

POS (positive) is marked if the adult is expressing consideration or affection for a child.

PRM (permissive) is marked if the adult is encouraging the child to choose for himself what he will do or allowing him to do something he wants to do.

SHTL (showing and telling) is marked when the adult is observed "teaching"--that is, when he acts like a teacher.

LSQU (listening and questioning) is marked if the teacher is either asking a question or listening to a pupil.

DO4 (doing for) is marked if the teacher does something for the child that he has been trying, or wants to do for himself.

CNTR (control) is used if the adult directs the child to do or stop doing something.

NEG (negative) is marked if the teacher scolds the child or punishes him.

4-5 PEER INTERACTION

If when the timer goes off the recorder observes the pupil paying attention to another pupil rather than to an adult, the first three items are left out and the recorder goes directly to the fourth item.

AGR (aggression) is marked if the pupil begins a fight with another child or acts in an unfriendly way--pushing, grabbing, etc.

INIT (initiation) is used if the pupil tries to get the attention of another child in a friendly manner.

COOP (cooperating) is marked if the child responds in a helpful manner to a classmate.

WTHD (withdrawal) is used if the pupil runs away or tries to avoid another pupil.

RST (resisting) is marked if the pupil is fighting back or refusing the initiation of another pupil.

If none of these apply, the item may be left blank. The behavior of the other pupil is recorded in the fifth item in exactly the same way.

6. CONTENT OF INTERACTION

This item is used whenever suitable to describe an adult-child contact.

CONTC (contact) is marked if at the moment the timer sounds, the adult and child are touching each other.

MTL (material) is marked if there is no actual touching, but some material is exchanged or used by both adult and pupil.

VRB (verbal) is marked if the exchange involves neither touching nor materials, but the adult or child is speaking.

7. INTERACTION STRUCTURE

The seventh item is used in adult-child or child-child contacts to indicate the sex and ethnic group of any person a child is in contact with.

OSOG (other sex, other group) is marked if the other person is of different sex from the child and of a different ethnic group.

OSSG (other sex, same group) is marked if the other person is of the opposite sex but the same ethnic group.

SSOG (same sex, other group) is marked when the other person is of the same sex, but of another ethnic group.

SSSG (same sex, same group) is marked if the other person is of the same sex and the same group.

8. LEVEL OF INVOLVEMENT

The eighth item is used in adult-child or child-material contacts to indicate how much attention the pupil is paying to what is going on.

COOP (cooperating) is marked if the pupil is "doing what he is supposed to."

DSTR (distracted) is marked when the pupil is watching or listening to someone or something while he is supposed to be doing something else.

RIS (responding to internal stimuli) is marked when the pupil does not seem to be aware of anyone else, but is daydreaming or otherwise withdrawn.

WOA (working on other activity) is marked when the child is seen concentrating on some activity other than the one he should be doing.

DSRP (disruptive) is marked when he is disturbing others.

9. ACTIVITY LEVEL

The ninth item records the level of physical activity of the child, and whether or not he is changing his location.

HIWL (high activity, with locomotion) is marked if the child is moving about the room vigorously.

MDWL (moderate activity, with locomotion) is marked when the child is moving about the room slowly.

HINL (high, no locomotion) is used if the child stays in one location while moving vigorously.

MDNL (moderate, no locomotion) is used when no locomotion is observed and the child moves slowly.

LOW is used when only slight arm or head movement is observed.

If no movement is observed, this item may be left blank.

10. ACTIVITY

The tenth item is used to record the kind of contact the child has with material.

FANT (fantasy) is marked if the pupil is acting out a fantasy or pretending.

DVG (divergent) is marked if the child is doing something which has a definite end that he seems to have decided upon himself.

CVG (convergent) is marked if an activity can only be done in a certain way which is specified by the teacher or the material.

WRK (work) is marked if the child's purpose is to do a useful task.

KIN (kinesthetic) is marked if the pupil's activity is repetitive or unplanned.

11. MANIFEST AFFECT

The eleventh item is used to record the pupil's strong feelings. If none occur it may be left blank.

POS (positive) is marked when the child laughs or seems especially happy.

NEG (negative) is used if the child cries or is clearly very unhappy.

EXHIBIT B

RECORDER'S GUIDE

GENERAL

As a PROSE recorder, you have two important jobs to perform. First, you must visit the classroom, observe the children, and record behaviors on the PROSE answer sheet. In addition, you must prepare for classroom visits beforehand and keep account of the answer sheets afterward.

PREPARATION

Before entering the classroom, you must arrange control cards for the class to be observed in proper order. Do this by matching the numbers in the upper right-hand corner of the control cards with the numbers on the sequence list. When all the cards are in order, place them in your notebook with a PROSE answer sheet behind each one. (Answer sheets should be statement side up.) Then, fill in the subject code number in the first seven spaces of the answer sheet just as it is punched on the control card. Make sure that the holes are lined up exactly with the numbers underneath.

EQUIPMENT

Bring the following things with you every day to the classroom:

1. Your notebook containing control cards and answer sheets for that class.
2. Your timer with its earplug.
3. One battery and a small screwdriver.
4. Three sharp pencils.

CLASSROOM PROCEDURE

When you first arrive, ask the teaching assistant for help in locating the pupils you are to observe. Once you have found a child, write down a brief description of him--his hair color, size, clothing or any other special features--to help you recognize him later.

If any of the children are not present, remove their control cards and answer sheets, and place them in the back of the notebook. Then, turn to the first child, start the timer, and begin coding.

You will record five statements (one for each click) under CYCLE A on the statement side of the answer sheet, and then you will fill out CYCLE A on the context side of the form. Turn to the next child and continue coding until all the children have been seen once. Then, take a fifteen-minute rest. After this break, turn again to the first child and begin again coding behaviors under CYCLE B. When CYCLE B is completed for all children, the visit is over.

GENERAL GUIDELINES

1. Try not to disturb the classroom. Do not encourage the children to watch you or play with you. If you try to ignore their attempts to get your attention, you will find that they soon forget your being there.

2. Try to follow the child you are coding as much as possible. If he wanders out of the room, keep track of him. If the children are in the playground, go outside also. If you follow, we will have a record of the child during rest time, snack time, gym, and free play.

3. If a child comes in late, after you have placed his sheets at the back of the notebook, but still during CYCLE A, place him after the others and code him last. If the child comes in during CYCLE B, regard him as absent.

4. If a child is not in sight for some reason, and you do not know where he is when you are to code him, go to the next child and return to him when he reappears. If he disappears after you have begun a cycle, continue timing him, leaving the statement side blank and coding "out of room" on the context side.

CLEANUP

After the visit, answer sheets must be gridded to identify you, the recorder, and the day of observation.

1. For each of the pupils you actually did observe, grid the following information before removing the forms from your notebook.

1. Subject Code has already been filled in. Check it.

2. Fill in, in the box provided, the Observer Code given to you during training. Then blacken the number in the space above.

3. In the column marked Round Number, blacken the number given to you during training. (In the winter and spring of 1969, the round number is 1.)

4. In the column marked Week Number, blacken the number of the week in which your visit was made. For example, during the first week, blacken "1" and during the third week, blacken "3".

5. In the column marked Day Number, blacken the following numbers:

   #1 if the observation is made on Monday

   #2 if on Tuesday

   #3 if on Wednesday

   #4 if on Thursday

   #5 if on Friday

6. The Sequence Number the number of the order in which you saw the child. For the first child you observed blacken the number "01," for the second child, blacken the number, "02" and so on until the last child. Note that there are two columns and both must be filled.

When you have finished marking the front of the answer sheet (work on only one at a time), remove it and the control card with it from the notebook. Then, grid the back of the sheet in exactly the same manner, using the holes in the control card to help you to grid the Subject Code. After you have checked all your marks, place the completed answer sheets in one of the envelopes labeled "completed forms." When all the answer sheets have been marked and placed in the envelope, label the envelope with your name and the name of the Head Start class, and circle the visit number.

2. For those pupils who were absent, place their answer sheets in the envelope marked "incomplete forms." Do not grid any more information on these sheets; they should have only the Subject Code gridded on one side. Label the envelope with your name, the Head Start class, the visit number, and the names of the absent children. These envelopes for incomplete forms may be used for five visits, but be sure to label each visit.

Note that if any of the children were seen for only one CYCLE, they should be treated as if they were absent.

When you have finished with the recording forms, take the control cards for the class and set these aside for the next visit.

SEQUENCE LIST

For the first visit to each class, put your subject control cards in this order.

2
7
4
5
1
3
6
9
14
15
11
10
12
13
8

For the second visit to each class, put your subject control cards in this order.

13
15
12
8
10
11
14
9
3
2
6
5
4
7
1

NOTE: This distribution is for a class of 15. Similar random distributions are prepared for classes of a different size.

EXHIBIT C

PROSE

Schedule of Visits by Four Recorders to Six Classes

Date	Recorders			
	1 Miss Berry	2 Miss Stewart	3 Mrs. Martin	4 Mrs. Batts*
2/26	A	B	C	
2/27	D	C	F	
2/28	C	A	D	
3/3	D	E	B	A***
3/4	F	A	D	C
3/5	B	F	A	D
3/6	E	D	C	F
3/7	School Holiday			
3/10	F	E	B	A***
3/11	B	C	A	E
3/12	A	F	E	B
3/13	E	D	F	C
3/14	C	B	E	F
3/17	Make-up**	Make-up	Make-up	E
3/18	Make-up	Make-up	Make-up	D
3/19	Make-up	Make-up	Make-up	B
3/20				
3/21	Viewing reliability tapes			
3/24				

A	Grant	AM)	
B	American Legion	AM)	AM = 8:15 - 11:15
C	St. Paul	AM)	
D	Grant	PM)	
E	American Legion	PM)	PM = 12:15 - 3:15
F	St. Paul	PM)	

*Mrs. Batts observes only those pupils absent on one or more previous visits by other recorders.

**Make-up means visiting a class which was not available on the scheduled day.

***Classroom observed from observation booth in Grant School.

EXHIBIT D

DATA CHECK LIST

Check for complete forms I for incomplete forms
Initial for make-up forms

Code Number	1 Miss Berry VISIT 1	1 Miss Berry VISIT 2	2 Miss Stewart VISIT 1	2 Miss Stewart VISIT 2	3 Mrs. Martin VISIT 1	3 Mrs. Martin VISIT 2
1011						
1021						
1031						
1041						
1051						
1061						
1071						
1081						
1091						
1102						
1112						
1122						
1132						
1142						
2151						
2161						
2171						
2181						
2191						
2201						
2211						
2221						
2231						
2241						
2252						
2262						
2272						
2282						
2292						

Summary of Pretest Populations for Measures of 3 1/2-Year-Olds

The pretest populations for the final forms of the instruments administered in the study are listed in Table 1. Considerable pretesting of earlier versions was carried out. The pretest populations for the earlier versions are not reported.

Moreover, special field tryouts were designed to provide additional instrument development information. Prior to the "dry run" in each of the four study sites, the feasibility of the use of the final instrument manuals by indigenous testers and of the administration of the entire battery was studied. A total sample of 120 3 1/2-year-olds, living outside of the target school districts, was drawn. This sample was similar to the expected longitudinal study group with respect to racial and SES characteristics. As a result of this feasibility study, minor revisions were made in the order of administration of the tests in the battery, in the wording of some of the manuals, and in the design of some of the answer sheets. The feasibility study also made possible a smoother operation of the local testing centers.

The information obtained through the feasibility study was verified in each of the four sites by a "dry run." Thus the final battery of tests was administered to an additional total of 60 subjects, again residing outside of the target school districts.

Table 1

MEASURES AND PRETEST POPULATIONS

(3 1/2-YEAR-OLDS)

Affectionateness (a Fels behavior rating scale used in several longitudinal studies) Several hundred low-income white, black (urban and rural)	(Shipman)
Brown IDS Self Concept Referents Test Several hundred low-income white, black (urban and rural)	(Shipman)
Child Cooperation Several hundred low-income white, black (urban and rural)	(Shipman)
Children's Auditory Discrimination Inventory 234 low-income (some middle-income) white, black (urban)	(Gordon)
Cooperative Preschool Inventory (Caldwell) Standardized on Head Start children	(Tanaka, Shipman)
Cooperative Preschool Inventory Supplement (Caldwell) (Includes items from the Preschool Inventory and the Developmental Test of Visual-Motor Integration) 28 3-year-olds (suburban) 59 4-year-olds (suburban) 113 5-year-olds (urban, rural, and suburban)	(Gordon)
ETS Matched Pictures Language Comprehension Task (New prepositional phrase items) 35 low-income white, black (small city) 100 low-income, 100 middle-income pre-K and K (urban)	(Bussis, Tanaka)
ETS Story Sequence Task, Part 1 100 low-income, 100 middle-income pre-K and K (urban)	(Tanaka)
Enumeration Task 12 low-income and middle-income white, black (small city)	(Chittenden, Tanaka)
First Day of School Question About 2,000 low-income white, black (urban and rural)	(Shipman)

(Table 1 continues)

Fixation Task 60 3 1/2-year-olds low- and middle-income white, black (small city); also used by other professionals	(Lewis)
Hess and Shipman Eight Block Sorting Task 160 black 4-year-olds and their mothers (40 each: upper-middle, upper-lower, lower-lower, and ADC) (urban); also used by other professionals with middle-income white and low-income black	(Shipman)
Hess and Shipman Etch-a-Sketch Interaction Task 160 black 4-year-olds and their mothers (40 each: upper-middle, upper-lower, lower-lower, and ADC) (urban); also used by other professionals with middle-income white (small city)	(Shipman)
Hess and Shipman Toy Sorting Task Several hundred low-income white, black (urban and rural); black 4-year-olds and their mothers (40 each: upper-middle, upper-lower, lower-lower, and ADC)	(Shipman)
Johns Hopkins Perceptual Test Over 200 4-year-olds through 9-year-olds (predominantly 4-year-olds through grade 1) low- and middle-income.	(Gordon)
Massad Mimicry Test 45 2 1/2-year-olds to 4 1/2-year-olds low-income white, black (small city)	(Massad)
Matching Familiar Figures Test 30 low-income white, black (daycare center, small city)	(Lewis, Ward)
Mischel (Delayed Reward) 160 low- and middle-income black, 200 4-year-olds low-income white, black (urban)	(Shipman)
Motor Inhibition Test 24 3-year-olds and 4-year-olds low-income (daycare center, small city); standardized by Maccoby; also used by other professionals with varied populations	(Ward)
Open Field Test 30 low-income white, black (small city and rural) to pretest scoring procedure	(Ward, Shipman)

(Table 1 continues)

Peabody Picture Vocabulary Test (Jungeblut)
30 low-income white, black (daycare center, small city); national standardization

Picture Completion Test, WPPSI (Dermen)
7 3 1/2-year-olds low-income white, black (urban); national standardization

Preschool Embedded Figures Test (Dermen)
20 low-income black (urban), about 55 3-, 4-, and 5-year-olds middle-income predominantly white (nursery school, small city)

Questions for Mothers (Shipman, Ryan)
30 low-income white, black mothers (small city)

Risk Taking I (Guessing Birds and Cats Task) (Lewis)
17 low-income black (daycare center, small city)

Risk Taking II (Guessing Games II) (Lewis)
12 low-income black (daycare center, small city)

Seguin Form Board (Gordon)
29 4 1/2-year olds, 83 5- and 6-year-olds low-income white, black (urban)

Sex Role Constancy (Emmerich, Goldman)
30 low- and middle-income white, 20 low-income black (small city)

Sigel Conceptual Style Sorting Task (Shipman)
500 low-income white, black (urban)

Spontaneous Numerical Correspondence (Chittenden, Tanaka)
20 low-income black (daycare center, small city)

TAMA General Knowledge Test (Ahrens, Tanaka)
30 low-income white, black (daycare center, small city)

Vigor I (Running) (Lewis)
30 low-income white, black (daycare center, small city)

Vigor II (Crank Turning) (Lewis)
30 low-income white, black (daycare center, small city)

APPENDIX B

WORKING PAPERS

Models of Continuity and Change in Development
Walter Emmerich

. . . . .

Protection of the Rights and Privacy of Human Subjects
Ann Jungeblut, for the Steering Committee

. . . . .

The Ghetto: Some Notes on Environmental Pressures and Educational Atmosphere
Fred Damarin, Irene Kostin

Models of Continuity and Change in Development*

Walter Emmerich

From a developmental standpoint, psychological functions are viewed as systems which grow, differentiate, become organized, and reorganized throughout childhood and beyond. These developmental processes are the central phenomena of our field, and research efforts are devoted in large part to bringing them into full view (Wohlwill, in preparation). There are three ingredients of research designs which greatly enhance our chances of discovering and understanding the basic phenomena of development. One such ingredient is longitudinal data, with repeated measures on the same persons taken at several age periods. While all psychological measurement deals with discrete behaviors assessed at a given point in time, development refers to processes that link behaviors over time and in the same persons.

A second desirable ingredient is assessment of multiple measures of the same construct. This is partially a matter of sampling a range of behaviors signified by the construct, but we also need to take account of behavioral differentiation in the course of development, representing orderly age variations in the way that underlying psychological functions are expressed. It is therefore important to derive multiple age-graded behavioral indicators of the same construct, and to assess the appropriate subset of indicators at each age period.

The ideas of behavioral organization and reorganization in development press us to carry this analysis one step further, and to add a third feature to developmental designs. Not only are the indicators of the same variables believed to change with age, but the very nature of the underlying constructs themselves and their interrelations may be subject to developmental transformation. Here we must move to a higher level of analysis, calling for multiple construct measurement at each age period. Ideally, this is done through multivariate designs which sample a domain comprehensively so that the most powerful configuration ordering the domain as a whole can be revealed at each age and compared across age periods.

---

*Paper presented as part of a "Symposium on Behavior Continuity and Change with Development" at the Biennial Meeting of the Society for Research in Child Development, Santa Monica, California, March 27, 1969.

I am suggesting that longitudinal, multimeasure, and multivariate strategies are required to lay bare the essential phenomena of our field. One reason for placing such emphasis upon this point is that we have become accustomed to thinking about these design characteristics in other terms. For example, there are certain statistical advantages of longitudinal over cross-sectional designs. However, if statistical efficiency or other methodological conveniences were the primary issue, it is doubtful that we would even turn to longitudinal or multivariate procedures in the first place, since they raise even greater technical and practical problems. In any event, the point is that these requirements are as much substantive as methodological.

Even if one grants that the essential phenomena of development are best revealed through a longitudinal multivariate approach, we still find ourselves facing a vast and mostly uncharted territory. At the simplest level, one can appreciate the problem by contemplating the sheer magnitude of information generated by such designs. More basic, however, is the question of how to deal with this wealth of information in sensible developmental terms.

It is an historical fact that broad developmental theories have not been very explicit about how they assimilate information produced by longitudinal multivariate studies. In the area of personality development, my own examination of the literature leads me to conclude that this situation is not, however, entirely chaotic (Emmerich, 1968). I find it possible to place personality theories into three broad groups. The first of these consists of classical stage theories which posit universal sequences of personality organization through which individuals pass at differing rates. The second group, which I call "differential," arises from a trait approach and deals with consistency and change over time in the structure of individual differences on dimensions derived from multivariate analyses. The third approach is ipsatively based, and considers the development of the individual's particular repertoire of behavioral attributes and their internal organization. I would emphasize that all three approaches do share a common interest in the developmental phenomena I have been discussing, and therefore all three call ideally for longitudinal multivariate designs. Each deals in its own way with such topics as continuity,

stability, sequentiality, transformation, structural complexity, and structural reorganization. However, despite these similarities in theoretical language, the three approaches differ markedly in how their constructs are coordinated to developmental data.

To illustrate this point, consider the concept of "stability" of individual differences over time. From a differential standpoint, the extent of stability of a behavior would be signified by the magnitude of the correlation between measures taken at two age periods, say at middle childhood and again at late adolescence. However, in a more classical stage conception, such as Loevinger's model of ego development (Loevinger, 1966), this same stability coefficient could have less meaning because stable individual differences are expected to occur only within a delimited period of ego development, and, indeed, should wash out as most individuals pass beyond the stage in which the behavior is most relevant. Perhaps the crux of the matter is that differential theories typically define individual differences in terms of trait frequencies or intensities, at least in the personality area, whereas stage theories treat individual differences as variations in rates of progression through normative sequences of development.

In their recent comparative analyses of developmental theories, both Baldwin (1967) and Langer (1969) emphasize that developmental theories rarely contradict one another in the sense of generating testable alternative interpretations of the same phenomena. As Baldwin puts it, "theories generally talk past each other" (Baldwin, 1967, p. 583). We naturally expect theories to differ when they attempt to explain different areas of behavior, or when they focus upon different behaviors within an area. I am suggesting that even when the same variables are assessed, theories will differ in how their constructs are coordinated to the parameters generated by comprehensive longitudinal multivariate designs.

This conclusion has important implications for strategies of model building in developmental research. First of all, it places a responsibility upon investigators working within a particular orientation to be very explicit about these translations. More interesting, however, is the possibility of an inductive strategy which considers all relevant parameters generated by developmental data, and then evaluates goodness of fit in relation to each of several alternative models.

Let me provide a partial illustration of how this might be done, using an extreme hypothetical example. Suppose we were able to accumulate multivariate information at several age periods from a large and broadly representative population. Imagine further that our set of variables had never before been studied in this systematic way. With these data in hand, we ask the following question: Of the many potential forms which development might take, which form best describes the actual trends? To answer this question, we proceed through a series of inductive steps. First, considering each measure as a discrete attribute, we ask whether, for our whole sample, there is an age shift between the first and second age periods. We make this test on all measures of the study, thereby sorting measures into two groups: age-related and non-age-related differential variables.

With regard to the set of age-related measures, we next ask the following question: Are the age changes in these measures in synchrony with one another in the sense that individuals who change rapidly on one measure also change rapidly on others, at least between the first two age periods? Put more generally, are the rates of change among measures so highly correlated that we are justified in saying that these variables are directed toward a common developmental end? If the empirical answer is yes, then we have evidence for a generalized unidirectional developmental trend. On the other hand, factor analysis of these change scores might reveal not a single large factor, but a small set of moderate factors representing independent clusters of change, or <u>multiple growth trends</u>. This idea of multiple growth trends is an important one, having analogues both in psychoanalytic (A. Freud, 1965) and Piagetian thinking (Flavell, 1963).

Carrying this same branch of the analysis one step further, we now ask how the age trends found between our first two time periods dovetail with those found between all other adjacent ages. If we find that all such mean changes are in the same direction, then we will have isolated monotonic age trends corresponding to what Loevinger (1966) calls "polar aspects" of development. However, if mean age changes between adjacent age periods themselves vary in <u>direction</u> as a function of age, then we will need to search for an alternative model. For example, if certain curves reach an asymptote and begin to decline at about the same age,

together with reductions in individual differences, then we might conclude that we are dealing with a "milestone" in development, in Loevinger's sense, which we would then want to coordinate to a sequential model.

There are several other important branchings and steps in this procedure, but what I have said will give you some idea of the general strategy. Essentially, it builds developmental phenomena in ordered steps, and at the end of each step the cumulative findings are evaluated in relation to an array of alternative developmental models. An interesting feature of this strategy is that it can lead to formulations not clearly implied by extant theories.

I have referred to a distinction between "classical" and "differential" models of development. This distinction has been useful because it summarizes two historical trends in the treatment of developmental information. However, our inductive strategy could lead to surprising but viable hybrids which combine certain features of both approaches.

One such combination is suggested by recent longitudinal studies of trait stability on dimensions of social behavior, such as activity-passivity, achievement, aggression, and dependency. Some of these have been long-term investigations (e.g., Kagan & Moss, 1962; Schaefer & Bayley, 1963), while others have been short-term studies (e.g., Emmerich, 1964, 1966), but all have revealed at least moderate stabilities on certain traits. However, because they arise from the differential tradition, these studies typically have not integrated stability findings with parallel changes in mean levels, a step called for by our inductive strategy. (There are notable exceptions, including the work of Witkin, Goodenough, and Karp (1967), where field independence is reported both to increase monotonically with age and to order stable individual differences between late childhood and late adolescence.)

Returning now to the earlier mentioned concept of differentiation, in the sense of repertoire proliferation with development, we might expect the age-specific phenotypic expressions of underlying genotypic traits to change with age (Emmerich, 1968). Indeed, the work of Kagan and Moss (1962) and Lewis (1967) suggests that when certain phenotypic expressions at different ages are properly coded, they will be found to belong to the same underlying construct. What is needed to clinch this argument is

evidence on mean trends, showing, perhaps, that each age-specific component of the construct fits a curvilinear age function. Also, this model seems to imply that maximum stability coefficients will occur between different phenotypic expressions precisely when the mean value of each phenotype reaches its peak.

Curiously, this particular combination of the differential and classical views subordinates stage to trait theory by treating curvilinear trends as age-specific phenotypes of underlying continuous dimensions. Put more broadly, this is a theory of transformation in the phenotypic expressions of invariant constructs; it does not bear on possible developmental reorganizations at the higher level of the constructs themselves, for which we have also found tentative evidence (Emmerich, 1964, 1966, 1968).

Despite this important limitation, however, there are certain additional findings in the literature on social development which may be clarified by this model. When a social motive such as dependency is broken down into its components and each component is assessed by systematic observation, the resulting intercorrelations typically are low, suggesting that dependency is not a unitary construct (Maccoby & Masters, in press; Mischel, 1968). Mischel (1968) has interpreted this state of affairs as partial evidence for the view that generalized traits do not exist in the area of social behavior. An alternative interpretation is suggested by the model of age-related changes in expressions of dependency. If the phenotypic manifestations of dependency emerge and decline in an age-related sequence, then high or even moderately positive intercorrelations would not be expected among some components within a particular age period. Indeed, as noted earlier in discussing the inductive strategy, this is precisely the kind of situation where intercorrelations among change scores, not raw scores, are most meaningful.

There have been a few studies which report age changes in the components of dependency behavior. In reviewing these studies, Maccoby and Masters (in press) note the presence of developmental trends when measures are differentiated according to the object of dependency. For example, in a study of preschool children, Heathers (1955) found a shift in which dependency toward adults decreased while dependency toward children increased during the same period. Also, Kohlberg and Zigler (1967) found (different)

age trends within each sex on the amount of verbal dependency expressed toward male and female experimenters. Such findings suggest, then, that certain differentiated components of social motives exhibit mean changes in development.

There is an interesting contrast between the kinds of specific measures just discussed and more global measures, such as those derived from ratings. In general, global rating studies do reveal a more unitary construct of dependency than studies based upon systematic observation procedures (Maccoby & Masters, in press; Mischel, 1968). Again, however, we are limited in our interpretation because rating studies do not typically integrate findings on correlational patterns with mean developmental trends. However, the model under consideration implies that the more specific our measures of a construct, the more likely we are to find mean shifts with age. Perhaps raters attend selectively to age-appropriate expressions of the trait in question when making their judgments. Since the subset of phenotypic expressions at any age period is not an accumulation of previous expressions, but is a substitution of one subset for another, the total score on the trait would balance out and be equalized across age periods, accounting for the lack of age trends on global scores. The mean values of the phenotypic components would be found to change with age, at least in so far as these changes are not blurred by the failure of judges to make appropriate discriminations among the components.

I have discussed continuity and change only in regard to the development of response systems, but the present approach also has implications for relationships between environmental events and developmental outcomes. When applications of the inductive procedures allow us to discern the course of development for a given response system, we are then in a good position to isolate environmental factors which determine, guide, or support the developmental behavioral trends. Indeed, when we can determine which developmental behavioral model is most appropriate, we already have gained some information bearing on the nature of environmental influences. Returning again to the example of an invariant psychological trait with developmental change in its phenotypic expressions, let us make the assumption (and it is only an assumption to illustrate one possible state of affairs) that specific environmental factors play a primary role in the

trait's development. We can now discern at least three distinct kinds of environmental influence that would be operative. First, there would be a certain age period when the trait is initially formed, signified when reliable individual differences first occur. This period represents a critical transition in the nature of the developmental process, perhaps analogous to that between embryological development and fetal growth. Next, there would be a set of influences which function to support, maintain, and perhaps even strengthen and amplify individual differences during subsequent age periods. This second set of supporting determinants need not be identical or even similar to those responsible for initial trait formation. Finally, superimposed upon but distinct from these generally supportive processess would be a series of age-graded environmental changes which function to alter phenotypic expressions. Here, environmental forces would be of such a nature as to lead to a differentiation of one or more components of the trait without cancelling out those influences which support underlying trait continuity.

To recapitulate, I have suggested that processes of continuity and change with age represent basic phenomena of developmental psychology, and that longitudinal, multimeasure, and multivariate studies are required to fully capture these processes. While our theories have not provided highly explicit rules for dealing with the information generated by comprehensive designs, they do suggest useful models which can be evaluated and revised through the application of an inductive strategy for analysis. Finally, I have explored ever so tentatively some implications flowing from the concept of differentiation in development, and have suggested that a model based upon this idea might prove useful in accounting for certain aspects of social development.

References

Baldwin, A. L. Theories of child development. New York: Wiley, 1967.

Emmerich, W. Continuity and stability in early social development. Child Development, 1964, 35, 311-332.

Emmerich, W. Continuity and stability in early social development, II: Teacher ratings. Child Development, 1966, 37, 17-27.

Emmerich, W. Personality development and concepts of structure. Child Development, 1968, 39, 671-690.

Flavell, J. J. Developmental psychology of Jean Piaget. Princeton, N. J.: Van Nostrand, 1963.

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

Heathers, G. Emotional dependence and independence in nursery school play. Journal of Genetic Psychology, 1955, 24, 169-179.

Kagan, J., & Moss, H. Birth to maturity. New York: Wiley, 1962.

Kohlberg, L., & Zigler, E. The impact of cognitive maturity upon the development of sex-role attitudes in the years four to eight. Genetic Psychology Monographs, 1967, 75, 89-165.

Langer, J. Theories of development. New York: Holt, Rinehart & Winston, 1969.

Lewis, M. The meaning of a response, or why researchers in infant behavior should be oriental metaphysicians. Merrill-Palmer Quarterly, 1967, 13, 7-18.

Loevinger, J. The meaning and measurement of ego development. American Psychologist, 1966, 21, 196-206.

Maccoby, E. E., & Masters, J. C. Attachment and dependency. In P. Mussen (Ed.), Carmichael's manual of child psychology. (3rd ed.) New York: Wiley, in press.

Mischel, W. Personality and assessment. New York: Wiley, 1968.

Schaefer, E. S., & Bayley, N. Maternal behavior, child behavior, and their intercorrelations from infancy through adolescence. Monographs of the Society for Research in Child Development, 1963, 28, 1-127.

Witkin, H. A., Goodenough, D. R., & Karp, S. A. Stability of cognitive style from childhood to young adulthood. Journal of Personality and Social Psychology, 1967, 7, 291-300.

Wohlwill, J. The age variable in psychological research. (In preparation)

Protection of the Rights and Privacy of Human Subjects

Ann Jungeblut

(for the Steering Committee)

In recent years, various segments of society have resoundingly raised the issue of invasion of privacy. The expressed concern, as it relates to data collection, ranges from certain federal and local government censuses through psychological testing and experimentation in behavioral research to the use of mechanical and electronic eavesdropping devices. Attention here will be focused on this concern as it relates to behavioral research.

The basic conflict in behavioral research (Ruebhausen & Brim, 1965, pp. 1189-90; Carlson, 1966, p. 2; Holmen & Docter, 1968, p. 13) stems from the confrontation of two values of society: the freedom of the individual to decide the extent to which, and the time and circumstances under which, he chooses to reveal or withhold from others his attitudes, beliefs, opinions, thoughts, feelings, and behavior; and the public good, the community's need to know, or the value "vigorously championed by most American citizens, the [scientific] right to know anything that may be known or discovered about any part of the universe" (Clark, 1967, p. 3).

Resolution of the conflict involves constant review of the purposes of gathering the information and the intended uses of the information. Such considerations led Willingham (1967, p. 2) to observe, "...it seems unlikely that the public makes clear distinctions between invasion of privacy in different fields for different purposes. Thus, concerns with respect to research and testing undoubtedly are enhanced by a backlash from electronic eavesdropping, super-secret intelligence activities, and other worrisome mechanisms of social control."

Discussing ethical bases for self-regulation of psychologists in assessment, Messick (1965, p. 140) concluded that, "It makes a difference why the information is being sought!...it seems imperative that we go beyond ethical absolutism...and espouse an 'ethics of responsibility,' in which pragmatic evaluations of the consequences of alternative actions form the basis for particular ethical decisions." In a similar vein, Bennett (1967, p. 9) noted: "Some psychologists wish to defend to the

death the principle of unlimited scientific inquiry. It seems more reasonable to assume that our profession will enjoy greater respect and freedom if it is practiced with a moderate degree of discretion and restraint." Psychologists should be rightfully proud of the American Psychological Association's record of early development of a code of ethics (1953, 1954, 1963).

In 1965 the National Advisory Health Council issued a resolution that resulted in a policy statement from the Surgeon General requiring prior review by institutional associates for all Public Health Service (1966a) supported projects. Assuredly much of the concern was directed toward medical research in which a procedure might induce a potentially harmful physiological condition in a subject. On the other hand, in a clarification of the policy (Public Health Service, 1966b), it was stated that although there is much social and behavioral research in which there is no personal risk to the subject, many investigations "involve procedures designed to alter the status of the individual as, for example, studies of human learning, social perception, or group effectiveness" (p. 2). The policy statement further stipulates that "the issues of concern are the fully voluntary nature of the participation of the subject, the maintenance of confidentiality of the information obtained from the subject, and the protection of the subject from misuse of the findings" (p. 1).

In response to the PHS policy, Educational Testing Service initiated a Committee on Prior Review of Research Projects. The jurisdiction of the Committee extends to "all psychological and educational investigation and research involving human subjects, which is proposed by staff members at ETS, for conduct independently or in conjunction with another agency" (Educational Testing Service, 1967, p. 1), not merely to those projects for which PHS support is sought. Emphasis in this ETS document is placed on review in reference to the appropriateness of the methods used to obtain informed consent. The report of the panel chaired by Clark for the Office of Science and Technology (Clark, 1967, pp. 23-26), covered the point of institutional responsibility and concluded that, "The institution should have the prerogative of setting up the monitoring machinery for insuring that the proposed procedures are acceptable" (p. 24). Hoch (1967)

offered one solution involving scaled dimensions of the value judgments in the use of human subjects to help "establish the degrees of freedom and the constraints that should characterize the proper use of human subjects in psychological research" (p. 20) that can be used on an institutional level.

Theoretical considerations about the concept of consent are well covered in the literature. Ruebhausen and Brim (1965, p. 1189) point out that, "Taken literally, the concept of consent would require that behavioral research refuse to engage in the probing of personality, attitudes, opinions, beliefs, or behavior without the fully informed consent, freely given, of the individual person being examined." However, they proceed to indicate several reasons why insistence on informed consent can be unrealistic in a research context. For example, the selectivity involved in consent would bias the sample, thus making it impossible to generalize the results; knowledge of and consent to measurement of subtle attitudes can distort results; and, in many instances, it is impossible to convey a full appreciation of the nature of a specific research project, the purposes to be achieved and the possible risks involved, because of their complexity, because they involve unknown factors, or because they are beyond the capacity of the understanding of the subject.

The report of the panel chaired by Clark (1967, p. 17), concludes that the individual claims the right to choose to what extent he will reveal his thoughts, feelings, and actions. Consent to participate in an experiment is the exercise of that choice and thereby satisfies the claim to privacy. However, Ruebhausen and Brim (1965, p. 1199) caution that, "consent to the revelation of private personality for one purpose, or under one set of circumstances, is not license to publish or use the information so obtained for different purposes or under different conditions." Westin (1967, p. 375) underscores this point by saying, "It should be recognized that consent to reveal information to a particular person or agency, for a particular purpose, is not consent for that information to be circulated to all or used for other purposes." Ruebhausen and Brim (1965, p. 1199) stipulate that varying degrees of consent must be recognized and also indicate that a complicating factor in consent is whether or not consent is freely given or coerced.

Ruebhausen and Brim (1965, p. 1200) also raise the question of the identification of the person whose consent must be obtained. In this connection, if the subject is not an adult, the legal principle seems to be quite clear--consent should be obtained from the guardian or parent of the subject. However, these authors feel that, ethically, even before the age of legal responsibility, a child also has the right to a private personality. In this case, "some form of prior consent to privacy probing should be obtained from both the parent and the respondent child."

On the other hand, the Clark panel (1967, p. 22) indicated that consent can be given by adult members of the community serving as surrogates for the child (e.g., the principal, the superintendent, or the school board) and need not necessarily be obtained from the parents. Tillery (1967), for example, reported that institutional rather than parental or pupil consent was obtained in project SCOPE. Furthermore, Ruebhausen and Brim (1965, pp. 1203-4) maintain that, "there are and will be many occasions in which conflict between the individual's claim to privacy and the larger community interest in research for the general good must be resolved--and the method of resolution must be an expression of community consensus." Such consensus can be expressed in laws, judicial decisions, or political constitutions as well as through the stated views of opinion leaders--leaders in government, community action groups, industry, labor, the professions, or the clergy.

In view of the fact that obtaining fully informed consent is often unrealistic, consent must be based on trust in the scientist and his institution (Clark, 1967, pp. 5, 22). The panel concluded that, "If the public is to feel secure about behavioral research and to feel free to participate in it, then it must gain a greater understanding of its objectives, methods, and values" (Clark, 1967, p. 29).

Once the data is collected with informed consent or community consensus, the issue of confidentiality becomes the focal point of concern. Anonymity is defined as the third state of privacy by Westin (1967, p. 31) and "occurs when the individual is in public places or performing public acts but still seeks, and finds, freedom from identification and surveillance." An important way to implement the concept of confidentiality in behavioral research is to design a study so that the responses of the

subjects can be anonymous (Ruebhausen & Brim, 1965, p. 1205; Clark, 1967, p. 17). Ruebhausen and Brim (1965, p. 1205) state and Roper (1969, p. 3) concurs that anonymity should be possible in all opinion surveys, but that the nature of the research might well require the ability to identify each subject with the data obtained from him. "This would of course be true in longitudinal studies--as of child growth and development--where respondents must be examined or interviewed a number of times, or in studies of several diverse sets of records which must be matched up to a particular individual" (Ruebhausen and Brim, 1965, p. 1205).

In personal communication with Ruebhausen and Brim (1965, p. 1205), Chein identified six possible levels of anonymity including the following: "the identity of the particular subject is known in conjunction with meaningful and interpretable data and these data are scrutinized from the point of view of interpreting some aspect of the individual or his behavior, but his identity is thereafter submerged in the collection of similar processes of interpretation for many subjects." At this level, confidentiality can be safeguarded through a number of means; for example: the integrity of the behavioral scientists; control techniques such as coding and locked files; and, finally, destruction of research data on completion of a project. Nevertheless, Ruebhausen and Brim (1965, p. 1206) note that, "It should be emphasized that neither the integrity of the scientist nor the technical safeguards of locks and codes can protect research data against a valid subpoena; such data are at present quite clearly subject to subpoena." Roper (1969, p. 4) emphasized the responsibility of survey organizations to educate the courts and researchers to prevent turning the names of survey respondents over to the courts. He concluded (p. 5) that, "If we don't hold absolutely firm and establish the precedent that our sources are inviolable, then we jeopardize the whole basis of our profession."

Confidentiality is also a matter to be dealt with in the reporting of research results. In referring to project SCOPE, a longitudinal study, Tillery (1967, p. 16) reported, "It is difficult to communicate some aspects of the findings and their value to the students whose continued support we need, without at the same time contaminating future data. This is a tricky balance, but it seems improbable that long-term cooperation

can be gained from students unless they have some feeling that the project is of value to them or to others who follow after them." The parallel between this situation and that in which informed consent is obtained from both parent or guardian and school boards in a longitudinal study of child development seems obvious. Moreover, Carlson (1966, p. 4) adds, "The issue in this case is not only our right as researchers to collect such data but, more important, our discretion and maturity of judgment in using them in reporting on the attitudes and the behavior of subgroups in our samples which might be identified and damaged by such revelations."

As will be seen from the preceding discussion, the existing literature on the subject is extensive. Although the literature initially appears to reflect a rather negative statement of the problem--invasion of privacy--it is obvious that much positive thought, many positive suggestions, and many helpful cautions have resulted from this effort. Nevertheless, perhaps it is timely to emphasize the positive statement of our concern--protection of privacy. Moreover, it seems appropriate to stress one reason for our attention that, unfortunately, often appears to be passed over too lightly. Basically, our effort to protect the privacy of the individual in behavioral research projects reflects a genuine concern and respect for each human being. If the concept of protection of privacy is uppermost in the mind of the researcher during the design and operational phases of a study, one would anticipate the chances of invasion of privacy of the subjects participating in that study would be very small indeed.

The following pages deal with means of protecting the privacy of those parents and children, school administrators and teachers, and community officers whose participation and cooperation are sought for the ETS-Head Start Longitudinal Study of Disadvantaged Children and Their First School Experiences.

The concept of protection of privacy of the individual--child, family, school administrator, teacher, community, and so on--has been uppermost in the minds of the staff of the longitudinal study since its inception. For example, in defining domains to be covered in the study, there was continual soul-searching among the staff on how to include all the areas necessary for the discovery of interactions, while at the same time avoiding those very personal areas the coverage of which would yield interesting, but not necessary, information.

Similarly, in designing and selecting the specific measurement instruments there was constant review and revision to ensure the privacy of all participants. Indeed there was both formal and informal review by ETS professional staff independent of the investigation during all phases--study design, proposal writing, and instrument development and selection. In addition, the final study design, as well as all instrumentation, was processed by the ETS Committee on Prior Review of Research Projects. Current and past members of the OEO Research Advisory Council and other members of the OEO and Head Start staff also provided considerable help in the area of review.

Some interesting points relevant to protection of privacy were anticipated during the first data collection phase of the longitudinal study. Many adult members of the study communities are participating as testers, raters, observers, and interviewers. Thus it was felt necessary to establish a policy whereby no staff member could interact, during these study operations, with any child or family related to or known to him. Other instances reflecting our concern for protection of privacy will be seen in the design of answer sheets. Once the field designated code is entered and checked, names may readily be removed. Another example is the policy stipulation that no child's test results will be scored within his home community.

Protection of privacy is also a great concern in the areas of data processing, storage, and analysis. Professional concern is reflected in such papers as that by Sawyer and Schechter (1968) and in such conferences as that on student information reporting, conducted by the Russell Sage Foundation in May 1969. In line with a point made by Ruebhausen and Brim (1965, p. 1206), our initial task was to design a unique and secure identification numbering system for subject coding. A procedure was used whereby lists of identification numbers were prepared by computer and issued to local field coordinators for assignment. The field assigned code number employs a checking digit that can be automatically verified at the point of keypunching. Moreover, all data destined for the data bank will be processed through a second coder program in which new identification numbers will be coded from the original field assigned numbers. In this way it will be impossible for <u>any</u> researcher to associate

specific responses with a particular study subject. In a longitudinal study such as ours, in which subjects will be tested, rated, observed, and interviewed a number of times, this coding system would seem to be the first important step toward anonymity for subjects.

There will, of course, be no access to the keys of the identification codes. Moreover, there will be no direct access to the data bank. All requests for analyses by authorized researchers will be handled through the office of the Data and Analysis Coordinator. As suggested by Ruebhausen and Brim (1965, p. 1206) it may even be possible at the end of the data collection phase of the study to destroy the master lists of subjects, their addresses and so forth, thus guaranteeing anonymity.

The confidentiality of all information collected will be preserved in all published reports. No reports of individual results on research instruments will be made to parents, school administrators, school boards, and the like. Moreover, no decisions concerning a study subject will be made on the basis of the results obtained on research instruments. On the other hand, appropriate medical information will be made accessible to parents. Similarly, reports of standardized achievement test results obtained in kindergarten through grade 3 will be made available to appropriate school personnel.

The examples cited of the steps taken to protect the privacy of subjects are in no way exhaustive. They are, however, presented as a sample of the active and continuing concern of the staff for the protection of the privacy of all subjects--child, family, school administrator, teacher, community, and so on--of the Longitudinal Study of Disadvantaged Children.

References

American Psychological Association. Ethical standards of psychologists. Washington: American Psychological Association, 1953.

American Psychological Association. Technical recommendations for psychological tests and diagnostic techniques. Psychological Bulletin, 1954, 51 (2, Part 2).

American Psychological Association. Ethical standards of psychologists. American Psychologist, 1963, 18, 56-60.

Bennett, G. K. Testing and privacy. In W. W. Willingham (Ed.), Invasion of privacy in research and testing. Proceedings of a symposium sponsored by the National Council on Measurement in Education. Published as a supplement to the Journal of Educational Measurement, 1967, 4 (1), 7-10.

Carlson, R. O. The issue of privacy in public opinion research. Public Opinion Quarterly, 1967, 31 (Spring), 1-8.

Carter, L. F. Psychological tests and public responsibility: Introduction. American Psychologist, 1965, 20 (2), 123-146.

Clark, K. E. (Chm.) Privacy and behavioral research. Washington: U. S. Government Printing Office, 1967.

Conrad, H. S. Clearance of questionnaires with respect to 'invasion of privacy,' public sensitivities, ethical standards, etc. American Psychologist, 1967, 22 (5), 356-359.

Educational Testing Service. Policy and procedure for safeguarding rights and welfare of human subjects in investigations. Princeton: Educational Testing Service, 1967.

Hoch, E. The privacy issue and a professional response at the departmental level. In W. W. Willingham (Ed.), Invasion of privacy in research and testing. Proceedings of a symposium sponsored by the National Council on Measurement in Education. Published as a supplement to the Journal of Educational Measurement, 1967, 4 (1), 17-22.

Holmen, M. G., & Docter, R. F. The testing industry and public policy. Los Angeles: Research Institute for Business and Economics, University of Southern California, 1965.

Messick, S. Personality measurement and the ethics of assessment. American Psychologist, 1965, 20 (2), 136-142.

Public Health Service. Investigations involving human subjects, including clinical research: Requirements for review to insure the rights and welfare of individuals. (Policy and Procedure Order #129 dated July 1, 1966.) Washington: Public Health Service Division of Research Grants, 1966. (a)

Public Health Service. Investigations involving human subjects, including clinical research: Requirements for review to insure the rights and welfare of individuals: Clarification. (Policy and Procedure Order #129 dated December 12, 1966.) Washington: Public Health Service Division of Research Grants, 1966. (b)

Roper, B. W. The responsibilities of the public opinion researcher to the public. Paper presented at the meeting of the American Association for Public Opinion Research, May 17, 1969.

Ruebhausen, O. M., & Brim, O. G., Jr. Privacy and behavioral research. Columbia Law Review, 1965, 65, 1184-1211.

Sawyer, J., & Schechter, H. Computers, privacy, and the national data center: The responsibility of social scientists. American Psychologist, 1968, 23 (11), 810-818.

Tillery, D. Seeking a balance between the right of privacy and the advancement of social research. In W. W. Willingham (Ed.), Invasion of privacy in research and testing. Proceedings of a symposium sponsored by the National Council on Measurement in Education. Published as a supplement to the Journal of Educational Measurement, 1967, 4 (1), 11-16.

Westin, A. F. Privacy and freedom. New York: Atheneum, 1967.

Willingham, W. W. (Ed.) Invasion of privacy in research and testing. Proceedings of a symposium sponsored by the National Council on Measurement in Education. Published as a supplement to the Journal of Educational Measurement, 1967, 4 (1), 1-31.

The Ghetto: Some Notes on Environmental Pressures and Educational Atmosphere*

Fred Damarin and Irene Kostin

Researchers and teachers in ghetto schools have recently provided a series of vividly-written, attention-getting memoirs (Eddy, 1965; Green and Ryan, 1964; Hentoff, 1966; Herndon, 1965; Kohl, 1967; Kozol, 1967; Moore, 1967). Because they have been enthusiastically reviewed by liberal leaning periodicals such as The New York Review of Books, the Saturday Review, The New Republic, Commentary, and Christian Century, and because they present a clear mutually reinforcing picture of the problems of ghetto education, they are likely to affect the educated layman's views for some time to come. They may form a context or background against which technical, statistical researches will be evaluated for relevance and seriousness of purpose.

We have surveyed these memoirs for descriptions of environmental pressures and atmosphere and their effects on children in the range of 4-9 years with special attention to circumstances that may retard learning. We have attempted to integrate these descriptions so as to develop variables and perspectives that may be of interest to those engaged in the longitudinal study.

Salient Features of the Ghetto

The work reviewed here reports chiefly on education in the Northern, urban, black "ghetto," described roughly as a place where large families occupy two- or three-room walk-up apartments that are rat infested, poorly repaired and, in the winter, inadequately heated. The father may be absent, the mother may work, and older children have frequently to care for the younger. On the streets outside, idle men occupy the front stoops of dwellings or congregate on street corners or in bars. Drug addicts can

---

*Ann Bussis, Marianne Amarel and Joyce Crossland are thanked for their thorough, critical assessment of an early version of this working paper.

sometimes be found on the roofs and in the hallways of the apartment houses as well as in the streets. Violence involving children and adolescents is frequent and the participation of adults in large scale disorders is becoming more common. The police are usually regarded as more of a threat than a help.

Characterizations like this serve literary purposes better than scientific ones. It is easy to recognize the essence of ghetto--it is the worst of the worst--but it is harder to decide to what extent the various families participating in the longitudinal study will be "ghettoized." Place of residence is not necessarily a good predictor of family cohesiveness, morale, and aspiration level. This observation leads to the first of a series of suggestions that are phrased as questions:

1. Do our social and family status variables cover so broad a spectrum of ghetto experiences that they can be assembled and scaled, classed or factored to discover whether the experience is unitary or multidimensional? Can this be done early enough in the research to provide guidance for subsequent investigation?

Some black families may have lived in the worst of the ghetto and may have escaped to a better neighborhood. Other families living in the ghetto may be new arrivals from the rural slums of the south. These differences in family history may condition their response to their present environments.

2. Will we learn enough about the family's previous history to test hypotheses about its relevance to adjusting in the present environment?

Suffering

As our sources describe the lives of ghetto families, the children endure chronic suffering. Certain personality traits are thereby emphasized and certain adaptive strategies pursued. Many aspects of this suffering require attention.

Chronic debilitation. Ghetto children seem likely to suffer chronically debilitating physiological states that reduce their attention span and interfere with the habits and attitudes that facilitate later learning, especially reading.

a. They may be inadequately-clothed and cold.

b. They may arrive at school after no breakfast at all, or after only toast and coffee.

c. They may share a bed with other children and, perhaps, seldom get enough sleep.

d. They may be ill; e.g., from chronic, low-grade infection.

This leads to the next question:

3. Will the children's dress, dietary habits, food preferences, hours of sleep and general health be reasssessed periodically throughout the study or only at the outset?

Sources of negative affect. The milieu of these children has an obvious capacity to invoke negative affect but it may be desirable to focus on specific sources (or press).

A. Parental discipline. Ghetto parents are said to rely almost exclusively on physical punishment to discipline their children, but there may be at least two reasons for this. They may follow a family social tradition stressing the efficacy of punishment as a means of social control. They may also lack the facilities required for subtler forms of control.

4. Will the family's attitudes about physical punishment and other control techniques be assessed? Will there be questions about grandparents' attitudes to various control techniques?

5. Will the availability of alternative techniques of control be assessed? Is isolation impossible because of crowded quarters? Are material rewards and punishments ineffective because there are few material resources? Is maternal responsibility so diffused that withdrawal of love is ineffective? Is there insufficient time, quiet, or privacy to "reason" with the child?

B. Neighborhood influences. The ghetto neighborhood contains real dangers for children and powerful aversive stimuli (garbage heaps, dope addicts) that may trigger innate reactions of disgust and aversion. The impact may vary considerably from one child to the next, however, as a function of the exact location of his home in the general neighborhood.

6. Will it be possible to code the environment of the home in ways that will identify children whose situation is objectively most

hazardous? Will it be possible to identify aversive stimuli that are especially obnoxious to children of this age and code their incidence in the neighborhood?

Varieties of negative affect. Since children who are equally exposed to such sources of press may differ enormously in affective response, it seems important to assess the nature, duration, and intensity of their typical feeling states on which particular emphasis might be placed.

7. Fear. How strong, persistent, and varied are the children's fears? Will it be possible to decide which of these are realistic? Will it be possible to learn how the child copes with fear? Does he limit his activities or try to learn as much as possible about the feared object?
8. Anger. Verbatim records of the conversations of these children suggest the presence of chronic states of anger that may emerge as heightened sensitivity to criticism from both peers and adults. Skin color is especially salient. How much will we learn about the things that make the child angry and about the ways in which he handles anger?
9. Contempt. Our observers report that ghetto children are exceptionally ready to offer insult--this paradoxical, given their sensitivity to criticism. Will there be any assessment of the things these children despise in others or of the number of scornful, mocking comments they may produce in school or on the playground?
10. Shame. Feelings of inferiority, discouragement, and shyness are frequently reported. Can they be assessed? Shame is an affect that applies particularly to the self. Coleman and his colleagues found that underachieving school children lacked confidence in their ability to produce positive effects on their environment, but how many of these same children are overconfident about the extraordinary ease with which they might produce negative, harmful effects?

Adaptive Strategies

Children survive in this world by adopting a variety of strategies to minimize their suffering. Some of these, like the special language of the ghetto, have been prepared in advance, by the community. Others may be more individualized and susceptible to correlational investigation. Among them:

11. Searching for novelty and stimulation. The child who arrives at school cold, hungry, sleepy, ill, angry, or frightened may well prefer active play--dancing and singing--to passive "academic" pursuits, because the activity keeps him awake and unmindful of physical distress. Will there be any attempt to test for the child's ability to remain intellectually alert even while he is physically passive?
12. Retreat into fantasy. In this context fantasy means passive daydreaming--private rehearsal of a few dramatic themes that are invested with great personal and emotional significance. It has no necessary connection with creative ability. Will the child's fantasy life be assessed for intensity and duration?

(Note that the combination of fantasy with strivings for novelty and stimulation may help to explain the teacher's complaints that ghetto children are excessively "unruly" and also distracted and inattentive in class. It may also explain why creative teachers turn to curricula that provide opportunities for the expression of fantasy in dramatic play.)

13. Self-reliance. These children are urged to assume adult responsibilities at an early age and, even if this were not the case, their environment is such that self-reliance will pay off. Pressures in this direction may produce too early a maturity, however. With white middle class children an instrument like the Vineland social maturity scale (which is administered to the mother) might be valuable in assessing self-reliance. It may work here too, but there is no way of knowing in advance. Will we learn anything about the behaviors these mothers prize as indices of self-reliance? Will we learn anything about the large fund of information that ghetto children acquire in the

course of daily living? While they may not be able to read Dick and Jane, they may know how to kill rats, how addicts take their drugs and how to play numbers. Will our tests reveal this?

14. Neighborliness. Ghetto families are said to be more willing to help each other than members of the white middle class. They react as parents toward all the neighborhood children, rather than toward just their own, and they are ready to help other families with a variety of minor adversities. Will we know to what extent our children's parents participate in such communal arrangements? Prosocial traditions of this sort may be an important positive influence especially in view of the countervailing negative affect that is present in the children's lives.

While we know little about the ways in which prosocial and antisocial motives interact at various stages of the child's development, Anna Freud's (1956) report of unusual cooperativeness among preschool children rescued from concentration camps suggests that prosocial motives emerge early.

The Failure of Educational Strategies

Advancement via education, a traditional escape route from the ghetto, has failed this generation of inhabitants. Society's prejudice against blacks--educated or otherwise--may explain some of this, but the schools themselves are clearly ineffective. Our sources offer the following bill of particulars:

a. An aging, neglected, vandalized physical plant;
b. Educational policies that discourage innovation;
c. Curricula that depend heavily on nonexistent reading skills and that feature content totally unrelated to the children's lives;
d. Homogeneous grouping by IQ in ways that destroy the initiative of the "dull" group;
e. Personnel policies that discourage friendship between the teachers and the pupils and their parents;
f. Inadequate facilities for handling children with behavior problems;

g. Continued interruption of class routine by administrative chores, PA systems, monitors, wandering children from other classes, noise, and fighting;

h. Incompetent teachers who "baby-sit" instead of teach, who are either too lax or too vindictively punitive, who overemphasize order and silence, and who discriminate in favor of some pupils.

15. Will we know how many of these criticisms apply to each school and each teacher in the study?

These characteristics of ghetto life and ghetto schools support the notion that educational efforts fail where they interfere with adaptive strategies that are needed for survival. Apart from transient Hawthorne effects, educational reforms may also fail unless the circumstances that produce the strategies are alleviated and the strategies themselves viewed with sympathetic understanding.

Social reform is mandatory, but this will require time and, in any case, lies beyond the specific mandate of the present research staff. Some mapping of the ghetto child's strategies for survival and some probing for their as yet untapped educational potential does seem to fall within our competence. Many of these children may have more going for them than anyone now knows.

Practical Suggestions

Having asked questions about the longitudinal study, we should also like to suggest means for evaluating and, perhaps, answering them. Note, first, that the authors of this paper are white as are the observers who provided our source material. Daniel Moynihan (1968) would express no surprise at this, for he recently said,

> So far as the social sciences are concerned it can be laid down that literary productivity on the subject of poverty will appear in inverse ratio to the incidence of poverty in the 'group' to which the social scientists happen to 'belong'...While more and more of the billowing literature on poverty and race relations in the United States centers on the conditions of black persons in

> a white society, less and less of it is actually the work of Negro scholars. This must be repeated: less. It was not ever thus.

Moynihan continues,

> At the least, then, what social science very much needs is a considerable widening of its ethnic, social, religious and regional base. When social scientists observing a given milieu find that their judgments as to its qualities and characteristics are similar to or convergent with the judgments of other social scientists actually drawn from the milieu in question, we will be entitled to a greater order of confidence in the results.

The substance of the present report should be submitted to the judgment of black social scientists or other black observers with firsthand experience of the realities we have attempted to describe.

Assuming their favorable response, how, then, to proceed? In our present thinking, interviews with the parents could be the single most important source for the information sought here. Among topics to be covered are: the residential and social history of the parents (Q. 2)*; the children's dietary habits, food preferences and sleep habits (Q. 3); the parents' attitudes and beliefs about control and especially about the use of techniques other than physical punishment (Q. 4,5); diffusion of maternal responsibility (Q. 5); loci of objective dangers in the neighborhood (Q. 6); the child's emotional reactions and states (Q. 7-10); and prosocial attitudes prevalent in the family and neighborhood (Q. 14). The use of audiovisual aids and other props to relieve the tedium of prolonged questioning is recommended. Sympathetic, compatible, interviewers are obviously mandatory.

The next most important source of data would be field observations of the child in the school situation. These should be broad enough to include the adequacy of the child's clothing, his actual food preferences (Q. 3), and a sufficient density of variables to represent affective states (Q. 7-10), as well as adaptive strategies (Q. 11-14). Observation

---

*This, and subsequent numbers in parentheses, refers to the 15 questions posed earlier.

of the child's milieu, possibly with assistance from local informants, will be needed to help assess the special "atmospheric" qualities of the neighborhood (Q. 6) and the schools (Q. 15).

Objective and projective personality tests might be used to assess the presence of physiological debilitation (Q. 3); types of affect (Q. 7-10); and the child's investment in various adaptive strategies (Q. 11-14). Since none of these tests will be easy to design, they should probably be attempted only after the staff has the "feel" of the real situation. Ideas for test formats will then occur spontaneously.

There is no reason to treat the majority of these variables as part of the longitudinal assessment battery; many of them might be recorded once a year at most. They are not a substitute for existing longitudinal variables but an interpretively useful addition to them. They may contribute to the understanding of differences in educational achievement and personality growth over time in ways that will seem enormously useful when the staff considers the problem of writing a final report.

References

Eddy, E. M. Walk the white line. New York: Doubleday, 1967.

In this book a sociologist discusses an important problem facing schools in the urban slums; namely, the conflict between the middle class teacher's expectations concerning the ways in which school children should behave and the actual behavior of most ghetto children. According to Dr. Eddy, it is the lack of fit between these expectations and the existing reality that leads to most of the problems in the ghetto schools. Her book includes several detailed illustrations of the non-adaptive ways in which teachers react to children who do not behave as expected; reactions which include excessive punitiveness, withdrawal, and merely passing the time without teaching. As may be expected, such reactions are rarely associated with effective teaching with the consequence that little education actually goes on in many classes.

Freud, Anna. Special experiences of young children particularly in times of social disturbance. In K. Soddy (Ed.), Mental Health and Infant Development. New York: Basic Books, 1956.

Greene, M. F., & Ryan, O. The schoolchildren. New York: Random House, 1964.

This book (especially part two) gives one a clear cut impression of the great variety of personalities and interests of ghetto children. We learn that these children, although differing greatly in personality, share many problems: how to deal with the omnipresent dope addicts in the neighborhood; how to earn money without having it stolen from you; how to kill the rats in your apartment. The gravity of the everyday problems facing these children helps us to understand why they do not devote much attention to the frequently dull and irrelevant subject matter presented to them in school.

Hentoff, N. Our children are dying. New York: Viking Press, 1966.

This book is about Elliot Shapiro, principal of P.S. 119 in Harlem, who, unlike many white principals in ghetto schools, enjoys a good relationship with the surrounding community (as well as with the teachers and pupils in his school). Some of the reasons for his success are given in this book: he encourages his teachers to experiment and try out new teaching methods; he tries to meet and, if he can, to help the people in the community; he formed a special class for problem children; and he was, perhaps, the first in New York City to introduce African studies and the history of the blacks in America into the curriculum.

After Shapiro took over as principal, there was substantial improvement in the achievement level of the students in his school although the majority of the children were still reading below grade level. To push up their achievement still further, Shapiro outlines additional measures (most of which cost money) that must be taken.

Herndon, J. The way it spozed to be. New York: Simon and Schuster, 1965.

In this book, Mr. Herndon describes his experiences as a beginning teacher in a ghetto school in Oakland, California. He describes, in particular, his efforts to teach two "slow classes" whom he found totally unresponsive to the regular curriculum. In a desperate effort to teach something to these classes, Mr. Herndon tried out a number of different techniques, the most successful of which were inspired by the students themselves. His book contains a harsh criticism of those who would stick tenaciously to traditional teaching methods and curriculum even though they have proved to be dismal failures in the past.

Kohl, H. 36 children. New York: The New American Library, 1967.

Like James Herndon, Herbert Kohl found that the curriculum he was required to teach only succeeded in boring his students and making them restless and rebellious. After closely observing and getting to know his students and with some trial and error, Mr. Kohl was able to devise a curriculum that captured his students' interest and enthusiasm so much so that some of them actually arrived an hour early to class.

In response to his stimulating teaching, some of Kohl's students were inspired to create literature themselves. There are many examples of their original work in this book.

Kozol, J. Death at an early age. Boston: Houghton Mifflin, 1967.

The author of this book, a new teacher in a Boston ghetto school, focuses on describing the attitudes and prejudices of the white teachers in his school toward their black students. Kozol found that while many of these teachers profess liberal attitudes in conversation, their actual classroom behavior, in contrast, revealed an underlying belief in the basic inferiority of their black students.

Kozol also points out and gives examples of the prejudice that is embedded in many of the textbooks used in his school. In an effort to counteract this, he brought in books of his own to read to his class, many of them by black authors. As a result of passing out a mimeographed copy of a poem by one of these authors, Kozol was fired from his job.

Moynihan, D. P. The professors and the poor. Commentary, 1968, 46, No. 2, 19-28.

Moore, A. Realities of the urban classroom. New York: Anchor Books, 1967.

This book includes descriptions made by trained observers of the actual events in several first-fourth grade classes in selected ghetto schools. In the best of these classes, the events focus primarily on the material to be learned with both teacher and students actively participating in the process. In the worst classes, virtually no learning is taking place as the teacher's attention is primarily directed towards disciplining the students in her class. Unfortunately, most of the classes described in this book lean towards the negative rather than the positive end of the scale.

Initial Community Data on Trenton, N. J.

Robert Althauser and Ann Ryan

The sign that greets the passing motorist at the city line reads, "Entering Historic Trenton." The proclamation that greets the city's residents passing through the doors of their downtown police headquarters reads:

> Whereas I have declared a local emergency exists in the City of Trenton:
>
> 1) No person under the age of 18 shall remain in or upon the public streets (after) 10 p.m. Sunday through Thursday and (after) 11 p.m. on Saturday and Sunday.
>
> 2) Local police and other authorities are directed to take any and all measures requisite to secure areas within the City of Trenton to prevent and deter actual threat or harm to persons or property....

The curfew was imposed in April 1968, after the riot which followed the death of Martin Luther King, and remains in effect to this day.

. . . . .

A community can be described on different levels, ranging from the block or neighborhood in which a child or parent resides; to the level of a public school district or census tract; to larger but fewer subdivisions of the city by social class, racial and ethnic characteristics, and age distributions; to the level of the corporate whole. At this stage in our study, we will confine the first half of this discussion to the city as a whole and the latter half to the block and school district level. Later reports based on data subsequently available (e.g., a 1968 census, community observation and family interview data) should be able to extend this description of Trenton to the other levels mentioned.

The question which should guide and constrain the wealth of descriptive material available in this and subsequent reports will remain constant; namely, what kind of place is Trenton, for growing children, working adults and for the organizations that try to meet the educational, political, economic and welfare requirements of the whole population?

Table 1

TRENTON POPULATION: 1960-1968 TRENDS, AGE, & RACIAL COMPOSITION

Part I Population Group	1960 Population No. Persons	1960 Population % Total	1968 Population No. Persons	1968 Population % Total	% Change 1960-68
White Population:					
Under 5 yrs.	6,435	5.6	3,696	3.3	- 42.6
5-18 yrs.	14,374	12.9	9,144	8.2	- 36.4
19-44 yrs.	30,988	27.0	17,415	15.9	- 43.9
45 yrs. & over	34,715	30.3	32,727	29.6	- 5.7
Sub-total White Pop.	86,512	75.8	62,982	57.0	- 37.2
Nonwhite Population:					
Under 5 yrs.	3,718	3.3	7,526	6.8	+102.4
5-18 yrs.	7,445	6.5	11,959	10.8	+ 60.6
19-44 yrs.	10,075	8.8	14,915	13.5	+ 48.0
45 yrs. & over	4,614	4.0	7,948	7.2	+ 72.2
Sub-total Nonwhite Pop.	25,852	22.6	42,348	38.3	+ 63.8
Puerto Rican Population:					
Under 5 yrs.	144	0.1	512	0.5	+255.5
5-18 yrs.	273	0.2	968	0.9	+254.6
19-44 yrs.	663	0.6	2,392	2.2	+260.8
45 yrs. & over	723	0.7	1,245	1 1	+ 72.1
Sub-total Puerto Rican Pop.	1,803	1.6	5,117	4.7	+183.8
Total Trenton Population	114,167	100.0	110,447	100.0	- 3.3

Part II Age Summary All Groups	1960 Population No. Persons	1960 Population % Total	1968 Population No. Persons	1968 Population % Total	% Change 1960-68
Under 5 yrs.	10,297	9.0	11,734	10.6	+ 14.0
5-18 yrs.	22,092	19.4	22,071	20.0	0.0
19-44 yrs.	41,726	36.6	34,722	31.4	- 16.8
45 yrs. & over	40,052	35.0	41,920	38.0	+ 4.6
Total Trenton Population	114,167	100.0	110,447	100.0	- 3.3

The material discussed in part I is drawn largely from informal interviews with people concerned with the government or welfare of the city either as employees of various organizations or as volunteer workers. In part II we base our discussion on data from such sources as the 1960 Census, records from the Trenton Public Schools, and an Environmental Survey report from the Trenton Bureau of Environmental Sanitation, issued in 1967.

I

Like many other major U. S. cities, Trenton appears in the throes of uncertain social change and general decline. Its viability as a city that can attract and hold individual and corporate members is in grave question. Its problems are both aggravated and partially mitigated by its status as the capital of New Jersey. The state provides employment in its offices for many of Trenton's residents and thus contributes to the economic maintenance of the city as a whole. But the substantial amount of land on which its buildings sit is thereby removed from the tax rolls of the city, and the voluntary contribution that it makes in lieu of muncipal taxation pales in comparison, no doubt, to the tax payments which industry or commerce would make when holding the same amount of land.

The single most dramatic indicator of both change and decline, in a time of increasing population, is the decline in the population of Trenton. The 1950 Census figure was 128,008. In 1960 it dropped to 114,167. A 1968 Census figure lowers to a total of 102,000; another estimate suggests 110,000. Both 1968 figures are disputed, because of the serious effect that a drop below the 100,000 mark will have on funds available from the federal government. These figures and additional data re. trends, age and racial composition are displayed in Table 1, which was constructed by the Trenton Department of City Planning.

There are several trends apparent in this table.

1. The figures reflect an exodus of whites, some 37% of the 1960 population having departed by 1968, and an influx of blacks and, especially, Puerto Ricans. But these newcomers have been insufficient to balance the numbers of departing whites. More concretely, the rate of departure of whites is double the rate of entry of blacks and Puerto

Ricans combined. If we work with the 1968 estimate of 110,000 for the moment, and assume that the current rates of entry and exodus will continue for the next three years, the population will level off at about 104,000 in 1971, and then begin to increase. (This result is approximate, based on graphic extrapolation of the 1960 and 1968 data.)

2. There is also a marked change in the age distribution of white residents. The last column in Table 1 reflects this only in part, because the basis of the percentage change shown is the relative number of whites in 1960 and 1968 in each age group. If the proportions in each group are calculated on the base of the white population alone (rather than on the total population as shown), and then compared for 1960 and 1968, a 12% increase in the proportion of the white population that is 45 years old and over becomes apparent. Equally important is the 9 to 10 percent decrease in the proportion of whites in the wage-earning age group, 19-44. Even the overall percentage change figures at the bottom of the table reflect the immobility of older people and the exodus of wage earners. In addition, these same overall figures reflect a sharp increase in the population of children under age 5, coming entirely from the nonwhite and Pureto Rican groups.

The implications of this data are clear and disturbing. Let us assume that there were no wage or job discrimination and that the incoming nonwhite and Puerto Rican population could replace the departing whites in jobs comparable to those of the whites. Even with this assumption, the overall decline in that proportion of the population most capable of earning wages and paying taxes, coupled with the overall increase in the youngest and oldest proportions of the population (who tend to require more family and municipal support than they contribute) is a source of economic and institutional decline for the city. If we now discard our previous assumption and introduce the effects of minority-group status in the market place, the situation becomes even more serious.

There are other signs of white exodus and black and Puerto Rican influx. Ecologically, the settlement of the black population centered about Battle Monument--an institution in itself--and is expanding from the center of the city. The Puerto Rican influx is of very recent origin and its impact is more difficult to define. There is evidence of the

exodus of white churches, more often Protestant than not, into the suburbs in pursuit of their parishioners. Those churches that have not been sold, as most are, to black congregations, increasingly minister to a commuting membership that has moved outside the city limits. When the new campus of Mercer County Community College is completed outside the city limits, the city proper will be left with no institution of higher learning. Though not necessarily evidence of migration out of Trenton itself, the turnover rate of students in some elementary schools is as high as 70% in one school year.

The character and signs of increasing social instability can be broken down into more specific categories: economic, political, educational, law and order, and housing and welfare conditions.

Economically, the city (as are other cities like it) is in trouble. As is true nationally, the unemployment rates for black adult males are sometimes twice as great but always greater than for white males. The city's financial problems are sufficiently serious that Trenton may well lose its Moody's Triple A rating and no longer be considered as a completely responsible borrower through the issuance of bonds. Taxes that have risen by more than 200 points and an increase of 124 in the staff of City Hall have not helped the situation. Long-touted by still unimplemented plans for a downtown shopping mall were recently set back when the S. Klein department store chain pulled out as the prospective second major store in the mall (Trenton Times, 7/3/69).

Politically, Trenton reflects the increasing polarization of white and nonwhite that one associates with the spate of recent, country-wide municipal elections in which "law and order" has been the key issue and candidates supported by lower-middle class and ethnic groups have been the winners. City Hall is dominated largely by the Italian-American community, which helped elect the present mayor, Carmen Armenti, in 1966. The former official, Mayor Holland, was defeated primarily because of his decision to move to a home within the ghetto and also because a large part of the Italian, Polish, Hungarian and other ethnic communities felt that he was unconcerned with their problems. The black vote and that of the liberal West Ward, where our study's Jr. 3 district is situated, were not enough to swing a victory for him. The political future of Trenton will

apparently remain in the hands of the white ethnic (and more conservatively voting) population. These families remain settled in their present residences, or at least within their areas of the city, to a greater extent than the upper middle class and presumably more "liberal" residents of the West side. We will see in the later data that particularly in the Robbins district, which is strongly settled by Italian-Americans, the percentage of nonwhite elementary students has leveled off in the last three years instead of showing continuing sharp increases.

Political integration of the nonwhite groups would ordinarily involve membership on various governing boards and in the body of municipal employees, yet we determined with some difficulty that only 32, or about 10%, of the 292 members of the Trenton police department are black. None holds a rank above sergeant. There are no blacks on the City Council or on the Mercer County Board of Chosen Freeholders. This despite the figure of some 40% of the total population which is nonwhite.

In addition to the matter of proportionate nonwhite representation in the governing of the city is the municipal government's response to the needs of the nonwhite population. The infrequent appointment of nonwhites to staff positions, like the employment of one black by the Mayor's office with a salary paid by the State Department of Community Affairs, does little to meet these needs. One observation and one incident perhaps serve to illuminate the unsatisfactory political situation in Trenton at present.

Dotted on several streets in the Jr. 5 district are recently erected modest public parks. They consist of wooden benches embedded in concrete and new trees placed in large concrete pots. Located in vacant lots, they extend less than 50 feet from the sidewalk area into these lots. There seems to have been no observation of, or consultation with, the groups who would use these parks--groups which almost always include children. The facilities appear quite new, rather durable, and almost completely unused. Mothers and grandparents still tend to follow the children to spots where "there is something to play with," or to remain on their own front steps.

Trenton's black population has formed a Political Action Committee and frequently takes to the streets to demonstrate. The most recent incident involved the tragic accidental death of one of the children of an A&S interviewer at a dangerous intersection in the Grant school district.

Requests made over a period of years for a traffic light at this dangerous corner where children frequently play had been ignored. After the death of the child and following a mass sit-down in the street that blocked traffic for a time, a traffic light has been promised but has not at this writing been installed.

The city's educational institutions similarly reflect the broader population change and turmoil. Even in the three school districts included for study which nominally have substantial white population, the lowest 1968 figure for the number of students from nonwhite groups is 45%. The faculty at the six elementary schools is predominantly white. The minimum difference between the percentage of nonwhite students and the percentage of nonwhite staff in any of the schools is 38%, the maximum 45%. The effects of this disparity on the educational process are not precisely definable. There is some evidence that overcrowding and particularly, the administrative policies and attitudes of the schools' principals are more important factors in defining the educational environment. Enrollment and staff figures do not in themselves convey the average classroom size. The teacher-pupil ratios for 1968 in the six elementary schools in our study all fall within the range of 1:23 to 1:27. The average classroom population is probably larger than this in most schools. There appears to be great variation in the manner in which school administrations and principals handle student "unrest." On the one hand, in one elementary school not in the study area the principal holds weekly meetings with class representatives to hear their complaints and suggestions. On the other, unassuaged racial tension in the high school helped contribute to the 1968 Trenton riot. There had been an explosive cafeteria free-for-all in December of 1967. The black students felt they were not listened to and not appointed in reasonable proportions to student-faculty committees. Neither their grievances nor the previous eruption had been explored or understood, when a black student leader requested a dismissal of classes for a day, to mourn the death of Martin Luther King. The administration hedged, indecisive and uncertain. The black students marched angrily and loudly out of the school, toward Battle Monument, breaking windows on the way. In the evening of the same day, the riot began at the Monument.

There is now a black Superintendent of Schools, but the high school continues to be a source of community-wide tension as evidenced by the trial of Reverend John Young, who is charged with entering the school last March with the intention of disrupting classes.

The maintenance of order and the detection and prosecution of crime are salient and controversial features of contemporary urban communities. Recent articles in the Trenton Times (6/25/69; 6/29/69) provide an unusually complete summary of Trenton's crime rates and the Police Department.

According to these articles, fewer than 10% of the property crimes in 1968 were cleared; investigated, that is, to the point of successful prosecution. There were 2,500 breaking and entering crimes (an increase of 23%); 1,081 larcenies (+33.6%); and 1,146 auto thefts (a decrease of 19.9%). The latter is a reflection of a city ordinance making it illegal to leave keys in the ignition of parked cars and the continuing curfew following the April 9, 1968 riot.

Crimes of violence also rose sharply: 16 murders (an increase of 23.1%); 512 robberies (30.9%); and 229 aggravated assaults (21.2%). One article reviews a survey of the police department by the N. J. State Law Enforcement Agency. This survey emphasized, in addition to the previous data, the seriousness of the ever-increasing amount of lesser crimes--non-atrocious assaults and larcenies of less than $50. Such crimes do not enter into the crime totals prepared by the state or F.B.I.

This same survey pointed up the need for more police and better training of those now on the force. It was noted that after a new policeman receives his basic training, there is no follow-up training except in firearms instruction. Of particular importance for community relations is the neglect in dissemination of information concerning recent court rulings on the rights of the accused. Such information as is distributed among the force has to be acquired by word-of-mouth and in locker-room conversation. The article dealing with this survey concludes with these comments: "Notwithstanding some of the shortcomings referred to, the Trenton department has a nucleus of capable, dedicated, well-trained and perceptive leaders and officers." It notes that the department has been

"...victimized by citizen apathy and political leadership that has failed to provide the tools, training and money necessary to supply efficient levels of service to the community."

It should not be surprising, in addition to the difficulties within the Police Department, to find that the police officers are caught up in the general white exodus. Indeed, the Policeman's Benevolent Association is currently appealing a city ordinance requiring city employees to live within the city limits. The City Council is reported as "wavering" on enforcement of this ordinance. Setting the ordinance aside would result in a decrease in the pressure to find qualified police officers from the inner city. On the other hand, there would also be less internal pressure to hire black officers and the loss of whatever advantage to community and racial harmony that may accrue from having officers that, because of residency, have a personal stake in the welfare of the community.

Not mentioned, so far, is the influence of organized crime in the life of Trenton. Some individual members of the department "opined" to one of the authors that there was little or no organized crime in the city, on virtually the same day as a Trenton Times article (7/2/69) appeared with the banner: "The Mob Treats Trenton 'Like One of the Family'." "Four of the nation's 24 Cosa Nostra families conduct illegal activities in the Trenton area, according to federal law enforcement agencies," the article begins. The impact of La Cosa Nostra on the life of the city is quite large in all probability, but, as law enforcement agencies know, the solid, court-worthy evidence is hard to come by.

Evidence pertaining to housing and welfare conditions suggests the existence of problems as serious as those previously discussed. There are claims that some of the worst slums are gone, that new state buildings replacing decrepit housing and several public housing developments represent an overall improvement in the situation. Nonetheless, there are vast stretches of rubble and ruins, a housing shortage that is acute, and increasing costs for inadequate housing. The usual rent in some areas is $90 a month for a sleeping room with no kitchen facilities and a shared or inadequate bath. One room with kitchen and bath usually costs about $120 in the ghetto.

As in other large cities, Trenton's residents are sometimes the prey of landlords who are anxious to keep the housing code from being enforced. The usual fine for violations of this code--$25--is considerably less than the probable average cost for relieving the conditions involved. The Welfare Department finds it hard to place families in housing, because the law requires that a dwelling be inspected before a family moves in. Landlords are sometimes reluctant to submit to this inspection.

It was charged by one of our informants that the County Welfare Board, headed by a political appointee, is extremely unsympathetic to the needs of poor people. The informant (white) runs an organization which maintains the only licensed day-care center in the city and which acts as an advocate for families and individuals in need of welfare aid. According to this informant, it is difficult without an advocate to obtain a telephone for a welfare recipient with a weak heart, or an humidifier or air conditioner for an asthma victim.

There is, in addition, a high turnover on the Welfare Board. Caseworkers receive low pay, have large case loads and little time to help a family get off welfare. There are also carefully documented cases which show that individuals have actually died prematurely or needlessly due to the slow response of the board. As is commonly complained by the poor in other cities, the County Board often fails to inform a working mother of certain kinds of assistance (e.g., baby-sitting) that would enable her to find work and to continue working.

II

The kind of material presented above is considerably more adequate as a means of understanding the broader problems facing Trenton and, perhaps, of anticipating the shape of the data which this study will yield than as a complete and accurate factual description of Trenton. The latter goal is impossible by means of interviews and observations short of an enormous amount of systematic survey work. It would require the examination of census and related data, where careful interpretation and adequate statistical manipulation can yield a more complete description. Of course, an

understanding of the patterns inherent in such descriptions is more elusive here than in material from interviews and observations. Such are the limitations of such data, to which we now turn.

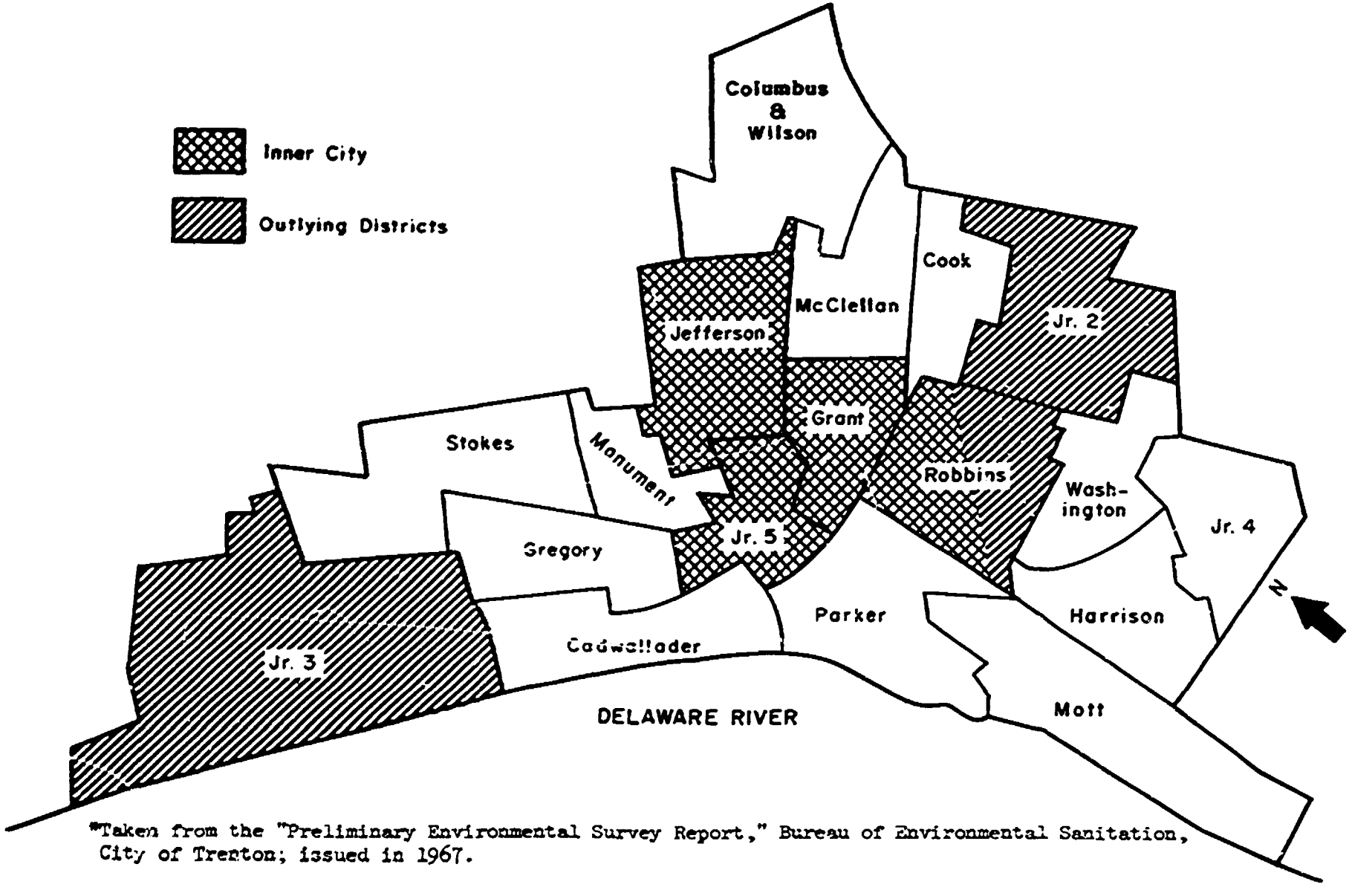
Descriptive statistical data of the sort subsequently discussed has certain problems which should be briefly noted. There is always the problem of data becoming out of date in the event of rapid change of the sort witnessed in Trenton. In some respects, the 1960 census data to be presented is primarily valuable as a baseline for 1968 census data largely unavailable at the time this report goes to press and for the Public School and Environmental Survey data also presented here. Equally troubling is the perennial lack of commonality in the boundaries of the ecological units in our study--Public School districts--and the census tract areas delineated by the U. S. Census. Sometimes it happens that there is great commonality, say between census tracts and political precincts or other units of interest. There is a serious absence of commonality here between census tract boundaries and school district boundaries. Since census data by census tracts is extremely detailed and varied, this is a serious loss of information useful in describing the study areas.

However, there is a limited amount of information--largely pertaining to housing conditions, the rental or ownership of dwelling units--which is available by individual blocks. The census data presented in Table 2 was developed by determining the census tract and block numbers of all blocks within each school district and summing up the information on blocks, often across several tracts, within each district. (See page 268.)

In addition, we face a problem in presenting a discussion of such data. Descriptive studies thrive on comparisons, but comparisons often require information on units not immediately under study. The greater the comparability of descriptive data on units of direct interest to other, larger units of reference, the more information that data can convey to the reader. Least desirable but easiest to come by, then, is simply an internal comparison of data by school districts. More desirable is an internal comparison supplemented by figures for the whole of Trenton or for some subdivision of Trenton larger than a school district. On the same level of desirability are comparisons involving all school

Figure 1

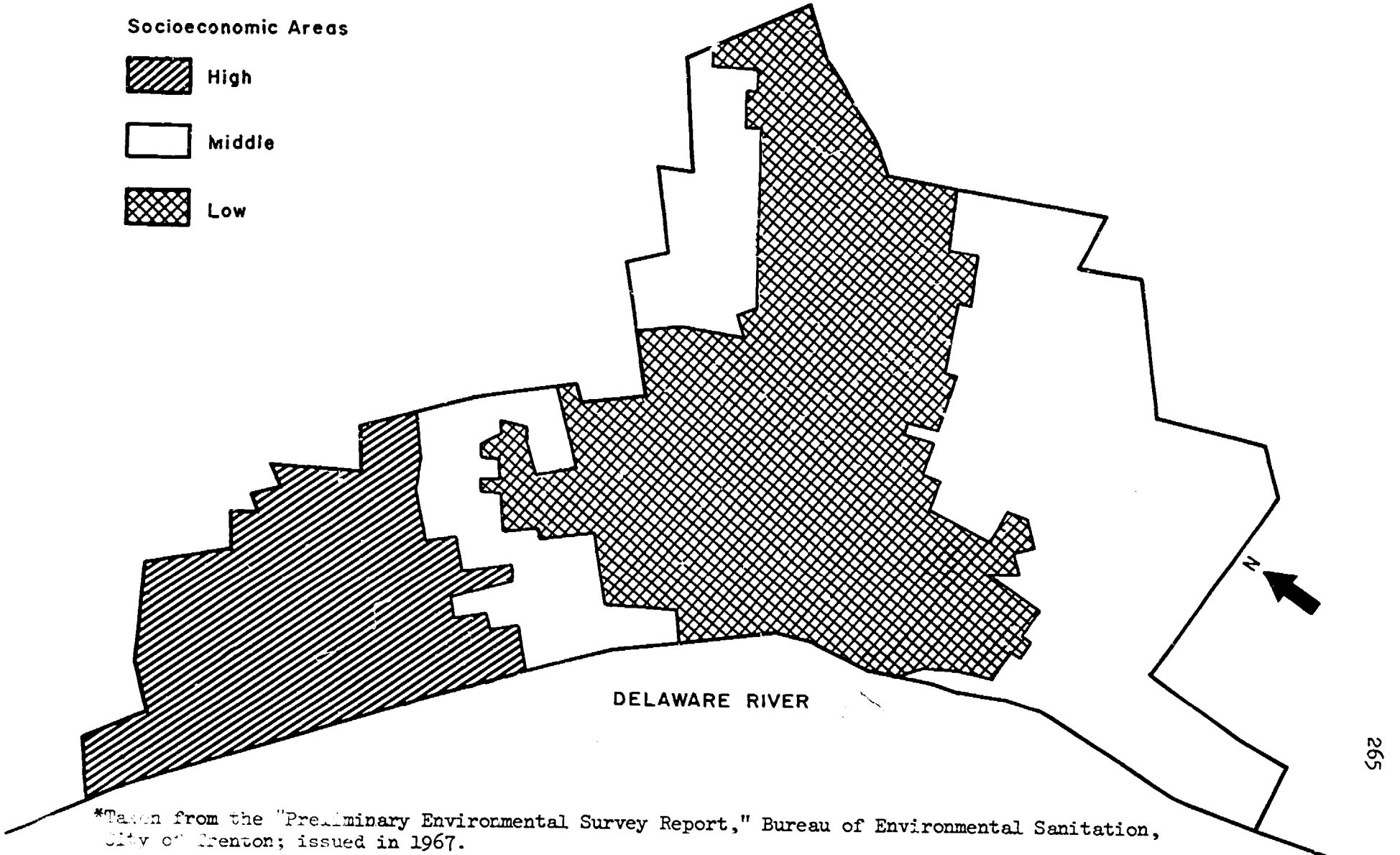
TRENTON, N.J.: PUBLIC SCHOOL DISTRICTS* WITH THOSE INCLUDED IN THE STUDY INDICATED BY SHADING.



*Taken from the "Preliminary Environmental Survey Report," Bureau of Environmental Sanitation, City of Trenton; issued in 1967.

Figure 2

TRENTON, N.J.: LOCATION OF THE HIGH, MIDDLE, AND LOW SOCIOECONOMIC AREAS IN 1965.*



*Taken from the "Preliminary Environmental Survey Report," Bureau of Environmental Sanitation, City of Trenton; issued in 1967.

districts, without regard to their inclusion in our study area. Internal comparisons with some totals for the whole of Trenton are as far as the present discussion goes. Even more desirable types of comparisons toward which future reports might point will be noted at the conclusion of this section.

Before launching into a discussion of statistical data on the school districts under study in Trenton, the location and approximate socioeconomic level of each can be viewed in Figures 1 and 2. On Figure 1, the districts under study are shaded. The darker shading indicates the districts--Grant, Jefferson, Jr. 5, and part of Robbins which lie essentially in the center of Trenton and are subsequently referred to as the inner city districts. The lighter shading indicates Jr. 2, Jr. 3, and the balance of Robbins which may be considered outlying districts and will be referred to as such. In part of the statistical comparisons among districts to come, the inner city and outlying districts display consistent differences.

By comparing Figures 1 and 2, a rough notion of the average socioeconomic level of the inner city and outlying districts can be had. Jr. 3 lies in the city's only "high" socioeconomic area, according to the Environmental Survey of the Bureau of Environmental Sanitation. (The survey was conducted in 1965 and published in 1967. No mention of criteria for the assignment of socioeconomic level is made.) Jr. 2 and about two-thirds of Robbins lie in the "middle" socioeconomic range, and the remaining districts--the inner city group--lie in the "low" socioeconomic areas.

As of 1960, nonwhites were concentrated in the inner city districts and the northwestern boundary of Robbins district. The Italian-American population, the largest of the ethnic groups living in Trenton, was heavily concentrated in Robbins and, to a lesser extent, in Jr. 2 and the northwestern half of Jefferson. Natives or descendants from the U.S.S.R., many of them probably Jewish, outnumbered the Italian-Americans in Jr. 3.

The figures for the location of nonwhites in occupied dwelling units in 1960 are shown in column 1 of Table 2 ("Some Characteristics of the Study's Sample Areas"). There is no doubt that these figures are sadly out of date and useful to us primarily as baseline figures for later data.

Note that these are percentages of dwelling units occupied by nonwhite, not percentage of the population that is nonwhite. Given the tendency for nonwhites to be renters, for larger nonwhite families to occupy fewer rooms and so occupy them more densely, these figures may be taken as conservative estimates, indeed the lower limits of the percentage of the population that is nonwhite. Note, too, that the inner city districts display percentages of nonwhite occupancy above the Trenton mean percentage (18%) for 1960; the outlying districts show percentages below the city average.

More recent data in Table 2 suggests inferences about the present distribution of nonwhite population, but it is data from the public schools. The relationships between nonwhite student enrollment, nonwhite occupancy of housing and nonwhite population figures are far from being isomorphic. It is useful, therefore, to compare the percentage of nonwhite occupancy and nonwhite student enrollment in 1960, for which year both figures are available. As can be seen by comparing the data in columns 1 and 2, it is usually the case (Robbins and the three inner city districts) for the percentage of nonwhite students enrolled to run far ahead of the percentage of nonwhites occupying dwelling units. No doubt, white parents may be sending their children to private or parochial schools, or they may be older and thus no longer sending children to school to the same extent as their younger, nonwhite, fellow residents.

The changes over the last 8 years in the percentage of nonwhite students enrolled in the elementary school in each district can be seen by comparing columns 2 and 3. For the inner city districts with already high percentages of nonwhite residents, there are only slight increases. The increase is also modest in Robbins, but among the remaining outlying districts the increase in nonwhites enrolled is dramatic. The changes in these figures over a longer period, from 1950 to the present, are shown in Figure 3. Note that while the increase in nonwhites in Jr. 2 and Jr. 3 is of more recent origin, the rate of increase in these districts is greater even than the rate of increase in the percentages for the inner city districts. Robbins, on the other hand, appears to be actually leveling off in the last three years. The movement of nonwhites into this district appears to have been firmly halted for the present.

Table 2

SOME CHARACTERISTICS OF THE TRENTON SAMPLE AREAS:
POPULATION, HOUSING OCCUPATION AND CONDITION, RACIAL COMPOSITION OF HOUSING,
ELEMENTARY SCHOOL ENROLLMENT AND ELEMENTARY SCHOOL TEACHING STAFF, AND ENVIRONMENTAL QUALITY

Public school district	Population[a]	(1) % of housing units occupied by nonwhite[a]	(2) 1960-61- % of students enrolled nonwhite[b]	(3) 1968-9- % of students enrolled nonwhite[b]	(4) 1968-9 % of first graders enrolled nonwhite[c]	(5) 1968-9 % of black teachers on staff[b]	(6) Characteristics of housing units 1960 % Sound	(7) % Dilapid[d]	(8) % Rented[e]	(9) 1965 % housing in good condition[f]	(10) 1965 ranking[f,g] among all 18 districts
Junior 2	9,510	0	1	45	45	6	93	0	20	81	7
Junior 3	8,762	2	11	66	66	21	100	0	34	96	1
Robbins	8,301	14	51	66	N.A.	24	82	6	53	33	5
Grant	4,659	42	92	95	96	45	55	12	78	12	22
Jefferson	8,194	36	82	96	95	53	68	9	45	9	18
Junior 5	4,109	56	90	98	100	60	92	12	73	7	14
TRENTON TOTALS:	114,167	18					84	4	42	43	

[a] From U. S. Census Block Data, 1960

[b] From Public Schools Records

[c] Memo from Albert E. Beaton, February 18, 1969

[d] Percentage of housing units in a "dilapidated condition"

[e] Percentage of housing units renter occupied.

[f] From Preliminary Environmental Survey Report, Bureau of Environmental Sanitation, Trenton, January, 1967

[g] The rank order of these districts among all 18 public school districts with respect to ten criteria: housing condition, refuse storage, rubble, unstacked lumber, abandoned autos, dilapidated sheds, fences, dogs, mosquito breeding sites, and weeds.

There would be some reason to expect that the racial composition of the first and second grades in these elementary schools might be proportionately more nonwhite than for the schools as a whole. The birth rate for the population in the lower socioeconomic areas where blacks are primarily situated is higher than for the other areas in Trenton, as elsewhere. However, a comparison of the figures in columns 3 and 4 shows virtually identical distributions by race in the first grade and in all grades, in 1968. It is the earlier grades, of course, whose racial and socioeconomic composition will define the environment most salient to the children in our study when they enter school.

Another salient feature of this environment for the black child is the proportion of the teaching staff that is black. Examination of the figures in column 5 suggests that about 50% of the teaching staffs from the inner city schools are black, in comparison to the over 90% nonwhite composition of the student bodies there. In the outlying districts these percentages of nonwhite staff and students are both smaller, but the same pattern prevails of greater proportions of nonwhite students relative to the proportion of black staff.

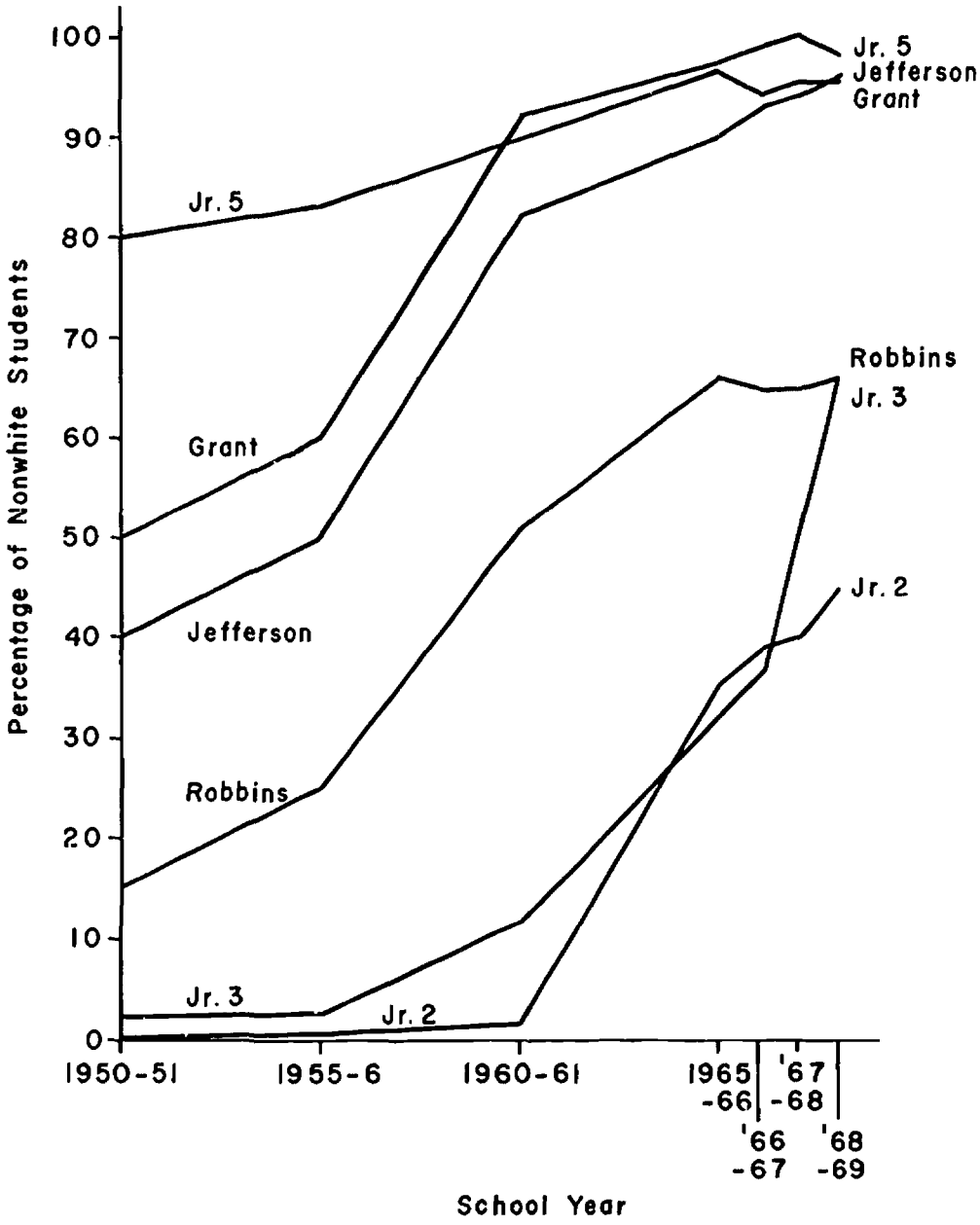
With respect to housing conditions, virtually the same pattern of differences between inner city and outlying districts previously observed again obtains. As the figures in columns 6 and 7 show, housing that is rated "sound" is considerably more prevalent in the outlying districts and above the rather high city average of 84% sound in two of the three outlying districts. Likewise the proportion "sound" is lower and the proportion "dilapidated"--the lowest of the three ratings of condition--greater than the city averages for the inner city districts.

Whether or not a dwelling unit is rented or owned is an important feature in the lives of parents and children: to depend upon a landlord more or less absentee, more or less responsive to the needs of the tenants and the statutes of the housing code; to choose how often and where else to move in search of better conditions. The figures in column 8 showing the percentage of occupied dwelling units which are rented are seemingly inconsistent with the previous differentiation of patterns for inner city and outlying districts. Thus there is a higher percentage of units rented

Figure 3

TRENTON, N.J.: PERCENTAGE OF NONWHITE ENROLLMENT IN THE SIX PUBLIC SCHOOL STUDY DISTRICTS.

1950-1968



in Robbins than in Jefferson in 1960. Nonetheless, the rank order correlation between districts ranked by percentage sound and percentage rented is virtually 1.00. Thus, these figures are not as inconsistent as would first seem. As rented housing traditionally tends to be less sound than self-owned, we would expect this sort of consistency.

In column 9 we show more recent (1965) data on the condition of housing in these districts. The differences between these figures and the census data in column 6 probably reflect two factors: possible changes in conditions in the five-year interval between them; and a more demanding application of the usual criteria of housing conditions. Differences between percentage sound and percentage in "good" condition in districts Jr. 2 and 3 are slight; more dramatic are the remaining districts, all of which had lower percentages of "sound" housing in 1960 than did the first two districts just mentioned. It seems likely that the modest differences in Jr. 2 and Jr. 3 reflect actual change in housing conditions over the interval, while the more dramatic differences reflect both considerable change and systematically more critical application of the usual standards. From the perspective of the later but more critical Environmental Study survey, it would seem that the 1960 census figures for percentage sound in the middle socioeconomic district of Robbins and the low socioeconomic areas of Jefferson, Grant and Jr. 5 overestimated the sound condition of housing in these four areas. Such an interpretation is based on the premise that unreliable ratings of condition would tend to cluster in those lower socioeconomic districts where ratings are more ambiguous.

Finally, column 10 displays the rankings of our six districts in comparison with all 18 city districts with respect to some ten criteria: housing condition, adequacy of refuse storage, the presence of rubble, unstacked lumber, abandoned automobiles, dilapidated sheds and fences, numerous dogs, weeds, and mosquito breeding sites. It can be seen that our sample includes both the "best" and "worst" environmentally rated districts in Trenton. The remainder of the districts are nicely distributed between these two extremes.

This statistical data collected from diverse sources has provided an initial and somewhat modest comparison of the housing and population

characteristics of the six school districts in our study with one another and, at times, with the city as a whole. The one emerging pattern which will probably be reinforced when more recent census and survey data becomes available has been the systematic differentiation of the inner city and outlying districts, with the former displaying poorer housing and other environmental conditons, together with above average percentages of non-white school enrollment and settlement. It is a familiar pattern for American cities; this data simply confirms its existence in Trenton.

But such description remains at relatively macroscopic levels, somewhat distant from the level of the individual block or of areas within school districts. Data from large ecological units like school districts remains vitally important, especially when similar data for smaller neighborhood or even family units is greatly inconsistent with the former data. One can imagine four quite distinct possibilities, say, with respect to housing conditions. We may study families living in blocks whose housing is in sound condition, where those blocks are situated within larger school districts whose housing on the whole is also sound. Likewise we will quite frequently deal with children from homes in a block of primarily deteriorated or dilapidated homes located in districts where housing is, overall, in similarly poor condition.

Of considerable interest and worthy of subsequent study are the possibilities just noted plus inconsistent combinations of sound housing conditions on the block level where the blocks are located in districts with generally poor housing or poor housing on blocks in districts with generally good housing conditions. Factors like housing conditions can operate on more than one level--here, the block and district levels. Their effects on the parents and children in our study may be consistent or quite possibly inconsistent, as we move from district to block, and perhaps to family unit levels. Only by studying all four combinations mentioned can the effects (known as "structural" or "compositional") of each level be distinguished. Such considerations as these are not only a program for the data analysis of community variables in relation to the other types of variables in this study, but they are also a worthy standard against which to hold later descriptions of the cities under study.

APPENDIX C

The ETS-Head Start Longitudinal Study and the "Westinghouse" Study

Samuel Ball

THE ETS-HEAD START LONGITUDINAL STUDY AND THE "WESTINGHOUSE" STUDY

Samuel Ball

A major upsurge of interest in preschool education began in 1964 when the federal government initiated the Head Start program. During the early years of Head Start, a major emphasis was on program development. When this first phase had become operational, a second important consideration--evaluation--became paramount.

In 1967, the Office of Economic Opportunity funded two major evaluational research studies on the effects of Head Start. One, popularly termed the Westinghouse study (1969), was an attempt to assess, as quickly as possible, the impact of Head Start. The second, the ETS-Head Start Longitudinal Study of Disadvantaged Children, was to be more comprehensive in scope and rigorous in design, and was to involve more extensive planning.

The Westinghouse Study

This study attempted to measure the extent to which first-, second-, and third-grade children who had attended Head Start preschool programs differed "in their intellectual and sociopersonal development from comparable children who did not attend." A major methodological problem not satisfactorily solved was that of ensuring the comparability of the control (non-Head Start) children. Such difficulties are inherent in studies whenever the assessment of a program's impact on children occurs long after the event has taken place.

The Westinghouse study also suffered in that it lacked breadth. This was candidly admitted in the summary to the published report which states:

> The very real limitation of our study should be established at once. The study did not address the question of Head Start's medical or nutritional impact. It did not measure its effect on the stability of family life. It did not assess its impact on the total community, or the schools, or the morale and attitudes of the children while they were in the program. It is therefore a limited and partial evaluation...

Again, of course, it should be emphasized that the study was undertaken with short notice and was carried through under strict time limitations.

Perhaps the most serious problem with the Westinghouse study (again one inherent in its "backward look" design) was the global treatment given both the Head Start program and the children who passed through the program. Actually, Head Start is a generic term covering a vast range of preschool classrooms including traditional Montessori, permissive English-style nursery classes, and eclectic modern American approaches. Yet, all of these were lumped together in the evaluation of the Head Start treatment. Similarly the children going to Head Start classrooms vary greatly in their backgrounds and characteristics. It would be surprising if they all were to benefit on the same dimensions to the same degree. A more sophisticated expectation would be that some Head Start programs benefit some sorts of children in certain ways, while other Head Start programs benefit different sorts of children in different ways. The Westinghouse report recognizes that this more sophisticated approach would be worth following but adds that it was "outside the scope of this study."

The Westinghouse study was an attempt to provide, in a relatively short time, some answers to some very broad questions concerning the merit of Head Start. Unfortunately, a proper understanding of a major research and assessment problem cannot be obtained without a great investment of time; and time was lacking in the Westinghouse study.

The ETS-Head Start Longitudinal Study of Disadvantaged Children

In the summer of 1969, ETS began seeking the 3 1/2-year-old children in four sites across the country who would become part of the six-year longitudinal study. These children have been measured on a large number of relevant variables before they experience such educational programs as Head Start. Of course, some of the children will later go to Head Start programs and some will not. All children will then be measured again. By this design, the impact of Head Start can clearly be assessed.

During each of its six years, the ETS study will obtain measures not only of the children's cognitive and personal-social development but also

of their physical health and physical growth. As well, information will be obtained about their families, communities, classrooms (preschools and regular schools), and about the educators with whom they come in contact. The impact of Head Start will be assessed, therefore, not only by its direct effects on children, but also through the way Head Start interacts with families, communities, and the school system. Never before has such a wealth of data been collected on such a large sample of children. The study will make available comprehensive answers to questions concerning the variety of Head Start programs and the direct and indirect effects they have on various types of children. Also, a great deal of information is bound to be uncovered about the growth and development of disadvantaged children and about their families, schools, and communities. Such information should prove useful to those interested in educational and social planning.

Reference

Westinghouse Learning Corporation & Ohio University. The impact of Head Start. An evaluation of the effects of Head Start experience on children's cognitive and affective development. April 1969, Contract OEO B89-4536, Office of Economic Opportunity.

APPENDIX D

PROJECT PERSONNEL

Steering Committee: Samuel J. Messick (Chairman), Albert E. Beaton, Walter Emmerich, Edmund W. Gordon, Winton H. Manning, Marshall P. Smith, Silva S. Tomkins, Melvin Tumin

Project Director: Scarvia B. Anderson
- Associate for Project Direction: Samuel Ball
- Assistant for Project Direction: Ann Jungeblut
- Assistant for Communications: Gray Sidwell
- Financial Consultants: Joseph Bianchini, Patricia E. Clyde
- Contract Consultants: Russell W. Martin, Charlotte Farley
- Public Relations Consultants: Thomas M. White, John P. Smith

Program Coordinator: Joseph L. Boyd, Jr.
- Associate for Field Coordination: Samuel Barnett
- Assistant for Program Coordination: Virginia Tillett
- Coordinator, Lee County, Alabama: Lida Campbell
  - Assistant: Mary Jo Frazier
  - Consultants and Monitors: Junius A. Davis, Roderick Ironside, Yvonne M. Rutledge
- Coordinator, St. Louis, Missouri: Bobby Westbrooks (acting)
  - Assistant: Flavious C. Johnson, Jr.
  - Consultants and Monitors: Chandra M. N. Mehrotra, Daniel P. Norton
- Coordinator, Portland, Oregon: Verna Shepherd
  - Assistant: Maxine Baker
  - Consultants and Monitors: Santelia A. Knight, Robert E. Lambert, George Temp
- Coordinator, Trenton, New Jersey: Conrad McLean
  - Assistant: Roberta K. Fitzgerald
  - Consultants and Monitors: Anne M. Bussis, Jean Orost

Measurement-Training Coordinator: Virginia Shipman
- Associate for Measurement Development: Ruth B. Ekstrom
- Associate for Measurement Operations:
- Assistant for Measurement-Training: Judith Grisham
- Professional Associates for Measurement-Training: Dolores Ahrens, Robert Althauser, Marianne Amarel, Nancy Ames, Scarvia B. Anderson, Samuel Ball, Gerald Bracey, Anne M. Bussis, Edward Chittenden, Rosalea Courtney, Diran Dermen, Walter Emmerich, Karla S. Goldman, George Gordon, Frances Green, Irwin Hyman, Sandra Landes, Michael Lewis, Carolyn Massad, Donald M. Medley, Jean Orost, Michael Rosenfeld, Esther Ann Ryan, Masako N. Tanaka, George Temp, Phyllis Ward, William C. Ward, Gita Wilder

Analysis Coordinator: Albert E. Beaton
- Professional Associates for Analysis, Data Handling: F. Reid Creech, Charles Hall, Harry Harman, Richard J. Martz, Frederick Toms, Jesse H. Webb, Victor E. Wichert